XXVIII

WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

LIBRO DE ACTAS

Ciudad de Buenos Aires – 9 y 10 de abril de 2026

UAI Universidad Abierta Interamericana

Universidad Abierta Interamericana Tecnología Informática UAI

RedUNCI

De Vincenzi, Marcelo
Workshop de investigadores en ciencias de la computación WICC 2026 : libro de actas / Marcelo De Vincenzi ; Compilación de Marcelo De Vincenzi. - 1a ed. - Ciudad Autónoma de Buenos Aires : Fundación Iberoamericana de Estudios Superiores, 2026.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-631-6904-02-7

1. Computación. 2. Tecnología Informática. 3. Tesis. I. De Vincenzi, Marcelo, comp. II. Título.
CDD 004



AUTORIDADES

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

AUTORIDADES UNIVERSIDAD ABIERTA INTERAMERICANA

Facultad de Tecnología Informática

Decano
Dr. Marcelo Edgardo De Vincenzi Zemborain

Vicedecano
Dr. Carlos Neil

Directora del Centro de Altos Estudios de Tecnología Informática
Dra. Claudia Pons

Director de la carrera de Ingeniería Bs As
Dr. Nicolás Battaglia

Directora de la carrera de Ingeniería Rosario
Dra. Silvia Poncio

AUTORIDADES REDUNCI

Coordinador Titular
Pesado Patricia (UNLP)

Coordinador Alterno
Eterovic Jorge (UNLaM)

Junta Directiva
Tugnarelli Mónica (UNER)
Panizzi Marisa (UNAHUR)
Aciti Claudio (UNICEN)
Brachetta Mariana (UChampagnat)
Larrea Martín (UNS)
Alfonso Hugo (UNLPam)
Guerrero Roberto (UNSL)
De Vincenzi Marcelo (UAI)

Miembro Honorario
De Giusti Armando (UNLP)

Secretarías
Secretaría Administrativa: Mariño Sonia (UNNE)
– Prosecretario Administrativo: Caballero Sergio (UNaM)
Secretaría Académica: Russo Claudia (UNNOBA)
– Prosecretaria Academica: Haderne Marisa (UNCUYO)
Secretaría de Ciencia y Técnica: Rodríguez Nelson (UNSJ)

– Prosecretaria de Ciencia y Técnica: Arroyo Marcelo (UNRC)
Secretaría de Asuntos Reglamentarios: Alfonso Hugo (UNLaPam)
Secretaría de Vinculación Tecnológica y Profesional: Bursztyn Andrés (UTN)
Secretaría de Congresos, Publicaciones y Difusión: Thomas Pablo (UNLP)
– Prosecretaría de Difusión: Malbernat Lucía (CAECE)
– Prosecretaría de Congresos WICC 2026: Haderne Marisa (UNCuyo)
– Prosecretaría de Congresos TEYET 2026: Mayer Roberto (UNdAv)
– Prosecretaría de Congresos CACIC 2026: Lugani Carlos (UNRN)

COMITÉ ACADÉMICO

UBA – Cs. Exactas	Brusco, Pablo Brusco
UN La Plata	Pesado, Patricia
UN Sur	Larrea, Martin
UN San Luis	Printista, Marcela
UNCPBA	Aciti, Claudio
UN Comahue	Grosso, Guillermo
UN La Matanza	Eterovic, Jorge
UN La Pampa	Alfonso, Hugo
UN Tierra del Fuego	Koremblit, Gabriel
UN Salta	Gil, Gustavo
UN Patagonia Austral	Gaetan, Gabriela
UN San Juan	Rodriguez, Nelson
UADER	Mengarelli, Jose Luis
UN Patagonia SJB	Buckle, Carlos
UN Entre Ríos	Tugnarelli, Mónica
UN Nordeste	Mariño, Sonia
UN Rosario	Casali, Ana
UN Misiones	Caballero, Sergio
UNNOBA	Russo, Claudia
UN Chilecito	Carmona, Fernanda
UN Lanús	Azcurra, Diego
UN Santiago del Estero	Figueroa, Liliana
Universidad de la Defensa	Cipriano, Marcelo
UN Litoral	Loyarte, Horacio
UN Rio Cuarto	Arroyo, Marcelo
UN Córdoba	Fridlender, Daniel
UN Jujuy	Herrera Cognetta, Analía
UN Rio Negro	Lugani, Carlos Fabian
UN Villa María	Prato, Laura

UN Lujan	Kaplan, Gladys
UN Catamarca	Poliche Maria Valeria
UN La Rioja	Molina, Miguel
UN Tres de Febrero	Oliveros, Alejandro
UN Tucumán	Luccioni, Griselda María
UNAJ	Morales, Martín
UN Chaco Austral	Zachman, Patricia
UN del Oeste	Massacesi, Alfredo
UN de Cuyo	Haderne, Marisa
UN deMardelPlata	Ríos, Carlos
UN deQuilmes	Pari, Denise
UN Hurlingham	Panizzi, Marisa
UNSAdA	Ramón, Hugo
UNSAM	Estayno, Marcelo
UNdAV	Mayer, Roberto
UNLZ	Serra, Diego
UTN	Bursztyn, Andres
UNGS	Fernandez, Enrique
U Morón	Chapperon, Gabriela
UAI	De Vincenzi, Marcelo
U Belgrano	Aguilera, Sergio
U kennedy	Albarracín, Mario
U Adventista del Plata	Bournissen Juan
UCAECE	Malbernat, Lucía
U Palermo	Alvarez Adriana
UCA Rosario	Grieco, Sebastián
U Salvador	Zanitti, Marcelo
U Aconcagua	Giordanini, Osvaldo
U Gastón Dachary	Ruidías, Hector Javier
UADE	Fernandez, Mary Auxiliadora
UCEMA	Guglianone, Ariadna
U Austral	Cosentino, Juan Pablo
U Atlántida Argentina	Rathmann, Liliana
UCA La Plata	Bertone, Rodolfo
ITBA	Bolo, Mario
U Champagnat	Brachetta, Mariana
U Mendoza	Asensio, Mariela
UMET	Aguiar, Costanza
Universidad de Chubut	Casco, Julio Cesar
Universidad Siglo 21	Conde, Sergio

COORDINADORES ÁREAS WICC 2026

Agentes y Sistemas Inteligentes

Marcelo Falappa (UNS)
Marcelo Errecalde (UNSL)
Franco Ronchetti (UNLP)
Daniel Pandolfi (UNPA)

Arquitectura, Redes y Sistemas Operativos

Luis Marrone (UNLP)
Orlando Micolini (UNC)
Alejandro Stankevicius (UNS)
Osvaldo Marianetti (UMendoza)

Computación Gráfica, Imágenes y Visualización

Martín Larrea (UNS)
María J. Abásolo (UNLP)
Juan Manuel Fernandez (UNLu)

Ingeniería de Software

Pablo Fillottrani (UNS)
Pablo Thomas (UNLP)
Fernanda Carmona (UNdeC)
Patricia Cristaldo (UTN FRCU)

Innovación en Sistemas de Software

Monica Tugnarelli (UNER)
Marisa Panizzi (UNAHUR)
Guillermo Feierherd (UNTDF)

Procesamiento Distribuido y Paralelo

Marcelo Naiouf (UNLP)
Marcela Printista (UNSL)
Javier Balladini (UNCOMA)

Tecnología Informática Aplicada en Educación

Alejandra Zangara(UNLP)
Gustavo Gil (UNSa)
Andres Bursztyn (UTN)

Procesamiento de Señales y Sistemas de TR

Oscar Bría (INVAP)
Fernando Tinetti (UNLP)
Nelson Rodriguez (UNSJ)

Bases de Datos y Minería de Datos

Laura Lanzarini (UNLP)
Norma Herrera (UNSL)
Andrea Villagra (UNPA)

Innovación en Educación Informática

Claudia Russo (UNNOBA
Lucía Malbernat (UCAECE)
Sonia Mariño (UNNE)

Seguridad Informática

Paula Venosa (UNLP)
Jorge Eterovic (UNLaM)
Mariana Brachetta (UCH)

Temas Interdisciplinarios: Esports

Fernando Puricelli (UNAHUR)
Santiago Medina (UNLP)

Tesis de Doctorado

Alejandra Cechich (UNCOMA)
Laura De Giusti (UNLP)
Marcela Printista (UNSL)
Martin Larrea (UNS)
Claudio Aciti (UNCPBA)
Marcelo De Vincenzi (UAI)
Claudia Russo (UNNOBA)
Marisa Panizzi (UNAHUR)
Enzo Rucci (UNLP)
Emanuel Fratti (UNdeC)
Nora Reyes (UNSL)
Elena Durán (UNSE)
Verónica Gil Costa (UNSL)
Waldo Hasperue (UNLP)

ÍNDICE

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

AGENTES Y SISTEMAS INTELIGENTES

Agente softbot para pruebas automáticas de GUI. 21
Agentes conversacionales modernos. 26
Análisis comparativo de modelos transformer para clasificación de texto en español. 29
Análisis de flujo vehicular utilizando redes neuronales. 33
Aplicación de técnicas de IA y modelos de lenguaje para la recuperación de información en colecciones documentales jurídicas. 38
Argumentación abstracta y análisis contextual de debates complejos. 43
Arquitectura inteligente para la gestión proactiva de la mora tributaria en administraciones subnacionales. 48
Arquitecturas de inferencia estocástica y modelado multicriterio para la vigilancia de la integridad biogeoquímica en sistemas lóticos antropizados. 53
Arquitecturas híbridas explicables para sistemas de apoyo a la decisión clínica: evolución desde el modelado predictivo. 56
Automatización no code con Make y n8n aplicada a PYMEs: arquitectura, costo total de propiedad y potencial educativo. 61
Chatbots e inteligencia artificial para la sistematización de reglas de glosa en lengua de señas argentina (LSA). 66
Darmic: dataset argentino de imágenes de resonancia magnética cardiaca. 71
Desarrollo de un chatbot aplicando principios de usabilidad para la orientación de ingresantes en el sitio web de la Facultad de Ciencias Exactas y Tecnologías. 75
Detección de enfermedades en ganado mediante visión por computadora e IoT en ambientes extensivos del NOA. 80
Detección eficiente y remota de salud cardiovascular mediante videos. 84
Dialysmart: sistema de apoyo a la toma de decisiones clínicas basado en inteligencia artificial. 89
Ecosistema computacional de simulación para la predicción y gestión de desastres ambientales. 94
El futuro de la logística es inteligente: sistema de soporte de decisiones con IA y lógica difusa para gestión de stock en la industria alimentaria. 98
Estimación personalizada no invasiva de hemoglobina mediante vision transformers: colaboración Argentina-Italia. 102
Estrategias avanzadas de optimización basadas en aprendizaje automático para sistemas de energías renovables. 107
Evaluación de espectrogramas en redes neuronales convolucionales para la detección de la enfermedad de Parkinson. 112
Física de la cognición y el modelo de topología oscilatoria informacional (TOI). 117
Hacia un paradigma unificado en optimización logística: desde la gestión de tráfico urbano hasta la agroindustria 4.0. 122
IA para asistir la calificación jurídica de medidas cautelares. Una arquitectura replicable desde la ciencia de datos aplicada al derecho. 125
IA y ML para la identificación de animales en entornos naturales. 130
Inteligencia artificial al servicio de la comunidad: un enfoque predictivo y personalizado. 134
Inteligencia artificial aplicada en el cuidado de la salud. 139
Inteligencia artificial y sistemas multiagentes: un marco para la transformación digital y la gestión colaborativa en las organizaciones. 144

Inteligencia computacional en ciudades inteligentes: modelo práctico aplicando tecnologías convergentes y predicción de erosión costera. 147
Inteligencia organizacional para gestionar estratégicamente las decisiones y la transformación digital en organizaciones regionales. 152
Interpretabilidad y métodos de entrenamiento de modelos de deep learning. 157
Machine learning aplicado a astronomía.....161
Marcos argumentativos interpretativos: formalizando la ambigüedad en el razonamiento computacional. 165
Mejorando la eficiencia de metaheurísticas mediante estrategias de inicializaciones basadas en machine learning.....170
Optimización del triaje médico y trazabilidad de prescripciones mediante la integración de agentes de IA y tecnología blockchain en el hospital Eleazar Herrera Motta. 174
Optimización no lineal entera asistido por redes neuronales en problemas de call centers.....179
Piña 4.0: sistema multiagente inteligente para la optimización hídrica y climática en el cultivo de piña mediante tecnologías LoRa e inteligencia artificial. 184
Predicción de cortes en la matriz energética argentina con IA.....188
Reconstrucción de trayectorias bovinas mediante seguimiento visual multiobjeto en escenarios extensivos.....192
Red neuronal convolucional preentrenada para clasificar partículas de arena. 195
Simulación inteligente de toma de decisiones mediante sistemas adaptativos basados en agentes. 200
Sistema inteligente de asistencia por voz para control de calidad en producción de tempeh: enfoque híbrido basado en grafos y modelos de lenguaje.....203
Un framework multi-agente basado en LLMs para la automatización del flujo de investigación en simulación de redes con NS-3.....208
Uso de simuladores para el entrenamiento de robots mediante aprendizaje por refuerzo. 214

ARQUITECTURA, REDES Y SISTEMAS OPERATIVOS

Análisis comparativo de arquitecturas CISC y RISC como soporte de bajo nivel para algoritmos de inteligencia artificial en centros de datos. 219
Arquitectura para la verificabilidad criptográfica de eventos basada en blockchain. 224
Cuyos: un sistema operativo para entornos académicos. 229
Gestión automatizada de cultivos basada en IoT. 234
Hacia un framework de decisión para arquitecturas de microservicios en aplicaciones móviles. 238
Laboratorios virtuales destinados a educación. 242
Microservicios en aplicaciones móviles: arquitectura, desafíos y buenas prácticas.....247
Tecnologías de IoT y visión artificial aplicadas a sistemas inteligentes.....250
UM-Cloud: infraestructura de IA on-premise para educación superior.....256
Visual CRANE. Implementación visual para el despliegue simplificado de aplicaciones contenerizadas en entornos locales. 260

BASES DE DATOS Y MINERÍA DE DATOS

Almacenamiento y recuperación para la gestión de datos masivos. 266
Bases de datos espaciales y espacio-temporales y su aplicación en gestión académica universitaria. 271
Diseño de un prototipo para la predicción de riesgo de deserción universitaria (PRDU) a partir de registros SIU-Guaraní. 275
Gestión e indexación de datos no estructurados: técnicas y aplicaciones. 280
Integración de métricas de calidad de datos en procesos de minería de datos: validación metodológica aplicada al caso INCUCAI. 285
Medición de inflación local con datos en línea y herramientas computacionales. 290
Metodología de análisis y mejora de la calidad de datos en sistemas de información. 295
Modelos generativos interpretables para diseño de novo de moléculas. 300
Pipeline integrativo de machine learning predictivo y genómica funcional para la caracterización de firmas moleculares en enfermedades complejas. 305
Sistemas inteligentes. Técnicas, modelos y aplicaciones. 310
Uso de técnicas de explotación de información para análisis de similitudes y diferencias entre las aguas del Río de la Reconquista y Matanza-Riachuelo. 315

COMPUTACIÓN GRÁFICA, IMÁGENES Y VISUALIZACIÓN

Análisis visual jerárquico para el manejo y gestión de incertidumbre aplicado a datos multidimensionales. 320
Aplicación de redes CNN para la detección de órganos reproductivos de kiwi en imágenes y videos. 325
Avances en experiencias de realidad extendida multisensorial accesible. 330
Integración de tecnologías inmersivas con emergentes. 335
La visualización comparativa en el análisis de conjuntos de datos complejos. 340
Realidad extendida aplicada a la exploración geológica: un enfoque inmersivo. 345

INGENIERÍA DE SOFTWARE

Administración del contexto de dominio organizacional para la colaboración entre múltiples agentes de inteligencia artificial y personas en proyectos de ingeniería software. 350
Analíticas permeables a la incertidumbre para la toma de decisiones en la ingeniería de software. 355
Aplicaciones móviles 3D con realidad virtual y realidad aumentada. 360
Aportes de la inteligencia artificial a la gestión del conocimiento en el ámbito empresarial. 364
Asistencia de un agente cognitivo en la elicitación de requerimientos. 369
Aspectos de ingeniería de software y bases de datos NoSQL para el desarrollo de sistemas de software y aplicaciones para dispositivos móviles. 374
Automatización de estimaciones. 381
Automatización de HuerTIC con framework de desarrollo basado en WS-CDL. 387
Automatización de la gestión de estacionamientos mediante visión por computadora e integración con Odoo. 392

Circuito paramétrico cuántico para la evaluación de riesgo de carteras de inversión con 110 qubits en hardware cuántico real. 397
Colaboración de los LLMs para construir modelos conceptuales. 402
Criterios y factores contextuales para la selección de recursos de design thinking en ingeniería de software: hacia una guía de buenas prácticas. 407
De la web colaborativa a la web semántica: recuperación y reuso inteligente de información en comunidades virtuales. 412
Desafíos en el diseño y la comunicación entre microfrontend. 417
Diseño y validación de un modelo de calidad basado en CMMI e ISO/IEC 25010 para gestión de riesgos de calidad derivada de la rotación de personal en los equipos de desarrollo. 422
El paradigma de la computación cuántica: estrategias de desarrollo, seguridad y formación ante la próxima revolución tecnológica. 426
El uso de NLP e IA para mejorar la calidad y extraer información de especificaciones en lenguaje natural. 431
Elicitación asistida por IA en IR. 434
Enseñanza de la gestión de proyectos con metodología BIM en la industria del petróleo y gas. 439
Estudio de la actividad de instalación de sistemas de software en PyMES de Argentina. 442
Evaluación de prácticas de ingeniería del software en organizaciones regionales: métricas, calidad y validación empírica. 447
Evaluación del cumplimiento de OWASP en productos software. 451
Formalización de heurísticas en ingeniería de requisitos mediante IA generativa. 456
Go como lenguaje de diseño: reducción de complejidad accidental en sistemas concurrentes. 461
Hacia la gestión y reuso de activos en sistemas de big data basado en CBR orientado a servicios. 465
Hacia un observatorio de tendencias en ingeniería de sistemas y de software en contextos de transformación digital. 470
Herramienta tecnológica para la gestión del conocimiento en laboratorios de informática forense. 475
Impacto de la gestión del conocimiento sobre la motivación de los/as trabajadores/as. 479
Incorporación de IA generativa en la reestructuración de escenarios. 484
Ingeniería de requerimientos acelerada con cómputo paralelo aplicada a idiomas indígenas. 489
Ingeniería de software para el desarrollo de sistemas inteligentes: operacionalización de requisitos éticos. 493
Ingeniería de software: desarrollo de métodos, técnicas y herramientas para la protección de bases de datos. 498
Innovación en desarrollo de software: validación de arquitectura backend-driven UI en la aplicación "+65 Agenda Salud". 504
Inteligencia artificial como andamiaje cognitivo: eficiencia y calidad estructural en ecosistemas de desarrollo híbridos. 509
La ingeniería de software aplicada al diseño de interfaces gráficas de usuario accesibles. 515
Mantenibilidad de software, un aporte de la academia a la industria. 521

Modelado y ejecución de sistemas reactivos distribuidos mediante redes de Petri no autónomas y microservicios. 526
Modelos de calidad y métodos de evaluación basados en ISO/IEC 25010 para simuladores clínicos en enfermería. 531
PetriNet Studio: plataforma SaaS para simulación, análisis y generación automática de código concurrente a partir de redes de Petri. 536
Red bayesiana para elegir un modelo de proceso de software. 540
Reentrenamiento del modelo de redes neuronales profundas (PhaseNet) con datos sismológicos locales para optimizar la detección de sismos en Argentina. 545
Reuso de sistemas big data basado en conocimiento contextual. 549
SiXAI: explicabilidad inmutable mediante lógica difusa y blockchain. 554
Synapsis: historia clínica digital y gestión multimodal para el seguimiento de pacientes con enfermedad de Parkinson. 559
Tecnologías para el desarrollo de software accesible, inteligente y de calidad. 563
Transformación de LMS a ALS utilizando una arquitectura de referencia. 568
Transformación digital. Tecnología en ciudades inteligentes sostenibles. IA en ingeniería de software, calidad de la IA. 573
Un enfoque unificado de verificación y validación de software basado en expresiones regulares. 577
Un estudio comparativo de especificaciones de requisitos de calidad de software generadas con copilotos de IA. 582
Uso de técnicas de ludificación en la gestión de proyectos de sistemas de información. 587
Uso del análisis de texto estructurado y no estructurado para la generación de material de entrenamiento para modelos de IA. 594
Validación y transformación de contratos inteligentes a Solidity. 600
Verificación y validación de software en la industria 5.0. 604

INNOVACIÓN EN EDUCACIÓN EN INFORMÁTICA

Adopción de RISC-V en la carrera de ingeniería en informática. 609
Desarrollo de un simulador educativo para la enseñanza de sistemas distribuidos en UNCAUS. 615
Didáctica de las ciencias de la computación: experiencias en distintos niveles educativos. 620
Enseñanza de la disciplina de ingeniería de requerimientos en carrera de grado en la era de la inteligencia artificial generativa. 625
Enseñanza de requerimientos experiencia con clientes reales. 629
Estrategias pedagógicas y desarrollo de recursos tecnológicos para la enseñanza de ciencias de la computación: avances del segundo año de un proyecto. 634
Herramientas de software utilizadas en la enseñanza de la programación desde una perspectiva de la ingeniería de software. 639
Integración de tecnologías emergentes para fortalecer el aprendizaje colaborativo y el desarrollo de competencias. 644
La IA generativa en la enseñanza introductoria de la programación: posibles transformaciones en la práctica docente. 649
La integración de la inteligencia artificial en la formación del ingeniero en sistemas:

desafíos, oportunidades y transformaciones en la práctica docente. 653
Modelo curricular AI-Native con trazabilidad N4 para la formación universitaria en programación. 656
Pensamiento computacional, pensamiento crítico y programación en la universidad: metodologías, herramientas y recursos para su enseñanza. 661
Predicción del riesgo de abandono de estudiantes en las carreras de ingeniería de la UTN FRC. 666
Punto tecnológico para el desarrollo de estrategias de accesibilidad comunicacional. 671
Sistema de tutorías híbridas para el acompañamiento de estudiantes en el inicio de la carrera ingeniería en sistemas de información. 677

INNOVACIÓN EN SISTEMAS DE SOFTWARE

Arquitecturas de software modulares y energéticamente eficientes para IoT. 683
Blockchain y ciudades inteligentes: posibilidades de implementación en la gestión pública local en Argentina. 686
De ERC-721 a soulbound: la ingeniería de CertChain. 691
Diseño de un metamodelo conceptual de infraestructura pública digital basado en ingeniería de software. 696
Fortaleciendo la conexión entre proyectos de software y el bienestar comunitario. 701
Infraestructura blockchain interoperable: integración de soluciones web 3.0 y web 2.0 mediante API REST chain-agnostic con soporte de privacidad. 706
Innovación en sistemas de software para automatización, eficiencia energética e integración de energías renovables en el centro de cómputos de la UTN La Plata. 710
Integración de asistencia inteligente a la terapia del habla. 715
Integración de la privacidad desde el diseño y los datos abiertos para fortalecer las plataformas digitales públicas. 720
Migem: módulo interactivo para la gestión de emergencias médicas. 725
Propuesta arquitectónica para interoperabilidad en salud mediante SSI, IPS-FHIR y ZK-Rollups. 729
Seguimiento automático de actividades en adultos mayores mediante geofencing y procesamiento local. 734
Sistema de respuesta a emergencias y gestión de riesgos en ciudades inteligentes debido al cambio climático. 740
Sistemas inteligentes para la gestión sostenible del recurso hídrico. 745
Tecnología y desarrollo turístico: aportes para la transformación digital de destinos locales. 750

PROCESAMIENTO DE SEÑALES Y SISTEMAS DE TIEMPO REAL

Caracterización de técnicas de sincronización horaria en sistemas distribuidos mediante GNSS y RTC de bajo costo para telemetría a escala global. 756
Detección de anomalías en telemetría satelital utilizando memoria temporal jerárquica. 760
Evaluación comparativa de JSON y CBOR para sistemas industriales. 765
Implementación de interfaces cerebro-computadora integrando biofeedback EEG en videojuegos. 772

Procesamiento computacional de series temporales para el estudio del caos y la aleatoriedad en sistemas físicos complejos. 779
Prótesis robótica basada en visión computacional y arquitectura distribuida. 784
Sistemas de tiempo real, sistemas embebidos y técnicas de simulación. 788

PROCESAMIENTO DISTRIBUIDO Y PARALELO

Adaptabilidad en la demanda para garantizar un uso eficiente de los recursos. 793
Análisis basado en simulación de sistemas de salud e infraestructuras de HPC. 797
Análisis de arquitecturas distribuidas cloud, fog y edge computing. 803
Aplicaciones distribuidas en arquitectura de microservicios. Desafíos técnicos y mejores prácticas. 810
Arquitectura IoT multiprotocolo para series temporales en el continuo borde-nube. 815
Arquitecturas multiprocesador en HPC: portabilidad, eficiencia energética y computación cuántica. 821
Ciencia computacional para la predicción y modelado en contextos de incertidumbre. 828
Fundamentos, construcción y evaluación de rendimiento de algoritmos paralelos en HPC. 833
HPC en la nube para el procesamiento de datos: simulación y despliegue en arquitecturas IaaS. 839
Marco de investigación para el modelado, procesamiento y nuevas tecnologías para la gestión de datos. 845
Marco de investigación y desarrollo a través de la integración de simulación, big data y ciencia de datos. 850
Rendimiento, resiliencia y soluciones inteligentes en aplicaciones de cómputo paralelo y distribuido. 855

SEGURIDAD INFORMÁTICA

Análisis de seguridad en dispositivos biométricos de control de acceso: caso de estudio ZMM220. 861
Análisis y detección de hábitos digitales orientados a la mitigación de fugas de datos personales. 866
Arquitectura híbrida para gestión dinámica de factores de acceso: integración de agentes de IA con WSO2 y Keycloak. 870
Atacantes en redes OT impacto funcional sobre PLC y procesos físicos. 875
Ciberdefensa desplegable. 880
Ciberseguridad en redes operacionales. 885
Detección de información de identificación personal (PII) en Argentina a través del modelos de inteligencia artificial optimizados para el español. 890
Disruptive compliance toolkit: framework para programas de compliance de ciberseguridad en contextos disruptivos. 895
Evaluación de vulnerabilidades y desempeño en redes IoT en un entorno experimental controlado. 900
Modelo de seguridad para internet de las cosas basado en inteligencia artificial distribuida y aprendizaje federado. 905

OTP-Vote: avances en seguridad, auditoría y verificabilidad end-to-end hacia su implementación. 911

TECNOLOGÍA INFORMÁTICA APLICADA EN EDUCACIÓN

Alfabetización digital y transformación digital: acciones universitarias para la sociedad. 917
Alfadatizando: las humanidades digitales como puente hacia una alfabetización digital crítica en datos e inteligencia artificial desde la educación abierta. 922
Análisis de la prosodia emocional japonesa mediante reconocimiento de emociones del habla (SER). 927
Aplicación de modelos de inteligencia artificial para la accesibilidad de contenido educativo. 930
Aportes a la transformación de la educación digital. Usabilidad e interacción inmersiva con tecnologías informáticas emergentes. 936
Apropiación de tecnologías digitales en Río Negro. Desafíos educativos en contextos de cambio socioeconómico y productivo. 941
Asistentes virtuales de enseñanza basados en generación aumentada por recuperación. 946
Autoeficacia en línea y grit como atributos predictivos en sistemas de aprendizaje adaptativo en educación superior virtual. 951
Data warehouse soft skills: modelo diagnóstico integral de competencias mediante business intelligence e inteligencia artificial. 955
Desarrollo de competencias en inteligencia artificial aplicada a la educación: resultados y prospectiva de un programa de formación. 959
Desarrollos tecnológicos para apoyar el aprendizaje y la enseñanza en grupos colaborativos. 964
Detección temprana del rendimiento estudiantil mediante analítica predictiva y vibe coding. 969
Diseño y validación de objetos de aprendizaje (OA) mediante el uso de modelos de inteligencia artificial generativa: un enfoque hacia la personalización de contenidos educativos en la educación superior. 974
Diseño, manufactura aditiva e integración de hardware de una plataforma bípeda modular para investigación en locomoción y control dinámico. 979
El prompt como unidad crítica de accesibilidad en la interacción con inteligencia artificial conversacional en educación superior. 984
Estrategias de ludoevaluación para la optimización del aprendizaje y reducción de ansiedad en ingeniería en sistemas. 989
Experiencias educativas innovadoras con tecnologías emergentes. 993
Framework para desarrollos de robótica educativa en espacios simulados. 999
Gamificación como catalizador del pensamiento computacional en carreras STEM: diseño, implementación y evaluación. 1000
Gestión del inteligente del conocimiento en la educación superior. 1010
Innovación y desafíos en la educación digital. 1015
Integración de inteligencia artificial generativa en la educación médica: un framework para el entrenamiento de competencias blandas y juicio clínico. 1021
Inteligencia artificial generativa (IAG): relevamiento al estado del arte de sobre las

transformaciones pedagógicas actuales y las plataformas educativas. 1025
Investigaciones emergentes sobre tecnologías digitales en escenarios educativos. 1029
La tecnología informática como mediadora del lenguaje oral en el nivel inicial: una experiencia de robótica educativa con Blue-Bot. 1035
Modelado y automatización de procesos para el aseguramiento de la calidad universitaria bajo ISO 21001:2018. 1040
Objetos de aprendizaje y metodologías activas: innovación pedagógica para la retención en carreras de ingeniería con planes de estudio por competencias. 1046
Optimización de un modelo de riesgo de abandono estudiantil mediante minería de datos en la UTN La Plata: análisis e integración de factores psicológicos y sociológicos. 1050
Pensar con datos: desarrollo de competencias en ciencia de datos para el nivel secundario. 1055
Predicción temprana de la deserción universitaria en primer año: un modelo multivariable basado en aprendizaje automático. 1060
Proyecto para la gestión de la información académica en las carreras de ingeniería de la Facultad Regional Córdoba de UTN. 1065
Reingeniería y expansión de una infraestructura CAVE de realidad virtual para simulación inmersiva multiusuario en ingeniería. 1070
Resultados proyecto democratizando educación y TIC, IDICE UNSJ expandida con curriculum integrado e internacionalizado con AI. 1074
Sistema inteligente de soporte a la retención estudiantil basado en minería de datos educaciónal en UNCAUS. 1080
SmartLab: innovación con tecnología para espacios educativos. 1085
Tecno pedagogía como motor de la transformación educativa. 1090
Transformación digital docente. Impacto y realidades en el nivel primario de la ciudad de Orán. 1096
Un abordaje de la enseñanza de algoritmos evolutivos en la materia inteligencia artificial en la carrera de ISI UTN FRC desde los ambientes de aprendizaje colaborativos. 1101
Videojuegos y analítica educativa en educación primaria: evaluación y desarrollo de propuestas pedagógicas. 1106
xSTEAM: un enfoque sistémico para el desarrollo de soluciones STEAM en contextos extremos. 1111

TEMAS INTERDISCIPLINARIOS: ESPORTS

Desarrollo del ecosistema de esports en la universidad: experiencias de formación, investigación y extensión en la Universidad Abierta Interamericana. 1117
Diplomatura en deportes electrónicos UNLP. 1122
Experiencia del taller de e-sports de la Universidad Nacional de Hurlingham. 1128
Grupo de interés en deportes electrónicos. 1132
Proyecto esport UNLP. 1134

REDES DE COOPERACIÓN CIENTÍFICA INTERNACIONALES

Acciones de colaboración en el Consorcio de Cloud Computing, Big Data y Emerging Topics durante 2025 (CCC-BD&ET). 1141
Consorcio UNI-UEAR - alianza universitaria Argentina europea para la transformación digital cursos de capacitación docente RedUNCI. 1147
Participación en el programa UNITA GEMINAE. 1151

TESIS DOCTORALES

Análisis visual de datos multidimensionales. 1154
Análisis visual de datos provenientes de registradores oculares. 1164
Análisis y desarrollo de representaciones generales de audio. 1173
Analíticos visuales con inteligencia artificial aplicados en ciencia marina. 1183
Behavior-driven microservice architecture: un marco metodológico para la identificación iterativa de microservicios en contextos ágiles greenfield. 1194
Definición de un modelo de clasificación de tráfico en redes definidas por software. 1203
Desarrollo de un modelo de elicitación de emociones a partir de las características de la música. Generación de un sistema recomendador. 1213
Desarrollo de un modelo de mirroring con indicadores y herramientas para el seguimiento y visualización de procesos educativos colaborativos mediados por tecnología digital en educación superior. 1218
Descubrimiento y diseño de microservicios asistido por IA a partir de requisitos textuales. 1222
Diagnóstico de competencias blandas mediante herramientas de business intelligence. 1232
Formulación de métricas para la evaluación integral de metodologías de gestión de proyectos. 1240
Gemelos digitales híbridos para la enseñanza de robótica. Un entorno distribuido en tiempo real. 1244
Marco de trabajo para la evaluación de experiencia de usuario en sistemas interactivos. 1254
Metodología integral de recomendación de revisores en revistas científicas basada en múltiples factores. 1258
Modelo combinado de co-training y aprendizaje por transferencia para clasificación de documentos, a partir de un análisis comparativo de modelos de aprendizaje semi-supervisados. 1268
Modelo de calidad para evaluación y mejora de visualizaciones de información narrativas. 1278
Predicción del comportamiento, clasificación semántica y ponderación de riesgos de incendios rurales. 1288
Semánticas argumentativas basadas en diálogo: fundamentos teóricos, representación y análisis formal. 1298
Televisión digital interactiva (TVDI) para reducir las alteraciones de la marcha en adultos mayores. 1308
Transhumamismos, experiencias e inteligencia artificial: articulaciones pedagógicas en ambientes virtuales en la Facultad de Humanidades, UNMDP. 1315

Agentes y Sistemas Inteligentes

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

AGENTE SOFTBOT PARA PRUEBAS AUTOMÁTICAS DE GUI

Lucas Martín Prado [1]
Carlos Casanova[1,3] y Fabiana Piccoli [2 y 3]
[1] Univ. Tecnológica Nacional, Facultad Regional Concepción del Uruguay, Concepción del Uruguay
[2] LIDIC- Univ. Nacional de San Luis, San Luis
[3] Universidad Autónoma de Entre Ríos, Concepción del Uruguay
Argentina
{pradol, casanovac}@frcu.utn.edu.ar
mpiccoli@unsl.edu.ar

RESUMEN

El proceso de testing es una actividad que involucra ejecutar el software o componente de interés para hallar errores y corregirlos, es decir, implica detectar y solucionar problemas en el software con el fin último de que el mismo se comporte como se espera. Este proceso cumple un rol fundamental a la hora de mejorar la fiabilidad, seguridad y rendimiento del software, incluso desde una etapa temprana del desarrollo. Al reducir la cantidad y probabilidad de errores en el sistema se mejora la usabilidad del mismo, aumentando a su vez el nivel de satisfacción del usuario que lo utiliza. En definitiva, el proceso de testing es vital para garantizar la calidad del software. Si bien para los niveles más bajos de abstracción del sistema desarrollado la industria adoptó rápidamente el uso de pruebas automáticas, lo cierto es que para pruebas que requieren un mayor grado de abstracción, como las pruebas guiadas por la interfaz gráfica de usuario (GUI, Graphical User Interface), las pruebas manuales siguen siendo la práctica más común.

En este trabajo proponemos una línea de investigación que tiene por objetivo último diseñar un agente inteligente del tipo softbot que realice pruebas automáticas de GUI mediante distintas estrategias y que pueda ejecutarse independientemente de la plataforma utilizada por el software bajo prueba. Para ello se propone la utilización de nuevas técnicas en el campo de la Visión Computacional e Inteligencia Artificial, de tal manera que el agente sea independiente del proceso de desarrollo del sistema y pueda, por lo tanto, realizar el proceso de testing sin los sesgos que surgen de la autoría del código a probar.

Palabras clave: *Testing, Pruebas Automáticas, Pruebas de GUI, Visión Computacional, Agentes Inteligentes*

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro del proyecto de investigación: "Analíticas basadas en incertidumbre para soportar la toma de decisiones en la ingeniería de software" (UTN).

1. INTRODUCCIÓN

La automatización de la tarea de testing es objeto de estudio por la importancia que tiene su aplicación en la industria y continúa teniendo un largo camino por delante [1]. El realizar pruebas manualmente es un proceso costoso, tedioso y propenso a errores por su naturaleza repetitiva y mecánica. Si bien la automatización de pruebas avanzó para los niveles más bajos de abstracción (por ejemplo: pruebas unitarias), las técnicas para automatizar pruebas de alto nivel (pruebas de sistema basadas en interfaz gráfica, o pruebas de aceptación) aún no están lo suficientemente avanzadas para tener una aceptación general en la industria [2].

Esto constituye un problema, pues este tipo de pruebas resultan un cuello de botella, retrasando el ritmo de desarrollo de las empresas y, aunque existen algunos casos de empresas de gran tamaño que utilizan algún grado de pruebas automáticas de estas características, suele limitar la velocidad (o la calidad, si se renuncia a hacer pruebas rigurosamente) con la que desarrollan sus productos de software, especialmente las PyMEs, donde la superposición de roles de desarrollo y testing implican una asignación de recursos mayor. Otro de los problemas asociados a las pruebas automáticas es su mantenimiento [3]. Para que una prueba automática tenga

utilidad, la misma tiene que estar actualizada, es decir, debe contemplar los cambios del sistema. Cuando un componente cambia, lo cual sucede a menudo en el desarrollo de software, y sobre todo en estos contextos en los que se busca un ritmo de trabajo constante y veloz, las pruebas automáticas asociadas a ese componente (o que necesiten interactuar con ese componente) pueden quedar obsoletas. A medida que se añaden nuevas funcionalidades y se modifican funcionalidades existentes, se requiere verificar y, de ser necesario, actualizar las mismas. Especialmente para las pruebas de un alto nivel de abstracción, como las pruebas guiadas por la interfaz gráfica de usuario, esta dificultad es determinante para decidir si se utilizan pruebas automáticas o manuales [4].

Las herramientas actuales que se basan en reconocimiento de imágenes para automatizar pruebas basadas en la GUI utilizan técnicas anticuadas que provocan un rendimiento menor al que es posible con la tecnología actual.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Las pruebas basadas en la interfaz gráfica de usuario hacen referencia a las pruebas de sistema realizadas a través de la GUI del sistema. A diferencia de pruebas de bajo nivel de abstracción, donde se puede determinar qué línea de código es la que provoca el error, estas pruebas con mayor nivel de abstracción solo responden qué funcionalidad presenta defectos en el sistema, por lo que requieren un posterior análisis del software bajo prueba (SUT, Software Under Test) para encontrar la causa del defecto. Existen múltiples técnicas para las pruebas basadas en GUI, a las cuales se las clasifica en distintas generaciones. La primera generación, también conocida como las pruebas basadas en GUI por coordenadas, utilizan las coordenadas exactas en la pantalla para interactuar con la GUI del SUT. Esta técnica se abandonó debido a que las GUI de los programas suelen cambiar con el tiempo, lo que provocaba que las pruebas precisen mucho esfuerzo de mantenimiento. La segunda generación, conocida también como pruebas basadas en GUI por componentes, es una técnica que aún a día de hoy continúa vigente y puede utilizarse para pruebas a nivel de sistema. Es una técnica de caja gris (el punto medio entre las pruebas de caja negra y de caja blanca) y utiliza propiedades de los componentes de la GUI para identificar e interactuar con ellos, lo cual le brinda mayor grado de robustez respecto de los cambios de la GUI, ya que pueden cambiarse la localización y la representación gráfica de la GUI sin necesidad de realizar mantenimiento del código de las pruebas. Sin embargo, como las pruebas utilizan información de los componentes de la GUI, cambios al código fuente del programa pueden causar que fallen ciertas pruebas. Otro problema de esta generación es que existen algunos componentes que no pueden ser probados, como, por ejemplo, un componente sobre el cual se dibuje la salida en tiempo de ejecución. La tercera generación, también conocida como pruebas visuales de GUI (VGT, Visual GUI Testing), es una técnica emergente que utiliza emparejamiento de plantillas (Template Matching) para identificar e interactuar con la GUI a través de los mapas de bits de las imágenes mostradas al usuario durante la ejecución. La interacción es realizada utilizando el puntero y el teclado, emulando la interacción de un usuario con el SUT. Esto permite que la técnica se utilice independientemente de la implementación o plataforma en la que se ejecute el sistema. Es una técnica de caja negra y no presenta las limitaciones mencionadas en la segunda generación.

La primera herramienta de VGT conocida es Triggers [5]. Sin embargo, debido a problemas de rendimiento, la herramienta fue descartada y la técnica no cobró relevancia. Con el avance tecnológico, los problemas de rendimiento se mitigaron y la siguiente herramienta de VGT en surgir fue Sikuli [6][7] (ahora llamada SikuliX), creada por investigadores del Instituto Tecnológico de Massachusetts (MIT, Massachusetts Institute of Technology). Esta herramienta de código abierto permite crear pruebas automáticas que utilizan lo que se muestra por monitor. La herramienta permite realizar acciones sobre algún componente del SUT utilizando una imagen determinada (por ejemplo, la imagen de un botón) para reconocerlo. Mediante emparejamiento de

plantillas también puede determinarse la presencia de algún componente visual para distinguir las pruebas superadas de las falladas. Mediante una secuencia de acciones y verificaciones se obtienen pruebas automáticas cuyo mantenimiento consiste en actualizar las imágenes si se produce algún cambio gráfico y modificar acciones si cambia el flujo de la aplicación, sin ser necesario el mantener el código en casos de mover la posición de los componentes (como sí lo es para las herramientas de primera generación).

JAutomate (ahora llamada EyeAutomate) es una herramienta comercial desarrollada por Swifting [8], la cual permite desarrollar pruebas de forma intuitiva y sencilla, incluso para programadores novatos. Además de contar con distintos modos para la creación de pruebas (programando manualmente o grabando las interacciones que realiza el usuario en una ejecución de prueba y traduciendo esos movimientos a código) y distintos algoritmos (buscando píxeles aleatoriamente, analizando contraste o zonas aledañas a los componentes de interés) también es importante mencionar que JAutomate puede integrarse y ejecutarse desde un servidor de integración continua, es decir, cumple lo necesario para poder utilizarse en contextos CI/CD.

Sin embargo, existe evidencia empírica [2][9][10] de pruebas que fallan sin motivo aparente, es decir, casos de falsos positivos en las pruebas. En la gran mayoría de los casos estos fallos se deben a un error en el módulo de reconocimiento de imágenes de la herramienta. Una estrategia utilizada para mitigar este problema es añadir redundancia en las pruebas de tal forma que, si falla el emparejamiento de plantillas, se ejecute la prueba nuevamente o se busque otra imagen durante la prueba. Otra estrategia para mitigar los fallos es incorporar tiempos de espera en la ejecución de las pruebas que le permitan a la herramienta disponer de más tiempo para reconocer las imágenes con éxito. Si bien estas estrategias reducen los falsos positivos, tienen el efecto de aumentar aún más el tiempo de ejecución de las pruebas.

En [10] se evalúa la aplicación de VGT en la industria, en el contexto de CI, y si bien la recepción de la técnica fue positiva, se identificaron ciertos inconvenientes en su aplicación: en primer lugar, la herramienta detecta fallas, pero no defectos, por lo que se logra identificar en qué interacciones ocurren problemas, pero no qué los causa. Esto implica que se debe invertir tiempo para realizar el análisis necesario que permita identificar el problema. Además, otro aspecto a considerar es que existen fallas detectadas con VGT que son difíciles de replicar manualmente, lo que significa invertir aún más tiempo en el proceso de verificación y análisis. Otro inconveniente detectado fue un alto porcentaje de falsos positivos. Si bien este problema es conocido y esperado, denota la necesidad de mejorar las herramientas para aplicar VGT de forma más eficiente. Aspectos positivos observados de utilizar esta técnica incluyen la ausencia de falsos negativos (no se superaron pruebas que debían fallar) y la detección de fallas existentes que llevaron a su posterior corrección. Uno de los defectos corregidos fue un problema de integración entre el SUT y su base de datos que, según los desarrolladores, no hubiese sido encontrado sin la utilización de VGT.

Existen estudios realizados que comparan y/o combinan técnicas de segunda y tercera generación: en [9] se evalúan y comparan estas dos generaciones utilizando 18 versiones alternativas de una aplicación, cada una con un fallo distinto, y se verifican los resultados. Entre las posibles versiones alternativas se encontraban: componentes extra, componentes faltantes, componentes invisibles, cambios de resolución, esconder componentes fuera del monitor, entre otras. Las técnicas de tercera generación buscan imitar a un usuario humano, por lo que al utilizar como guía las imágenes del monitor fallan las pruebas donde los componentes son invisibles, no existen, o están escondidos (tal como le pasaría a una persona). Si bien esto es un efecto beneficioso, pues al simular la interacción humana son un mejor indicador de la respuesta esperada de un usuario final, estas técnicas tienden a presentar mayor cantidad de falsos positivos para las pruebas de sistema (fallan pruebas que deberían de haber sido superadas). Por el contrario, las técnicas de segunda generación pueden superar las pruebas donde los componentes son invisibles o están

escondidos (pruebas que un humano fallaría), pues interactúan directamente con el componente a nivel de código, por lo que reportan un mayor número de falsos negativos en las pruebas de aceptación. Las herramientas que utilizan técnicas de segunda generación tienen problemas para detectar defectos que una persona detectaría fácilmente, como, por ejemplo, la ausencia de un componente en pantalla. La conclusión de este estudio es que resulta beneficioso combinar ambos tipos de técnicas para complementar sus limitaciones. Por otro lado, en [11] se utilizan técnicas de Machine Learning junto con técnicas de segunda y tercera generación para una aplicación desarrollada utilizando CI/CD, con resultados positivos.

3. RESULTADOS OBTENIDOS/ESPERADOS

Uno de los objetivos de esta línea de investigación es diseñar un modelo de visión computacional que permita reconocer los componentes de la GUI de aplicaciones que se ejecutan en distintas plataformas. Mediante la utilización de nuevas técnicas en el campo de la Visión Computacional (CV, Computer Vision), principalmente los avances en reconocimiento de imágenes [12], pueden mitigarse los efectos negativos derivados de realizar pruebas de GUI por reconocimiento de imágenes, sin depender de utilizar atajos que dependen de la plataforma de implementación del software desarrollado, como lo son las pruebas que utilizan código de los componentes. Con ese fin se ha creado un dataset de imágenes con etiquetas para los distintos componentes GUI encontrados en páginas web y aplicaciones móviles. Además, se están entrenando distintos modelos de detección de objetos para reconocer dichos componentes.

A su vez, utilizando técnicas de Inteligencia Artificial [13], como lo es el Aprendizaje por Refuerzo (RL, Reinforcement Learning) [14][15] o la Computación Evolutiva[16][17], es posible construir un agente que sea independiente del proceso de desarrollo del sistema y pueda, por lo tanto, realizar el proceso de testing sin los sesgos que surgen de la autoría del código a probar, permitiendo una reducción en el esfuerzo requerido para tareas de testing manual.

A futuro se planea:

- Formular un modelo conceptual de exploración de GUI y un lenguaje de especificación de acciones.
- Formular diversas estrategias de testing, las cuales constituirán el programa del agente, dependiendo de la tarea de testing requerida (por ejemplo, pero no excluyentemente, exploratorio, guiado por casos de uso, *fuzz testing*, identificación rápida de defectos, testing de requisitos no funcionales, etc.) y arquitecturas solidarias (esto es, sensores y actuadores).
- Implementar el agente en un prototipo de software y realizar pruebas exhaustivas.

4. FORMACIÓN DE RECURSOS HUMANOS

El resultado esperado más importante respecto a la formación de recursos humanos es el desarrollo de una tesis de doctorado. Además, para la confección del dataset de imágenes de componentes GUI se contó con el trabajo de 10 becarios de investigación y se espera que la temática sirva de plataforma para continuar con la formación de becarios alumnos y graduados que quieren iniciarse en la investigación.

5. BIBLIOGRAFÍA

[1] N. Alshahwan, M. Harman, and A. Marginean, "Software testing research challenges: An industrial perspective," in 2023 IEEE Conference on Software Testing, Verification and Validation (ICST), pp. 1–10, 2023.

[2] E. Alegroth, R. Feldt, and L. Ryrholm, "Visual gui testing in practice: challenges, problems and limitations," Empirical Software Engineering, vol. 20, 06 2014.

[3] E. Alegroth, R. Feldt, and P. Kolström, "Maintenance of automated test suites in industry: An empirical study on visual gui testing," Information and Software Technology, vol. 73, 02 2016.

[4] E. Alegroth and R. Feldt, "On the long-term use of visual gui testing in industrial practice: a case study," Empirical Software Engineering, vol. 22,pp. 2937–2971, Dec 2017.

[5] R. Potter, “Triggers: guiding automation with pixels to achieve data access,” 1993.
[6] T. Yeh, T.-H. Chang, and R. Miller, “Sikuli: Using gui screenshots for search and automation,” pp. 183–192, 10 2009.
[7] T.-H. Chang, T. Yeh, and R. Miller, “Gui testing using computer vision,” vol. 3, pp. 1535–1544, 04 2010
[8] E. Alegroth, M. Nass, and H. Olsson, “Jautomate: A tool for system and acceptance-test automation,” pp. 439–446, 03 2013
[9] E. Alegroth, Z. Gao, R. Oliveira, and A. Memon, “Conceptualization and evaluation of component-based testing unified with visual gui testing: An empirical study,” in 2015 IEEE 8th International Conference on Software Testing, Verification and Validation (ICST), pp. 1–10, 2015.
[10] E. Alegroth, A. Karlsson, and A. Radway, “Continuous integration and visual gui testing: Benefits and drawbacks in industrial practice,” in 2018 IEEE 11th International Conference on Software Testing, Verification and Validation (ICST), pp. 172–181, 2018.
[11] F. Macchi, P. Rosin, J. M. Mervi, and L. Turchet, “Image-based approaches for automating gui testing of interactive web-based applications,” in 2021 28th Conference of Open Innovations Association (FRUCT), pp. 278–285, 2021.
[12] F. Sultana, A. Sufian, and P. Dutta, A Review of Object Detection Models Based on Convolutional Neural Network, pp. 1–16. Singapore: Springer Singapore, 2020.
[13] S. J. Russell and P. Norvig, Inteligencia Artificial: un enfoque moderno. No. 04; Q335, R8y 2004., Pearson, 2004.
[14] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction. The MIT Press, second ed., 2018.
[15] A. G. Barto, R. S. Sutton, and C. W. Anderson, “Looking back on the actor–critic architecture,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 1, pp. 40–50, 2021.
[16] Kelly, Stephen, and Jory Schossau. "Evolutionary computation and the reinforcement learning problem." Handbook of Evolutionary Machine Learning. Singapore: Springer Nature Singapore, pp. 79-118, 2023.
[17] Stanton, Adam, and Alastair Channon. "Incremental neuroevolution of reactive and deliberative 3d agents.", pp. 341-348, 2015

AGENTES CONVERSACIONALES MODERNOS

Germán Lescano[1], Pablo Santana-Mansilla[1], Gregorio Tkachuk[1], Noelia Yolde, Luz Coria, Eliana Proserpio, César Ponce-Padilla, Angélica Carabajal, Benjamín Caloso

[1]Instituto de Investigación en Informática y Sistemas de Información (IIISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)
{gelescano, psantana, gregorio}@unse.edu.ar; {noe.yolde, luzcoria025, elianaproserpio, ponce.padilla.c, carabajal.angelicamaria, benjacaloso92}@gmail.com

RESUMEN

Los agentes conversacionales son sistemas diseñados para imitar el diálogo humano a través de voz, texto o gestos, y se clasifican en tres tipos principales: chatbots, avatares virtuales y robots físicos. Su uso se ha expandido en áreas como la salud, la educación y la atención al cliente. Los avances en la Inteligencia Artificial y el aprendizaje profundo han impulsado la creación de modelos conversacionales más avanzados, especialmente los Modelos de Lenguaje de Gran Escala (LLMs), capaces de generar texto contextual y coherente. Los LLMs actualmente son usados como base por los agentes conversacionales.

Desde la perspectiva de la Interacción Humano-Computador (HCI), los agentes conversacionales representan una nueva generación de interfaces inteligentes que integran lenguaje natural, razonamiento y adaptabilidad. Las líneas de investigación más relevantes incluyen el fine-tuning e instruction tuning para la especialización en dominios concretos; la generación aumentada por recuperación (RAG), que combina búsqueda y generación neuronal; el prompt engineering para modular estilo y comportamiento; el aprendizaje por refuerzo con retroalimentación humana (RLHF); y la multimodalidad, que integra texto, voz e imagen.

Un tópico de relevancia es la incorporación del componente afectivo para lograr interacciones más empáticas. En este marco, el proyecto se propone investigar principalmente el desarrollo de agentes inteligentes afectivos, sin excluir otras líneas complementarias vinculadas al aprendizaje profundo, la personalización y la mejora técnica de los sistemas conversacionales, con el fin de contribuir al desarrollo de una nueva generación de agentes más útiles, confiables y humanos.

Palabras clave: *Agente Conversacional; Deep Learning; Informática Conversacional; Chatbot, LLM, RAG, HCI, Computación Afectiva; RLHF; Fine Tuning; Prompt Engineering.*

CONTEXTO

En este trabajo se presenta el proyecto "Agentes conversacionales modernos", acreditado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT-UNSE) para el período 2026-2027.

1. INTRODUCCIÓN

Los agentes conversacionales son sistemas diseñados para imitar las conversaciones humanas mediante distintos canales de comunicación, tales como la voz, el texto e incluso expresiones faciales y gestos [1], [2], [3]. Estos sistemas pueden clasificarse en tres grandes tipos: chatbots sin cuerpo, avatares virtuales con representación corporal o facial, y robots físicos [1]. Actualmente se observa una marcada tendencia en el uso de agentes conversacionales como punto de contacto inicial en sectores como la hospitalidad, la banca, el entretenimiento y la salud [1], [2]. A pesar del progreso tecnológico y del valor agregado que ofrecen en términos de presencia social y disponibilidad, diversos estudios señalan que los resultados aún no alcanzan el

nivel de efectividad esperado [1]. Entre las posibles causas se encuentra la dificultad de los agentes para reproducir comportamientos comunicativos humanos complejos, así como para adaptar sus respuestas a distintos contextos, estilos conversacionales o necesidades del usuario [1].

En los últimos años, los avances en Inteligencia Artificial (IA) y aprendizaje profundo (deep learning) han impulsado el desarrollo de modelos conversacionales más sofisticados, capaces de generar texto de manera contextual y coherente. En particular, los Modelos de Lenguaje de Gran Escala (Large Language Models, LLMs) han revolucionado el campo de la informática conversacional al permitir que los sistemas comprendan y produzcan lenguaje natural con alto grado de fluidez y flexibilidad. Estos modelos, entrenados con grandes volúmenes de datos textuales, estiman la probabilidad de secuencias lingüísticas —símbolos, tokens o frases completas— y generan respuestas mediante técnicas de optimización basadas en redes neuronales profundas [4]. En la actualidad, estos modelos de lenguaje son usados como base por los agentes conversacionales.

El desafío actual no solo consiste en producir contenido gramaticalmente correcto o informativamente relevante, sino en controlar y adaptar la generación de lenguaje a distintos objetivos comunicativos, dominios de conocimiento o perfiles de usuario.

Desde la perspectiva de la HCI, estos avances reflejan una evolución del interés inicial por las interfaces hacia un enfoque centrado en la experiencia interactiva entre humanos y sistemas inteligentes [5]. Rogers [5] caracteriza este campo como un espacio interdisciplinario donde convergen la ciencia de la computación, la ingeniería del software, la psicología, la ciencia cognitiva, el diseño y las ciencias sociales. En este marco, los agentes conversacionales representan una forma avanzada de interfaz que combina procesamiento del lenguaje natural, razonamiento semántico, generación contextual y adaptabilidad.

La Informática Conversacional, como campo interdisciplinario, integra los aportes de la Inteligencia Artificial, el Procesamiento de Lenguaje Natural, la Ciencia Cognitiva, el Análisis Conversacional y la Ingeniería del Conocimiento [9]. En este contexto, el uso combinado de técnicas de deep learning, mecanismos de recuperación aumentada y estrategias de personalización ofrece un terreno fértil para la investigación aplicada. El objetivo no es únicamente lograr que los agentes comprendan o respondan, sino que puedan razonar, adaptarse, aprender y comunicarse de forma contextualizada, constituyendo así una nueva generación de sistemas conversacionales más útiles, versátiles y humanos. Un aspecto importante en esta dirección es incorporar aspectos afectivos para lograr interacciones más naturales y empáticas [6], [7], [8].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO/OBJETIVOS

En este proyecto se realizarán contribuciones vinculadas con el desarrollo de agentes conversacionales en los cuales se apliquen técnicas de fine-tuning, RAG, prompt engineering, RLHF, multimodalidad y adaptación contextual.

Para la presente investigación se estableció el objetivo general de desarrollar y evaluar modelos de agentes conversacionales inteligentes que integren capacidades afectivas, adaptativas y contextuales mediante técnicas de aprendizaje profundo, fine-tuning y RAG, con el fin de mejorar la calidad, naturalidad y efectividad de la interacción humano-computadora.

Los objetivos específicos que se establecieron son los siguientes:

- Revisar críticamente los enfoques existentes sobre modelos de lenguaje, sistemas afectivos, técnicas de Deep learning, fine-tuning, RAG y mecanismos de control conversacional.
- Implementar un módulo de detección afectiva.

- Integrar un mecanismo de adaptación contextual (historial, preferencias, dominio).
- Integrar una arquitectura RAG para mejorar precisión y coherencia.
- Evaluar de la efectividad del sistema en términos de naturalidad, aceptación y calidad de la interacción.
- Publicar avances y resultados parciales de los resultados conseguidos.

3. RESULTADOS ESPERADOS

Con la concreción de este proyecto se esperan obtener los siguientes resultados principales:

- Revisión sistemática sobre agentes conversacionales, incluyendo modelos de lenguaje, sistemas afectivos, técnicas de deep learning, fine-tuning, RAG y mecanismos de control conversacional, identificando limitaciones y oportunidades.
- Arquitectura de agente conversacional que integre representaciones semánticas (embeddings), módulos afectivos (reconocimiento y expresión de emociones) y mecanismos de RAG para acceso a información contextual o factual.
- Un prototipo funcional de agente conversacional siguiendo la arquitectura propuesta.
- Evaluación experimental del rendimiento del agente conversacional implementado.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de este proyecto está constituido por tres docentes investigadores formados. A través del proyecto, se contribuirá a la formación de estudiantes: cinco integrantes desarrollarán sus Trabajos Finales de Graduación de licenciatura y uno realizará su trabajo final de posgrado en el marco de una especialización en informática educativa.

5. BIBLIOGRAFÍA

[1] M. M. E. Van Pinxteren, M. Pluymaekers, and J. G. A. M. Lemmink, "Human-like communication in conversational agents: a literature review and research agenda," Journal of Service Management, vol. 31, no. 2, pp. 203–225, 2020, doi: 10.1108/JOSM-06-2019-0175.

[2] X. Yang, M. Aurisicchio, and W. Baxter, "Understanding affective experiences with conversational agents," Conference on Human Factors in Computing Systems - Proceedings, pp. 1–12, 2019, doi: 10.1145/3290605.3300772.

[4] E. Voita, R. Sennrich, and I. Titov, "Language Modeling, Lexical Translation, Reordering: The Training Process of NMT through the Lens of Classical SMT," in 2021 Conference on Empirical Methods in Natural Language Processing, Punta Cana, República Dominicana, 2021, pp. 8478–8491. doi: 10.18653/v1/2021.emnlp-main.667.

[5] Y. Rogers, HCI Theory Contemporary: Clasical, Modern and Contemporary. Pensilvania, Estados Unidos, 2012.

[6] Z. Callejas, R. López-Cózar, N. Ábalos, and D. Griol, "Affective conversational agents: The role of personality and emotion in spoken interactions," in Conversational Agents and Natural Language Interaction: Techniques and Effective Practices, D. Perez-Marin and I. Pascual-Nieto, Eds., IGI Global, 2011, pp. 203–222. doi: 10.4018/978-1-60960-617-6.ch009.

[7] R. W. Picard, Affective Computing. Cambridge, MA: Massachusetts Institute of Technology., 1997.

[8] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," IEEE Trans Affect Comput, vol. 1, no. 1, pp. 18–37, 2010, doi: 10.1109/T-AFFC.2010.1.

[9] T. Nishida, "Introduction," in Conversation Informatics: An Engineering Approach, T. Nishida, Ed., John Wiley & Sons, 2007.

ANÁLISIS COMPARATIVO DE MODELOS TRANSFORMER PARA CLASIFICACIÓN DE TEXTO EN ESPAÑOL

Raúl Paz-Zannini[1], Germán Lescano[1], Lilia Palomo[1]
[1]Facultad de Ciencias para la Innovación y el Desarrollo (FCID)
Universidad Católica de Santiago del Estero (UNSE)
{raulleduardo.pazzannini, germanezequiel.lescano, lilia.palomo}@ucse.edu.ar

RESUMEN

En los últimos años, los modelos basados en la arquitectura Transformer han marcado un hito en el desarrollo del Procesamiento del Lenguaje Natural (PLN), al posibilitar una representación contextual más precisa de las secuencias lingüísticas. Su capacidad para capturar dependencias semánticas a diferentes escalas ha redefinido el enfoque tradicional en múltiples tareas del área. A pesar del éxito de estos modelos en inglés, existe una necesidad de evaluar su desempeño en contextos hispanohablantes, donde los recursos lingüísticos y los modelos optimizados son aún limitados.

Este estudio se centrará en explorar, comprender y comparar el desempeño de modelos Transformer preentrenados tales como como BETO, mBERT, RoBERTa, en tareas de clasificación de texto en idioma español. Para realizar comparaciones objetivas, se considerarán métricas estándar como precisión, recall, F1-score y eficiencia computacional.

Entre los resultados esperados se anticipa que los modelos Transformer entrenados específicamente en español (como BETO) superarán en rendimiento a los modelos tradicionales.

Este análisis tiene implicancias directas para la elección de modelos en aplicaciones reales que requieren procesamiento de texto en español, como asistentes virtuales, moderación de contenidos y análisis automático de documentos.

Palabras clave: *Transformer; PLN; BETO; mBERT; RoBERTa.*

CONTEXTO

En este trabajo se presenta el proyecto "Análisis comparativo de modelos Transformer para clasificación de textos en español", acreditado y financiado por la facultad de Ciencias para la Innovación y el Desarrollo (FCID) de la Universidad Católica de Santiago del Estero para el período 2026.

La FCID impulsa la implementación de proyectos que apuntan a incentivar la investigación desde las cátedras promoviendo la interacción vertical y horizontal entre ellas, y posibilitando a los docentes obtener resultados que puedan ser aplicados en las aulas.

1. INTRODUCCIÓN

La clasificación de texto es una tarea supervisada que asigna etiquetas a secuencias textuales, y requiere corpus etiquetados, técnicas de preprocesamiento (tokenización, normalización) y modelos de aprendizaje automático.

Los modelos Transformers utilizan mecanismos de atención para procesar secuencias de texto de forma paralela, a diferencia de modelos recurrentes como LSTM (Long Short-Term Memory). Esto permite una mayor eficiencia y capacidad para capturar dependencias de largo alcance [1].

Si bien, actualmente, los modelos de lenguaje basados en la arquitectura Transformer han mostrado un desempeño sobresaliente en tareas de Procesamiento de Lenguaje Natural (PLN), la mayoría de las

investigaciones y modelos pre-entrenados se centran en el idioma inglés, lo que plantea interrogantes respecto a su aplicabilidad y rendimiento en español [2].

Entre los modelos más relevantes para esta investigación se encuentran:

- BERT (Bidirectional Encoder Representations from Transformers) [3].
- BETO: versión de BERT entrenada exclusivamente en corpus en español [4].
- mBERT: modelo multilingüe que incluye el español entre sus lenguas [4].
- XLM-RoBERTa: modelo robusto y multilingüe entrenado en grandes cantidades de datos en varios idiomas [5].
- ALBETO: versión liviana y eficiente de BERT en español [6].

El desarrollo y evaluación de modelos Transformer para la clasificación de texto en español reviste una doble relevancia, tanto tecnológica como lingüística. Por un lado, fomenta la creación de herramientas más inclusivas y eficaces en uno de los idiomas con mayor número de hablantes a nivel global; y por otro, su implementación tiene implicancias prácticas directas en múltiples áreas, incluyendo la mejora de sistemas automatizados de atención al cliente, el análisis semántico de redes sociales, la automatización de procesos judiciales y la detección de discursos de odio en entornos digitales.

Asimismo, el análisis comparativo de modelos permite avanzar hacia una optimización inteligente de recursos computacionales, especialmente relevante en contextos de computación en el borde (edge computing), donde es necesario balancear rendimiento con limitaciones de hardware. Esta línea de investigación facilita la selección de arquitecturas más livianas y adaptables sin sacrificar precisión.

Desde una perspectiva interdisciplinaria, el estudio establece puentes con diversos campos. En el área de bases de datos, se exploran representaciones semánticas del lenguaje mediante embeddings y técnicas avanzadas de recuperación de información. En el ámbito del análisis numérico, se aplican algoritmos iterativos de optimización, cálculo de gradientes, funciones de activación, técnicas de reducción de dimensionalidad (como PCA o t-SNE), así como mecanismos de validación numérica para evaluar métricas de rendimiento.

Los modelos Transformer, concebidos para capturar relaciones complejas entre elementos de una secuencia sin necesidad de recurrencia explícita, han demostrado ser especialmente eficaces tanto en el tratamiento de datos numéricos secuenciales (como series temporales), como en datos lingüísticos, donde el orden y el contexto desempeñan un rol fundamental.

En este marco, la presente investigación se centra en la comparación empírica de diversos modelos Transformer pre-entrenados aplicados a tareas de clasificación de texto en español, con el fin de determinar sus fortalezas, limitaciones y posibles ámbitos de aplicación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO/OBJETIVOS

En este proyecto se realizarán contribuciones vinculadas con la evaluación de modelos Transformers preentrenados para realizar tareas de clasificación usando como entrada textos en español, ámbito en el que se identifica escasa experimentación.

Para la presente investigación se estableció el objetivo general de comparar el rendimiento de diferentes modelos Transformer preentrenados en tareas de clasificación de texto en español.

Los objetivos específicos que se establecieron son los siguientes:

- Comprender las bases fundacionales del modelo Transformer.

- Revisar el estado del arte en modelos Transformer aplicados al PLN en español.
- Seleccionar corpus adecuados en lengua española con etiquetas clasificatorias.
- Implementar un pipeline de preprocesamiento y entrenamiento.
- Evaluar los modelos utilizando métricas.
- Analizar los resultados obtenidos y establecer recomendaciones sobre su aplicabilidad en distintos contextos.

3. RESULTADOS ESPERADOS

La presente investigación tiene como propósito generar aportes significativos en los planos técnico, pedagógico y social.

Desde la perspectiva técnica, se prevé:

- Elaborar un ranking comparativo del desempeño de diversos modelos basados en arquitecturas Transformer aplicados a tareas de clasificación de textos en lengua española, evaluando su eficacia mediante métricas estandarizadas y criterios metodológicos rigurosos.
- Formular recomendaciones prácticas para la selección de modelos en función de variables clave tales como el tipo de tarea, el tamaño y características del corpus, y la disponibilidad de recursos computacionales. Estas orientaciones estarán dirigidas a investigadores, desarrolladores y profesionales que trabajan en el ámbito del Procesamiento de Lenguaje Natural (PLN) en español.

En el plano pedagógico, se espera que los resultados obtenidos contribuyan a la actualización y fortalecimiento de los contenidos curriculares de las asignaturas Bases de Datos III y Análisis Numérico, correspondientes a la carrera de Ingeniería en Informática de la Facultad de Ciencias para la Innovación y el Desarrollo de la Universidad Católica de Santiago del Estero, promoviendo la incorporación de avances recientes en inteligencia artificial y aprendizaje automático.

En el ámbito social, la investigación aspira a contribuir a la democratización del acceso y uso de la inteligencia artificial en lengua española, mediante la producción de conocimiento riguroso, claro y accesible que oriente a usuarios y desarrolladores en la selección fundamentada de modelos adecuados a sus necesidades.

Finalmente, se proyecta que los resultados sean difundidos en ámbitos académicos y científicos, favoreciendo la consolidación de un grupo de investigación especializado, la formación de nuevos investigadores y el fortalecimiento de las competencias investigativas en estudiantes de grado.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de este proyecto está constituido por tres docentes investigadores. A través del proyecto, se contribuirá a la formación de estudiantes de grado en tareas vinculadas con investigación.

5. BIBLIOGRAFÍA

[1] Vaswani Ashish, Shazeer Noam, Parmar Niki, Uszkoreit Jakob, Jones Llion, Gomez Aidan N., Kaiser Lukasz, Polosukhin Illia (2017) "Attention is all you need," in Proc. 31st Conf. Neural Inf. Process. Syst., 2017, pp 6000-6010

[2] Lescano, Germán; Rodríguez, Juan Manuel; Costaguta, Rosanna (2025) “Deep Learning Techniques for Text-based Emotional Response Generation: A Systematic Review”, IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2025.3638957

[3] Devlin, J., et al. (2018). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding."

[4] Cárdenas, J., & Chiruzzo, L. (2020). "BETO: Spanish BERT for Text Classification."

[5] Conneau, A., et al. (2020). "Unsupervised Cross-lingual Representation Learning at Scale.", arXiv.

[6] Cañete, J., Donoso, S., Bravo-Marquez, F., Carvallo, A., & Araujo, V. (2022). ALBETO and DistilBETO: Lightweight Spanish Language Models. In Proceedings of the Thirteenth Language Resources and Evaluation Conference (pp. 4291–4298). European Language Resources Association.

Análisis de Flujo Vehicular utilizando Redes Neuronales

Diego A. Godoy[1,3,4,6], Javier Delucca[1,5], Lucas G. Kucuk[1,4], Jonathan M. Schuster [1,4] Nadia Freaza[1,5], Enrique M. Albornoz[2], Lucas M. Jardín[3]
[1]Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.) Universidad Gastón Dachary
[2]Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional sinc(i), UNL-CONICET, Santa Fe, Santa Fe
[3]Municipalidad de Posadas, Provincia de Misiones. Secretaria de Movilidad Urbana
[4]Instituto de Investigación en Regiones Transfronterizas Inteligentes (IRTI). Universidad Gastón Dachary
[5]Instituto de Investigación de Arquitectura Regional (IDEAR). Universidad Gastón Dachary
[6]Universidad Fuerzas de Seguridad de la Provincia de Misiones.

diegodoy@citic.ugd.edu.ar, javier.delucca.3789@gmail.com, lukucuk@citic.ugd.edu.ar, jonathanmschuster@gmail.com, freazanadia@gmail.com, emalbornoz@sinc.unl.edu.ar, mov.urb.lucasjardin@gmail.com

Resumen

Este trabajo propone el desarrollo de un prototipo de sistema de análisis de flujo vehicular para la Ciudad de Posadas, Misiones. El sistema construido permite la detección de sentido, giro y conteo de vehículos dado un video tomado con una cámara de vigilancia o un dron. El prototipo está basado en una red neuronal YoloV8 para la detección de los vehículos, ByteTrack para el tracking y los datasets VisDrone y COCO. Los videos utilizados son de una intersección de dos avenidas con mucha congestión que sirven de ingreso al casco céntrico de la ciudad. El sistema permite definir zonas de ingreso y egreso para contar los vehículos que ingresan al nudo y la dirección que toman, de tal manera que permite ver cuáles son los giros más utilizados cuando los vehículos ingresan a la intersección. La posibilidad de replicar esta solución en otros municipios es muy factible ya que todo el sistema ha sido construido con software libre y solo requiere de videos capturados con cualquier dispositivo.

Palabras claves: Análisis de Flujo Vehicular, Visión Artificial, YoloV8.

Contexto

Este trabajo presentado está enmarcado en el proyecto de investigación denominado "Tecnologías para Desarrollos Sostenibles de Ciudades Inteligentes apoyados por machine learning", registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) con el número Código IP A12001/22 (extendido hasta 2026) y radicado en el Centro de Investigación en Tecnologías de la Información y Comunicaciones de dicha universidad y el recién creado Instituto de Investigación en Regiones Trasfronterizas Inteligentes (IRTI) de UGD. Además, participan la Municipalidad de la Ciudad de Posadas y el Instituto de Investigación de Arquitectura Regional (IDEAR) también de la UGD. Todo esto para conseguir un trabajo interdisciplinario y con posibilidades de trasferencia.

Introducción

En el organigrama de la Municipalidad de Posadas se encuentra la Secretaría de Movilidad Urbana la cual se encarga de la Planificación, regulación y fiscalización en materia de movilidad y tránsito, en el ámbito de la ciudad. Una de sus funciones es la de realizar estudios del flujo vehicular y análisis de tráfico [1]. Como sabemos la seguridad vial comprende un conjunto de estrategias y medidas aplicadas para garantizar una movilidad segura para conductores y peatones. Sin embargo, esta tarea se complejiza cada vez más con el crecimiento del tránsito y las ciudades, lo que trae aparejado el aumento de riesgo de accidentes, congestión vial y contaminación, por lo cual es necesario contar con herramientas automatizadas que permitan hacer el análisis del flujo vehicular.

En experiencias previas, la Municipalidad realizaba estos conteos mediante agentes ubicados en las esquinas, quienes registraban manualmente cada vehículo y la dirección tomada. Sin embargo, este procedimiento resultaba sumamente dificultoso en jornadas de alto caudal vehicular y temperaturas elevadas propias de la región.

Si bien existen muchos softwares capaces de realizar análisis de flujo vehicular, como ejemplos podemos citar a Aimsun Live [2] es una solución de gestión predictiva del tráfico en tiempo real, que utiliza datos históricos y en vivo para simular y monitorear una red y proporcionar pronósticos inmediatos del estado del tráfico. Otra opción, son los sistemas de flujo de tráfico de Hikvision [3] monitorean las densidades y tasas de tráfico, procesan datos originales y muestran el estado del tráfico.

Asimismo, existen trabajos académicos como los desarrollados por Silvero et al. [4] en donde solo realizan un conteo vehicular y necesitan realizar un intenso trabajo de preparación de la imagen. Otros como Zhang et al. [5] presentan un sistema para detectar y contabilizar el número de vehículos en vídeos de vigilancia de tráfico basado en la Red Convolucional Basada en Regiones Rápidas (Fast R-CNN).

Para el desarrollo del prototipo que se presenta en este trabajo se realizaron los siguientes pasos utilizando las siguientes herramientas:

- Construcción del prototipo de detección El prototipo está basado en una red neuronal YoloV8 [6] para la detección de los vehículos, ByteTack [7] para el tracking y los dataset COCO [8] y visDrone [9].
- Configuración mediante código Python las zonas de detección, los contadores y pesos de las redes neuronales preentrendas.
- Configuración de la Fuentes de video: Los videos utilizados son tomados por un dron o provistos por el sistema 911 de la policía de Misiones.
- Ejecución el algoritmo de detección de flujo vehicular.
- Generar el video aumentado con los resultados de la detección y el tracking.
- Generar un reporte para su posterior análisis en formato de

tabla donde se cuenten los vehículos que ingresado por una zona y salido por otra.

En cuanto a las características del equipo utilizado, el mismo cuenta con un Procesador I3 con 16GB de RAM, disco SSD de 500 GB y una Placa de Video Nvidia GTX 1660 Súper. Se utilizó Ubuntu 20.04 como sistema operativo.

Línea de Investigación

En esta línea de investigación se han planteado diversos objetivos. El objetivo general consiste en desarrollar un prototipo que permita analizar el tráfico en intersecciones de calles y avenidas mediante técnicas de visión artificial. Este desarrollo, combinado con un enfoque interdisciplinario, permitirá detectar las vías y nodos estratégicos más conflictivos de la ciudad, así como las conexiones, ingresos y egresos más relevantes del flujo vehicular. A partir de las mediciones obtenidas se generará información clave para el diseño de estrategias de prevención de siniestros viales, el mejoramiento de la infraestructura y la señalización, el ajuste de la semaforización inteligente y otros aportes a la planificación de la movilidad urbana. Como objetivos específicos se plantean los siguientes:

- Estudiar el estado del arte de las soluciones existentes que utilizan visión artificial.
- Analizar la detección de vehículos y sentido de giro en intersecciones.
- Identificar, mediante un enfoque interdisciplinario, las vías, nodos críticos y determinar los horarios pico y valle de flujos vehiculares para detectar los nodos más conflictivos de la ciudad.
- Determinar los requerimientos del prototipo para detección para el conteo de vehículos que pasan por las distintas zonas, diseñar un prototipo que permita la gestión y de las zonas a partir de un video.
- Finalmente definir un conjunto de pruebas para verificar el funcionamiento del prototipo.

Resultados

En esta sección se detallan los avances realizados. En la Figura 1 se puede ver la interfaz de un video de diez minutos capturado con un dron de la intersección de las Avenidas Maipú y Uruguay en la ciudad de Posadas.

Por cada mano y cada arteria se definen cuadros o zonas de entrada y salida por donde se van a contar los vehículos que las atraviesan. Por ejemplo, la zona roja situada en parte superior central se muestran dos autos que están ingresando por la entrada en la parte izquierda y la Entrada Nro:0 y a la derecha el contador donde se identifica de que zona de entrada provienen los vehiculas que salan por la salida roja. En este caso 35 autos provienen de la entrada Verde Nro1 y 8 de Amarilla Nro:2. En caso del valor 26 de Entrada roja corresponde a falsos positivos ya un vehículo de la entrada de un determinado color sería imposible que saliera por salida del mismo color ya que no están permitidos los giros en “U”. Además, no se

detectaron salidas desde la entrada Azul Nro3 hacia la salida roja.

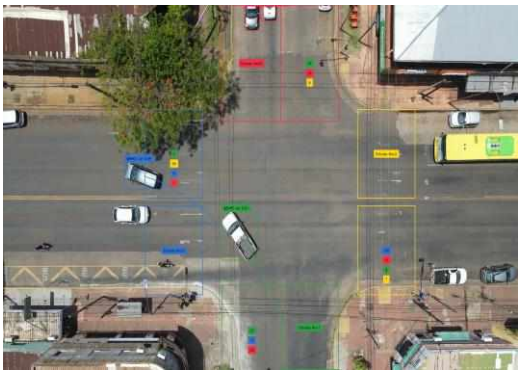


Fig. 1. Video tomado desde un dron aumentado con el análisis del flujo vehicular.

En la Figura 2 se pude ver la implementación, pero esta vez utilizando un recorte de quince minutos de una cámara del sistema 911 ubicada en la entrada del barrio Itaembé Guazú de la ciudad de Posadas. En este caso ha sido necesario definir polígonos y rectángulos de distintos tamaños y formas para poder intersecar la mayor cantidad de trayectorias de vehículos que pasan por las zonas de entrada y salida.



Fig. 2. Video aumentado de una cámara de vigilancia del sistema 911 de Policía de Misiones

En la Figura 3 se puede ver la interfaz desarrollada Utilizando una librería grafica de Python tkinter [10].

En la misma se pueden especificar todos los parámetros del modelo y de la detección (como ser ruta del archivo preentrenado de Yolo y la del video a analizar, el umbral de confianza y de IOU, la duración de los rastros en los frames y si se desea guardar el video de salida aumentado y verlo en pantalla mientras se realiza el conteo y detección de giros.), así como ver las clases que pueden ser detectadas según lo cargado en la red.



Fig. 3. Interfaz gráfica para el análisis del Flujo Vehicular.

Formación de Recursos Humanos

El equipo encargado del desarrollo de este trabajo se encuentra formado Un Doctor en Tecnología de la Información y Comunicación y un Doctor en Ciencia y Tecnología, Un Magíster en Tecnologías de la Información. Hay cuatro estudiantes en instancia de realización de trabajo final de la carrera de Ingeniería en Informática de la UGD. También participan del Proyecto. Arquitectos e Ingenieros Civiles con Formación de Posgrado. Además, el número de trabajos finales de carrera finalizados en este proyecto es de nueve y en curso de dos.

Bibliografía.

[1] R. P. P. E. S. &. M. W. R. Roess, Traffic engineering. (4th ed.), Prentice Hall, 2011.

[2] AIMSUM, «Aimsun-live,» 2024. [En línea]. Available: https://www.aimsun.com.

[3] Hikvision, «Hikvision Traffic-Flow-Systems,» 2025. [En línea]. Available: https://www.hikvision.com/es/products/ITS-

Products/Traffic-Flow-Systems/Traffic-Flow-Detection/.

[4] F. Silvero, G. H. Devincenzi y G. R. Di Rado, «Conteo vehicular mediante visión artificial,» 2022.

[5] K. L. F. G. X. L. a. G. W. Z. Zhang, «Vision-based vehicle detecting and counting for traffic flow analysis,» de *International Joint Conference on Neural Networks (IJCNN)*, Vancouver, 2016.

[6] Ultralitics, «YoloV8,» 2024. [En línea]. Available: https://docs.ultralytics.com/es/models/yolov8/.

[7] Y. a. S. P. a. J. Y. a. Y. D. a. W. F. a. Y. Z. a. L. P. a. L. W. a. W. X. Zhang, «ByteTrack: Multi-Object Tracking by Associating Every Detection Box,» *Proceedings of the European Conference on Computer Vision (ECCV),* 2022.

[8] T. Y. M. M. B. S. H. J. P. P. R. D. .. &. Z. C. L. Lin, «Microsoft coco: Common objects in context.,» *Computer vision–ECCV 2014: 13th European conference,* 2014.

[9] P. a. W. L. a. D. D. a. B. X. a. F. H. a. H. Q. a. L. H. Zhu, «Detection and Tracking Meet Drones Challenge,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2021.

[10] M. Roseman, Modern Tkinter, 2025.

“Aplicación de técnicas de IA y modelos de lenguaje para la recuperación de información en colecciones documentales jurídicas”

Osvaldo Spositto[1], Julio Bossero[1], Viviana Ledesma[1], Lorena Matteo[1], Sebastián Quevedo[1], Carlos Colombain[1], Luis Busnelli[2], Laura Conti[2], Gastón Procopio[1], Cecilia Gargano[1], Gerardo Frega[2], Sergio García[3].
[1]Universidad Nacional de La Matanza. Departamento de Ingeniería e Investigación Tecnológicas. {spositto, jbossero, vledesma, gprocopio, lmatteo, cgargano, ccolombain, squevedo}@unlam.edu.ar
[2] Universidad Nacional de La Matanza. Departamento Derecho y Ciencia Política. {lbusnelli, lconti, gfrega}@unlam.edu.ar
[3] Palacio de Tribunales. Departamento Judicial de Morón. Alte. Brown. Piso 4. Morón. sergiogabriel.garcia@pjba.gov.ar

RESUMEN

Este trabajo presenta la aplicación de técnicas de *Inteligencia Artificial* (IA) para la recuperación de información y el análisis documental en contextos de alta complejidad, en particular en el ámbito judicial. Se investiga el uso de *Modelos de Lenguaje de Gran Escala* (LLM) combinados con *Generación Aumentada por Recuperación* (RAG) y *IA Explicable* (XAI) para el procesamiento automatizado de documentos jurídicos. La propuesta contempla la preparación y estructuración de corpus legales, la vectorización semántica y la recuperación contextual de normativa y antecedentes, a fin de generar respuestas fundamentadas en evidencia verificable. Asimismo, incorpora trazabilidad, coherencia normativa y control del contexto, evaluando la integración de protocolos como el *Model Context Protocol* (MCP) para reforzar transparencia y auditabilidad. El enfoque se orienta a asistir al operador jurídico sin sustituir el juicio humano.

Palabras clave: LLM, RAG, XAI, Recuperación de Información Jurídica.

CONTEXTO

El equipo de trabajo ha desarrollado en los últimos años diversos proyectos orientados al procesamiento automatizado de grandes volúmenes de documentos y la recuperación eficiente de información. Entre ellos se encuentran: *Justicia Tecnológica: Reforzando Experticia con Redes Neuronales para una Toma de Decisiones más Eficiente* (C262/PII, 2024–2025) y *Optimización de un Sistema de Recuperación de Información para Expedientes Digitalizados en la UNLaM: Integración de un Algoritmo K-Means Mejorado y Exploración de Técnicas Avanzadas* (C268, 2025–2026) [1]. Ambos aportaron avances en clasificación, estructuración documental y métodos de acceso mejorado a grandes corpus

Actualmente, el grupo participa del proyecto PROINCE 2026–2027 *“Desarrollo de una Plataforma basada en Agentes Inteligentes para el Análisis Automatizado de Expedientes Judiciales*”, que incorpora LLM, recuperación semántica y mecanismos de trazabilidad en tareas de análisis documental automatizado.

En estas investigaciones participan

equipos de los Departamentos de Ingeniería e Investigaciones Tecnológicas (DIIT) y de Derecho y Ciencia Política de la Universidad Nacional de La Matanza (UNLaM), junto con el apoyo técnico de la Subsecretaría de Tecnología Informática del Poder Judicial de la Provincia de Buenos Aires.

1. INTRODUCCIÓN

La incorporación de la IA en el ámbito jurídico ha despertado un interés creciente en la literatura especializada, impulsado por el aumento exponencial del volumen normativo y jurisprudencial y por la necesidad de optimizar los procesos de análisis y toma de decisiones. Como señalan Parra Sepúlveda y Concha Machuca [2], el acelerado desarrollo tecnológico —y, en particular, el avance de la IA— está produciendo transformaciones profundas en las estructuras jurídicas, generando oportunidades relevantes, pero también tensiones para los sistemas normativos. Entre los principales desafíos se destacan la adaptación del derecho a realidades emergentes, la necesidad de marcos regulatorios flexibles y la protección efectiva de los derechos fundamentales frente a decisiones automatizadas.

En este contexto, las técnicas de aprendizaje automático (*machine learning* ML) y de procesamiento del lenguaje natural (*natural language processing* NLP) se consolidan como herramientas estratégicas para la gestión masiva del acervo normativo y jurisprudencial. Según las directrices de la Comisión Europea [3], estas tecnologías permiten procesar grandes volúmenes de datos, estructurar expedientes, recuperar antecedentes relevantes e identificar regularidades útiles para el análisis jurídico, especialmente en escenarios donde la complejidad documental supera la capacidad humana de revisión exhaustiva. Esta línea de investigación se inscribe en una tradición académica consolidada del campo de la IA y el derecho. Como destaca Casanovas [4], el área ha evolucionado desde la informática jurídica y los sistemas expertos hacia enfoques centrados en la representación del conocimiento, el razonamiento legal y la construcción de estructuras argumentativas, habilitando análisis de mayor granularidad sobre el derecho positivo y sus interpretaciones.

No obstante, la aplicación de estas tecnologías en dominios sensibles como el judicial exige un marco de "*IA fiable*" centrado en la persona. Las políticas europeas [5] subrayan la necesidad de garantizar trazabilidad, transparencia y supervisión humana significativa, de modo que la automatización opere como apoyo —y no como reemplazo— del juicio profesional. Como advierte Cotino [6], la opacidad y los sesgos de los sistemas de IA pueden afectar principios estructurales del derecho, como la igualdad ante la ley, la motivación de los actos y las garantías del debido proceso.

Frente a estos desafíos, la investigación actual converge hacia enfoques de XAI, arquitecturas híbridas y mecanismos de recuperación de información que permitan fundamentar los resultados en evidencia documental verificable. En esta línea, Verheij [7] sostiene que los sistemas de IA aplicados al derecho deben articular razones, pruebas y contraargumentos de manera transparente, reforzando el rol del operador jurídico y promoviendo una integración gradual y controlada de estas tecnologías en el ámbito judicial.

El presente trabajo presenta los avances derivados de estos desarrollos, focalizados en el diseño de una arquitectura basada en LLM y RAG para apoyar la recuperación contextualizada de información en expedientes extensos.

2. LÍNEA de INVESTIGACIÓN y DESARROLLO

La línea de investigación principal del grupo se centra en la aplicación de técnicas de IA para la recuperación de información en grandes colecciones documentales jurídicas.

Dentro de este marco se desarrollan herramientas basadas en LLM y agentes inteligentes para analizar y contextualizar documentos complejos. En el campo del NLP, han demostrado una capacidad significativa para modelar relaciones semánticas complejas y manejar la ambigüedad inherente al lenguaje natural. Entrenados bajo el paradigma de modelos fundacionales [8], estos sistemas permiten contextualizar significados e identificar dependencias de largo alcance en documentos extensos. Asimismo, su utilización como componente central de agentes inteligentes facilita la recuperación de información relevante incluso cuando esta se encuentra implícita, una característica especialmente relevante para dominios de alta complejidad como el jurídico [9].

A partir de esta línea general, el trabajo aborda los siguientes desarrollos:

a) Preparación y estructuración del corpus jurídico.

Se implementan procesos de limpieza, normalización lingüística y segmentación del texto (*chunking*), orientados a organizar el contenido relevante de resoluciones judiciales. Este proceso contempla la anonimización de los documentos y la diferenciación de secciones clave —tales como hechos del caso, normativa citada y fundamentación del fallo— con el objetivo de asegurar la calidad de las etapas posteriores de recuperación y análisis semántico [8–10].

b) Vectorización semántica de documentos jurídicos.

Cada segmento textual es transformado en una representación vectorial densa (*embedding*) mediante modelos de vectorización semántica, tales como *sentence-transformers* u otros modelos especializados. Esta representación permite medir la similitud conceptual entre fragmentos del corpus y habilita búsquedas contextuales que superan la coincidencia puramente léxica, constituyendo un componente esencial para la recuperación de información relevante [9–10].

c) Desarrollo de una arquitectura de Generación Aumentada por Recuperación (RAG).

La arquitectura propuesta se basa en la recuperación de fragmentos normativos, jurisprudenciales y fácticos relevantes a partir del corpus vectorizado, con el fin de construir un contexto jurídico confiable que alimente al LLM. Las consultas se formulan mediante *prompts* instructivos que inducen al modelo a razonar sobre la evidencia recuperada y a fundamentar sus respuestas mediante referencias explícitas a las fuentes documentales utilizadas [2, 8, 9].

d) Integración de mecanismos de explicabilidad.

El sistema incorpora técnicas de XAI como componente transversal, orientado a garantizar la trazabilidad y coherencia del razonamiento generado. Entre los mecanismos considerados se incluyen la estructuración lógica de las respuestas, la identificación explícita de los fragmentos documentales utilizados y la verificación de consistencia normativa entre la respuesta generada y el material jurídico recuperado, permitiendo evaluar la correspondencia entre las salidas del prototipo y los fundamentos empleados en las decisiones judiciales [8, 10].

La arquitectura propuesta se basa en el paradigma de RAG, integrando LLM con mecanismos explícitos de recuperación semántica de información jurídica. Este enfoque permite condicionar la generación de respuestas a evidencia documental verificable, aspecto central

para la auditabilidad y confiabilidad de los sistemas de apoyo a la decisión en el dominio jurídico.

El sistema se concibe como una arquitectura funcional basada en agentes inteligentes especializados, donde cada componente encapsula una función cognitiva específica dentro del proceso de análisis jurídico automatizado. El flujo de procesamiento se inicia con la conformación de un corpus jurídico, integrado por sentencias, normas y dictámenes, que es sometido a tareas de limpieza, normalización y segmentación, generando unidades textuales adecuadas para su representación semántica.

Los segmentos resultantes son transformados en representaciones vectoriales densas (*embeddings*) que capturan relaciones de similitud contextual relevantes para el razonamiento jurídico. Estas representaciones se almacenan en una base de datos vectorial, permitiendo la recuperación eficiente de información contextualizada ante consultas formuladas por el usuario.

Ante una solicitud de información jurídica, el sistema ejecuta un proceso de recuperación semántica que identifica los fragmentos más relevantes del corpus en función de su similitud con la consulta. Este conjunto reducido de antecedentes se incorpora como contexto de entrada al modelo de lenguaje, condicionando la generación de la respuesta sobre la base de fuentes documentales concretas, en línea con el enfoque propuesto por Lewis [11].

La salida del modelo incluye el texto generado junto con referencias explícitas a las fuentes utilizadas y metadatos de trazabilidad que permiten reconstruir el proceso de recuperación. Sobre este resultado se aplican mecanismos iniciales de explicabilidad, orientados al análisis de la coherencia interna de la respuesta y su consistencia con el material jurídico recuperado [12].

Desde el punto de vista operativo, los distintos agentes del sistema pueden interactuar de manera desacoplada e iterativa, especialmente en los procesos de recuperación y análisis de explicabilidad. En todos los casos, la arquitectura se concibe como una herramienta de asistencia al operador jurídico, manteniendo la supervisión humana como principio rector del diseño.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera que el desarrollo de este trabajo aporte evidencias concretas sobre la aplicación de arquitecturas basadas en LLM al dominio jurídico, manteniendo un equilibrio entre eficiencia técnica y rigor normativo. En particular, los resultados obtenidos y esperados incluyen:

- La definición de un flujo metodológico para la preparación y estructuración de corpus jurídicos extensos, compatible con arquitecturas basadas en LLM y mecanismos de gestión de contexto como el MCP.
- Una evaluación exploratoria de técnicas de vectorización semántica aplicadas a documentos judiciales, analizando su impacto en la recuperación contextual de información dentro de flujos RAG.
- El diseño y validación preliminar de una arquitectura RAG para el análisis de expedientes, orientada a la generación de respuestas fundamentadas en evidencia documental recuperada.
- La incorporación de mecanismos iniciales de explicabilidad (XAI) para analizar la coherencia, trazabilidad y consistencia normativa de las respuestas generadas, en concordancia con los requerimientos

de transparencia del dominio jurídico.

- La obtención de evidencias empíricas preliminares sobre el potencial de estas tecnologías como herramientas de asistencia para la lectura y análisis de expedientes judiciales complejos, sin sustituir el rol del operador jurídico.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por 14 integrantes, de los cuales 11 son investigadores formados y 3 son investigadores en formación, trabajando en el área de la recuperación de la información y la IA. Además, 4 integrantes están cursando actualmente el Doctorado en Ingeniería con Mención en Tecnología de la Información en UNLaM.

5. BIBLIOGRAFÍA

[1] Spositto, et. al. (2025). Líneas de investigación complementarias en el marco de proyectos de Aplicación de Tecnologías de Inteligencia Artificial. Disponible https://sedici.unlp.edu.ar/handle/10915/182261.

[2] Parra Sepúlveda, D. and Concha M., R. (2021) "Inteligencia artificial y derecho. Problemas, desafíos y oportunidades", *Vniversitas*, 70, pp. 1–25. doi:10.11144/Javeriana.vj70.iadp.

[3]Comisión Europea. (2018). Inteligencia artificial para Europa (Comunicación COM. 237 final). EUR-Lex. Recuperado de. Recuperado de: https://eur-lex.europa.eu/legal-content/es/txt/pdf/?uri=celex:52018DC0237

[4] Casanovas, P. (2010). Inteligencia artificial y derecho: a vuelapluma. Teoría & Derecho. Revista de Pensamiento Jurídico, (7), 203–221. https://teoriayderecho.tirant.com/index.php/teoria-y-derecho/article/view/248/244

[5] Comisión Europea. (2019). Generar confianza en la inteligencia artificial centrada en el ser humano (Comunicación COM (2019) 168 final). EUR-Lex. Recuperado de: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52019DC0168

[6]Cotino, L. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho. Revista Catalana de dret públic, (58), 29–48. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=7005055&orden=0&info=link

[7]Verheij, B. (2020). Artificial Intelligence as Law. Artificial Intelligence and Law, 28, 181–206. Disponible en: https://doi.org/10.1007/s10506-020-09267-0

[8]Corredera, J. C. (2023). Inteligencia artificial generativa. In Anales de la Real academia de Doctores (Vol. 8, No. 3, pp. 475-489). Disponible en: https://www.rade.es/imageslib/publicaciones/articulos/V8N3%20-%2001%20-%20ED%20-%20CASAR.pdf

[9]Wang, Y., Pan, Y., Su, Z., Deng, Y., Zhao, Q., Du, L., Luan, T. H., Kang, J., & Niyato, D. (2025). Large model based agents: State-of-the-art, cooperation paradigms, security and privacy, and future trends. arXiv. https://arxiv.org/abs/2409.14457.

[10]Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2024). A comprehensive overview of large language models. Preprint en arXiv. https://arxiv.org/abs/2307.06435.

[11]Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. arXiv preprint arXiv:2005.11401.

[12] Priya Desai. (2021). A Comprehensive Survey on Explainable Artificial Intelligence (XAI): Challenges, Opportunities, and Future Research Directions. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, *2*(2), 1-10. https://doi.org/10.63282/3050-9262.IJAIDSML-V2I2P101

Argumentación Abstracta y Análisis Contextual de Debates Complejos

Irene M. Coronel[1] Gerardo I. Simari[1] Maximiliano C. D. Budán[2]

[1] Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS) & Instituto de Ciencias e Ingeniería de la Computación (ICIC UNS-CONICET)

[2] Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero & CONICET

{irene.coronel, gis}@cs.uns.edu.ar, mcdb@unse.edu.ar

RESUMEN

La argumentación abstracta se ha consolidado como un formalismo adecuado para modelar el razonamiento no monótono y conflictos entre información heterogénea. Sin embargo, los marcos clásicos presentan limitaciones al representar debates complejos en los que los argumentos se relacionan con múltiples temas, poseen distinta relevancia contextual o requieren ser evaluados localmente. Esta línea de investigación propone extensiones a los modelos de argumentación abstracta orientadas a enriquecer la representación y evaluación de argumentos mediante información adicional, tales como etiquetas temáticas y funciones de valoración, incorporando mecanismos de análisis contextual. El objetivo es permitir un razonamiento más estructurado y sensible al contexto, adecuado para escenarios con información potencialmente contradictoria y semánticamente diversa. Los desarrollos teóricos incluyen la definición de marcos argumentativos enriquecidos, el estudio de propiedades semánticas relevantes y el análisis del impacto de la focalización temática en la aceptación de argumentos. Estos aportes sientan bases conceptuales para la evaluación diferenciada de argumentos y el soporte a procesos de toma de decisiones asistida, sin limitarse a esquemas de clasificación puramente automáticos. Finalmente, se discuten posibles líneas de aplicación del enfoque en dominios complejos de análisis de evidencias, tales como la ciberseguridad, donde la interpretación de eventos requiere considerar explicaciones alternativas y justificar conclusiones en contextos dinámicos.

Palabras clave: Argumentación abstracta, Razonamiento no monótono, Focalización temática, Valoración de argumentos.

CONTEXTO

Esta línea de investigación se desarrolla en el marco del proyecto "Herramientas de razonamiento automatizado para el desarrollo de sistemas socio-técnicos inteligentes" (PGI 24/N057) dirigido por Gerardo I. Simari, financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional del Sur (UNS) para el período 2023-2026 y radicado en el Departamento de Ciencias e Ingeniería de la Computación.

1. INTRODUCCIÓN

Los sistemas argumentativos han demostrado ser una herramienta fundamental dentro de la inteligencia artificial para modelar razonamiento no monótono, gestionar información inconsistente y analizar conflictos entre distintas piezas de conocimiento (Dung 1995). En particular, los marcos de argumentación abstracta y los sistemas argumentativos estructurados han sido ampliamente utilizados en dominios como el soporte a la toma de decisiones (Amgoud 2009), el razonamiento legal (T. a. Bench-Capon 2003, T. J. Bench-Capon 2003), los sistemas de recomendación (Chesñevar 2004) y los sistemas multi-agente (Parsons 2002, Rahwan 2008), entre otros.

Una de las principales capacidades de estos enfoques radica en la representación explícita de relaciones de ataque, soporte y agregación entre argumentos, permitiendo justificar conclusiones a partir de información potencialmente contradictoria proveniente de múltiples fuentes (Verheij 1995, Cayrol 2005, Prakken 2005). Sin embargo, en escenarios complejos del mundo real, los debates rara vez se desarrollan sobre un único tema homogéneo. Por el contrario, los argumentos suelen involucrar múltiples dimensiones conceptuales, distintos grados de relevancia contextual y efectos que no necesariamente se propagan de manera uniforme a lo largo de todo el sistema argumentativo.

En este contexto, los marcos argumentativos clásicos presentan limitaciones para mantener el foco del razonamiento en los temas centrales de una discusión, así como para delimitar el alcance de los efectos positivos y negativos que los argumentos ejercen entre sí. La falta de mecanismos explícitos para capturar nociones como cercanía temática, centralidad o relevancia contextual puede conducir a análisis excesivamente globales, en los que argumentos marginales influyen de manera desproporcionada sobre las conclusiones alcanzadas (Walton 2012, Baroni 2014, Hunter 2021).

En consecuencia, emerge un problema conceptual y formal: ¿cómo extender los marcos argumentativos clásicos para representar explícitamente la estructura temática y la relevancia contextual de los argumentos, de modo que la influencia entre ellos refleje su verdadera proximidad conceptual y no únicamente su conectividad estructural? Sin mecanismos que delimiten el alcance de la interacción argumentativa, los modelos existentes pueden generar evaluaciones excesivamente globales, donde efectos locales se propagan sin distinción a través de todo el sistema, afectando conclusiones que deberían permanecer conceptualmente independientes

La línea de investigación aquí presentada busca abordar estas limitaciones mediante el desarrollo de extensiones de los marcos argumentativos existentes, incorporando información adicional asociada a los argumentos, como etiquetas temáticas y funciones de valoración. Estos enriquecimientos permiten modelar de manera más precisa la estructura interna de los debates y habilitan formas de razonamiento contextual y localizado, en las que la aceptación de un argumento depende no sólo de las relaciones estructurales de ataque o defensa, sino también de su vinculación con los temas del dominio considerados como relevantes (M. C. Budán 2017, Escañuela Gonzalez 2021, Gonzalez 2021).

Desde un punto de vista teórico, esta línea explora la incorporación de conceptos provenientes de áreas afines, como la topología y la noción de vecindad, para caracterizar la proximidad entre argumentos y analizar el impacto de la focalización temática en distintas semánticas de aceptabilidad (M. C. Budán 2015, M. C. Budán 2013). Desde una perspectiva aplicada, estos desarrollos sientan las bases para el diseño de sistemas de razonamiento capaces de asistir procesos de toma de decisiones complejos, proporcionando evaluaciones más diferenciadas y justificables, especialmente en contextos complejos donde intervienen múltiples criterios, fuentes de información y posibles interpretaciones.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación se articula en torno a varios ejes complementarios que abordan, desde una perspectiva teórica y aplicada, el enriquecimiento del razonamiento argumentativo en contextos complejos.

En particular, se investigan los siguientes ejes:

- **Extensión de marcos de argumentación abstracta mediante información adicional.** Este eje se centra en el estudio y desarrollo de extensiones a los marcos abstractos, incorporando meta-información asociada a los argumentos con el fin de

enriquecer su evaluación sin abandonar el nivel de abstracción del formalismo. En particular, se analizan mecanismos de etiquetado y funciones de valoración que permiten capturar aspectos contextuales relevantes del dominio, habilitando formas de evaluación argumentativa más flexibles y sensibles al contexto.

- **Focalización temática y razonamiento contextual.** Un segundo eje aborda la necesidad de mantener el foco del razonamiento en los temas centrales de una discusión. Para ello, se investiga la organización de argumentos en torno a temas o dimensiones conceptuales, permitiendo analizar ataques y defensas de manera localizada y contextual. Este enfoque busca limitar la influencia de argumentos marginales y modelar de forma más precisa el alcance efectivo de los efectos argumentativos en debates multi-temáticos.
- **Análisis de relevancia, centralidad y valoración de argumentos.** Este eje estudia criterios para evaluar la importancia relativa de los argumentos dentro de un sistema argumentativo, considerando tanto su posición estructural como su relación con los temas centrales del dominio. Se analizan nociones de cercanía, centralidad y agregación, así como su impacto en distintas semánticas de aceptabilidad, con el objetivo de jerarquizar la información y refinar las conclusiones obtenidas.
- **Proyección hacia sistemas de soporte a la toma de decisiones y dominios aplicados.** Finalmente, la línea contempla la aplicación de los desarrollos teóricos en el diseño de sistemas de razonamiento que asistan procesos de toma de decisiones complejos. Estos sistemas buscan proporcionar evaluaciones argumentativas más diferenciadas y justificables, facilitando la interacción entre mecanismos automáticos y analistas humanos en dominios caracterizados por información abundante, heterogénea y potencialmente contradictoria.

Estos ejes conforman una línea de trabajo integrada orientada a enriquecer los procesos de razonamiento argumentativo mediante mecanismos estructurados y contextuales. En conjunto, buscan complementar los enfoques clásicos de evaluación argumentativa, proporcionando herramientas conceptuales para el análisis diferenciado de argumentos y el soporte a procesos de toma de decisiones en escenarios complejos.

3. RESULTADOS OBTENIDOS y ESPERADOS

En el marco de esta línea de investigación se han desarrollado y formalizado extensiones del enfoque clásico de argumentación abstracta, orientadas a representar debates reales caracterizados por la presencia de múltiples tópicos y relaciones contextuales entre los argumentos.

Los resultados obtenidos hasta el momento incluyen la incorporación de información temática asociada a los argumentos, así como la introducción de nociones de vecindad y comunidad argumentativa, que permiten modelar interacciones locales y patrones colectivos dentro del debate (Coronel 2024). Esto hace posible refinar la evaluación de los argumentos manteniendo el marco abstracto original, y han demostrado ser adecuados para capturar dinámicas que no pueden ser representadas mediante marcos abstractos tradicionales.

El trabajo en curso se orienta a profundizar y extender este enfoque mediante el estudio de mecanismos de debilitamiento de ataques dependientes del contexto temático, así como el diseño de funciones de valoración que permitan priorizar argumentos en función de información contextual adicional, sin requerir el acceso a su estructura interna. Estas extensiones buscan incrementar la expresividad del marco manteniendo su nivel de abstracción.

Asimismo, se prevé avanzar en el análisis de propiedades teóricas relevantes de estos

marcos abstractos enriquecidos, así como en el estudio de su comportamiento frente a cambios en la información disponible. Como objetivo general, esta línea de investigación apunta a contribuir al desarrollo de modelos de razonamiento argumentativo abstracto más expresivos y explicables, capaces de complementar mecanismos automáticos de detección existentes mediante decisiones mejor justificadas. En coherencia con estos objetivos, en trabajos futuros se considera de interés estudiar el comportamiento del enfoque propuesto en escenarios de análisis de eventos complejos, particularmente en dominios como la ciberseguridad, donde la información disponible suele dar lugar a interpretaciones alternativas y potencialmente engañosas. En este contexto, el razonamiento argumentativo permite modelar y evaluar hipótesis relacionadas con conclusiones benignas y maliciosas en función de su coherencia contextual y su interacción con el resto de la evidencia, favoreciendo una mayor resiliencia frente a explicaciones débiles o insuficientemente respaldadas cuando el entorno global indica riesgo. Asimismo, el enfoque ofrece una base conceptual para priorizar casos y alertas que requieren atención humana, organizando el análisis en función de la fuerza argumentativa y la relevancia contextual de la información, sin pretender reemplazar los mecanismos automáticos de detección existentes. Estos esfuerzos se alinean con trabajos previos que exploran el uso de modelos argumentativos para el análisis de amenazas y la atribución en ciberseguridad (Nunes, y otros 2016, Leiva, y otros 2019), y abre la posibilidad de evaluar su aplicabilidad sobre conjuntos de datos reales ampliamente utilizados en la literatura académica y el desarrollo de herramientas prácticas utilizadas en la industria.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación se desarrolla en el marco del tópico más general de Razonamiento bajo Incertidumbre que se lleva adelante en el *Grupo de Investigación en Ingeniería Cognitiva y Bases de Datos* (ICBD), el cual es dirigido por Gerardo I. Simari y forma parte del Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) de la Universidad Nacional del Sur, con colaboración del Instituto de Ciencias e Ingeniería de la Computación (ICIC), instituto de doble dependencia del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y la UNS. El grupo posee sólidas conexiones de colaboración con el exterior, en particular con grupos de investigación que estudian temas afines en instituciones de Estados Unidos (Syracuse University y Northwestern University), Italia (Università della Calabria), España (IIIA-CSIC Barcelona), Inglaterra (University of Oxford y Alan Turing Institute), Austria (TU Vienna), entre otros, como así también a nivel nacional, como la Universidad de Buenos Aires (UBA), Universidad Nacional de Entre Ríos (UNER) y Universidad Nacional de Santiago del Estero (UNSE) y Universidad Nacional de San Luis (UNSL).

Esta línea de investigación forma parte del núcleo central de la tesis doctoral de Irene M. Coronel, tesista del DCIC UNS dirigida por Gerardo I. Simari y Maximiliano C. D. Budán y becaria CONICET. Se prevé que la misma será continuada durante el período posdoctoral y sumará a futuro a otros tesistas (de grado o posgrado) que se interesen en los diferentes ejes descriptos.

5. BIBLIOGRAFÍA

Amgoud, Leila & Prade, Henri. 2009. «Using arguments for making and explaining decisions.» *Artificial Intelligence*, 173(3-4), 413-436.

Baroni, Pietro, Massimiliano Giacomin, and Beishui Liao. 2014. «On topology-related properties of abstract argumentation semantics. A correction and extension to dynamics of argumentation systems: A division-based method.» *Artificial Intelligence (212)* 104-115.

Bench-Capon, Trevor JM. 2003. «Persuasion in practical argument using value-based argumentation frameworks.» *Journal of Logic and Computation* 13(3), 429-448.

Bench-Capon, Trevor, and Giovanni Sartor. 2003. «A model of legal reasoning with cases incorporating theories and values.» *Artificial Intelligence* 150(1-2), 97-143.

Budán, M. C., Cobo, M. L., Martinez, D. C., & Simari, G. R. 2017. «An Approach to Characterize Topic-Centered Argumentation.» *Proc. PRIMA.* Springer International Publishing. 459-467.

Budán, M. C., Lucero, M. G., Chesñevar, C., & Simari, G. R. 2015. «Modeling time and valuation in structured argumentation frameworks.» *Inf. Sci. (290)* 22-44.

Budán, M. C., Lucero, M. J. G., & Simari, G. R. 2013. «Modeling reliability varying over time through a labeled argumentative framework.» *Proc. WL4AI (26).*

Cayrol, Claudette, and Marie-Christine Lagasquie-Schiex. 2005. «On the acceptability of arguments in bipolar argumentation frameworks.» *Proc. ECSQARU.* Springer Berlin Heidelberg. 378-389.

Chesñevar, Carlos I., Maguitman, Ana G., and Simari, Guillermo. 2004. «A first approach to argument-based recommender systems based on defeasible logic programming.» *Proc. NMR* 109-117.

Coronel, I. M., Gonzalez, M. G. E., Martinez, D. C., Simari, G. I., & Budán, M. C. 2024. «Neighborhood-based argumental community support in the context of multi-topic debates.» *International Journal of Approximate Reasoning (170)* 109189.

Dung, Phan Minh. 1995. «On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games.» *Artificial intelligence* 77(2), 321-357.

Escañuela Gonzalez, M. G., Budán, M. C., Martínez, D. I., Cobo, M. L., & Simari, G. R. 2021. «Focusing the argumentative process: neighborhood-based semantics in abstract argumentation.» *Proc. CLAR.* Springer International Publishing. 20-39.

Gonzalez, M. G. E., Budán, M. C., Simari, G. I., & Simari, G. R. 2021. «Labeled bipolar argumentation frameworks.» *JAIR (70)* 1557-1636.

Hunter, A., Polberg, S., Potyka, N., Rienstra, T., & Thimm, M. 2021. «Probabilistic argumentation: A survey.» En *Handbook of Formal Argumentation (2)*, 397-441.

Leiva, M. A., Simari, G. I., Simari, G. R. & Shakarian, P. 2019. «Cyber Threat Analysis with Structured Probabilistic Argumentation.» *Proc. AI³@AIIA**, 50-64.

Nunes, E., Shakarian, P., Simari, G. I. & Ruef, A. 2016. «Argumentation models for cyber attribution.» *Proc. IEEE/ACM ASONAM*, 837-844.

Parsons, S., and Wooldridge, M.. 2002. « Game theory and decision theory in multi-agent systems.» *Autonomous Agents and Multi-Agent Systems* 5(3), 243-254.

Prakken, Henry. 2005. «Coherence and flexibility in dialogue games for argumentation.» *Journal of logic and computation 15(6)* 1009-1040.

Rahwan, I., and Larson, K.. 2008. «Mechanism design for abstract argumentation.» *Proc. AAMAS (pp. 1031-1038). IFAAMAS.*

Verheij, B. 1995. «Accrual of arguments in defeasible argumentation.» *Proceedings of the Second Dutch/German Workshop on Nonmonotonic Reasoning.* 217-224.

Walton, D. 2012. «Story similarity in arguments from analogy.» *Informal Logic* 32(2):190-221.

ARQUITECTURA INTELIGENTE PARA LA GESTIÓN PROACTIVA DE LA MORA TRIBUTARIA EN ADMINISTRACIONES SUBNACIONALES

[1] BENEDETTO Marcelo Gabriel, [2] MARTÍNEZ, Roxana

[1] Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos - Monseñor Tavella 1424 – Concordia, Entre Ríos (3200) - Tel.: +54(0345)4231433 - marcelo.benedetto@uner.edu.ar

[2] Facultad de Tecnología Informática - Universidad Abierta Interamericana - Avenida San Juan 951 - 10º piso Capital Federal (C1147AAU) - Tel.: (+54) 11 4363-3500 - Roxana.Martinez@uai.edu.ar

RESUMEN

La mora tributaria constituye un problema estructural para las administraciones tributarias subnacionales debido a su impacto directo en la sostenibilidad financiera, la equidad fiscal y la eficiencia del gasto público.

En la práctica, la gestión de la mora continúa sustentándose en procedimientos predominantemente reactivos y reglas estáticas, con fuerte dependencia de tareas manuales, lo que limita la efectividad institucional frente a contextos económicos dinámicos y contribuyentes heterogéneos.

Paralelamente, la disponibilidad de grandes volúmenes de datos y los avances en analítica e inteligencia artificial han ampliado el potencial de las administraciones tributarias para anticipar el incumplimiento, segmentar contribuyentes por perfiles de riesgo y automatizar procesos, con mejoras reportadas en eficiencia y control.

Este trabajo presenta una línea de investigación orientada al diseño de una arquitectura inteligente integral basada en Big Data e Inteligencia Artificial para gestionar la mora de manera proactiva en administraciones subnacionales.

La propuesta combina modelos predictivos supervisados, segmentación automatizada no supervisada y mecanismos de automatización de seguimiento y trazabilidad, integrados en una solución modular adaptable a cambios normativos.

El proyecto se enmarca en la tesis doctoral "Desarrollo de un modelo de Big Data e IA para mejorar la eficiencia en la gestión de la cobranza de las Administradoras Tributarias" (Doctorado en Informática, UAI), constituyendo su núcleo metodológico y empírico.

Se discuten contexto, motivación, fundamentos técnicos, líneas de desarrollo, resultados esperados e impacto en la formación de recursos humanos, vinculando la propuesta con la Transformación Digital del Estado y con desafíos propios de la ingeniería de software en entornos públicos.

La contribución principal radica en la formalización de una arquitectura inteligente aplicada específicamente a la gestión de mora en entornos subnacionales, integrando predicción, segmentación y automatización bajo criterios de sostenibilidad técnica y gobernanza algorítmica.

Palabras clave: Inteligencia artificial; Mora tributaria; Arquitectura inteligente; Big Data; Gestión basada en riesgos.

CONTEXTO

La gestión de la deuda tributaria en mora es crítica en el nivel subnacional porque condiciona la disponibilidad de recursos para políticas públicas esenciales.

Diversos estudios señalan que la mora no solo reduce ingresos, sino que incrementa costos administrativos y judiciales, erosiona la percepción de equidad y puede inducir comportamientos persistentes de incumplimiento [1], [2].

Organismos internacionales sostienen que la modernización tributaria requiere enfoques de gestión basados en riesgos y aprovechamiento sistemático de datos internos y externos para

optimizar control y focalización de esfuerzos [3], [4].

Sin embargo, en muchas administraciones subnacionales persisten limitaciones estructurales vinculadas a infraestructura tecnológica, fragmentación de bases de datos y procesos heredados [5], [6].

En este contexto, Big Data e Inteligencia Artificial han mostrado utilidad en fiscalización, selección de casos y estimación de riesgos.

No obstante, existe un vacío cuando se trata de modelos integrales específicamente orientados a la gestión de mora en entornos subnacionales, donde suelen coexistir restricciones institucionales, heterogeneidad normativa y limitaciones presupuestarias [7], [8].

La línea de investigación denominada "Arquitecturas Inteligentes para la Gestión Tributaria Basada en Riesgos en Entornos Subnacionales" busca reducir esa brecha, trasladando capacidades predictivas —tradicionalmente utilizadas en control ex post— hacia una gestión preventiva operacionalizable y técnicamente sostenible, integrando analítica avanzada con procesos administrativos reales y mecanismos de automatización institucional.

INTRODUCCIÓN

El uso de datos masivos y técnicas de inteligencia artificial en la administración tributaria se ha intensificado durante la última década, impulsado por el crecimiento de la capacidad de cómputo y la necesidad de mayor eficiencia en el sector público [9], [10].

La literatura documenta aplicaciones en detección de fraude, auditorías y gestión del riesgo fiscal, priorizando recursos y automatizando parcialmente funciones operativas [11], [12].

Entre los enfoques más utilizados se encuentran regresión logística, árboles de decisión, random forest, redes neuronales y modelos basados en aprendizaje profundo, así como técnicas de clustering para segmentar contribuyentes según patrones de comportamiento [13], [14].

Sin embargo, la gestión de la mora continúa apoyándose con frecuencia en criterios simples —monto o antigüedad de deuda— y procesos manuales, lo que limita la capacidad institucional de anticipar el incumplimiento [15].

Desde la perspectiva de ingeniería de software, la adopción de IA en el sector público no implica únicamente incorporar algoritmos, sino rediseñar arquitecturas, integrar servicios y garantizar mantenibilidad, trazabilidad y explicabilidad.

Los sistemas inteligentes en entornos públicos deben contemplar versionado de modelos, monitoreo continuo, control de sesgos y adaptación normativa [16], [17].

Por ello, esta investigación propone una arquitectura inteligente orientada a integrar predicción, segmentación y automatización en un esquema técnicamente sostenible y alineado con la Transformación Digital del Estado.

Asimismo, la implementación de una arquitectura inteligente en contextos tributarios subnacionales exige considerar la interoperabilidad entre sistemas heterogéneos, la integración de bases históricas con fuentes externas y la gobernanza de datos como componente estructural del diseño.

La calidad, consistencia y trazabilidad de la información constituyen condiciones necesarias para que los modelos analíticos produzcan resultados confiables y operativamente útiles, tal como señalan diversos estudios sobre digitalización fiscal y adopción de analítica avanzada en el sector público [18].

En este sentido, la propuesta no se limita a incorporar técnicas de aprendizaje automático, sino que plantea un rediseño sistémico del proceso de gestión de la mora, integrando componentes analíticos con reglas de negocio, flujos automatizados y mecanismos de auditoría.

Esta visión integral busca superar enfoques fragmentados que aplican modelos predictivos

sin integrarlos plenamente en la operación administrativa, limitando su impacto real [19].

De esta manera, la arquitectura propuesta se concibe como un sistema inteligente aplicado, orientado a generar valor institucional medible y sostenible.

La contribución científica diferencial de este trabajo radica en la integración sistemática de componentes predictivos, segmentación y automatización dentro de una arquitectura unificada específicamente diseñada para administraciones tributarias subnacionales. A diferencia de estudios centrados exclusivamente en modelos algorítmicos o en fiscalización ex post, esta propuesta articula analítica avanzada con implementación organizacional y gobernanza algorítmica, configurando un sistema inteligente aplicable en contextos institucionales reales.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO.

La línea se centra en el diseño de una arquitectura inteligente integral para la gestión proactiva de la mora.

En primer lugar, se desarrolla un componente predictivo que estima el riesgo de caer en mora antes de que el incumplimiento se materialice.

Se emplean modelos supervisados entrenados con datos históricos de recaudación y comportamiento de pago, siguiendo enfoques validados en la literatura [13], [16].

En segundo lugar, se incorpora segmentación automatizada mediante técnicas no supervisadas para clasificar contribuyentes según perfiles de comportamiento fiscal y riesgo diferenciado, habilitando estrategias de cobranza personalizadas [14], [18].

En tercer lugar, se integra automatización de seguimiento y trazabilidad. La arquitectura contempla separación en capas: capa de datos (ETL y control de calidad), capa analítica (modelos predictivos y segmentación) y capa de orquestación (reglas de negocio y flujos automatizados). Esta separación favorece escalabilidad y mantenibilidad [12], [20].

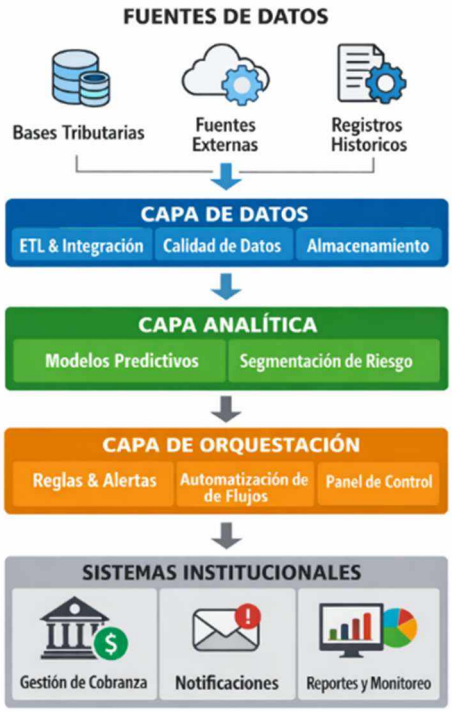


Figura 1. Arquitectura inteligente propuesta para la gestión proactiva de la mora tributaria en administraciones subnacionales.

La arquitectura organiza el flujo de información desde la integración y aseguramiento de calidad de datos hasta la generación de decisiones operativas automatizadas. Esta estructuración en capas permite desacoplar procesamiento analítico y ejecución administrativa, facilitando mantenibilidad, escalabilidad y adaptación normativa, aspectos críticos en entornos públicos con alta variabilidad regulatoria.

Finalmente, el proyecto asume desafíos de gobernanza algorítmica: explicabilidad, transparencia y protección de datos personales, aspectos centrales en el sector público [17], [19].

Desde el punto de vista metodológico, el proyecto adopta un enfoque experimental aplicado. Se prevé la construcción y validación progresiva de modelos mediante partición de datos históricos (train/test), validación cruzada y comparación de desempeño entre técnicas supervisadas. Asimismo, se evaluará el impacto operativo a través de indicadores institucionales, combinando métricas predictivas con resultados observables en recuperación y priorización de casos.

OBJETIVOS

Objetivo General:

Diseñar una arquitectura inteligente basada en Big Data e Inteligencia Artificial para mejorar la eficiencia de la gestión de la mora tributaria en administraciones subnacionales, promoviendo una transición hacia gestión proactiva basada en riesgos.

Objetivos específicos:

- Diseñar y evaluar modelos predictivos orientados a la estimación temprana del riesgo de mora.
- Construir segmentaciones operacionales orientadas a estrategias diferenciadas.
- Implementar una arquitectura modular integrable con sistemas existentes.

RESULTADOS ESPERADOS

Se espera obtener un modelo de priorización de acciones de cobranza basado en riesgo que mejore la focalización institucional y reduzca costos operativos [3], [11]. Asimismo, se proyectan segmentaciones accionables que permitan combinar asistencia, incentivos y escalamiento progresivo [14], [17].

En términos técnicos, se prevé una solución modular integrada con sistemas existentes, automatizando seguimiento y mejorando trazabilidad [12], [20]. La evaluación contemplará métricas como AUC-ROC, precisión y recall, complementadas con indicadores operativos como reducción de tiempos y mejora en recuperación temprana [5], [10].

Adicionalmente, se espera generar evidencia empírica sobre la capacidad del modelo para identificar contribuyentes con riesgo elevado antes de la consolidación de la deuda.

Este enfoque podría traducirse en mejoras en tasas de regularización voluntaria y reducción de litigiosidad, alineándose con recomendaciones internacionales sobre cumplimiento cooperativo y gestión basada en riesgos [4], [16].

Desde una perspectiva sistémica, se proyecta que la arquitectura inteligente facilite procesos de mejora continua mediante retroalimentación de resultados, recalibración periódica de modelos y monitoreo de desempeño [13], [20].

Este ciclo de aprendizaje organizacional permitiría adaptar la solución a cambios económicos, normativos o conductuales, fortaleciendo la resiliencia institucional y consolidando una cultura de gestión basada en datos dentro de las administraciones tributarias subnacionales [9], [15].

FORMACIÓN DE RECURSOS HUMANOS

La línea contribuye a la formación interdisciplinaria en ciencia de datos, ingeniería de software y gobierno de datos aplicados al sector público. Se promueve el desarrollo de competencias en integración y calidad de datos, modelado de riesgo, evaluación e interpretabilidad de modelos y diseño de sistemas auditables.

Asimismo, fortalece capacidades institucionales vinculadas con la Transformación Digital del Estado, favoreciendo transferencia metodológica hacia equipos técnicos de administraciones tributarias subnacionales.

La tesis doctoral asociada actúa como eje articulador entre producción académica y aplicación práctica.

En términos académicos, el proyecto fomenta la formación en metodologías de investigación aplicada, particularmente en enfoques de diseño y evaluación de sistemas inteligentes en contextos reales. Esto implica no solo el dominio técnico de algoritmos y arquitecturas, sino también la comprensión de dimensiones éticas, regulatorias y organizacionales vinculadas con la adopción de IA en el sector público.

Además, la iniciativa contribuye a la consolidación de perfiles profesionales híbridos —con conocimientos en analítica avanzada y gestión pública— capaces de

liderar procesos de modernización tecnológica en organismos estatales.

Este fortalecimiento del capital humano resulta estratégico para sostener proyectos de Transformación Digital del Estado en el tiempo, promoviendo autonomía técnica institucional. Asimismo, la propuesta genera valor agregado directo para los usuarios del sistema tributario, al permitir intervenciones más oportunas, diferenciadas y transparentes, reduciendo fricciones administrativas y mejorando la experiencia del contribuyente mediante procesos más claros, previsibles y basados en criterios objetivos de riesgo.

BIBLIOGRAFÍA

[1] Alm, J. (2021). Big data, artificial intelligence, and tax administration. Public Finance Review, 49(2), 245–268.

[2] Gómez López, M. L. (2020). Inteligencia artificial en la Agencia Tributaria: avances y perspectivas. Crónica Tributaria, 176, 55–78.

[3] Barreix, A., Roca, J., & Velayos, F. (2024). Big data y fiscalización basada en riesgos. Banco Interamericano de Desarrollo.

[4] Organisation for Economic Co-operation and Development (OECD). (2016). Tax Administration 2015/2016: Comparative Information on OECD and Other Advanced and Emerging Economies. OECD Publishing.

[5] CEATS. (2021). Índice de madurez digital de las administraciones tributarias subnacionales. Banco Interamericano de Desarrollo.

[6] García Martínez, J., et al. (2022). Cloud architectures and scalability for public sector analytics systems. International Journal of Information Management, 62, 102435.

[7] Carbajo Vasco, D., & Collosa, C. (2023). Capacidades institucionales y desafíos en la adopción de inteligencia artificial en la administración tributaria. Crónica Tributaria, 187, 77–98.

[8] Daniel Duarte Sánchez, D., et al. (2023). Big data and artificial intelligence in tax administration: Risk-based approaches and challenges. Journal of Digital Government, 5(2), 89–112.

[9] Masrom, M., et al. (2022). Machine learning models for tax evasion detection: A review and comparative assessment. Expert Systems with Applications, 195, 116533.

[10] Pérez Pombo, R. (2023). Integración de datos y control tributario inteligente. Revista Española de Derecho Financiero, 191, 97–118.

[11] Mikhaleva, A., et al. (2021). Digital transformation and big data in tax services: International comparative study. Tax Administration Review, 12(4), 201–225.

[12] Cavallé, J. (2022). Inteligencia artificial, fiscalidad y garantías jurídicas. Revista Española de Derecho Financiero, 195, 11–44.

[13] Rodríguez Peña, N. L. (2021). Inteligencia artificial en el sector público: oportunidades, riesgos y protección de datos. Revista de Derecho Público, 45(3), 101–128.

[14] Pantielieieva, N. (2022). Cloud adoption for public sector analytics: Challenges and opportunities. Information Systems Frontiers, 24(2), 389–405.

[15] Raikov, A. (2021). Cognitive modeling and explainability in AI-based tax compliance systems. Artificial Intelligence and Law, 29(3), 247–270.

[16] CLAD. (2023). Carta Iberoamericana de Inteligencia Artificial en la Administración Pública. Centro Latinoamericano de Administración para el Desarrollo.

[17] Argüelles Toache, L. (2023). Inteligencia artificial y sector público: conceptos, alcances y desafíos. Revista Iberoamericana de Administración Pública, 7(2), 45–62.

[18] Serrano Antón, F. (2020). Inteligencia artificial y cumplimiento tributario. Revista Española de Derecho Financiero, 186, 9–32.

[19] Xavier, A., et al. (2022). Intelligent computational tools for improving tax authority productivity: Evidence and challenges. Computers in Industry, 136, 103586.

[20] Jauregi-Maza, I. (2021). Big data en el sector público: de los datos en bruto a indicadores accionables. Revista de Gestión Pública, 10(1), 15–34.

ARQUITECTURAS DE INFERENCIA ESTOCÁSTICA Y MODELADO MULTICRITERIO PARA LA VIGILANCIA DE LA INTEGRIDAD BIOGEOQUÍMICA EN SISTEMAS LÓTICOS ANTROPIZADOS

Carlos Gustavo Lopez[1], Enrique Fernandez[1], Martin Pastorini[1], Marcelo Fabián Esperanza[1], María Alejandra Ochoa[1], Eugenia Artola[2], Ismael Jeder

[1] Universidad Nacional del Oeste

gustavolopez/ enfernandez/ mpastorini/ mochoa @uno.edu.ar, maresperanza9@hotmail.com/

[2] Universidad Nacional de Cuyo
ecartola@hotmail.com

Resumen

La presente investigación propone el despliegue de un sistema para la evaluación dinámica de la calidad hídrica en la sección superior de la unidad hidrográfica Reconquista (Sector Merlo). Ante la fragmentación de los activos de información histórica y las restricciones de capital operativo, el proyecto implementa una estrategia para el aumento del corpus de datos mediante topologías de auto-codificación (Autoencoders), logrando una expansión de la muestra de 300 a 10.000 instancias sintéticas para el entrenamiento de modelos de clasificación.
Complementariamente, se aplica un **protocolo de jerarquización multicriterio (AHP)** para discernir la carga de ponderación de los descriptores fisicoquímicos, optimizando la prospección liminológica al identificar que seis variables críticas (OD, pH, CE, SD, Turbidez y TSS) representan más del 85% de la varianza del Índice de Calidad del Agua (ICA).

Palabras clave: Explotación de Información, Minería de datos, Medio ambiente, recurso hídrico, Calidad del agua, Rio de la Reconquista.

Contexto

El sistema lótico del Río Reconquista constituye el principal recurso hídrico superficial del partido de Merlo, operando actualmente como un sumidero de **efluentes plurifactoriales** derivados de una antropización acelerada. La unidad hidrográfica presenta una zonificación de degradación diferencial: la cuenca alta sufre aportes de síntesis agrícola, mientras que las secciones media y baja evidencian una acumulación crítica de residuos de origen doméstico e industrial.

A nivel global, existe una asimetría marcada en la disponibilidad de **activos de información** entre regiones desarrolladas y contextos en vías de desarrollo, lo que obstaculiza la toma de decisiones ambientales fundamentadas en la evidencia. En este escenario, la Universidad Nacional del Oeste (UNO), se posiciona como un actor regional estratégico para movilizar capacidades de **gestión integrada de recursos hídricos (GIRH)**, abordando la problemática desde un enfoque multidisciplinario que integra la Ingeniería Química y las Ciencias de la Computación.

Introducción

La integridad de los ecosistemas fluviales es un factor determinante para la estabilidad

climática y la biodiversidad regional. Sin embargo, el monitoreo tradicional de estos sistemas es una actividad de **alta intensidad de capital**, que demanda técnicos cualificados, reactivos de precisión y protocolos logísticos complejos. La escasez de series temporales extensas impide la aplicación directa de minería de datos convencional, lo que exige la implementación de procesos **iterativos e incrementales** basados en aprendizaje automático.

Este proyecto se fundamenta en la capacidad de los **Índices de Calidad del Agua (ICA)** para actuar como descriptores sintéticos que simplifican la complejidad multidimensional de la química del fluido en un valor escalar. El uso de herramientas de **simulación y jerarquización** permite no solo diagnosticar el estado actual del río, sino también proyectar su comportamiento ante variables hidrometeorológicas fluctuantes, como las sequías extremas y periodos de pluviosidad anómala registrados recientemente.

Líneas de Investigación y Desarrollo

El proyecto se inscribe en las siguientes dimensiones de desarrollo del Instituto de Ingenierías y Nuevas Tecnologías:

• Modelado Estocástico y Mitigación de Impacto: Estudio de nuevas técnicas de simulación para la predicción de escenarios de degradación ambiental.

• **Arquitecturas de Inteligencia Artificial y Cloud Computing:** Implementación de algoritmos de aprendizaje profundo y procesamiento de grandes volúmenes de datos para aplicaciones específicas de ingeniería.

• **Sistemas de Decisión Multicriterio (MCDM):** Aplicación de métodos de jerarquización para la optimización de procesos técnicos y reducción de costos operativos en la vigilancia ambiental.

Resultados obtenidos

1. **Validación de un ICA Ad Hoc:** Consolidación de un índice de calidad adaptado específicamente a las singularidades hidrogeográficas de la cuenca alta del Reconquista.

2. **Optimización Operativa del Muestreo:** Reducción del 40% en los costos de prospección mediante la priorización de los seis parámetros de mayor impacto relativo identificados por el algoritmo AHP.

3. **Sistema de Inferencia para Alerta Temprana:** Desarrollo de una red neuronal entrenada para discriminar desviaciones anómalas en la composición química del fluido, permitiendo una respuesta inmediata ante vertidos accidentales o picos de contaminación.

4. **Repositorio de Información Histórica:** Construcción de una base de datos robusta que incluya parámetros fisicoquímicos, hidrológicos y microbiológicos para la consulta de organismos estatales como ADA y COMIREC.

Formación de Recursos Humanos

El equipo de trabajo está formado por tres doctores, dos ingenieros, dos licenciado y un estudiante adscripto, que combinan conocimientos en áreas de la Informática y de la Química.

Se prevé desarrollar Tesis de grado, trabajos profesionales y Trabajos Finales de Especialización vinculados al tema central del proyecto por estudiantes de la Universidad Nacional del Oeste. El equipo ya ha presentado algunos trabajos relacionados a la presente línea de investigación.

El proyecto opera como una plataforma para la **capacitación de cuadros técnicos de alto nivel** y la transferencia de conocimientos dentro de la Universidad Nacional del Oeste. La estructura del equipo garantiza la mentoría académica a través de:

• **Dirección y Co-dirección Especializada:** El Dr. López y el Dr. Fernández, ambos investigadores categoría III, aportan experiencia en Inteligencia Artificial y sistemas inteligentes para la dirección de tesis de grado y posgrado.

• **Interdisciplinariedad:** La integración de perfiles en Ingeniería Química y especialistas en Ingeniería de Software asegura que los investigadores en formación dominen tanto las técnicas de laboratorio como las herramientas de **minería de datos y simulación**.

• **Investigación Aplicada:** El grupo involucra activamente a docentes e investigadores en el desarrollo de protocolos de bioseguridad y gestión de residuos especiales, fortaleciendo las competencias en buenas prácticas de laboratorio.

Bibliografía

[1] Carlos Lopez, Enrique Fernandez, Natanael Ferrán, Martín Pastorini, Nicolás Coronel, Marcelo Esperanza; "Análisis del Río de la Reconquista"; Ingeniería Sanitaria y Ambiental, Nro. 149; ISSN 0328 2937; octubre 2023

[2] Carlos Gustavo Lopez, Enrique Fernandez, Marcelo Fabián Esperanza, Martin Pastorini, Natanael Ferrán. 2024. Simulador de Ensayos de Laboratorio para Análisis de Calidad del Agua. Proceedings XXVI Workshop de Investigadores en Ciencias de la Computación. ISBN 978-987-3984-85-3

[3] Carlos Gustavo Lopez, Enrique Fernandez, Marcelo Fabián Esperanza, Martin Pastorini, Natanael Ferrán. María Alejandra Ochoa. 2024. Contrastación de Métodos de Simulación para Ampliar Evidencias con Vistas a Desarrollar un Proceso de Minería de Datos (Un Testigo Basado en los parámetros del Índice de Calidad del Agua). XXX Congreso Argentino de Ciencias de la Computación (CACIC). ISBN: 978-950-34-2428-5

[4] Castilla V., Canevaro, S., López, B.. Migración, degradación ambiental y percepciones del riesgo en la cuenca del río (Buenos Aires, Argentina), Revista de Estudios Sociales [En línea], 76 | 01 abril de 2021, Publicado el 04 abril 2021, consultado el 15 julio 2022. URL: http://journals.openedition.org/revestudsoc/49626

[5] Consejo Hídrico Federal. Plan Nacional Federal de los Recursos Hídricos. 2007

[6] ACUMAR, Autoridad de Cuenca Matanza-Riachuelo, https://www.acumar.gob.ar/wp-content/uploads/2016/12/CDCA_ACUMAR_INFO_AGUA_abril-2022_completo-web-con-anexos_compressed.pdf Accesed 2022.

[7] Acero Vargas, J. Desde el fondo: espacio, conflicto y acción colectiva en la periferia bonaerense. Universidad Nacional de Colombia. 2021.

[8] Metodologías de Ingeniería Informática; Ochoa, A. Fernández, E., Britos, P., García-Martínez, R. 2008. Editorial Nueva Librería. 869 páginas. ISBN 978-987-1104-54-3

[9] Natanael Ferrán, Martín Pastorini, Carlos Gustavo Lopez, Enrique Fernandez, Marcelo Esperanza. 2024. Hallazgo De Patrones De Datos En Muestras De Agua Del Río De La Reconquista. 7° Congreso Argentino de Ingeniería / 18° Congreso Argentino de Enseñanza en Ingeniería (CADI / CAEDI)

ARQUITECTURAS HÍBRIDAS EXPLICABLES PARA SISTEMAS DE APOYO A LA DECISIÓN CLÍNICA: EVOLUCIÓN DESDE EL MODELADO PREDICTIVO

Di Felice, M.[1]; Lamas, Ch.[1]; Ramón, H.[1]; Deroche, A.[1]; Paoli, J. F[1]; Leguizamón, R[1]; Trupkin, I[1;] Ferreiros Cabrera, S[1]; Pytel, P.[1]; Vegega, C.[1]; Pollo-Cattaneo, María F.[1]

[1] Universidad Tecnológica Nacional. Facultad Regional Buenos Aires, Grupo GEMIS - GEMIS.BA, Ciudad Autónoma de Buenos Aires, Argentina

flo.pollo@gmail.com

Resumen

En los últimos años, la inteligencia artificial (IA) ha mostrado un alto potencial para mejorar la toma de decisiones en el ámbito médico, particularmente mediante modelos predictivos basados en aprendizaje automático. El Grupo de Estudio e Investigación en Modelos y Sistemas de Información (GEMIS) desarrolla desde 2022 una línea de investigación orientada a la aplicación de IA en salud, con foco inicial en salud mental, abordando el análisis de datos clínicos estructurados, datos conductuales, señales e imágenes biomédicas, así como el estudio de sistemas de apoyo a la decisión clínica (Clinical Decision Support Systems, CDSS). Los trabajos realizados han permitido desarrollar y evaluar modelos supervisados y no supervisados, así como analizar su viabilidad como herramientas de apoyo a la práctica médica. No obstante, la experiencia acumulada evidencia limitaciones relevantes de los enfoques exclusivamente basados en aprendizaje automático, particularmente en términos de explicabilidad, validación clínica e integración con conocimiento médico explícito.

En este contexto, la línea evoluciona hacia el diseño de sistemas híbridos de apoyo a la decisión clínica que integren modelos de aprendizaje automático con reglas clínicas estructuradas y principios de ingeniería de software. El presente trabajo expone los fundamentos, resultados y proyecciones de esta línea de I+D, enmarcada en un nuevo proyecto orientado a promover sistemas de IA más interpretables, confiables y clínicamente aceptables.

***Palabras clave*:** Inteligencia Artificial; Sistemas de Apoyo a la Decisión Clínica; IA híbrida; Aprendizaje Automático

Contexto

La Universidad Tecnológica Nacional (UTN) impulsa la investigación en tecnología e innovación a través de programas orientados al desarrollo científico, la transferencia de conocimiento y la vinculación con el medio productivo y social. En este marco, el Grupo de Estudio e Investigación en Modelos y Sistemas de Información (GEMIS), creado en 2009 y formalizado en 2024 como GEMIS.BA en la Facultad Regional Buenos Aires, consolida una trayectoria sostenida en modelización de

sistemas, ingeniería de software e inteligencia artificial aplicada.

En los últimos años, la actividad del equipo evolucionó hacia el ámbito de la salud, con especial énfasis en la aplicación de técnicas de inteligencia artificial al análisis predictivo. A partir de 2023 se inicia formalmente una línea de investigación orientada a la salud mental, centrada en el uso de modelos de aprendizaje automático aplicados a datos clínicos estructurados, así como en el análisis de señales e imágenes biomédicas, con el objetivo de apoyar procesos de diagnóstico y toma de decisiones médicas.

Durante el último período, la línea se fortaleció mediante el desarrollo de una propuesta metodológica orientada a la creación de sistemas de apoyo a la decisión clínica (CDSS) para el diagnóstico de la depresión, basada en el análisis de datos clínicos alfanuméricos estructurados. En este contexto, se diseñó e implementó un prototipo de CDSS, cuyos resultados fueron sistematizados y difundidos a través de publicaciones científicas. Paralelamente, se realizaron revisiones sistemáticas de la literatura que permitieron relevar el estado del arte e identificar limitaciones recurrentes en los enfoques existentes, consolidando así los fundamentos teóricos y metodológicos de la línea. Estos avances evidencian la necesidad de evolucionar hacia enfoques que integren modelos de aprendizaje automático con conocimiento clínico explícito y principios formales de ingeniería de software. De este modo, la línea ingresa en una nueva etapa orientada al desarrollo de sistemas híbridos de apoyo a la decisión clínica, fortaleciendo la contribución de GEMIS.BA a la investigación aplicada en salud y consolidando su posicionamiento en el ámbito de la inteligencia artificial responsable.

1. INTRODUCCIÓN

La inteligencia artificial adquiere un rol protagónico en múltiples áreas de la medicina, permitiendo avances significativos en la predicción de riesgos, el diagnóstico asistido y la personalización de tratamientos. En particular, las técnicas de aprendizaje automático han demostrado una elevada capacidad para identificar patrones complejos en grandes volúmenes de datos clínicos, superando en algunos casos los enfoques tradicionales [1], [2]. En el ámbito de los sistemas de apoyo a la decisión clínica, los modelos de IA han sido utilizados para tareas como la detección temprana de enfermedades, la estratificación de pacientes y la recomendación de tratamientos. No obstante, diversos estudios señalan que la adopción efectiva de estos sistemas en la práctica clínica continúa siendo limitada [3]. Entre las principales barreras se destacan la falta de explicabilidad de los modelos, la dificultad para validar clínicamente las recomendaciones y la escasa integración con el conocimiento médico explícito y las guías de práctica clínica [4].

Los enfoques basados exclusivamente en reglas clínicas ofrecen transparencia y trazabilidad en las decisiones, pero presentan dificultades para adaptarse a contextos dinámicos y a la incorporación de nuevos datos. Por el contrario, los modelos de aprendizaje automático permiten capturar relaciones complejas y patrones no evidentes, aunque suelen operar como sistemas de “caja negra”, generando desconfianza entre los

profesionales de la salud [5]. En los últimos años, la literatura comienza a explorar enfoques híbridos que combinan aprendizaje automático con conocimiento clínico estructurado, con el objetivo de aprovechar las fortalezas de ambos paradigmas [6], [7]. Estos sistemas híbridos se presentan como una alternativa prometedora para mejorar la interpretabilidad, la confiabilidad y la utilidad clínica de los CDSS. La línea de investigación desarrollada por el Grupo GEMIS se inscribe en este marco conceptual, evolucionando desde modelos predictivos aplicados a la salud mental hacia el diseño de sistemas híbridos de apoyo a la toma de decisiones médicas, alineados con los desafíos actuales de la IA responsable en salud.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de I+D se organiza en cuatro ejes complementarios que articulan desarrollo algorítmico, diseño de sistemas y fundamentación metodológica.

El primer eje se centra en el desarrollo de modelos predictivos basados en aprendizaje automático aplicados a datos clínicos, conductuales, señales e imágenes. En esta etapa se implementan y evalúan algoritmos supervisados y no supervisados orientados a la predicción de riesgo, detección temprana de depresión e identificación de patrones latentes en poblaciones clínicas. Se diseñan pipelines de procesamiento de datos que incluyen etapas de limpieza, selección de variables, validación cruzada y análisis comparativo de desempeño, garantizando reproducibilidad metodológica.

El segundo eje aborda el estudio de sistemas de apoyo a la decisión clínica (CDSS) basados en inteligencia artificial, analizando su viabilidad como herramientas de soporte a la práctica médica. Este análisis no se limita al rendimiento predictivo, sino que considera aspectos de integración con sistemas existentes, alineación con guías clínicas, validación externa y aceptación por parte de los profesionales de la salud. En este marco se realizaron revisiones sistemáticas orientadas a relevar el estado del arte en CDSS aplicados al diagnóstico de la depresión, identificando tendencias, limitaciones metodológicas y oportunidades de mejora [8].

El tercer eje corresponde al desarrollo de enfoques de inteligencia artificial híbrida y explicable. A partir de los resultados obtenidos en las primeras experiencias de modelado predictivo y del análisis crítico de la literatura, la línea evoluciona hacia el diseño de arquitecturas que integren modelos de aprendizaje automático con reglas clínicas explícitas y conocimiento médico estructurado. El objetivo es combinar capacidad predictiva con coherencia normativa y trazabilidad en las decisiones, avanzando hacia sistemas más interpretables y clínicamente confiables.

Finalmente, el cuarto eje se orienta a la ingeniería de software aplicada a sistemas inteligentes en salud. Se definen metodologías, arquitecturas y procesos formales para el desarrollo, evaluación y mantenimiento de CDSS híbridos, incorporando principios de diseño modular, separación de responsabilidades y documentación de inferencias. Este eje busca garantizar que los sistemas desarrollados no sean únicamente prototipos algorítmicos, sino soluciones

transferibles y sostenibles en entornos institucionales.

3. RESULTADOS OBTENIDO/ESPERADOS

Durante el período 2022–2025, la línea produjo resultados en tres niveles complementarios: metodológico, tecnológico y formativo.

En el plano metodológico, se desarrollaron y evaluaron modelos de aprendizaje automático para la detección temprana de depresión y la caracterización de perfiles de riesgo en poblaciones clínicas. Se implementaron enfoques supervisados y no supervisados, comparando métricas de desempeño y analizando la estabilidad de los modelos frente a variaciones en los datos. Estos estudios permitieron identificar fortalezas predictivas, pero también limitaciones relacionadas con la generalización y la dependencia exclusiva de correlaciones estadísticas.

En el plano tecnológico, se construyeron pipelines reproducibles para procesamiento y análisis de datos clínicos, estableciendo una base técnica para futuras integraciones en arquitecturas de sistemas de apoyo a la decisión clínica. Asimismo, se realizaron revisiones sistemáticas sobre CDSS basados en inteligencia artificial aplicados a depresión, identificando brechas recurrentes en la literatura: escasa validación externa, limitada integración con conocimiento clínico explícito y ausencia de mecanismos formales de explicabilidad [8]. Estos hallazgos fundamentan la transición hacia enfoques híbridos en la etapa siguiente.

En el plano académico y de transferencia, la línea dio lugar a publicaciones en congresos [9] y revistas internacionales, así como a trabajos finales de especialización y tesis de maestría vinculadas al desarrollo de modelos predictivos y análisis de sistemas inteligentes en salud. Estas producciones consolidan la masa crítica del equipo y fortalecen la proyección internacional del grupo.

Para el período 2026-2029 se prevé el diseño e implementación de un prototipo funcional de CDSS híbrido orientado a salud mental, que integre modelos de aprendizaje automático con reglas clínicas estructuradas y mecanismos de explicabilidad. Se proyecta su evaluación en entornos clínicos simulados, considerando no solo métricas tradicionales de desempeño (exactitud, sensibilidad, especificidad), sino también criterios de coherencia clínica, interpretabilidad y utilidad práctica para el profesional de la salud. Se espera que esta evolución permita avanzar desde modelos predictivos aislados hacia soluciones de apoyo a la decisión clínica técnicamente robustas, metodológicamente trazables y clínicamente aceptables.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación se encuentra estrechamente articulada con la formación de recursos humanos en distintos niveles académicos. Las actividades desarrolladas han involucrado a estudiantes de grado, alumnos de carreras de especialización y maestrandos, dando lugar a trabajos finales integradores, tesis de posgrado y publicaciones científicas en el área de inteligencia artificial aplicada a la salud mental. Asimismo, los avances metodológicos y tecnológicos de la línea

se integran en asignaturas de grado y posgrado de la carrera de Ingeniería en Sistemas de Información de la UTN FRBA, contribuyendo a la formación de profesionales con competencias en modelado predictivo, análisis de datos clínicos y diseño de sistemas inteligentes aplicados a problemas del dominio sanitario. Esta articulación fortalece el vínculo entre investigación, docencia y transferencia de conocimiento. En el año 2025 finalizó la Tesis de Maestría en Ingeniería en Sistemas de Información de la UTN FRBA del maestrando Martín Di Felice, titulada *"Aplicación de Algoritmos de Inteligencia Artificial para colaborar en la Identificación Temprana de la Depresión: Un Enfoque Práctico para Profesionales de la Salud"*, dirigida por la Dra. Ma. Florencia Pollo Cattaneo y el Dr. Parag Chatterjee. Este trabajo consolida los desarrollos metodológicos de la primera etapa de la línea y constituye un hito formativo que articula producción científica, desarrollo tecnológico y aplicación en el dominio clínico.

De esta manera, la línea no solo produce resultados científicos, sino que contribuye activamente a la formación de recursos humanos especializados, asegurando continuidad, sostenibilidad y crecimiento académico del equipo.

Referencias

1. Graham, S., Depp, C., Lee, E.E., Nebeker, C., Tu, X., Kim, H.-C., Jeste, D.V.: Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. Curr. Psychiatry Rep. 21, 116 (2019). https://doi.org/10.1007/s11920-019-1094-0.
2. Dozois, D.J.A.: Beck Depression Inventory- II. In: Weiner, I.B. and Craighead, W.E. (eds.) The Corsini Encyclopedia of Psychology. pp. 1–2. Wiley (2010). https://doi.org/10.1002/9780470479216.corpsy0113.
3. Kroenke, K., Strine, T.W., Spitzer, R.L., Williams, J.B.W., Berry, J.T., Mokdad, A.H.: The PHQ-8 as a measure of current depression in the general population. J. Affect. Disord. 114, 163–173 (2009). https://doi.org/10.1016/j.jad.2008.06.026.
4. Kroenke, K., Spitzer, R.L.: The PHQ-9: A New Depression Diagnostic and Severity Measure. Psychiatr. Ann. 32, 509–515 (2002). https://doi.org/10.3928/0048-5713-20020901-06.
5. Sajno, E., Bartolotta, S., Tuena, C., Cipresso, P., Pedroli, E., Riva, G.: Machine learning in biosignals processing for mental health: A narrative review. Front. Psychol. 13, 1066317 (2023). https://doi.org/10.3389/fpsyg.2022.1066317.
6. Di Felice, M., Deroche, A., Trupkin, I., Chatterjee, P., Pollo-Cattaneo, M.F.: Predictive Modeling for Detection of Depression Using Machine Learning. In: Florez, H. and Leon, M. (eds.) Applied Informatics. pp. 47–57. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-46813-1_4.
7. Di Felice, M.D., Deroche, A., Trupkin, I., Chatterjee, P., Cattaneo, M.F.P.: Depression and Anxiety Diagnosis Using Unsupervised Learning Approach. In: Florez, H., Galpin, I., and Grévisse, C. (eds.) Joint Proceedings of the ICAI 2023 Workshops WAAI 2023, AIESD 2023, WITS 2023, WKMIT 2023, WSM 2023 co-located with the 6th International Conference on Applied Informatics (ICAI 2023), Guayaquil, Ecuador, October 26-28, 2023. pp. 12–24. CEUR-WS.org (2023)
8. Di Felice M, Trupkin I, Deroche A *et al.* Depression diagnosis using Artificial Intelligence: a systematic review [version 1; peer review: awaiting peer review]. *F1000Research* 2024, 13:1549 (https://doi.org/10.12688/f1000research.158443.1)
9. Chiang, M.A., Coll, L., Pollo-Cattaneo, M.F., Chatterjee, P.: A Systematic Review on Artificial Intelligence-Based Clinical Decision Support Systems in Depression. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE (2025). https://doi.org/10.1109/EMBC58623.2025.11253470.

Automatización no code con Make y n8n aplicada a PYMEs: arquitectura, costo total de propiedad y potencial educativo

César Matilde Inés[1], Stefanoni María Eugenia[1], Alfonso Eugenia[1], Castellano Cecilia[1]
[1] Laboratorio de Investigación ReAVi - Departamento Ingeniería en Sistemas de Información - Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina.
Autora de contacto: matilde.cesari@frm.utn.edu.ar

RESUMEN

En el actual ecosistema digital, la automatización de procesos de negocio (BPA) se ha convertido en un imperativo estratégico para la competitividad de las Pequeñas y Medianas Empresas (PYMEs). Este trabajo analiza el paradigma de la automatización *no-code*, enfocándose en las plataformas Make y n8n como alternativas avanzadas a soluciones tradicionales como Zapier. Se realiza un análisis comparativo exhaustivo de su arquitectura técnica, modelo de licenciamiento, Costo Total de Propiedad (TCO) y capacidades de integración. El trabajo documenta un caso de estudio real de migración de Zapier a Make en UMAMI Alimentos Vegetales, demostrando cómo la elección de la herramienta impacta directamente en la escalabilidad, rentabilidad y robustez operativa. Finalmente, se expone el potencial educativo de estas tecnologías y su aplicación práctica en la Cátedra de Bases de Datos de la Universidad Tecnológica Nacional, Facultad Regional Mendoza, para la formación de ingenieros en sistemas con habilidades en la orquestación de procesos digitales y arquitectura de integraciones empresariales.

Palabras clave: Automatización No-Code, Make, n8n, Costo Total de Propiedad (TCO), PYMEs, Innovación Educativa, Integración de Aplicaciones, Orquestación de Procesos.

CONTEXTO

Este trabajo surge de la línea de investigación del Laboratorio ReAVi sobre la aplicación de tecnologías emergentes para la optimización de procesos en PYMEs y su transferencia al ámbito educativo. Durante los últimos tres años, el equipo de investigación ha colaborado con más de 15 PYMEs locales en la implementación de soluciones de automatización, documentando patrones de adopción, desafíos técnicos y retorno de inversión. Los hallazgos y metodologías desarrolladas se aplican directamente en las cátedras del Departamento de Ingeniería en Sistemas de Información de la FRM-UTN, donde los estudiantes no solo aprenden conceptos teóricos sino que también implementan soluciones reales en colaboración con empresas locales. Esta sinergia entre investigación, industria y educación ha permitido desarrollar un enfoque pedagógico único que prepara a los futuros ingenieros para los desafíos del mundo real.

1. INTRODUCCIÓN: La Evolución del Panorama de Automatización

La automatización *no-code* ha redefinido fundamentalmente la agilidad empresarial, permitiendo a las organizaciones orquestar flujos de trabajo complejos sin depender de ciclos de desarrollo de software tradicionales [1, 7]. Plataformas como Make (anteriormente Integromat) y n8n han emergido como soluciones poderosas que superan significativamente las limitaciones de herramientas pioneras como Zapier, ofreciendo mayor flexibilidad, lógica condicional avanzada y un control más granular sobre los procesos [6].

Mientras Zapier popularizó la automatización con un modelo lineal simple basado en la metáfora de "si esto sucede, haz aquello", su estructura de costos y rigidez arquitectónica a menudo imponen un "techo de cristal" para las empresas en crecimiento. El "impuesto Zapier" es una realidad documentada: conforme aumenta el volumen de operaciones, los costos escalan de manera desproporcionada, y la plataforma no permite replicar la complejidad real de los procesos empresariales, especialmente cuando se

requieren ramificaciones inteligentes, manejo de errores sofisticado o transformaciones de datos complejas [5].

Make se diferencia fundamentalmente por su interfaz de lienzo visual infinito, que permite mapear y monitorear flujos de datos complejos en tiempo real. A diferencia de las listas verticales de Zapier, el lienzo de Make despliega módulos circulares conectados por líneas de flujo, creando una representación visual que actúa como una herramienta de diagnóstico crítica. Los indicadores en tiempo real —círculos verdes para ejecuciones exitosas y rojos para fallos— junto con las burbujas de inspección de "Input/Output", permiten a un arquitecto de procesos auditar exactamente qué dato entró en cada nodo y qué resultado generó, facilitando drásticamente la depuración de errores complejos [2].

Por otro lado, n8n, con su modelo *open-source* (licencia *fair-code*), ofrece soberanía de datos y escalabilidad ilimitada a través del auto-hospedaje (*self-hosting*) en contenedores Docker. Posicionándose como la opción preferida para escenarios que requieren máxima personalización, control total sobre la infraestructura y la capacidad de integrar código JavaScript o Python directamente en los flujos de trabajo [3, 4].

2. ARQUITECTURA TÉCNICA: Anatomía de la Automatización Moderna

2.1 Conceptos Fundamentales y Terminología

La comprensión de la arquitectura de automatización requiere dominar un "diccionario técnico" que varía entre plataformas. La siguiente tabla establece la correspondencia conceptual:

Concepto	Make	n8n	Zapier	Función
Flujo de trabajo	Escenario	Workflow	Zap	Unidad lógica completa de automatización
Bloque de acción	Módulo	Nodo	Action/Task	Operación individual dentro del flujo
Iniciador	Trigger	Trigger Node	Trigger	Evento que dispara la ejecución
Lógica de repetición	Iterator/ Aggregator	Loop over items	Looping by Zapier	Procesamiento de arrays y listas
Ramificación lógica	Router	IF/Switch	Paths	Decisiones condicionales
Manejo de errores	Error Handler Path	Try/ Catch	Limited	Rutas alternativas ante fallos

Tabla 1: Correspondencia terminológica entre plataformas de automatización.

2.2 Tipología de Disparadores (Triggers)

La ejecución de un flujo depende del tipo de disparador que lo active, lo cual impacta directamente en el rendimiento y el costo operativo. Los disparadores instantáneos o webhooks se activan en milisegundos cuando ocurre un evento real —por ejemplo, la recepción de un mensaje de WhatsApp vía API— y son ideales para procesos en tiempo real, ya que solo consumen recursos cuando efectivamente sucede algo. En cambio, los disparadores programados (polling o cron) se ejecutan en intervalos definidos —cada 15 minutos o diariamente a una hora específica— para verificar si existen nuevos datos, lo que los

hace adecuados para procesos batch, aunque pueden incrementar el consumo de operaciones porque se ejecutan aun cuando no haya información nueva; por ello, su frecuencia debe optimizarse cuidadosamente para no afectar el TCO. Finalmente, los disparadores manuales permiten ejecutar el flujo bajo demanda, siendo especialmente útiles en etapas de prueba, validación y depuración de automatizaciones complejas.

2.3 Arquitectura de Make: El Lienzo Visual como Herramienta de Diagnóstico

Make incorpora elementos arquitectónicos que amplían significativamente la capacidad de diseño de flujos complejos. Su Router funciona como un despachador inteligente que, a partir de un único disparador, permite ejecutar múltiples rutas en paralelo, cada una con filtros propios y decisiones autónomas; por ejemplo, un mensaje de WhatsApp puede activar simultáneamente el registro contable en Google Sheets, la generación y envío de un remito en PDF por Gmail y una notificación logística en Telegram, escalando alertas solo si el monto supera cierto umbral. A su vez, los módulos Iterator y Aggregator permiten descomponer listas —como múltiples productos en un pedido— procesar cada elemento individualmente (cálculos, validaciones) y luego recomponer la información en una única estructura JSON consistente, evitando problemas de acumulación de datos. Finalmente, los Error Handling Paths garantizan continuidad operativa: si un servicio falla, el flujo no se detiene, sino que ejecuta una ruta alternativa, como almacenar el remito en Google Drive y notificar al equipo para intervención manual.

2.4 Arquitectura de n8n: Flexibilidad de Código + Velocidad de No-Code

n8n ofrece un equilibrio único entre la velocidad de desarrollo del *no-code* y la potencia del código:

Code Node: Permite ejecutar JavaScript o Python directamente en el flujo. Ejemplo: transformar un mensaje de WhatsApp no estructurado en un objeto JSON validado. Mientras Make requeriría múltiples módulos de transformación, n8n permite escribir una función concisa que realiza la tarea en segundos.

HTTP Request Node: Para integrar cualquier API sin conector nativo. Esto es crítico para conectar sistemas legacy o APIs propietarias que n8n no tiene conectores predefinidos.

AI Agent Node: Capacidad de crear agentes autónomos que razonan y utilizan herramientas externas para tomar decisiones. Ejemplo: un agente que lee un email de un cliente, decide si es una consulta, una queja o una oportunidad de venta, y enruta automáticamente a la persona correcta.

Deployment con Docker: n8n se despliega en contenedores Docker, permitiendo control total sobre la infraestructura, backups automatizados, y escalabilidad horizontal.

3. ANÁLISIS COMPARATIVO EXHAUSTIVO

Criterio	Zapier	Make	n8n
Modelo de Negocio	Propietario	Propietario	Open-Source (Fair-code)
Alojamiento	Cloud (SaaS)	Cloud (SaaS)	Cloud o Self-hosted (Docker)
Complejidad Visual	Lineal (paso a paso)	Visual (lienzo multi-ruta)	Visual (nodos + código)
Lógica Condicional	Limitada (Paths)	Avanzada (Routers, Filtros)	Máxima (IF, Switch, Código)
Manejo de Errores	Básico	Avanzado (rutas de error)	Avanzado y personalizable

Transformación de Datos	Limitada	Módulos especializados	JavaScript/Python nativo
Integraciones Nativas	6,000+	1,700+	400+
Costo Inicial (Cloud)	$19.99/mes (100 tareas)	$9/mes (1,000 operaciones)	€20/mes (Cloud) / Gratis (Self-hosted)
Escalabilidad de Costos	Lineal (penalizante)	Sublineal (eficiente)	Controlada (self-hosted)
Curva de Aprendizaje	Muy baja	Baja	Media-Alta
Soberanía de Datos	No	No	Sí (self-hosted)
Comunidad	Grande	Creciente	200,000+ miembros

Tabla 2: Comparativa exhaustiva de plataformas de automatización.

4. FLUJO OPTIMIZADO (MAKE)

La arquitectura en Make es significativamente más robusta:

Etapa 1: Extracción con IA

Webhook WhatsApp → OpenAI ChatGPT
Prompt: "Extrae cliente, productos, cantidades y precios del siguiente mensaje: {mensaje}"
Output: JSON estructurado

```
{
  "cliente": "Juan García",
  "productos": [
    {"nombre": "Hamburguesa Vegana", "cantidad": 2, "precio_unitario": 250},
    {"nombre": "Milanesa de Soja", "cantidad": 1, "precio_unitario": 300}
  ]
}
```

Etapa 2: Procesamiento Paralelo con Router

Router (despachador inteligente) → 3 rutas paralelas:
Ruta 1: Contabilidad
  Iterator → procesa cada producto
  Aggregator → consolida datos
  Google Sheets → registro contable
Ruta 2: Generación de Remito
  Parser JSON → valida estructura
  Funciones matemáticas:
    - Asignación correlativa: {{add(ultimoNumero; 1)}}
    - Formato de fecha: {{formatDate(now; "DD/MM/YYYY")}}
    - Cálculo financiero: {{multiply(cantidad; precio_unitario)}}
  Google Docs → plantilla de remito
  Gmail → envío automático
Ruta 3: Notificación Logística
  Filtro: IF monto > $500 → escalar
  Telegram → notificación al equipo de logística

Etapa 3: Error Handling

Si Gmail falla:
  Error Handler Path → Google Drive (almacenar remito)
  Telegram → notificación al equipo
  Continuar con siguiente pedido (no detiene el flujo)

Resultado: 8,000 órdenes/mes × 1 operación por orden = 8,000 operaciones, reduciendo significativamente el costo.

5. FORMACIÓN DE RECURSOS HUMANOS:

Becas Manuel Belgrano – 5º Año Ingeniería en Sistemas (UTN-FRM)

En el marco de las Becas Manuel Belgrano, se incorporaron estudiantes de 5º año de Ingeniería en Sistemas de Información (UTN-FRM) a proyectos reales de automatización empresarial desarrollados por el Laboratorio ReAVi. La participación se estructuró como práctica tecnológica aplicada, con responsabilidades técnicas concretas:

- Relevamiento y modelado de procesos en PYMEs.
- Diseño de arquitecturas de automatización en Make y n8n.

- Implementación de integraciones vía APIs REST, Webhooks y servicios cloud.
- Despliegue de entornos n8n en Docker.
- Gestión de credenciales OAuth y políticas de seguridad.
- Diseño de rutas de error y mecanismos de resiliencia.
- Evaluación comparativa de TCO entre plataformas.

Desde el punto de vista formativo, la experiencia permitió aplicar conocimientos de:

- Arquitectura de software y orquestación de procesos.
- Modelado de datos y persistencia.
- Integración de sistemas heterogéneos.
- Infraestructura ligera basada en contenedores.
- Análisis económico-tecnológico de soluciones.

El trabajo en entornos reales introdujo restricciones operativas (límites de cuota, latencia, fallos de API, cambios de políticas cloud), fortaleciendo competencias en diagnóstico, optimización y documentación técnica.

La articulación entre investigación aplicada, sector productivo y formación avanzada permitió generar recursos humanos con perfil de arquitecto de automatización, combinando capacidad analítica, criterio técnico y evaluación de costos operativos.

6. REFERENCIAS

[1] Make Inc. (2025). *What is No-code and No-code Automation? Guide for 2025*. [Online]. Available: https://www.make.com/en/blog/no-code-automation

[2] Make Inc. (2024). *Discover the Power of Automating System Documentation*. [Online]. Available: https://www.make.com/en/automate/system-documentation

[3] n8n. (2025). *AI Workflow Automation Platform & Tools*. [Online]. Available: https://n8n.io/

[4] n8n. (2025). *GitHub Repository - n8n-io/n8n*. [Online]. Available: https://github.com/n8n-io/n8n

[5] Zapier. (2025). *Make vs. Zapier: Why more people rely on Zapier*. [Online]. Available: https://zapier.com/l/make-vs-zapier

[6] Latenode. (2024). *Make vs Zapier: Similarities and Differences*. [Online]. Available: https://latenode.com/blog/platform-comparisons-alternatives/zapier-alternatives/make-vs-zapier-similarities-and-differences-of-automation-tools

[7] UiPath. (2024). *What is Robotic Process Automation - RPA Software*. [Online]. Available: https://www.uipath.com/rpa/r

Chatbots e Inteligencia Artificial para la Sistematización de Reglas de Glosa en Lengua de Señas Argentina (LSA)

Luis Alberto Olguin, Gustavo Conturzo, Fabrizio Amaya, Raúl Oscar Klenzi, María Isabel Masanet

Instituto de Informática / Departamento de Informática /Facultad de Ciencias Exactas Físicas y Naturales / Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas", Rivadavia, San Juan, Teléfonos: 4260353, 4260355 Fax 0264-4234980, Sitio Web: http://www.exactas.unsj.edu.ar
e-mail: {lolguinunsj, gconturzo, fabrizioamaya88, rauloscarklenzi, mimasanet} @gmail.com

RESUMEN

Este documento presenta una síntesis de una incipiente investigación sobre el uso de sistemas de IA generativa (Claude AI, ChatGPT y DeepSeek) en la creación de un sistema bidireccional de interpretación entre el español escrito y la glosa de la Lengua de Señas Argentina (LSA). Mediante la aplicación de una **ingeniería de prompt iterativa**, se ha desarrollado un conjunto inicial de **30 reglas de traducción directa (español → Glosa LSA)** y **25 reglas de traducción inversa (Glosa LSA → español)**.

El sistema integra fundamentos lingüísticos de autores clave como Burad y Massone.

Los hallazgos preliminares sugieren que los modelos de IA poseen capacidad para identificar patrones implícitos y procesar literatura técnica especializada, estableciendo bases para la democratización de herramientas de accesibilidad para la comunidad sorda. La validación empírica con intérpretes de LSA constituye trabajo futuro.

Palabras clave: Machine Learning, Deep Learning, LLM, LSA.

CONTEXTO

En diciembre de 2025 concluyó el proyecto CICITCA-UNSJ ***"Machine Learning y Deep Learning aplicado a la Lengua de Señas Argentina (LSA)"*** llevado adelante en el trienio 2023-2025 en dependencias del Departamento de Informática (DI) y del Instituto de Informática (Idel) de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFN) dependiente de la Universidad Nacional de San Juan (UNSJ). Las metas alcanzadas, además del aprendizaje y conocimiento adquirido en el mencionado proyecto por los investigadores, posibilita dar continuidad a la tarea investigativa en varias líneas, una de las cuales permitió presentar en una nueva convocatoria para el bienio 2026-2027 el proyecto, aún sin evaluar, ***"Desarrollo de un sistema de interpretación y animación automática de Lengua de Señas Argentina mediante IA y síntesis de movimiento 3D".***

Con el permanente objetivo de lograr una mayor inclusión de la comunidad sorda en la actual era digital, la temática que hace inicialmente de puente entre una y otra propuesta, donde la LSA que es una lengua natural, viso-espacial y ágrafa, con una gramática que difiere sustancialmente del español, es aquella que torna indispensable el surgimiento de una estructura intermedia denominada GLOSA.

- **Necesidad de la Glosa:** La lengua de señas no posee un sistema de escritura (ágrafa) por lo que se hace necesario utilizar una herramienta de transcripción lingüística, Glosa LSA, para representar

de forma escrita la lengua de señas. Consiste en escribir palabras de la lengua oral (generalmente en mayúsculas) que equivalen al concepto de la seña, y sus rasgos no manuales (expresiones faciales y corporales), representados mediante líneas superiores (Peluso, 2020).

- **Desafío Tecnológico:** Se estima que para el año 2050, más de 2500 millones de personas tendrán pérdida auditiva (Organización Mundial de la Salud, 2024). La automatización de la traducción mediante glosas es fundamental para crear herramientas de comunicación escalables y equitativas.
- **Vacío de Investigación:** Si bien existen trabajos internacionales que dan cuenta del uso de LLMs para obtener glosas para ASL, hasta donde hemos relevado, no se ha documentado previamente investigaciones que utilicen ingeniería de prompts iterativa para la sistematización de reglas de glosa específicamente en Lengua de Señas Argentina (LSA).

1. INTRODUCCIÓN

La Lengua de Señas Argentina (LSA) es la lengua natural de la comunidad sorda en el país y constituye un medio fundamental para la comunicación, la educación y el acceso a la información. Sin embargo, la producción y disponibilidad de materiales educativos y herramientas tecnológicas accesibles continúa siendo limitada. Si bien existen avances en modelos de reconocimiento y clasificación de señas, aún falta una línea de investigación y desarrollo enfocada en la **síntesis automática de señas**, es decir, la **generación de movimiento en un avatar animado a partir de datos lingüísticos o glosas estructuradas**.

El proyecto a desarrollar en el actual bienio 2026-2027 consideramos, permitirá abordar ese desafío mediante la integración de técnicas de **visión por computadora, aprendizaje automático y animación 3D**. La idea central consiste en establecer un flujo de trabajo que permita transformar:

> **texto/glosa → secuencia de landmarks (cuerpo + manos) → animación realista en un avatar 3D** (y su camino inverso).

Este tipo de herramientas permitiría crear recursos educativos digitales accesibles para sordos o hipoacúsicos, y desarrollar sistemas de comunicación aumentativa entre personas sordas y oyentes. Por otro lado, se generaría un corpus animado sin necesidad de filmar intérpretes de LSA.

Todo esto favorecerá la estandarización y preservación de la lengua de señas. Asimismo, constituye un aporte novedoso a nivel local, ya que actualmente no existen soluciones de síntesis animada para LSA basadas en codificación de movimiento corporal y posición de las manos.

Como se mencionó, el presente trabajo se centra en una fracción intermedia de tan ambiciosa propuesta, que contempla una interpretación bidireccional de texto español a glosa y viceversa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El desarrollo del sistema se basa en un paradigma de **ingeniería inversa lingüística** y co-creación iterativa con múltiples sistemas de IA.

2.1 Corpus y Procesamiento Multimodal

Esta fase inicial de la propuesta, utiliza un corpus heterogéneo que incluye:

- **Fuentes Teóricas:** "La glosa" de Viviana Burad (2008) y estudios morfológicos de Massone & Martínez (2012).
- **Material Pedagógico:** Apuntes de cursos de LSA y manuales de gramática.

- **Procesamiento de Imagen:** Debido a las limitaciones de tamaño de archivos en versiones gratuitas, se capturaron páginas de apuntes como imágenes PNG. La capacidad multimodal (OCR) de los chatbots permitió extraer reglas de estos formatos analógicos.

2.2 Fases del Desarrollo Iterativo

El proceso de construcción del prompt siguió un modelo incremental de iteraciones:

2.2.1 Extracción de Reglas Base:

Identificación de reglas iniciales para la conversión español → Glosa LSA.

2.2.2 Sistematización Bidireccional:

Generación de reglas inversas manteniendo coherencia teórica.

2.2.3 Integración Morfológica:

Incorporación de procesos derivacionales complejos (clasificadores, nominalización, incorporación numeral).

2.3 Arquitectura del Sistema de Reglas

La investigación identificó una asimetría inherente entre las direcciones de traducción, lo que explica la diferencia en el número de reglas.

2.3.1 Asimetría Estructural y Cognitiva

La traducción hacia la Glosa LSA requiere más reglas (30) que la traducción hacia el español (25) debido a la naturaleza de ambas lenguas:

Español: Marca obligatoriamente género, número, tiempo y artículos.
LSA: Elimina redundancias, priorizando el uso del espacio y el gesto.

Seguidamente se explicitan las 25 reglas utilizadas en esta instancia de traducción desde la Glosa-LSA al Español. Siendo S: Sujeto, O: Objeto y V: Verbo tenemos:

Tabla 1- NIVEL 1: Reorganización Sintáctica

Código	Descripción	Ejemplo
RI1	S-O-V → S-V-O (orden español)	JUAN LIBRO LEER → Juan lee libro
RI2	CD posponer al verbo	REGALO DAR → dar regalo
RI3	CI con preposición "a"	DAR ANA-SN → dar a Ana

Tabla 2 - NIVEL 2: Restauración Gramatical

Código	Descripción	Criterio de Decisión
RI4	Restaurar artículos definidos	Si entidad específica/conocida → el/la
RI5	Restaurar artículos indefinidos	Si entidad nueva/genérica → un/una
RI6	Restaurar preposiciones obligatorias	CI → "a", locativo → "en", etc.
RI7	Concordancia de género/número	Según sustantivo núcleo

Tabla 3 - NIVEL 3: Conversión Pronominal

Código	Descripción	Ejemplo
RI8	PRO1/PRO1pl → yo/nosotros	PRO1 COMER → yo como
RI9	PRO2/PRO2pl → tú,usted/ustedes	PRO2 ESTUDIAR → tú estudias
RI10	PRO3/PRO3pl → él,ella/ellos,ellas	PRO3 TRABAJAR → él/ella trabaja

Tabla 4 - NIVEL 4: Conjugación Temporal

Código	Descripción	Ejemplo
RI11	YA + VERBO → pretérito perfecto simple	YA COMER → comí
RI12	FUTURO + VERBO → futuro simple	FUTURO COMER → comeré
RI13	VERBO (sin marca) → presente	COMER → como/comes/come

Tabla 5 - NIVEL 5: Puntuación y Modalidad

Código	Descripción	Ejemplo
RI14	___int → signos de interrogación + wh al inicio	__int PRO2 COMER QUÉ → ¿Qué comes?
RI15	___neg/___af/___excl → puntuación	_excl QUÉ BUENO → ¡Qué bueno!
RI16	Doble verbo LSA: restaurar orden español	COMER QUERER → quiere comer
RI17	NO-HAY → conjugar TENER/HABER apropiadamente	NO-HAY → no tengo/no hay
RI18	Posesivos LSA: restaurar orden español	CASA MÍA → mi casa

Tabla 6 - NIVEL 6: Reglas Morfológicas Inversas (RI19-RI25)

Código	Descripción	Ejemplo	Contrapartida
RI19	**Reconocer señas lexicalizadas**	VIERNES → viernes	← R26
RI20	**Expandir iniciales geográficas**	S-J-N → San Juan	← R24
RI21	**Des-nominalizar con contexto**	ELEGANTE NOMINALIZADOR → elegancia	← R25
RI22	**Interpretar clasificadores**	VASO CL:cilíndrico → vaso alto	← R27
RI23	**Interpretar clasificadores instrumentales**	TOMAR CL:pinza → tomar con pinza	← R28
RI24	**Interpretar clasificadores de dimensión**	SUPERFICIE CL:redonda → superficie redonda	← R29
RI25	**Expandir incorporación numeral específica**	SEMANA TRES-INCORPORADO → tres semanas PERSONA TRES-INCORPORADO -> tres personas	← R30

3. RESULTADOS OBTENIDOS Y ESPERADOS

3.1 Comparación de Procesos

Tabla 7

Proceso	Español → Glosa LSA (30 Reglas)	Glosa LSA → Español (25 Reglas)	Razón de la Diferencia
Artículos	R2: Eliminación selectiva.	RI4-RI5: Restauración contextual.	Eliminación simple vs. decisión compleja.
Modismos	R20-R23: Detección y transformación.	RI16-RI18: Reconocer y expandir.	Detección compleja vs. expansión directa.
Morfología	R24-R30: Análisis derivacional.	RI19-RI25: Interpretación y expansión.	Glosa LSA → Español es más directo sistemáticamente.

3.2 Algoritmos de Decisión y Validación

Para la traducción inversa (Glosa LSA → Español), se desarrollaron algoritmos lógicos para resolver ambigüedades:

- **Algoritmo de Artículos:** Determina si usar artículos definidos o indefinidos basándose en si la entidad es específica o genérica en el contexto.
- **Algoritmo de Tiempo Verbal:** Si aparece el marcador YA, se usa pretérito; si aparece FUTURO, se usa el tiempo futuro; si no hay marcador, se aplica el presente.
- **Algoritmo de Doble Verbo:** Invierte el orden de LSA (infinitivo + conjugado) al orden del español (conjugado + infinitivo).

3.3 Acciones Perspectivas Futuras

Como línea de trabajo futuro, en el marco del proyecto 2026-2027, se contempla una etapa de validación empírica con intérpretes de LSA para consolidar un ground truth que sirva de referencia para evaluar las reglas generadas por la IA.

3.4 Conclusión

La metodología de co-creación con IA indica que los modelos actuales poseen capacidad para identificar patrones implícitos en la literatura lingüística técnica y convertirla en especificaciones computacionales funcionales. Este enfoque no solo optimiza drásticamente los tiempos de desarrollo, sino que permite la creación de herramientas de traducción bidireccional que respetan la complejidad gramatical y morfológica de la LSA, facilitando el camino hacia una sociedad más equitativa.

4. FORMACIÓN DE RECURSOS HUMANOS.

En el contexto del proyecto concluido a fines del 2025 está próximo a defenderse el trabajo final de grado en LCC titulado "*Del Lenguaje Natural a la Lengua de Señas: Un Enfoque Basado en LLMs*". Así mismo y atento a la nueva convocatoria de proyectos se ha incorporado tres alumnos de las carreras del departamento Informática en condiciones de llevar adelante sus correspondientes trabajos finales de grado. Es interesante destacar que, en la medida de poder avanzar en futuras acciones en línea con lo presentado en la nueva convocatoria, se interactuará con intérpretes de LSA que serán los validadores de lo que se genere como traducción bidireccional automática GLOSA-LSA-> Español.

5. BIBLIOGRAFÍA

[1] Burad, V. (2008). *La glosa: un sistema de notación para la lengua de señas*. Publicado en CulturaSorda.org.

[2] CulturaSorda.org. Apuntes del curso de LSA y material pedagógico.

[3] Díaz Rodríguez (2024). *Aplicación de técnicas de Procesamiento de Lenguaje Natural (PLN) para la traducción del español a la glosa de la Lengua de Señas Colombiana (LSC).*

[4] Hernández Cruz (2019). *Propuesta de un traductor de texto a glosa para la Lengua de Señas Mexicana (LSMx) mediante un módulo de análisis gramatical.*

[5] Massone, M. I., & Martínez, R. A. (2012). *Procesos morfológicos derivacionales en la Lengua de Señas Argentina.*

[6] Organización Mundial de la Salud. (2024). *Sordera y pérdida de la audición*.

[7] Peluso Crespi, L. (2020). *La escritura y los sordos. Entre representar, registrar/grabar, describir y computar*.

[8] Perea-Trigo et al. (2023). *A rule-based translation from written Spanish to Spanish Sign Language glosses.*

[9] Porta, J. (2014). *Sistema automático basado en reglas para glosas LSE.*

[10] Sign4ALL (2023). *Proyecto de la Universidad de Alicante para la interpretación bidireccional entre español y Lengua de Señas Española (LSE).*

DARMIC: Dataset argentino de imágenes de resonancia magnética cardiaca

Sandra Serafino[1], Federico Cintora[123], Lucas Benjamín Cicerchia [1], Hugo Ramón [1], Facundo Argento[1], Luis Prats[1], Ezequiel Bruno[1], Elías Uribe Curapil[1]

[1]Instituto de Investigación y Transferencia en Tecnología (ITT)
Centro asociado a la Comisión de Investigaciones Científicas (CIC)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

[2] Centro de Tomografía Computada y Resonancia Magnética del Noroeste de la Provincia de Buenos Aires (Sanatorio Junín)

[3]Sociedad Argentina de Cardiología (SAC)

{sandra.serafino, federico.cintora, lucas.cicerchia, hugo.ramon, facundo.argento}@itt.unnoba.edu.ar

{laprats, ebruno, eauribecurapil}@comunidad.unnoba.edu.ar

Resumen

Este proyecto de investigación interdisciplinario surge a partir de la consulta de especialistas en diagnóstico por imágenes de resonancia magnética cardíaca. El proceso de segmentación de estas imágenes para cuantificación de cavidades y el cálculo de variables de relevancia clínica son muy importantes a la hora de interpretar el estudio y generar el reporte por parte del médico informante. Por otra parte, el análisis cuantitativo del movimiento y la deformación miocárdica a partir de RMC han emergido en los últimos años como elemento fundamental para la detección precoz de múltiples patologías, con sensibilidad superior al análisis subjetivo o semicuantitativo realizado en la práctica habitual. El problema que enfrentan los médicos es la complejidad, el tiempo de desarrollo de cada uno de estos estudios y la variabilidad inter observador. El objetivo principal de esta propuesta es generar un conjunto normalizado de imágenes de diagnóstico de enfermedades del corazón que cumpla con ser de pacientes argentinos, multicéntrico, multimarca, multi operador y multi diagnóstico (patológicos y normales), con la finalidad de entrenar redes neuronales que permitan predecir, cuantificar y calificar datos que sirvan para optimizar las herramientas de diagnóstico clínico y mejorar la calidad de la información disponible en la consulta del paciente.

Palabras claves: Dataset. Imágenes médicas; Procesamiento digital; Inteligencia Artificial; CNN.

Contexto

Las líneas de investigación a describir se enmarcan en el proyecto de investigación: ***Técnicas y algoritmos avanzados para la optimización de procesos: seguridad, sostenibilidad e innovación***, con lugar de trabajo en el Instituto de Investigación y Transferencia en Tecnología (ITT), presentado en la convocatoria de Subsidios a la Investigación Bianuales (SIB) 2025, aprobado y financiado por la Secretaría de Investigación de la Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA). Este proyecto propone el desarrollo de metodologías innovadoras para el procesamiento de señales (imagen, audio, texto) con un enfoque multimodal, apalancado por tecnologías emergentes como inteligencia artificial (IA), aprendizaje automático (ML), robótica, IoT y computación de alto rendimiento.

La obtención de imágenes de resonancia magnética cardiaca de diferentes centros médicos de todo el país es posible gracias al marco regulatorio del Comité de Ética de la universidad (COENOBA) y de la Sociedad Argentina de

Cardiología (SAC), quienes además apoyan la participación de las instituciones y expertos en el área.

En la actualidad DARMIC es parte también del proyecto “IA argentina en RM cardíaca”, de la Convocatoria CIN (Consejo Interuniversitario Nacional) - UNIUEAR (Alianza Universitaria Argentina Europea para la Transformación Digital) de Proyectos de Desarrollo Tecnológico y Social (PDTS).

1. Introducción

Desde hace más de una década, el Instituto de Investigación y Transferencia en Tecnología trabaja en proyectos vinculados al análisis automatizado de imágenes y visión artificial en distintos campos de aplicación como son el agro, cámaras de seguridad, e imágenes médicas con base en técnicas de automatización, procesamiento digital, sensores, y, redes neuronales.

Las enfermedades cardiovasculares son la principal causa de morbimortalidad en la población adulta en todo el mundo. Las imágenes cardiovasculares han hecho un gran aporte en las últimas décadas al manejo de estas patologías. Su utilidad no se limita al diagnóstico, sino que también son muy importantes para pronóstico, seguimiento y guía de tratamiento. Entre las diversas modalidades diagnósticas disponibles en cardiología, como el ecocardiograma Doppler, la tomografía computarizada, la medicina nuclear y la resonancia magnética; esta última destaca por sus significativas ventajas. La resonancia magnética no emplea radiación ionizante, ofrece una excelente definición tanto espacial como temporal, y posibilita la caracterización tisular. Por consiguiente, se ha consolidado como el método de elección para el diagnóstico y seguimiento de numerosas patologías.

Las imágenes obtenidas en resonancia magnética cardíaca (RMC) deben ser procesadas y analizadas para extraer información clínicamente relevante. La complejidad creciente de las imágenes implica mayor dificultad en la interpretación[1]. Dada esta complejidad y la gran cantidad de parámetros a analizar, las técnicas basadas en IA, tienen gran potencialidad para aportar soluciones con aplicabilidad clínica[2][3].

Consolidar un conjunto de datos normalizados en base a estudios de resonancia magnética cardíaca, de diferentes centros médicos de todo el país, obtenidos por diferentes expertos, utilizando diferentes equipos y abarcando diferentes patologías es fundamental para sustentar el desarrollo de nuevas herramientas de inteligencia artificial para el apoyo diagnóstico[4][5]. Esta iniciativa no solo demuestra el compromiso con la investigación y la innovación dentro del ámbito científico-tecnológico nacional, sino que también garantiza la relevancia y la precisión de los resultados para la población local. La disponibilidad de datos nacionales en el entrenamiento de la IA no solo mejora la exactitud de los diagnósticos, sino que también fortalece la confianza en la aplicación de nuevas herramientas o modelos en cualquier institución de salud del país.

2. Líneas de investigación y desarrollo

Habiendo presentado el problema que se aborda en este proyecto, el equipo de trabajo se centra en:

- Analizar y desarrollar una herramienta para la recolección de imágenes de resonancia magnética cardíaca.
- Analizar y desarrollar los procesos necesarios para la anonimización de los datos digitales obtenidos de acuerdo al marco normativo respectivo (Protocolo).
- Generar modelos de representación y registro de los datos anonimizados de acuerdo a ubicación geográfica, equipo (marca, modelo, potencia), datos antropométricos, edad y sexo, patología.
- Normalizar el conjunto de datos de modo que resulten aplicables al entrenamiento de diferentes modelos de inteligencia artificial, como redes neuronales convolucionales, transformers, modelos fundacionales.
- Desarrollar herramientas de soporte al diagnóstico médico con base en el conjunto de datos generado.
- Mantener e incrementar DARMIC.

3. Resultados obtenidos/esperados

El nivel de implementación de este proyecto, en base a las líneas mencionadas anteriormente, se indica a continuación:

1- Herramienta de recolección

Se desarrolló una herramienta web para la generación de DARMIC. El acceso se realiza mediante el login de un usuario registrado y vinculado a alguno de los centros médicos adheridos al Proyecto. Una vez habilitado el operador puede subir los estudios realizados, los que se someten a la primera etapa del pipeline de anonimización. Los datos quedan vinculados en una base de datos y los estudios guardados en un soporte digital protegido. En la actualidad esta herramienta cuenta con 35 usuarios registrados correspondientes con 41 centros médicos de 12 provincias, que han aportado ya más de 2500 estudios de resonancia magnética.

2- Proceso de anonimización

El proceso de anonimización de datos se enmarca en el protocolo "PROTOCOLO - Cuantificación automática en RM cardiaca V2", aprobado el Comité de Ética de la UNNOBA y el Comité de Ética de la Sociedad Argentina de Cardiología (SAC).

Los archivos son generados por diferentes equipos resonadores y almacenados en formato DICOM[6]. Cada estudio comprende una cantidad diferente de archivos, los que dependen del tipo de estudio. Se aplica un pipeline de 3 etapas de anonimización a cada conjunto de archivos por estudio. La primera etapa anonimiza los nombres originales de los estudios subidos por los especialistas. La segunda etapa anonimiza y/o elimina todos aquellos archivos que no corresponden al formato DICOM o MAT. La última etapa implica la conversión del formato DICOM al formato NRRD[7], con un mínimo de datos relevantes para el cálculo de variables significativas para el diagnóstico médico.

A la fecha se ha generado la versión 1 de DARMIC, con un total de 1800 estudios anonimizados, y se encuentra ya lista para su registro la versión 2 sobre un total de 2500 estudios.

3- Registro y representación

Este proceso implicó la generación de un modelo de base de datos que permitiera el registro relacional de los estudios de resonancia magnética cardiáca como su tabulación en base tipo de resonador (marca, modelo, potencia, ubicación geográfica del paciente, datos antropométricos, edad y sexo, y patología. El dataset cuenta hoy con estudios de 16 patologías diferentes además de estudios normales.

4- Normalización y entrenamiento de datos

Dentro de esta línea de trabajo, por un lado se analizan, prueban e implementan técnicas de procesamiento digital automático relacionadas directamente a las imágenes y las máscaras de ventrículos y miocardio dibujadas por los especialistas médicos para la determinación del diagnóstico resultante. Esta normalización se lleva a cabo considerando las necesidades de los posibles modelos computacionales automáticos que luego podrían utilizarlas. En particular se han entrenado hasta ahora modelos de redes neuronales convolucionales como UNET[8], y transformers como SEGFORMER[9], entre otros.

5- Herramientas de soporte diagnóstico

Se espera que el uso de DARMIC pueda ayudar a generar soluciones informáticas basadas en IA para segmentación automática de cavidades cardíacas y análisis del movimiento y la deformación miocárdica en estudios de RMC[10]. Y que estas herramientas permitan la generación, también automática, de informes clínicos preliminares precisos y reproducibles, con el beneficio de optimizar el diagnóstico clínico y mejorar la calidad de la información disponible en la consulta del paciente. De esta manera, apuntamos a mejorar la productividad, calidad y eficiencia operativa de los servicios que cuenten con este desarrollo.

6- DARMIC incremental

Se espera que DARMIC siga creciendo, tanto en cantidad de estudios como en la incorporación de nuevos equipos, provincias y patologías, que ayuden a incrementar la variabilidad de datos, en pos de nuevos y mejores resultados.

4. Formación de recursos humanos

Este proyecto cuenta a la fecha con la participación de 3 alumnos avanzados de grado de la carrera de Ingeniería en Informática, dependiente de la Escuela de Tecnología de la universidad, formados en este tipo de trabajos y técnicas, con postulación a futuras becas de investigación o especialización de postgrado. Se espera poder sumar a otros alumnos a lo largo del proyecto.

El equipo de trabajo es interdisciplinario, participando en diferentes instancias, médicos cardiólogos, físicos, especialistas en medicina nuclear e informáticos.

4. Bibliografía y referencias

1. Rajiah et al.; Cardiac MRI: State of the Art; Radiology 2023 307(3):https://doi.org/10.1148/radiol.223008.
2. Sermesant, M., Delingette, H., Cochet, H. et al.; Applications of artificial intelligence in cardiovascular imaging; Nat Rev Cardiol 2021 18: 600–609.
3. Argentiero A. et al.; The Applications of Artificial Intelligence in Cardiovascular Magnetic Resonance—A Comprehensive Review; J. Clin. Med. 2022 11: 2866.
4. A. Curiale, F. Colavecchia, G. Mato; Automatic quantification of the LV function and mass: a deep learning approach for cardiovascular MRI; Comp Meth Prog Biomed 2019 169: 37-50
5. M. A. Morales, W. J. Manning, y Reza Nezafat; Present and Future Innovations in AI and Cardiac MRI; Radiology 2024 310:1.
6. NEMA. (n.d.). DICOM - Digital Imaging and Communications in Medicine. Recuperado de https://www.dicomstandard.org/
7. Kindlmann, G. (2019). NRRD: Nearly Raw Raster Data file format. University of Chicago / Teem Project. Recuperado de https://teem.sourceforge.net/unrrdu/index.html
8. Castellaccio, A., Arostegui, N. A., Jiménez, M. P., Tapia, D. Q., Zurita, M. B., & Galván, E. V. (2025). Inteligencia artificial en la imagen cardiovascular mediante resonancia magnética. Radiología, 67(2), 239-247.
9. Mateu, M., Olveres, J., Escalante-Ramírez, B., Raboisson, M. J., Miró, J., & Duong, L. (2025, July). Automatic Segmentation of the Left Ventricle Through the Cardiac Cycle in Pediatric Echocardiography Videos Using SegFormer Architecture. In 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1-5). IEEE.
10. Sun, X., Yin, Y., Yang, Q., & Huo, T. (2023). Artificial intelligence in cardiovascular diseases: Diagnostic and therapeutic perspectives. European Journal of Medical Research, 28(1), 1–12.
11. Al-Kaisey, A. M., Adhikari, S., & Voskoboinik, A. (2023). Artificial intelligence and ECG: A new frontier in cardiac diagnostics and prevention. Journal of Clinical Medicine, 12(3), 1005.

Desarrollo de un chatbot aplicando principios de Usabilidad para la orientación de ingresantes en el sitio web de la Facultad de Ciencias Exactas y Tecnologías

Luz Coria, Eliana Proserpio, Germán Lescano, Pablo Santana-Mansilla

Instituto de Investigación en Informática y Sistemas de Información (IIISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)
Avda. Belgrano (S) 1912 – Santiago del Estero CP 4200

{luzcoria025, elianaproserpio}@gmail.com , {gelescano, psantana}@unse.edu.ar

Resumen

En la actualidad, los chatbots constituyen una de las aplicaciones tecnológicas en crecimiento y muchos de ellos incorporan inteligencia artificial.

La utilidad de los chatbots radica en la capacidad de ofrecer atención continua, mejorar la experiencia del usuario, automatizar tareas repetitivas y garantizar la disponibilidad del servicio sin limitaciones de horario[3]

Sin embargo, estudios recientes muestran que muchos usuarios aún no se sienten satisfechos al interactuar con chatbots y, en ocasiones, prefieren recurrir a la atención humana, lo que revela que su efectividad no depende únicamente de la tecnología subyacente, sino de la calidad de la experiencia de uso.

En este contexto, el presente trabajo tiene como objetivo investigar principios de usabilidad aplicables al diseño de chatbots, con el fin de seleccionar un conjunto de criterios considerados necesarios para garantizar interacciones eficaces, eficientes y satisfactorias. A partir de dicho análisis, estos principios de usabilidad serán tenidos en cuenta para el desarrollo de un chatbot para el sitio web de la Facultad de Ciencias Exactas y Tecnologías (FCEyT), con el propósito de facilitar el acceso a la información académica y administrativa a los estudiantes ingresantes.

Para validar la propuesta, se realizará una etapa de evaluación mediante encuestas de post-interacción, permitiendo medir la percepción de los estudiantes ingresantes sobre el prototipo de chatbot .

Palabras clave: *Chatbot, Usabilidad, Ingresantes, FCEyT, Inteligencia Artificial*

Contexto

En este artículo se presenta una de las líneas de investigación del proyecto “Agentes Conversacionales Moderno”, acreditado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT - UNSE) para el período 2026-2027.

Introducción

La Inteligencia Artificial (IA) se refiere a sistemas o máquinas capaces de imitar ciertas capacidades de la inteligencia humana con el propósito de resolver problemas específicos. En la actualidad, la IA continúa incrementando su presencia en la industria y en la vida cotidiana. En este contexto, los chatbots constituyen una de las aplicaciones tecnológicas en crecimiento, y muchos de ellos incorporan inteligencia artificial. Los chatbots que incorporan IA utilizan técnicas de procesamiento del lenguaje natural para

permitir la interacción con los seres humanos a través de texto[1]

En particular, los chatbots se definen como programas capaces de interactuar con los usuarios mediante lenguaje natural. Son asistentes virtuales que pueden entablar un diálogo de computadora a humano ya sea escribiendo texto o mediante la síntesis de voz. Aparecieron por primera vez en la década de 1960 con el objetivo inicial de los desarrolladores de "hacer creer" a los usuarios que se comunicaban con una persona real [2].

La utilidad de los chatbots radica en la capacidad de ofrecer atención continua, mejorar la experiencia del usuario, automatizar tareas repetitivas y garantizar la disponibilidad del servicio sin limitaciones de horario[3]. De esta manera, se convierten en un recurso estratégico para optimizar procesos y garantizar una atención más eficiente y personalizada. Las empresas desean incorporarlos principalmente porque permiten reducir costos operativos, atender a múltiples usuarios simultáneamente y recolectar datos relevantes sobre las necesidades y comportamientos de los clientes, información que luego puede utilizarse para diseñar estrategias de mejora continua.

Si bien los chatbots se presentan como una herramienta versátil y en auge, capaz de responder consultas de manera inmediata, personalizar la interacción y recopilar datos para optimizar los servicios, estudios recientes muestran que muchos usuarios aún no se sienten satisfechos al interactuar con chatbots. Teniendo en cuenta que los usuarios muchas veces prefieren recurrir a la atención humana, queda claro que la efectividad de los chatbots no depende únicamente de la tecnología subyacente, sino de la calidad de la experiencia de uso [4]

Como parte de un proceso de análisis preliminar para identificar las causas por las que las personas no se sienten satisfechas con los chatbots, se implementó la técnica de brainwriting, mediante la cual los participantes generaron ideas de manera colaborativa y anónima. La actividad contó con la participación de cinco personas de la FCEyT de la Universidad Nacional de Santiago del Estero (UNSE): dos profesionales del área de informática y tres integrantes de las áreas administrativas de la facultad. La pregunta guía de la sesión fue: ¿Por qué algunas personas prefieren no interactuar con chatbots?. Entre las respuestas se destacaron términos vinculados a distintos ejes de análisis. Entre los más frecuentes se encuentran la edad de los usuarios, la tecnología, el diseño y la interfaz del chatbot. Asimismo, se mencionaron la incorporación de inteligencia artificial, la confianza que el sistema debe generar, y su capacidad de dar respuesta a las consultas planteadas. También surgieron factores negativos como la frustración experimentada por algunos usuarios, así como otros aspectos asociados a la experiencia, tales como la satisfacción obtenida, la calidad de las respuestas, el mantenimiento del sistema, la disponibilidad de opciones y la eficacia de la comunicación.

De este análisis se desprende que, para que los chatbots sean realmente útiles y adoptados por los usuarios, es fundamental que la interacción con ellos sea clara, rápida y satisfactoria. En este sentido, la usabilidad constituye un marco adecuado para su evaluación, ya que permite medir factores que influyen directamente en la percepción y la confianza de los usuarios: eficacia, eficiencia y satisfacción [5]

Como antecedentes de este trabajo pueden citarse las siguientes investigaciones.

En [6] analizan el diseño de asistentes digitales en el ámbito universitario, destacando la personalización, el uso de un lenguaje acorde al contexto institucional y la claridad en la organización de la información como factores clave para evitar la frustración del estudiante.

De este estudio se rescata especialmente la importancia de estructurar la información de manera clara y consistente, emplear terminología propia del entorno universitario y brindar retroalimentación comprensible al usuario, aspectos que influyen directamente en

la eficacia, eficiencia y satisfacción. Estos principios aportan un marco conceptual que orienta decisiones de diseño en el desarrollo del chatbot, particularmente en lo referido a la claridad comunicacional y al acceso ágil a información. No obstante, el trabajo se centra en un asistente de planificación y no en un sistema conversacional, por lo que no aborda específicamente los desafíos propios de la interacción mediante lenguaje natural.

En [7] se enfatiza la importancia de los factores humanos a través de la guía GCCD (Guidelines for Chatbot Conversational Design), la cual promueve la transparencia sobre las capacidades del bot y el manejo empático del error para construir confianza. De este antecedente se toma la incorporación explícita de dimensiones como naturalidad, emocionalidad y transparencia en el diseño conversacional, las cuales influyen directamente en la percepción de confianza y satisfacción del usuario. Estos aportes resultan relevantes para el presente trabajo, ya que fundamentan decisiones de diseño orientadas a mejorar la experiencia de los usuarios y contribuyen a fortalecer los componentes de usabilidad durante la interacción.

En [8] proponen dos tecnologías orientadas a mejorar la usabilidad y la experiencia de usuario en chatbots textuales: una checklist de evaluación, y un conjunto de 21 patrones de diseño para guiar su desarrollo desde etapas tempranas. Ambas herramientas fueron validadas experimentalmente y demostraron contribuir a la calidad de la interacción humano–computadora, constituyéndose en un antecedente relevante para el diseño y evaluación de chatbots. De este estudio se rescata especialmente el aporte de patrones vinculados con la conservación del contexto, la retroalimentación y estado del sistema, la robustez ante entradas inesperadas, la comprensión de errores gramaticales y el uso de botones de respuesta rápida, dado que estos inciden directamente en la eficacia y eficiencia de la interacción. Estos aportes ofrecen un marco metodológico específico para sistemas conversacionales, que orienta el diseño del chatbot propuesto en el presente trabajo. Sin embargo, el estudio valida sus herramientas principalmente con desarrolladores e ingenieros de software, sin centrarse en la evaluación con estudiantes universitarios como usuarios finales, aspecto que la presente investigación busca abordar.

El estado del conocimiento actual confirma que, si bien la tecnología necesaria para el desarrollo de chatbots se encuentra ampliamente disponible y en constante evolución, la usabilidad de estos sistemas continúa siendo un problema abierto en términos de diseño. En consecuencia, resulta fundamental investigar criterios de usabilidad que orienten el desarrollo de soluciones conversacionales capaces de garantizar eficacia, eficiencia y satisfacción.

Líneas de Investigación, Desarrollo e innovación

El presente trabajo se enmarca en la línea de investigación sobre Agentes Conversacionales Modernos, con especial énfasis en el análisis de criterios de usabilidad reportados en la literatura científica. A partir de dicha revisión, se definirá un conjunto propio de criterios de diseño de usabilidad que orientará el desarrollo del prototipo. El objetivo principal es estudiar cómo la incorporación sistemática de estos criterios influye en la percepción de los usuarios respecto de la eficacia, eficiencia y satisfacción durante la interacción con el sistema.

La hipótesis que guía esta investigación sostiene que un chatbot que integra el conjunto de criterios de usabilidad definidos en este trabajo será percibido como eficaz, eficiente y satisfactorio por los usuarios.

Para validar esta hipótesis, se desarrollará un prototipo de chatbot como caso de aplicación en el ámbito universitario. Este actuará como instrumento experimental que permitirá evaluar empíricamente la percepción de eficacia, eficiencia y satisfacción por parte de los usuarios en un contexto real.

De este modo, el chatbot no constituye el fin del trabajo, sino un medio para contrastar los criterios de usabilidad propuestos y analizar su impacto en la experiencia de usuario.

Resultados esperados y Objetivos

Este trabajo pretende, desde la perspectiva del proyecto de investigación en el que se halla inmerso revisar mecanismos de control conversacional y evaluación de la calidad y efectividad de la interacción humano-computador.

Desde la perspectiva puntual de este trabajo, se espera obtener la identificación y selección de un conjunto de criterios de usabilidad aplicables al diseño de agentes conversacionales. Este resultado permitirá establecer lineamientos claros que orienten el desarrollo del chatbot, asegurando que contemple los componentes de eficacia, eficiencia y satisfacción del usuario, considerados fundamentales para favorecer interacciones claras, confiables y satisfactorias.

A partir de estos criterios definidos, se espera obtener el desarrollo de un prototipo funcional de chatbot conforme a los lineamientos establecidos. Dicho prototipo será evaluado por una muestra de estudiantes ingresantes mediante una encuesta post-uso, diseñada para medir la eficiencia, eficacia y satisfacción del usuario

A partir del análisis de los datos obtenidos mediante estadística descriptiva, se espera determinar en qué medida el desempeño del chatbot, diseñado con los criterios seleccionados, se refleja en los componentes de la usabilidad: eficacia, eficiencia y satisfacción, buscando validar que los usuarios perciban al chatbot como eficaz en la resolución de consultas, eficiente en términos de tiempo y esfuerzo, y satisfactorio en cuanto a claridad y experiencia de uso. Finalmente, si los indicadores de eficiencia, eficacia y satisfacción resultan favorables, se espera respaldar la hipótesis planteada, demostrando que un chatbot diseñado siguiendo principios de usabilidad es percibido por los usuarios como un sistema usable y acorde con los criterios establecidos.

El objetivo general de este trabajo es desarrollar un prototipo funcional de chatbot dirigido a ingresantes de la Facultad de Ciencias Exactas y Tecnologías (FCEyT), aplicando criterios de usabilidad.

Para alcanzar estos objetivos, primero se investigará y definirá los criterios de usabilidad que guiarán el desarrollo del prototipo de chatbot. Luego, se desarrollará un prototipo de chatbot aplicando los criterios de usabilidad definidos. Por último, se evaluará la percepción de los usuarios (ingresantes) sobre el prototipo de chatbot, mediante la medición de la eficacia, eficiencia y satisfacción del usuario.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto por dos docentes investigadores formados, y dos integrantes estudiantes. Los estudiantes están desarrollando su trabajo final de grado para obtener el título de Licenciado en Sistemas de Información en el marco de esta investigación.

Referencias

[1] Espinosa-Luna, B. H., Castillo-Oliva, J., Montañez-Díaz, B. A., & Mendoza-De los Santos, A. (2023). Implementación de un chatbot basado en modelo de lenguaje de inteligencia artificial para responder preguntas frecuentes de estudiantes universitarios. *Revista Científica de Sistemas e Informática*, *3*(2), e570. https://doi.org/10.51252/rcsi.v3i2.570

[2] Dimitriadis, G. (2020a). Evolution in Education: Chatbots. 3.
Dimitriadis, G. (2020b). Evolution in Education: Chatbots. Homo Virtualis, 3(1), 47. https://doi.org/10.12681/homvir.23456

[3] Campines Barría, F. J. (2024). Chatbots como estrategias de marketing de servicios:

una perspectiva desde la experiencia del cliente. *Revista Colón Ciencias, Tecnología y Negocios*, *11*(2), 1-20. https://doi.org/10.48204/j.colonciencias.v11n2.a5438

[4] Brandtzaeg, P. B., & Følstad, A. (2019). Why People Use Chatbots. *Interacting with Computers*, *31*(3), 317–335.

[5]International Organization for Standardization. (1998). Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability (ISO Standard No. 9241-11:1998).

[6] Bartel, L., Ochs, M., Hirmer, T., & Henrich, A. (2024). Design Principles for a Study Planning Assistant in Higher Education. Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval, 243-253. https://doi.org/10.1145/3627508.3638327

[7] Silva, G. R. S., & Canedo, E. D. (2024). Human Factors in the Design of Chatbot Interactions: Conversational Design Practices. Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems, 1-12. https://doi.org/10.1145/3702038.3702083

[8] Mafra, M. G. S., Nunes, K., Rocha, S., Braz Junior, G., Silva, A., Viana, D., Silva, W., & Rivero, L. (2024). Proposing Usability-UX technologies for the design and evaluation of text-based chatbots. Journal on Interactive Systems, 15(1), 234-251. https://doi.org/10.5753/jis.2024.3856

Detección de enfermedades en ganado mediante visión por computadora e IoT en ambientes extensivos del NOA

Marcos D. Aranda[1,2], Ivanna Lazarte[1], Paola Beltramini[1], Eduardo Cano[1], Juan Pablo Moreno[1], Daniel Villagrán[1], Soledad Bustos[1], Sergio Gallina[1]
maranda@tecno.unca.edu.ar
(1) Universidad Nacional de Catamarca. Facultad de Tecnología y Ciencias Aplicadas.
(2) Universidad Siglo 21. Decanato de Ciencias Aplicadas.

AREA TEMATICA: Agentes y Sistemas Inteligentes.

RESUMEN

La detección temprana de enfermedades en sistemas de ganadería extensiva continúa siendo un desafío debido a la dispersión espacial del ganado, la limitada infraestructura tecnológica y la escasa posibilidad de observación continua. En este trabajo se presenta una línea de investigación orientada al desarrollo de sistemas no invasivos para el monitoreo sanitario bovino, basada en la integración de visión por computadora, dispositivos IoT de bajo consumo e inteligencia artificial.
La línea propone el análisis automático del movimiento y del comportamiento animal a partir de secuencias de video adquiridas en condiciones reales de campo, complementadas con información sensorial proveniente de collares inteligentes. A partir de la detección, el seguimiento visual y la extracción de métricas cinemáticas, se buscan identificar patrones asociados a estados normales y patológicos.
Los resultados parciales obtenidos validan la factibilidad técnica del enfoque y constituyen la base para el desarrollo de sistemas inteligentes de detección temprana de enfermedades, orientados a mejorar el bienestar animal y apoyar la gestión ganadera en entornos rurales con conectividad limitada.

Palabras Claves: Visión por computadora, Seguimiento multiobjeto, Ganadería de precisión, Monitoreo locomotor bovino.

CONTEXTO

La ganadería extensiva en el Noroeste Argentino (NOA) se caracteriza por sistemas productivos de gran escala, baja densidad animal y limitadas capacidades de monitoreo sanitario continuo. En este contexto, la detección tardía de enfermedades impacta negativamente en el bienestar animal, la productividad y la sostenibilidad del sistema productivo.
Los métodos tradicionales de control sanitario, basados principalmente en inspecciones visuales manuales, presentan limitaciones en términos de frecuencia, objetividad y escalabilidad, especialmente en sistemas extensivos. Frente a esta problemática, las tecnologías de visión por computadora, sensores embebidos e inteligencia artificial emergen como alternativas viables para el monitoreo remoto, automatizado y no invasivo del estado sanitario del ganado.
Esta línea de investigación surge como una evolución natural de trabajos previos orientados al monitoreo animal mediante IoT y visión artificial, incorporando progresivamente el análisis del estado de salud como eje central.

1. INTRODUCCIÓN

La ganadería de precisión ha emergido en los últimos años como un enfoque orientado a la incorporación de tecnologías digitales para el monitoreo continuo, objetivo y automatizado de animales de producción, con el fin de mejorar el bienestar animal, optimizar la toma de decisiones y aumentar la eficiencia productiva. En este marco, la visión por computadora y el aprendizaje automático se han consolidado como herramientas clave para

la detección, identificación y seguimiento de animales en entornos productivos reales, permitiendo el análisis cuantitativo del comportamiento sin necesidad de dispositivos invasivos [1], [2].

En sistemas de ganadería extensiva, como los predominantes en el Noroeste Argentino (NOA), el monitoreo sanitario continuo presenta desafíos adicionales asociados a la gran dispersión espacial del ganado, la limitada infraestructura tecnológica y la dificultad de realizar observaciones individuales frecuentes. En este contexto, las alteraciones locomotoras —particularmente la cojera— han sido identificadas como uno de los indicadores tempranos más relevantes de enfermedad, debido a su impacto directo sobre el bienestar animal y la productividad del sistema [3], [4].

Diversos estudios han demostrado que el análisis del movimiento animal a partir de trayectorias visuales permite extraer métricas cinemáticas relevantes, tales como distancia recorrida, velocidad promedio y tiempo de actividad, las cuales pueden correlacionarse con estados de salud, comportamiento y bienestar [2], [5]. En particular, reducciones sostenidas en la velocidad o cambios en los patrones de desplazamiento han sido asociadas con la presencia de patologías incipientes, incluso antes de que se manifiesten signos clínicos evidentes [6].

Los avances recientes en modelos de detección basados en redes neuronales profundas, como la familia YOLO, han habilitado aplicaciones de seguimiento visual automático en tiempo real, incluso en escenarios complejos con variaciones de iluminación, escala y oclusión parcial [7]. Complementariamente, el uso de algoritmos de seguimiento multiobjeto ha permitido mantener la identidad temporal de los animales y reconstruir trayectorias individuales de forma robusta, sentando las bases para el análisis longitudinal del comportamiento [8].

Si bien gran parte de los desarrollos reportados en la literatura se concentran en sistemas de ganadería intensiva o estabulada, su adaptación a contextos extensivos continúa siendo limitada y constituye un desafío científico-tecnológico relevante, especialmente en regiones semiáridas de América Latina [4], [9]. En este escenario, la integración de visión por computadora con dispositivos IoT de bajo consumo y estrategias de procesamiento local (Edge AI) surge como una alternativa prometedora para el diseño de sistemas autónomos, escalables y adecuados a entornos con conectividad restringida [10].

En este contexto, el presente trabajo se inscribe en una línea de investigación orientada al desarrollo de sistemas no invasivos de monitoreo sanitario bovino, basados en la detección y el seguimiento visual automático, el análisis cinemático del movimiento y la integración progresiva de información sensorial. La línea articula investigación aplicada, desarrollo tecnológico y validación experimental en condiciones reales de campo, y constituye la base para la evolución hacia sistemas inteligentes de detección temprana de enfermedades y apoyo a la gestión ganadera en sistemas extensivos.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La línea de I+D tiene como objetivo general desarrollar y validar un sistema automático de monitoreo sanitario bovino, capaz de identificar patrones de comportamiento y movimiento asociados a estados normales y patológicos, con énfasis en la detección temprana de enfermedades en sistemas extensivos.

Los objetivos específicos de la línea incluyen:

- Diseñar sistemas de detección y seguimiento automático de bovinos mediante técnicas de visión por computadora en escenarios reales de campo.
- Reconstruir trayectorias individuales a partir del seguimiento visual multiframe, manteniendo la identidad temporal de cada animal.
- Extraer métricas cinemáticas relevantes (distancia recorrida, velocidad promedio, tiempo activo) como indicadores indirectos del estado locomotor y sanitario.

- Analizar patrones diferenciales de comportamiento entre animales sanos y aquellos con signos visibles o incipientes de enfermedad.
- Sentar las bases para el desarrollo de modelos inteligentes de clasificación del estado sanitario, integrando información visual y sensorial.

La línea articula investigación aplicada, desarrollo tecnológico y validación experimental, con proyección hacia soluciones de ganadería de precisión y Edge AI.

3. METODOLOGÍA GENERAL

El enfoque metodológico combina técnicas de visión por computadora, seguimiento multiobjeto, análisis cinemático y sistemas embebidos de bajo consumo.

Las secuencias de video se adquieren mediante cámaras fijas y drones en establecimientos ganaderos reales. Sobre estas imágenes se aplican modelos de detección basados en redes neuronales profundas, entrenados específicamente para las condiciones ambientales del NOA.

El seguimiento visual permite reconstruir trayectorias individuales, a partir de las cuales se calculan métricas cinemáticas utilizadas como indicadores de comportamiento. De forma complementaria, se integran datos provenientes de collares inteligentes con conectividad IoT, orientados a enriquecer el análisis y favorecer la detección temprana de eventos sanitarios relevantes.

La evaluación del sistema se realiza mediante análisis comparativos y validaciones preliminares en campo, priorizando métricas de consistencia, estabilidad y sensibilidad del enfoque propuesto.

4. RESULTADOS PARCIALES

Como resultados parciales de esta línea, se desarrolló un sistema de detección y seguimiento visual automático de bovinos, evaluado en escenarios reales de producción. A partir de las trayectorias reconstruidas, se calcularon métricas cinemáticas que permitieron identificar diferencias consistentes entre animales con comportamiento normal y aquellos con signos visibles de alteraciones locomotoras. Estas diferencias se manifestaron principalmente en patrones de baja movilidad, menor velocidad promedio y trayectorias más restringidas.

Los resultados obtenidos, presentados previamente en congresos nacionales, constituyen una validación preliminar del enfoque y evidencian el potencial del análisis visual automático como herramienta para la caracterización temprana del estado sanitario bovino.

5. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación se desarrolla en el ámbito de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca, con la participación de docentes-investigadores de las carreras de Ingeniería en Informática e Ingeniería Electrónica.

El equipo está integrado por investigadores con formación de posgrado (doctorado y maestría), docentes investigadores categorizados y docentes en formación, promoviendo un entorno interdisciplinario que articula informática, electrónica y sistemas embebidos.

Asimismo, la línea contribuye activamente a la formación de recursos humanos mediante la participación de tesistas de posgrado y estudiantes de grado, quienes intervienen en tareas de diseño, implementación y validación experimental, fortaleciendo su formación temprana en actividades de investigación aplicada.

6. BIBLIOGRAFÍA

[1] H. Meng, L. Zhang, F. Yang, L. Hai, Y. Wei, L. Zhu y J. Zhang, “Livestock biometrics identification using computer vision approaches: A review,” *Agriculture*, vol. 15, no. 1, p. 102, 2025. doi:10.3390/agriculture15010102.

[2] P. Garcia, J. Smith y A. Lee, “Real-time animal tracking with deep learning: A case study on cattle monitoring,” *Sensors*, vol. 24, p. 6486, 2024. doi:10.3390/s2406486.

[3] M. D. Aranda, P. I. Beltramini, J. E. Cano, L. D. Villagrán, J. P. Moreno, S. H. Gallina y O. A. Herrera Conegliano, “IoT aplicado a la ganadería extensiva,” *Revista Argentina de Ingeniería*, vol. 9, no. 17, pp. 106–113, 2021.

[4] A. Bewley, Z. Ge, L. Ott, F. Ramos y B. Upcroft, “Simple online and realtime tracking,” en *Proc. IEEE ICIP*, 2016, pp. 3464–3468.

[5] C.-Y. Wang, A. Bochkovskiy y H.-Y. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” arXiv:2207.02696, 2022.

[6] M. D. Aranda, J. Redolfi y G. Araguás, “Aplicación de visión por computadora en ganadería de precisión en Catamarca,” en *XX Reunión de Trabajo en Procesamiento de la Información y Control (RPIC)*, 2023.

[7] M. D. Aranda, J. Redolfi y G. Araguás, “Análisis del recorrido de bovinos mediante seguimiento visual automático,” en *RPIC*, 2025.

[8] M. D. Aranda et al., “Técnicas para la optimización de recursos en la detección de objetos en tiempo real,” en *XXVI Workshop de Investigadores en Ciencias de la Computación (WICC)*, 2024.

DETECCIÓN EFICIENTE Y REMOTA DE SALUD CARDIOVASCULAR MEDIANTE VIDEOS

Pedro Barrera (ITBA), María Paula Bonomini (ITBA), Cristian Mateos (ITBA & UNICEN)

e-mail: cmateos@itba.edu.ar

CONTEXTO

Las líneas de trabajo puntualizadas debajo corresponden al proyecto titulado "Detección eficiente y remota de salud cardiovascular mediante videos", el cual se ejecuta en el Instituto Tecnológico de Buenos Aires. En este contexto se desarrolla la tesis doctoral del primer autor.

RESUMEN

Las enfermedades cardiovasculares (ECV) son una causa principal de mortalidad, evidenciando la urgencia de métodos de monitoreo más eficientes y accesibles que los tradicionales, limitados por costos, logística y riesgos de contagio en centros de salud. La detección remota de la salud cardiovascular mediante videos surge como una solución innovadora, alineada con los paradigmas de la medicina personalizada, preventiva y de precisión [1]. Esta tecnología facilita mediciones frecuentes en el entorno del paciente, siendo la información recopilada vital para la medicina de precisión, permitiendo un mejor entendimiento, seguimiento y prevención de las ECV [2].

Sin embargo, existen pocos algoritmos robustos, una escasez de datos para su entrenamiento y validación, y cuestiones tecnológicas por resolver relativas a dispositivos de captura, condiciones de medición y factores humanos [3]. Asimismo, empujar la medición al usuario final implica desafíos adicionales dados por los requerimientos computacionales de tales algoritmos, creando la necesidad de explorar nuevos paradigmas de cómputo intensivo como Edge AI [14].

El proyecto desarrolla algoritmos para la detección remota de parámetros cardiovasculares mediante videos sobre arquitecturas Edge AI. El objetivo es superar los obstáculos tecnológicos actuales, contribuyendo con herramientas que permitan un monitoreo cardiovascular más accesible, continuo, eficiente y personalizado.

1. INTRODUCCIÓN

La fotopletismografía (PPG) es una técnica basada en el principio de Lambert-Beer para estimar los cambios en el volumen sanguíneo midiendo la luz que atraviesa o se refleja en la piel. El latido cardíaco provoca la expansión y contracción rítmica de arteriolas y capilares superficiales, generando una componente pulsátil (AC) superpuesta a un nivel constante (DC) correspondiente a los tejidos no pulsátiles. Sin embargo, esta técnica requiere el uso de sensores de contacto colocados sobre la piel, lo que limita su aplicabilidad ya sea por temas de comodidad como por accesibilidad a oximetros de pulso, limitando su utilización para monitoreo continuo no invasivo en los hogares de manera masiva.

A partir de esta técnica surge rPPG, que conserva los principios fisiológicos, pero sin necesidad de sensores de contacto. En su lugar, emplea cámaras para captar la luz reflejada por la piel y, mediante algoritmos específicos, aísla las diminutas fluctuaciones ópticas asociadas al pulso [4][5].

La monitorización continua y no invasiva de la presión arterial (BP) es de gran interés para la detección temprana y el manejo de la hipertensión, una condición que afecta a una porción significativa de la población mundial. En los últimos años, la rPPG ha surgido como una técnica prometedora para la estimación de la BP sin necesidad de contacto. A partir de la señal rPPG extraída, se pueden derivar diversas características del pulso, como el tiempo de tránsito del pulso (PTT) o características morfológicas de la propia onda rPPG, que han demostrado tener una correlación con los niveles de BP [5,4]. La monitorización remota y continua de la presión arterial mediante métodos sin

contacto, especialmente utilizando rPPG, ha avanzado considerablemente gracias a algoritmos avanzados basados en técnicas modernas de IA. Se han utilizado ampliamente redes neuronales convolucionales (CNN), que permiten extraer patrones espacio-temporales desde señales de video facial para estimar la presión arterial. Schrumpf et al. [6] reportaron errores absolutos medios (MAE) inferiores a 5 mmHg mediante el uso de CNN en escenarios controlados. Por otro lado, se han propuesto modelos de Transformers, atractivos por su habilidad para capturar relaciones temporales complejas de largo alcance en comparación con métodos tradicionales como CNN [4]. En la mayoría de estos enfoques, los parámetros clave utilizados para la estimación de presión arterial incluyen el PTT, así como características morfológicas específicas extraídas directamente de la señal rPPG.

La rPPG puede captarse con una amplia gama de sensores: cámaras RGB convencionales, módulo infrarrojo cercano (NIR, por sus siglas en inglés) dedicados, sistemas híbridos RGB-NIR, matrices multiespectrales, térmicas o incluso cámaras de alta velocidad y profundidad. Cada alternativa aporta ventajas particulares de sensibilidad o inmunidad a la iluminación, como demuestra el uso combinado de RGB y NIR para reforzar la señal pulsátil en condiciones difíciles [10]. Sin embargo, el mayor interés práctico recae sobre las cámaras RGB de los teléfonos móviles, porque están integradas en miles de millones de dispositivos. Este potencial masivo se ve frenado por la gran variabilidad entre marcas y modelos: tamaño de píxel, óptica, filtros IR y algoritmos internos de posprocesado (balance de blancos, compresión, reducción de ruido) alteran la señal rPPG, y la mayoría de los estudios evalúan únicamente un dispositivo, impidiendo comparar métodos y extrapolar el desempeño a la práctica real. Ciertamente, las últimas revisiones subrayan la urgencia de protocolos "cross-device" y de homogeneizar el proceso de adquisición para obtener resultados reproducibles [11] [12] y dar un primer paso hacia la estandarización y evaluación sistemática de estos algoritmos. La carencia de datasets públicos multidispositivo agrava el problema: bases abiertas como PURE, UBFC-rPPG o VIPL-HR se grabaron con una sola cámara y en escenarios controlados, de modo que no cubren la diversidad de sensores ni de condiciones cotidianas, lo cual dificulta un benchmark externo riguroso y retrasa la adopción clínica [12]. Este problema se intensifica en la detección de presión sanguínea, donde la carencia de bases de datos públicas es aún mayor. Para atacar esto, cada estudio utiliza una base de datos propia privada, que además suele ser capturada con un único dispositivo móvil.

En resumen, la detección de presión sanguínea y otros parámetros cardiovasculares muestran potencial, pero su traducción sanitaria exige (i) bases de datos heterogéneas y abiertas; (ii) pipelines de pre-procesado y calidad de señal estandarizados; y (iii) diseños de estudio que cumplan AAMI/ESH/ISO con muestras representativas y múltiples dispositivos.

En contraste, la detección automática de fibrilación auricular (FA) mediante rPPG se encuentra aún en una etapa más incipiente. En la actualidad, existe una considerable abundancia de bases de datos públicas y ampliamente utilizadas de electrocardiogramas (ECG) para la detección automática de FA, tales como la MIT-BIH Arrhythmia Database [7], que incluye diversos registros etiquetados con múltiples arritmias incluyendo la fibrilación auricular, y la base de datos del PhysioNet/CinC Challenge 2017, diseñada específicamente para el desarrollo competitivo de algoritmos de detección de FA mediante ECG [8]. No obstante, la disponibilidad de bases de datos abiertas específicas para detección de FA mediante señales rPPG es notablemente limitada, lo que restringe significativamente el desarrollo y la validación de algoritmos robustos para esta modalidad. Una solución potencial para esta limitación es el uso del aprendizaje por transferencia, en donde algoritmos entrenados previamente con las amplias bases de datos ECG existentes se adaptan posteriormente a señales obtenidas mediante rPPG. Este enfoque ha sido implementado en aplicaciones comerciales, como FibriCheck, una

aplicación móvil que utiliza señales de vídeo facial para detectar FA, cuyos modelos fueron validados con señales ECG [9]. Sin embargo, la literatura actual no detalla qué parámetros específicos del ritmo cardíaco fueron empleados para garantizar la transferibilidad efectiva de estos modelos a señales rPPG, abriendo una clara oportunidad de investigación. Resulta de gran interés identificar cuáles parámetros específicos del ritmo cardíaco (variabilidad RR, entropía del ritmo cardíaco y medidas basadas en diagramas de Poincaré) pueden ser transferidos efectivamente entre señales ECG y rPPG, permitiendo así aprovechar las grandes bases de datos ECG existentes para acelerar el desarrollo y la validación de algoritmos en señales rPPG.

En trabajos anteriores de los autores, se abordó la detección automática de FA usando únicamente parámetros derivados del ritmo cardíaco extraídos desde señales ECG mono derivación (derivación I) obtenidas mediante dispositivos wearables y utilizando 3 bases de datos distintas, alcanzando una precisión cercana al 90% [17][18]. Dado el éxito obtenido mediante la aplicación exclusiva de parámetros del ritmo cardíaco, como por ejemplo índices basadas en entropía, variabilidad del ritmo o diagramas de Poincaré, resulta de gran interés explorar la extensión y aplicación directa de estos algoritmos sobre señales obtenidas mediante técnicas rPPG, evaluando así su capacidad para detectar FA y aprovechando plenamente las grandes bases de datos ECG existentes para acelerar el desarrollo y la validación de algoritmos robustos en esta nueva modalidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto desarrolla algoritmos rPPG para el monitoreo remoto de salud cardiovascular, con foco en presión sanguínea y arritmias (FA), atendiendo a los requerimientos computacionales asociados a su ejecución eficiente en dispositivos de usuario final mediante técnicas de Edge AI. Se persiguen las siguientes líneas:

Construcción de algoritmos de monitoreo remoto de presión arterial: Diseñar e implementar algoritmos para obtener la presión arterial a través de videos faciales, aplicando algoritmos de procesamiento de señales de IA.

Construcción de algoritmo de monitoreo remoto de FA: Diseñar e implementar algoritmos robustos a cambios de dispositivos de captura para a partir de la señal de rPPG obtener índices basados en el ritmo cardiaco, que luego serán utilizados para alimentar un clasificador de FA.

Evaluación de la transferibilidad de parámetros del ritmo cardíaco de electrocardiogramas (ECG) a rPPG para detectar fibrilación auricular: Analizar la variación de los distintos parámetros de irregularidad del ritmo cardíaco que son utilizados para la detección de fibrilación auricular (FA) al ser capturados por ECG vs rPPG, para evaluar si modelos entrenados con ECG pueden aplicarse eficazmente a datos de rPPG dependiendo de qué parámetros de entrenamiento utilizaron.

Estudio comparativo de métodos de captura de señal para rPPG: Realizar un análisis comparativo de las señales rPPG capturadas mediante distintos dispositivos móviles (Android, iOS) y webcams, evaluando cómo la calidad, resolución y procesamiento interno de cada dispositivo afecta a la señal reconstruida y a los parámetros cardiovasculares derivados.

Diseño de estrategia centrada en Edge AI para la ejecución del algoritmo: Diseño de estrategia de ejecución en el Edge utilizando arquitecturas Edge AI [14], que proveen técnicas para aumentar las capacidades computacionales de dispositivos móviles de usuario que ejecutan los algoritmos de detección, y así lograr una detección en tiempo real. Estas técnicas van desde particionado de los algoritmos y paralelización [15] hasta uso de clústers ad-hoc de dispositivos móviles y fijos en el entorno local de ejecución [16].

Metodológicamente, para el monitoreo remoto de BP, se realizará un análisis exhaustivo orientado a evaluar la correlación entre parámetros derivados directamente de la señal rPPG (como el PTT, y características morfológicas específicas de la onda de pulso) y mediciones clínicas validadas

obtenidas mediante un tensiómetro certificado. A partir de estos datos, se implementarán algoritmos de machine learning tradicional, como XGBoost y métodos de Deep Learning, explorando diversas arquitecturas como CNNs, modelos basados en atención como Transformers, y combinaciones híbridas de estos métodos. El desempeño predictivo y la precisión clínica de cada algoritmo serán rigurosamente evaluados y comparados mediante métricas estándar reconocidas internacionalmente, como el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE), para determinar su viabilidad clínica real. Los algoritmos comenzarán desarrollándose sobre la base de datos pública vital videos [13].

En cuanto a la detección de FA, el proyecto se concentra en evaluar la transferibilidad de parámetros derivados del ritmo cardíaco entre señales ECG e rPPG. Para ello, se realizará una extracción y comparación sistemática de diversos índices cuantitativos del ritmo cardíaco, como la variabilidad RR, la entropía del ritmo y análisis gráficos mediante diagramas de Poincaré. A partir de esto, se investigará cómo modelos previamente entrenados con señales ECG pueden adaptarse eficazmente a señales rPPG dependiendo qué atributos se utilizan en el algoritmo. Este análisis será validado mediante métodos estadísticos, para establecer qué parámetros presentan mayor transferibilidad entre modalidades de captura. Para esto, se utilizará la base de datos OBF, una base de datos que posee videos obtenidos simultáneamente por cámaras RGB y NIR para pacientes con y sin fibrilación auricular [19].

En cuanto a la metodología para medir la efectividad de las técnicas de Edge AI, se hará uso de métricas de la computación distribuida, preferentemente throughput, o en este contexto, la cantidad de frames de video que es posible procesar con los algoritmos de rPPG por segundo. También, se considerarán métricas de consumo de batería, por cuestiones ligadas a usabilidad.

3. RESULTADOS ESPERADOS

Se espera identificar las variables críticas y más robustas para la detección de fibrilación auricular (FA) mediante rPPG. Se obtendrá una caracterización profunda sobre qué parámetros del ritmo cardíaco, tales como variabilidad RR, entropía de ritmo cardíaco y parámetros derivados del análisis gráfico con diagramas de Poincaré, poseen mayor estabilidad y transferibilidad desde ECG hacia rPPG. Esto permitirá establecer recomendaciones precisas sobre qué variables deberían considerarse en futuros algoritmos de detección automática de FA en videos.

Asimismo, se explorarán paradigmas de computación recientes para la ejecución de IA en dispositivos de usuario final, en particular Edge AI, que apunta a romper con esquemas de ejecución tradicionales dispositivo-Cloud y apuntan a aprovechar capacidades de hardware en el entorno local de ejecución. Esto trae aparejadas ventajas como el procesamiento real-time y la posibilidad de operar bajo desconexión.

Por último, se prevé el desarrollo de algoritmos propios, validados y robustos, para la estimación remota de presión arterial y detección de FA. Para presión arterial, se espera obtener modelos basados en técnicas avanzadas como XGBoost y algoritmos de deep learning (CNN, Transformers e híbridos), rigurosamente evaluados mediante métricas clínicas estándar (MAE, RMSE) para determinar su precisión y viabilidad clínica. En cuanto a FA, se espera validar la efectividad del aprendizaje por transferencia desde señales ECG hacia rPPG, identificando qué parámetros permiten una adaptación más efectiva y robusta.

4. REFERENCIAS

1. *Hood, L., & Auffray, C. (2013). Participatory medicine: A driving force for revolutionizing healthcare. Genome Medicine, 5(12), Article 110.*
2. *Kraus, S., Schiavone, F., Pluzhnikova, A., & Invernizzi, A. C. (2021). Digital transformation in healthcare: Analyzing the current state-of-*research. Journal of *Business Research, 123, 557-567.*
3. *Debnath, U., & Kim, S. (2025). A comprehensive review of heart-rate measurement using remote*

photoplethysmography and deep learning. BioMedical Engineering OnLine, 24, 73.

4. *Chen, W., Yi, Z., Lim, L. J. R., Lim, R. Q. R., Zhang, A., Qian, Z., Huang, J., He, J., & Liu, B. (2024). Deep learning and remote photoplethysmography powered advancements in contactless physiological measurement. Frontiers in Bioengineering and Biotechnology, 12, Article 1420100.*
5. *Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind-source separation. Optics Express, 18(10), 10762-10774.*
6. *Cheng, H., Xiong, J., Chen, Z., & Chen, J. (2023). Deep learning-based non-contact iPPG signal blood-pressure measurement research. Sensors, 23(12), Article 5528.*
7. *Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine, 20(3), 45-50.*
8. *Clifford, G. D., Liu, C., Moody, B., Lehman, L.-W. H., Silva, I., Li, Q., Johnson, A. E., & Mark, R. G. (2017). AF classification from a short single-lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017. Computing in Cardiology, 44, 1-4.*
9. *Proesmans, T., Mortelmans, C., Van Haelst, R., Verbrugge, F., Vandervoort, P., & Vaes, B. (2019). Mobile phone–based use of the photoplethysmography technique to detect atrial fibrillation in primary care: Diagnostic accuracy study of the FibriCheck App. JMIR mHealth and uHealth, 7(3), e12284.*
10. *Kim, S., Lee, K., Hwang, J., Lee, H., & Lee, E.-C. (2024). Enhanced remote photoplethysmography using hybrid RGB and near-infrared cameras. Journal of Electronic Imaging, 33(6), 063058.*
11. *Lee, R. J., Sivakumar, S., & Lim, K. H. (2024). Review on remote heart-rate measurements using photoplethysmography. Multimedia Tools and Applications, 83, 4.*
12. *Liu, I., Ni, S., & Peng, K. (2020). Enhancing the robustness of smartphone photoplethysmography: A signal quality index approach. Sensors, 20(7), 1923.*
13. *Toye, P.-J. (2023). Vital Videos: A dataset of face videos with PPG and blood-pressure ground truths. arXiv Preprint arXiv:2306.11891.*
14. *Singh, R., & Gill, S. S. (2023). Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 3, 71-92.*
15. *Guo, X., Pimentel, A. D., & Stefanov, T. (2023). Automated exploration and implementation of distributed CNN inference at the edge. IEEE Internet of Things Journal, 10, 5843-5858.*
16. *Hirsch, M., Mateos, C., & Majchrzak, T. A. (2025). Exploring smartphone-based edge-AI inferences using real testbeds. Sensors, 25(9), 2875.*
17. *Barrera, P. L., Vecino Schandy, L. G., Bonomini, M. P., Mateos, C., Hirsch, M., Graña, L. R., & Liberczuk, S. (2023). A machine-learning approach for atrial fibrillation detection in telemonitored patients. Actas del Congreso Argentino de Bioingeniería, 36-45.*
18. *Bonomini, M. P., Mateos, C., Hirsch, M., Vecino Schandy, L., Barrera, P., Grana, L., & Liberczuk, S. (2023). Plataforma de telemonitoreo IoT basada en algoritmos de detección de patologías cardíacas. JAIIO, 9(5), 163-175.*
19. *Li, X., Sidlauskas, K., Pouladzadeh, P., Gabbouj, M., & Iosifidis, A. (2018). The OBF Database: A large face video database for remote physiological signal measurement and atrial fibrillation detection. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 242–249.*

DIALYSMART: SISTEMA DE APOYO A LA TOMA DE DECISIONES CLÍNICAS BASADO EN INTELIGENCIA ARTIFICIAL

Gluk, D.[1]; Goldfarb, M.[1]; Goldfarb, N.[1]; Lamas, C.[1]; Trupkin, I[1]; Di Pinto, D.[1]; López, L.[1]; Vegega, C.[1]; Pollo-Cattaneo, M. F.[1]; Olivas, J. A.[2]

[1] GEMIS - GEMIS.BA
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
Ciudad Autónoma de Argentina. Buenos Aires. Argentina.

[2] Soft Management of Internet and Learning (SMILe)
Universidad de Castilla-La Mancha.
España.

cinthia.vegega@gmail.com

RESUMEN

El crecimiento del uso de los Sistemas de Soporte a la toma de Decisiones basados en Inteligencia Artificial (IA) en dominios críticos complejos e imprecisos, ha intensificado la necesidad de procesos de desarrollo estructurados, transparentes y trazables que combinen datos como conocimiento para la toma de decisiones.

El Grupo GEMIS, desde el año 2022, desarrolla una línea de investigación orientada a la aplicación de IA en salud, con foco inicial en la salud mental y extendiéndose en el año 2025 al ámbito pediátrico. En particular, abordando el dominio complejo de la hemodiálisis pediátrica, desarrollando un Sistema de Soporte a las Decisiones Clínicas basado en IA que combina el conocimiento experto en nefrología con los datos existentes en el dominio.

En este contexto, la línea de investigación se orienta al desarrollo de sistemas híbridos de apoyo a la decisión clínica que integran modelos de Aprendizaje Automático con reglas clínicas estructuradas que son representadas a través de técnicas y herramientas propias de la Ingeniería del Conocimiento. Este trabajo presenta los alcances de dicha línea de I+D, los resultados alcanzados y los objetivos actuales y proyectados, en el contexto de un nuevo proyecto destinado a impulsar sistemas de IA más interpretables, confiables y clínicamente aceptables.

***Palabras clave*:** Inteligencia Artificial; Sistema de Soporte a la Toma de Decisiones Clínicas; Ingeniería del Conocimiento; Hemodiálisis Pediátrica.

CONTEXTO

La Universidad Tecnológica Nacional (UTN) fomenta programas de desarrollo científico y transferencia de conocimiento, impulsando de esta manera a la investigación en tecnología e innovación tecnológica. Es en este marco, que el Grupo de Estudio e Investigación en Modelos y Sistemas de Información (GEMIS), creado en 2009 y formalizado

en 2024 como GEMIS.BA en la Facultad Regional Buenos Aires, refuerza una línea de investigación centrada en la modelización de sistemas y el desarrollo de soluciones basadas en Inteligencia Artificial (IA). Actualmente, una de sus principales líneas de trabajo es la IA aplicada al ámbito de la salud, en donde se desarrollan proyectos en áreas de salud mental y clínica médica.

Durante el último año, la presente línea se fortalece mediante el desarrollo de una propuesta metodológica orientada a la creación de sistemas de apoyo a la toma de decisiones clínicas. En este contexto, se desarrolla un Sistema Clínico de Soporte a las Decisiones (Clinical Decision Support System en inglés, CDSS) para la prescripción de diálisis pediátrica. Este sistema es desarrollado en el marco de la materia "Proyecto Final" de quinto año de la carrera de Ingeniería en Sistemas de Información y pertenece a un caso de estudio de la tesis de doctorado que se encuentra en desarrollo, bajo esta misma línea, llamada "Modelo de Proceso para el Aprovechamiento Inteligente de Datos Masivos guiado por Ingeniería del Conocimiento". Asimismo, se realizan revisiones sistemáticas de la literatura que permiten relevar el estado del arte tanto de la hemodiálisis pediátrica como de las metodologías, técnicas y herramientas de la Ingeniería del Conocimiento, evidenciando la necesidad de integrar el Aprendizaje Automático con el conocimiento propio de los expertos clínicos, creando de esta manera sistemas híbridos de apoyo a la toma de decisiones.

Esto lleva al interés y la cooperación del grupo GEMIS.BA con el grupo de investigación Soft Management of Internet and Learning (SMILe), que desarrolla su trabajo dentro del campo de la aplicación de la Computación Suave, los Sistemas de Apoyo a la Decisiones, el Análisis de Sentimientos y Opiniones y el Análisis Inteligente de Datos en el ámbito de la Universidad de Castilla-La Mancha (España). Esta cooperación se realiza a través de las actividades vinculadas sobre la Ingeniería del Conocimiento y el Análisis Inteligente de Datos.

1. INTRODUCCIÓN

Los Sistemas de Soporte a la Toma de Decisiones basados en IA utilizan datos y conocimientos para apoyar a los tomadores de decisiones en las organizaciones a resolver problemas complejos e imprecisos o problemas mal estructurados [1], manteniendo una alta demanda tanto en investigación como en aplicaciones prácticas. Recientes análisis bibliométricos [2, 3], muestran un crecimiento en las publicaciones y citaciones en los últimos años en diferentes dominios tales como salud, logística, finanzas, entre otros. En particular, los CDSS son una herramienta que contribuye a mejorar la prestación de servicios de salud a través de la asistencia al personal médico en la toma de decisiones relacionadas con tratamientos, diagnósticos o seguimientos clínicos, combinando conocimiento médico basado en evidencias con la información del paciente [4]. La adopción de dichos sistemas beneficia a los pacientes mediante la reducción de errores médicos, mejora del nivel de calidad de la atención y optimización de los recursos de salud [5].

La línea de investigación desarrollada por el Grupo GEMIS en colaboración con el Grupo SMILe, avanza en este marco de trabajo con el fin de desarrollar CDSS basados en IA. Para ello, la combinación de la Ingeniería del Conocimiento y el Aprendizaje Automático permite complementar las fortalezas y debilidades

de ambos, abriendo nuevas posibilidades para la organización del conocimiento y contribuyendo a la creación de Sistemas Inteligentes Híbridos [6]. De esta manera, se logran objetivos complejos combinando la inteligencia humana con la IA de forma de alcanzar resultados superiores a los que cada una de ellas podría haber logrado en forma separada [7]. La participación humana permite evitar los errores o fallos que provocaría un sistema que solo trabaje bajo la IA [8].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de I+D se estructura en los siguientes ejes:

- **Sistemas de Apoyo a la Toma de Decisiones Clínicas basado en IA:** Modelado de un CDSS basado en IA aplicado a la prescripción de hemodiálisis pediátrica, dado que este tipo de diálisis requiere una manipulación cuidadosa de parámetros tales como: acceso vascular, tipo y tamaño del dializador, composición del dializado y caudales de sangre y dializado. Además, la monitorización del paciente, la evaluación y el manejo de complicaciones. Es por esto que es un tratamiento de alta complejidad, que debe realizarse en unidades especialmente equipadas y dentro de hospitales pediátricos, bajo la supervisión permanente de un nefrólogo pediatra [9].

- **Inteligencia Artificial Híbrida:** Diseño de una arquitectura híbrida que integre modelos de Aprendizaje Automático junto con reglas clínicas elicitadas de los expertos y modeladas a través de la utilización de técnicas de la Ingeniería del Conocimiento. La Ingeniería del Conocimiento se define en la norma ISO/IEC 5392 [10], como la "disciplina dedicada a la adquisición de conocimiento a partir de expertos del dominio y otras fuentes de conocimiento, y a su incorporación en una base de conocimiento".

3. RESULTADOS OBTENIDO/ESPERADOS

Entre los resultados obtenidos se destaca el desarrollo de un prototipo en el marco de la cátedra "Proyecto Final" de quinto año de la carrera Ingeniería en Sistemas de Información de la UTN Facultad Regional Buenos Aires (FRBA) del sistema híbrido de apoyo a la toma de decisiones clínicas denominado "DialySmart". La forma de razonamiento de este sistema es un motor de inferencia borrosa [11], dado que de esta manera se puede modelar en forma adecuada el razonamiento clínico de los expertos en un dominio complejo caracterizado por la presencia de incertidumbre e imprecisión. Mientras que la base de conocimientos que procesa dicho modelo posee reglas difusas capaces de capturar y expresar el conocimiento humano representadas como sentencias condicionales de la forma "Si (condiciones) entonces (acciones)." El conocimiento fue obtenido de dos expertas nefrólogas pediátricas.

Como objetivos en curso y futuros, se plantea en el marco del proyecto "Inteligencia Artificial Híbrida para el desarrollo de Sistemas de Apoyo a la Toma de Decisiones en el ámbito médico" (2026–2029), el uso de técnicas de educción, conceptualización y formalización de conocimientos para ampliar la base de conocimiento incorporando nuevas reglas de razonamiento. Una vez incorporadas estas reglas, se espera evaluar dicho sistema en

diferentes escenarios clínicos simulados, analizando su desempeño para poder implementarlo en un entorno real. Asimismo, se desea que la interfaz gráfica sea traducida al idioma inglés para que el sistema sea bilingüe.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación está estrechamente vinculada a la formación de recursos humanos, promoviendo la formación de capacidades relacionadas con la IA de nuevos investigadores. Los trabajos que se desarrollan involucran estudiantes de grado y posgrado de la carrera de Ingeniería en Sistemas de Información de la UTN FRBA, permitiendo el desarrollo de trabajos finales de carrera, trabajos finales de especialización, tesis de posgrado y publicaciones científicas. También se incorporan becarios BINID dentro de esta línea de trabajo. Asimismo, existe una tesis de doctorado que se encuentra en desarrollo que propone un modelo de procesos para el análisis inteligente de datos guiado por Ingeniería del Conocimiento. En [12] se presenta el modelo propuesto junto con ejemplos de casos de uso, que demuestran su generalización y aplicación.

Asimismo, como resultado del año 2025 se publica el prototipo de CDSS confeccionado por los estudiantes de grado en el congreso CONAIISI 2025 [13]. A partir de este proyecto, los estudiantes han aprobado “Proyecto Final” por lo cual se han recibido de Ingenieros en Sistemas de Información.

5. BIBLIOGRAFÍA

1. Panda, S. (2021). AI in Decision Support Systems. SSRN eJournal; Elsevier SSRN: Rochester, NY, USA; ISSN 2994-3981. http://dx.doi.org/10.2139/ssrn.5250480
2. Muñoz-Vilela, A. J., Vellón-Flores, V. I., Solano-Armas, T. & Ramos-la-Rosa, P. E. (2025). Decision support systems: An overview at global research trends. Revista Científica De Sistemas E Informática, 5(2), e914. https://doi.org/10.51252/rcsi.v5i2.914
3. Guan L., Merigó J., Beydoun G. (2025). 40 years of Decision Support Systems: A bibliometric analysis. Decision Support Systems. Volume 194,114469. https://doi.org/10.1016/j.dss.2025.114469
4. Osheroff, J., Teich, J., Levick, D., Saldana, L., Velasco, F., Sittig, D., Rogers, K., Jenders, R. (2012). Improving Outcomes with Clinical Decision Support: An Implementer's Guide, Second Edition. Chicago, IL: Healthcare Information and Management Systems Society (HIMSS). ISBN 9780984457731. https://doi.org/10.4324/9781498757461
5. Gupta P., Ramachandran, A., Keerthi, A., Dave, P., Giridhar, S., Kallapur, S., Saikia, A. (2021). An Overview of Clinical Decision Support System (CDSS) as a Computational Tool and Its Applications in Public Health, Applications in Ubiquitous Computing, Springer, Cham, pp. 81–117. https://doi.org/10.1007/978-3-030-35280-6_5
6. Dellermann D., Calma A., Lipusch W., Weigel N.S.T., P. Ebel P. (2021). The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems. Hawaii Int. Conf. Syst. Sci., Hawaii, USA. https://doi.org/10.48550/arXiv.2105.03354
7. Dellermann D., Ebel P., Söllner M., Leimeister J. M. (2019). Hybrid Intelligence, Business & Information Systems Engineering, 61 (5), 637-64. https://doi.org/10.48550/arXiv.2105.00691
8. Kamar E. (2016). Directions in hybrid intelligence: complementing AI systems with human intelligence. Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16). AAAI Press, 4070–4073. https://dl.acm.org/doi/10.5555/3061053.3061219
9. Alonso Melgar A., Fijo López-Viota, J. (2022). Hemodiálisis Pediátrica. Asociación Española de Pediatría. Protocolos, 1:459-76. ISSN 2171-8172. https://static.aeped.es/25_hemodialisis_0_18076ae576.pdf
10. ISO/IEC 5392:2024 (2024) Information technology — Artificial intelligence — Reference architecture of knowledge engineering. https://www.iso.org/standard/81228.html
11. Zadeh, L. A. (1965). Fuzzy Sets, Information and Control, 8(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X
12. Vegega, C., Pollo-Cattaneo, M.F., Olivas, J.A. (2025). Examples of the Use of Knowledge Engineering for a Better Exploitation of Massive Company Data. 26th International Conference, ICAI 2024, Held as Part of the World Congress in Computer Science, Computer Engineering and Applied Computing, CSCE 2024. Artificial Intelligence and Applications. Communications in Computer and Information Science, vol 2252.

Springer, Cham. Pags. 398-408. ISBN 978-3-031-86623-4. https://doi.org/10.1007/978-3-031-86623-4_36

13. Gluk, D.; Goldfarb, M.; Goldfarb, N.; Lamas, C.; Trupkin, I. (2025) Sistema experto de apoyo a la toma de decisiones para el tratamiento de hemodiálisis en niños. 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información. CONAIISI 2025 (en prensa).

ECOSISTEMA COMPUTACIONAL DE SIMULACIÓN PARA LA PREDICCIÓN Y GESTIÓN DE DESASTRES AMBIENTALES

Miguel Méndez-Garabetti[1,2,3,4], Rodrigo Elgueta[1], Nicolás Ruiz Ahumada[4]

[1] Universidad Nacional de Cuyo, Facultad de Ingeniería.
[2] Universidad del Aconcagua, FCSA.
[3] Universidad Argentina de la Empresa, Instituto de Tecnología (INTEC), Facultad de Ingeniería y Ciencias Exactas.
[4] Universidad Nacional de Cuyo, Instituto Tecnológico Universitario.

miguel.mendez@uncuyo.edu.ar

RESUMEN

Los incendios forestales representan un desafío socio-ambiental crítico a nivel global, exigiendo herramientas predictivas cada vez más sofisticadas frente a escenarios de riesgo extremo. En respuesta a esta creciente necesidad, la presente línea de investigación propone un cambio de paradigma en la modelización de la propagación del fuego mediante el desarrollo de un ecosistema computacional de próxima generación. El enfoque trasciende las limitaciones históricas de los simuladores aislados al articular tres ejes estratégicos: la creación de un framework concurrente de multi-simulación cooperativa, la calibración inteligente de modelos de alta fidelidad física mediante Inteligencia Artificial Explicable (XAI), y la aceleración algorítmica profunda para entornos de Computación de Alto Rendimiento (HPC). Esta convergencia tecnológica distribuida no solo persigue una mejora sustancial en la precisión predictiva y la mitigación de la incertidumbre computacional, sino que busca proveer una infraestructura de software escalable, eficiente y transferible para revolucionar la toma de decisiones en la gestión de emergencias y la planificación territorial.

Palabras clave: simulación de incendios, HPC, multi-simulación, computación evolutiva, Cell2Fire, WRF-Fire.

CONTEXTO

La presente línea de investigación se desarrolla de manera sinérgica a través de tres proyectos en curso, radicados en distintas instituciones académicas bajo la dirección del Dr. Miguel Méndez-Garabetti. Cabe destacar que cada uno de estos proyectos es evaluado y financiado de forma independiente por su respectiva casa de estudios, convergiendo metodológicamente en un ecosistema tecnológico unificado. Cada proyecto aborda una dimensión específica y complementaria del problema:

1. **Universidad Nacional de Cuyo (Facultad de Ingeniería):** El proyecto *"Framework de multi-simulación para predicción de propagación de incendios forestales"* (Proyecto Nº 80020240200027UN, aprobado mediante Resolución RRE 1935/2025 de la Secretaría de Investigación, Internacionales y Posgrado - SIIP) constituye el núcleo integrador de la arquitectura general, siendo responsable del acoplamiento operativo y espacial de los distintos simuladores predictivos.

2. **Universidad Argentina de la Empresa (INTEC/FAIN):** Mediante el proyecto *"Sistema de calibración automática para mejorar la precisión de modelos de incendios forestales mediante inteligencia artificial y cómputo evolutivo"* (Proyecto Nº A26T15–NR, resolución en trámite), se desarrolla la capa de optimización algorítmica y el ajuste paramétrico automatizado basado en IA.

3. **Universidad del Aconcagua (FCSA):** El proyecto *Optimización de la herramienta Cell2Fire [5] para su ejecución en entornos de computación de alto rendimiento (HPC)* (resolución en trámite) se centra en la optimización computacional de bajo nivel y la paralelización para garantizar la escalabilidad de los modelos celulares.

1. INTRODUCCIÓN

Los incendios forestales constituyen desastres ambientales de magnitud creciente, exacerbados por el cambio climático y la constante expansión de la interfaz urbano-forestal. La modelización predictiva de su propagación espacial y temporal resulta una herramienta crítica para la toma de decisiones, la planificación de recursos y la respuesta operativa ante emergencias [4]. Históricamente, el desarrollo de simuladores ha transitado desde modelos empíricos bidimensionales, como BEHAVE [1] y FARSITE, fundamentados teóricamente en las ecuaciones de propagación superficial de Rothermel (1972), hasta sistemas avanzados basados en dinámica de fluidos computacional.

Actualmente, el estado del arte en la predicción de incendios presenta una marcada dicotomía tecnológica. Por un lado, los modelos de alta fidelidad física, como

WRF-Fire y FIRETEC, acoplan bidireccionalmente la dinámica atmosférica con el comportamiento del fuego superficial . Estos sistemas logran capturar fenómenos complejos como la convección térmica y la generación de vientos locales [2]. Sin embargo, su altísimo costo computacional y la exigencia de ajustar cientos de parámetros termodinámicos limitan severamente su viabilidad para generar predicciones operativas en tiempo real o ensayar múltiples escenarios. Por otro lado, los modelos basados en propagación discreta y autómatas celulares, como fireLib y Cell2Fire, ofrecen una notable flexibilidad espacial y gran velocidad de ejecución iterativa, pero adolecen de la robustez física necesaria para representar comportamientos extremos impulsados por la atmósfera [5].

A esta fragmentación tecnológica se suma el problema inherente de la incertidumbre predictiva. Los errores en la resolución de los datos de entrada (cartografía de combustibles, topografía, meteorología local) y las simplificaciones matemáticas de cada algoritmo generan discrepancias significativas entre las simulaciones aisladas y la dinámica real del fuego. Investigaciones previas han demostrado que los enfoques de simulación estática individual son insuficientes para lidiar con entornos dinámicos, evidenciando que la implementación de metodologías algorítmicas para la reducción de incertidumbre como la ejecución de escenarios concurrentes guiados por heurísticas evolutivas mejora sustancialmente la calidad de las predicciones [3].

Frente a este escenario, la presente línea de investigación propone superar las limitaciones de los modelos individuales mediante una arquitectura de software integral basada en la multi-simulación y la optimización distribuida. En lugar de depender de un sistema monolítico, se plantea la construcción de un ecosistema computacional donde distintos frentes metodológicos resuelven cuellos de botella específicos: la integración cooperativa de los simuladores bajo un mismo framework concurrente, la calibración inteligente de los modelos de base física mediante técnicas de Inteligencia Artificial (SAEA/XAI), y la aceleración algorítmica de bajo nivel (MPI/OpenMP) para escalar los modelos celulares. Esta convergencia permitirá generar proyecciones multimodelo precisas y procesables en tiempo real sobre infraestructuras de Computación de Alto Rendimiento (HPC).

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El desarrollo metodológico y tecnológico del ecosistema se estructura a través de tres líneas de trabajo concurrentes, las cuales ya exhiben avances sustanciales en su ejecución:

Línea 1: Multi-simulación y Orquestación Arquitectónica (UNCuyo): Enfocada en el diseño de un entorno concurrente capaz de unificar la ejecución de diversos simuladores de código abierto. A la fecha, se ha logrado completar con éxito el acoplamiento operativo de las dos primeras herramientas de simulación predictiva, estableciendo los protocolos de comunicación de bajo nivel necesarios para la sincronización de matrices de estado espaciales y la consolidación de polígonos de propagación en una única salida de consenso.

Línea 2: Calibración Inteligente y Optimización Evolutiva (UADE): Orientada al desarrollo del framework ECHO (*Evolutionary Coupled High-fidelity Optimization*). Esta línea implementa algoritmos evolutivos combinados con modelos sustitutos basados en Inteligencia Artificial (SAEA/XAI) para el ajuste paramétrico automatizado del modelo físico WRF-Fire, minimizando drásticamente el costo computacional inherente a su alta fidelidad termodinámica.

Línea 3: Aceleración HPC de Modelos Celulares (UDA): Centrada en la reescritura computacional y paralelización profunda. Como logro preliminar consolidado, se ha optimizado exitosamente la ejecución de la herramienta Cell2Fire mediante directivas MPI y OpenMP. Esto ha habilitado su despliegue nativo en arquitecturas multi-nodo (clústeres HPC), garantizando la viabilidad de realizar proyecciones a gran escala con tiempos de procesamiento rigurosamente acotados.

3. RESULTADOS OBTENIDOS Y ESPERADOS

El objetivo rector de este esquema de investigación distribuida es la consolidación de un sistema de predicción de incendios de última generación, capaz de mitigar la incertidumbre inherente a los modelos de simulación aislados.

En cuanto a los resultados preliminares obtenidos, se destaca la validación empírica de la escalabilidad de Cell2Fire en infraestructuras de supercómputo y la implementación de la primera etapa del framework unificador, evidenciada en la integración funcional de los dos simuladores iniciales.

Respecto a los resultados esperados para las etapas subsiguientes, el proyecto proyecta:

- La integración definitiva del sistema ECHO, proveyendo parámetros óptimos para la inicialización dinámica de WRF-Fire dentro del ecosistema.
- La incorporación de módulos matemáticos dedicados a la cuantificación formal de la incertidumbre (UQ) sobre las salidas del ensamble predictivo multimodelo.
- La formulación de metodologías rigurosas de evaluación comparativa (*benchmarking*), utilizando telemetría satelital y datos históricos de incendios reales para validar la precisión espacial del framework.

- La elaboración de protocolos de transferencia tecnológica orientados a facilitar el acceso del software a entidades gubernamentales y agencias de protección ambiental, asegurando una reducción medible en los tiempos de planificación y respuesta operativa ante escenarios críticos.

4. RESULTADOS ESPERADOS/OBTENIDOS

El objetivo rector de este esquema de investigación distribuida es la consolidación de un ecosistema predictivo de última generación, capaz de mitigar la incertidumbre inherente a los modelos aislados y operar con tiempos de respuesta viables para la gestión de emergencias ambientales. Los avances consolidados y las proyecciones a futuro se estructuran de la siguiente manera:

Resultados Preliminares Obtenidos: A la fecha, el avance coordinado de las líneas de investigación ha permitido materializar hitos tecnológicos fundamentales para la viabilidad de la arquitectura propuesta:

- *Despliegue y escalabilidad HPC:* Se ha validado empíricamente la escalabilidad de la herramienta Cell2Fire en infraestructuras de supercómputo. La optimización del código fuente mediante directivas de memoria compartida y paso de mensajes (OpenMP/MPI) ha evidenciado una reducción drástica en los tiempos de ejecución para escenarios bidimensionales de alta resolución espacial.
- *Prueba de concepto de integración concurrente:* Se ha implementado la fase inicial del framework unificador, logrando el acoplamiento funcional de los dos simuladores base. Esto contempló la definición de interfaces estandarizadas para la sincronización dinámica de variables de estado (meteorología, humedad del combustible) y la superposición de matrices topográficas.
- *Arquitectura de modelos sustitutos:* Se ha diseñado la estructura algorítmica preliminar del sistema SAEA/XAI, estableciendo los modelos de aprendizaje automático que actuarán como emuladores paramétricos de la física térmica subyacente en WRF-Fire.

Resultados Esperados a Mediano y Largo Plazo: La convergencia final de los tres proyectos radicados proyecta impactos sustanciales en múltiples dimensiones:

- *Dimensión Tecnológica y Metodológica:* Se espera la integración definitiva del sistema ECHO, el cual funcionará como un motor de inferencia capaz de suministrar configuraciones iniciales óptimas para WRF-Fire en tiempos computacionales reducidos. Asimismo, se consolidará la incorporación de módulos matemáticos dedicados a la Cuantificación de Incertidumbre (UQ), permitiendo que el ensamble multimodelo no solo entregue polígonos deterministas de propagación, sino mapas dinámicos de probabilidad de riesgo.
- *Dimensión de Validación Científica:* Se formularán protocolos rigurosos de *benchmarking*. Esto implicará contrastar las salidas predictivas del framework integral contra telemetría satelital retrospectiva y bases de datos históricas de incendios forestales reales, permitiendo demostrar con rigor estadístico la superioridad del enfoque cooperativo frente a las herramientas aisladas.
- *Dimensión de Transferencia e Impacto Territorial:* El proyecto culminará con la elaboración de manuales y protocolos de transferencia tecnológica dirigidos a entidades gubernamentales y agencias de protección civil (sistemas de manejo del fuego). La liberación del ecosistema bajo esquemas de código abierto (*open-source*) garantizará el acceso democratizado a herramientas de alta fidelidad, aportando una infraestructura de software crítica para la planificación preventiva y la respuesta operativa ante escenarios de desastre.

5. FORMACIÓN DE RECURSOS HUMANOS

Los proyectos involucran una extensa red de formación de investigadores. El equipo cuenta con la dirección del Dr. Miguel Méndez-Garabetti y una estructura que agrupa a investigadores noveles, tesistas de posgrado y más de 25 estudiantes de grado y pregrado en la Facultad de Ingeniería y el Instituto Tecnológico Universitario de la UNCuyo (integrados al proyecto 2025-2027), junto con estudiantes y docentes investigadores de la UADE y la UDA. Las temáticas de investigación promueven la generación de tesis de posgrado y proyectos de graduación enfocados en simulación paralela, inteligencia artificial explicable y arquitecturas HPC.

REFERENCIAS

[1] Patricia L. Andrews. Behave: fire behavior prediction and fuel modeling system-burn subsystem, part 1. *Gen. Tech. Rep. INT-194. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Research Station. 130 p.*, 194, 1986.

[2] Jan Mandel, Jonathan D Beezley, Adam K Kochanski, Volodymyr Y Kondratenko, and Minjeong Kim. Recent advances and applications of wrf-sfire. *Natural Hazards and Earth System Sciences*, 11(4):1083–1101, 2011.

[3] M. Méndez-Garabetti. *Método de Reducción de Incertidumbre Basado en Algoritmos Evolutivos y Paralelismo Orientado a la Predicción y Prevención de Desastres Naturales*. PhD thesis, Universidad Nacional de San Luis, 2020.

[4] J. R. y otros Molina. Incendios forestales: Fundamentos y aplicaciones para su predicción. *Revista de Ciencias Forestales*, 15(2):45–60, 2010.

[5] C. Pais, J. Carrasco, D. L. Martell, A. Weintraub, and D. L. Woodruff. Cell2fire: A cell based forest fire growth model. *arXiv preprint arXiv:1905.09317*, 2019.

El Futuro de la Logística es Inteligente: Sistema de Soporte de Decisiones con IA y Lógica Difusa para Gestión de Stock en la Industria Alimentaria

Ignacio Bufarini[1], Césari Matilde[2], Stefanoni María Eugenia[2]

[1]Alumno de Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina. ignacio.bufarini@alumnos.frm.utn.edu.ar

[2]Docentes cátedra Inteligencia Artificial, Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina. matilde.cesari@frm.utn.edu.ar

RESUMEN

El presente documento detalla la ejecución y los resultados de las Prácticas Profesionales Supervisadas (PPS) desarrolladas por los estudiantes de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional (UTN) - Facultad Regional Mendoza. Este proyecto representa una convergencia estratégica entre la investigación de vanguardia del **Laboratorio ReAVi – PID 8759,** y la necesidad de aplicaciones pedagógicas reales en el aula. Al integrar modelos de lenguaje de última generación con estructuras de conocimiento formal, el sistema trasciende el entorno experimental para convertirse en una herramienta de asistencia directa, mitigando la brecha entre la teoría académica y la práctica tecnológica en el aprendizaje de sistemas complejos

1. CONTEXTO DE LA INVESTIGACIÓN

La gestión de inventarios en las Pequeñas y Medianas Empresas (PYMEs) del sector alimentario enfrenta desafíos críticos que impactan directamente en la seguridad alimentaria y la rentabilidad empresarial. Según estudios recientes, el 65% de las PYMEs alimentarias argentinas experimentan quiebres de stock frecuentes, mientras que el 43% reporta problemas en la trazabilidad de productos (INDEC, 2023).

La problemática se agrava considerando que los sistemas tradicionales de gestión se caracterizan por su naturaleza pasiva: registran datos históricos sin capacidad predictiva ni de análisis semántico. Esta limitación resulta particularmente crítica en la cadena de suministros alimentaria, donde factores como la frescura del producto, desvíos de temperatura y cumplimiento normativo requieren evaluaciones complejas que escapan a la lógica binaria convencional.

La empresa UMAMI Alimentos Saludables, PYME dedicada a la producción de alimentos fermentados, representó el caso de estudio ideal por presentar estas problemáticas en su máxima expresión: quiebres de stock de productos clave como Shio Koji y Tempeh, pérdida de trazabilidad de lotes frescos, y dependencia total del juicio humano para decisiones críticas de despacho.

2. OBJETIVOS TEÓRICOS Y EXPERIMENTALES

2.1. Objetivo General

Desarrollar y validar un Sistema de Soporte de Decisiones (SSD) basado en Inteligencia Artificial que transforme la gestión pasiva de inventarios en un sistema inteligente con capacidad de razonamiento semántico y evaluación de riesgo mediante lógica difusa.

2.2. Objetivos Teóricos

- **O1.** Investigar la aplicabilidad de grafos de conocimiento como estructuras de razonamiento semántico para la trazabilidad alimentaria.
- **O2.** Analizar la efectividad de modelos de lenguaje grandes (LLM) como traductores de lenguaje natural a consultas estructuradas en dominios especializados.
- **O3.** Evaluar modelos de lógica difusa para la cuantificación de riesgos operativos en cadena de suministros.
- **O4.** Diseñar arquitecturas híbridas que integren bases de datos relacionales y grafos de conocimiento manteniendo consistencia transaccional.

2.3. Objetivos Experimentales

- **E1.** Implementar un prototipo funcional utilizando MariaDB como

base de datos transaccional y Neo4j como motor de grafos.

- **E2.** Integrar el modelo qwen2.5-coder:14b mediante Ollama y LangChain para procesamiento de lenguaje natural.
- **E3.** Desarrollar un motor de inferencia difusa basado en el modelo Mamdani para evaluación de riesgos.
- **E4.** Validar el sistema mediante pruebas controladas en operaciones reales de UMAMI durante un período de 90 días.
- **E5.** Medir indicadores de rendimiento: reducción de quiebres de stock, precisión en trazabilidad y tiempo de respuesta del sistema.

3. METODOLOGÍA Y DESARROLLO EXPERIMENTAL

3.1. Arquitectura del Sistema

El proyecto adoptó una arquitectura de "cerebro dual" donde coexisten dos paradigmas de almacenamiento complementarios. La base de datos relacional MariaDB (v10.11) preserva la integridad transaccional y el cumplimiento de propiedades ACID para operaciones críticas. Paralelamente, Neo4j (v5.0) actúa como motor de razonamiento semántico, permitiendo consultas complejas de trazabilidad y relaciones multidimensionales.

La sincronización entre ambos sistemas se implementó mediante un middleware basado en eventos asíncronos, garantizando consistencia eventual con latencias promedio de 150ms. Esta arquitectura permitió mantener la "verdad legal" en tablas estructuradas mientras habilitaba consultas semánticas avanzadas sobre el grafo.

3.2. Componente de Inteligencia Artificial

El procesamiento de lenguaje natural se centralizó en el modelo qwen2.5-coder:14b, seleccionado por su especialización en generación de código y consultas estructuradas. La integración se realizó mediante Ollama como motor de inferencia local y LangChain como framework de orquestación.

El sistema implementa un pipeline de tres etapas:

1. **Extracción de Entidades:** Identificación automática de entidades críticas (cliente, producto, cantidad)
2. **Generación de Consultas:** Traducción de intenciones a consultas SQL y Cypher
3. **Validación Regulatoria:** Verificación de cumplimiento de normativas antes de ejecución

La precisión en la traducción alcanzó un 94.3% en pruebas con 200 muestras de lenguaje natural, superando las expectativas iniciales del 80%.

3.3. Motor de Lógica Difusa

Se implementó un sistema de inferencia difusa mediante el modelo Mamdani, con tres variables de entrada:

- **Frescura del Lote:** {viejo, medio, fresco} basado en días de producción
- **Severidad de Desvío:** {baja, media, alta} medida en desviaciones estándar de temperatura
- **Tiempo de Exposición:** {corto, mediano, largo} en horas

Las reglas de inferencia se definieron mediante consulta a expertos de la industria, generando 27 reglas combinadas. El sistema procesa estas entradas mediante funciones de pertenencia triangulares y trapezoidales, generando salidas cuantificadas que se integran al grafo como atributos de riesgo.

3.4. Experimentos y Validación

El sistema se sometió a cuatro fases experimentales:

Fase 1 (Días 1-30): Configuración inicial y carga histórica de datos. Se migraron 1,247 transacciones del sistema legacy.

Fase 2 (Días 31-60): Operación paralela con sistema tradicional. Se registraron 89 operaciones de prueba comparando resultados entre sistemas.

Fase 3 (Días 61-75): Operación exclusiva con supervisión humana. Se procesaron 156 transacciones reales con monitoreo continuo.

Fase 4 (Días 76-90): Operación autónoma completa. Se analizaron 203 transacciones sin intervención humana directa.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto constituyó una experiencia formativa integral en el contexto de las Prácticas Profesionales Supervisadas (PPS) de la UTN FRM. El desarrollo implicó la adquisición y aplicación de competencias en múltiples dominios tecnológicos:

4.1. Competencias Técnicas Desarrolladas

- **Inteligencia Artificial:** Implementación práctica de LLMs, grafos de conocimiento y sistemas difusos.
- **Bases de Datos:** Diseño y gestión de arquitecturas híbridas relacionales-grafos.
- **Ingeniería de Software:** Desarrollo basado en microservicios y patrones de sincronización.
- **Ciencia de Datos:** Análisis de datos operativos y modelado de indicadores de riesgo.

4.2. Competencias Interdisciplinarias

- **Dominio Aplicado:** Comprensión profunda de procesos logísticos y normativas alimentarias.
- **Investigación Aplicada:** Metodología experimental en contextos industriales reales.
- **Transferencia Tecnológica:** Comunicación efectiva entre mundo académico y sector productivo.

4.3. Impacto Formativo Adicional

El proyecto generó material didáctico replicable:

- Guía de implementación de sistemas híbridos de gestión de conocimiento
- Reproducible pipeline de procesamiento de lenguaje natural para dominios especializados
- Framework de evaluación de riesgo mediante lógica difusa adaptable a otros sectores

5. RESULTADOS Y DISCUSIÓN

5.1. Indicadores Cuantitativos

Métrica	Sistema Tradicional	Sistema Propuesto	Mejora
Quiebres de Stock	12.3% mensual	2.1% mensual	83% reducción
Tiempo Trazabilidad	45 minutos	3 segundos	99.3% reducción
Precisión Despacho	87%	98.7%	13.4% mejora
Alertas Preventivas	0	23 mensuales	Nuevo indicador
Errores Humanos	15 mensual	2 mensual	86.7% reducción

5.2. Análisis Cualitativo

El sistema demostró capacidad para manejar escenarios complejos previamente no atendidos:

Escenario 1: Detección temprana de lote de Tempeh con riesgo de vencimiento combinado con desvío de temperatura. El sistema generó alerta automática 72 horas antes del evento crítico, permitiendo reprogramación de despachos.

Escenario 2: Identificación de patrón de demanda estacional para Shio Koji mediante análisis histórico en el grafo, generando recomendación de stock óptimo que redujo desperdicio en 67%.

Escenario 3: Bloqueo preventivo de despacho de lote afectado por incidente logístico, manteniendo trazabilidad completa y permitiendo respuesta auditada ante autoridades sanitarias.

5.3. Limitaciones y Trabajo Futuro

Si bien los resultados son prometedores, se identificaron áreas de mejora:

- Necesidad de mayor granularidad en variables difusas para productos con ciclos de vida específicos

- Integración con sistemas externos de proveedores para visibilidad end-to-end
- Expansión a múltiples locaciones y gestión distribuida

6. BIBLIOGRAFIA

[1] Li, D., Qu, H., & Wang, J. (2023, November). A survey on knowledge graph-based recommender systems. In 2023 China Automation Congress (CAC) (pp. 2925-2930). IEEE. https://ieeexplore.ieee.org/abstract/document/10450693

[2] Zadeh, L. A. (2011). A note on Z-numbers. *Information sciences*, *181*(14), 2923-2932. https://www.sciencedirect.com/science/article/pii/S0020025511001046

[3] Torres Murillo, L. V., Hidalgo Urrea, J. E., & Álvarez Paz, D. M. (2024). Retos y oportunidades de la inteligencia artificial en la gestión de la cadena de abastecimiento en la industria de alimentos. . https://repository.universidadean.edu.co/items/cde75ab8-c699-4d24-bc43-a41a7303388a

[4] Zhu, X., Li, Q., Cui, L., & Liu, Y. (2024). Large language model enhanced text-to-sql generation: A survey. *arXiv preprint arXiv:2410.06011*. https://ieeexplore.ieee.org/abstract/document/11198017/

[5] INDEC. (2023). *Encuesta Nacional de PYMEs - Sector Alimentario*. Buenos Aires: Instituto Nacional de Estadística y Censos. https://www.indec.gob.ar

[6] Neo4j. (2023). *Graph Database Patterns for Supply Chain Management*. Technical White Paper. **Link de Acceso:** https://neo4j.com/use-cases/supply-chain-management/

Lal, M. (2015). *Neo4j graph data modeling*. Packt Publishing Ltd. http://136.175.10.10:8090/ebook/pdf/Neo4j_Graph_Data_Modeling.pdf

[7] Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies*, *7*(1), 1-13. https://www.sciencedirect.com/science/article/abs/pii/S0020737375800022

[8] Ollama. (2024). *Running LLMs Locally: Best Practices and Performance Optimization*. Documentation Guide. https://docs.ollama.com/

[9] LangChain. (2024). *Building Applications with Large Language Models*. Developer Documentation. https://docs.langchain.com/oss/python/langchain/overview

[10] UTN FRM. (2023). *Lineamientos de Prácticas Profesionales Supervisadas*. Reglamento Académico.

Estimación Personalizada No Invasiva de Hemoglobina mediante Vision Transformers: Colaboración Argentina-Italia

Brenda Carolina Martinez[1]; Giovanni Dimauro[2]; Nunzia Lomonte[2]; Alicia Mon[1]
[1]Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903 - San Justo (CP 1754) Buenos Aires, Argentina
[2] Department of Computer Science
Università degli studi di Bari Aldo Moro
Piazza Umberto I, 1, 70121 Bari BA
brenmartinez@alumno.unlam.edu.ar; giovanni.dimauro@uniba.it; nunzia.lomonte@uniba.it; amon@unlam.edu.ar

RESUMEN

El presente artículo describe los avances de un proyecto de investigación internacional desarrollado en colaboración entre la Universidad Nacional de La Matanza (Argentina) y la Universidad de Bari "Aldo Moro" (Italia), enfocado en la detección no invasiva de anemia mediante técnicas de inteligencia artificial. El proyecto propone un sistema innovador para estimar niveles de hemoglobina a partir de imágenes de la conjuntiva palpebral, utilizando Vision Transformers (ViT) con un enfoque de doble ajuste fino que permite personalización para cada paciente.

Durante la estancia de investigación en Bari, se trabajó con un dataset de 436 imágenes de conjuntiva de 109 pacientes del Instituto Tumori "Giovanni Paolo II", logrando resultados excepcionales: R^2 de 0.94, precisión del 98% y un error absoluto medio de 0.25 g/dL. La metodología desarrollada representa un avance significativo en medicina de precisión, ofreciendo una alternativa accesible y de bajo costo para el monitoreo continuo de hemoglobina, especialmente beneficiosa para pacientes con enfermedades hematológicas que requieren seguimiento frecuente.

Este trabajo de colaboración internacional no solo ha generado resultados científicos de alto impacto, sino que también ha permitido el intercambio de conocimientos y metodologías entre ambas instituciones, fortaleciendo los lazos académicos y abriendo nuevas líneas de cooperación en el área de inteligencia artificial aplicada a la salud.

***Palabras clave**: Inteligencia Artificial, Detección de Anemia, Medicina de Precisión, Vision Transformer, Imagenología Médica, Medicina Personalizada.*

1. CONTEXTO

El presente estudio se desarrolla en el marco del Proyecto TNE-DeSK (Developing Shared Knowledge) [TNE-DeSK Proyect 2024], financiado a través del programa NextGenerationEU, iniciativa impulsada por la Comisión Europea para la recuperación y fortalecimiento de los sistemas de investigación e innovación de los países miembros [European Commission, 2024]. Este Proyecto es liderado por la Universidad de Bari "Aldo Moro", Italia, en tanto que la Universidad Nacional de La Matanza (UNLaM) aporta su colaboración con la participación de estudiantes de Ingeniería Informática realizando una estancia de investigación. El Proyecto forma parte de una línea de investigación sobre el diagnóstico no invasivo de anemia que el grupo de investigación italiano viene desarrollando desde 2020, con múltiples publicaciones en revistas de alto impacto.

El proyecto forma parte de una línea de investigación más amplia sobre diagnóstico no invasivo de anemia que el grupo italiano viene desarrollando desde 2020, con múltiples publicaciones en revistas de alto impacto.

La anemia constituye un problema de salud pública global de gran magnitud. De acuerdo con la Organizacion mundial de la Salud, más de mil millones de personas en el mundo presentan algún grado de anemia [World Health

Organization, 2024]. Estudios epidemiológicos de alcance internacional han confirmado que la prevalencia de anemia continúa siendo elevada tanto en países desarrollados como en regiones de bajos recursos, afectando especialmente a mujeres en edad reproductiva, niños y pacientes con enfermedades crónicas [Stevens et al., 2013; Stevens et al. 2022]. La carga global de la enfermedad no solo implica consecuencias clínicas, como fatiga, disminución del rendimiento físico y cognitivo y complicaciones cardiovasculares, sino también un impacto económico significativo en los sistemas de salud [McLean et al., 2009].

En este contexto, la necesidad de métodos diagnósticos accesibles, frecuentes y de bajo costo resulta particularmente relevante para poblaciones vulnerables y pacientes con patologías oncohematológicas que requieren monitoreo continuo de los niveles de hemoglobina. El desarrollo de herramientas no invasivas basadas en inteligencia artificial surge como una alternativa prometedora frente a los métodos tradicionales basados en extracción sanguínea, especialmente en entornos con recursos limitados o cuando se requieren mediciones repetidas [Stevens et al., 2022].

Actualmente, se está trabajando con un dataset clínico de imágenes de conjuntiva recolectado en colaboración con el Instituto Tumori "Giovanni Paolo II" de Bari, un hospital especializado en oncohematología. Además, se contó con equipamiento especializado de captura de imágenes diseñado específicamente para este proyecto, que garantiza condiciones estandarizadas de iluminación y resolución, permitiendo asegurar calidad y reproducibilidad en los datos utilizados para el entrenamiento de modelos de inteligencia artificial.

2. INTRODUCCIÓN

La anemia es una condición médica que afecta a más de mil millones de personas en el mundo, según la Organización Mundial de la Salud. Se caracteriza por niveles bajos de hemoglobina en sangre, lo que compromete el transporte de oxígeno a los tejidos y puede causar fatiga, debilidad y complicaciones graves si no se detecta y trata oportunamente.

El diagnóstico tradicional de anemia requiere análisis de sangre invasivos que resultan costosos, incómodos para los pacientes y poco prácticos para monitoreo frecuente. Esto representa un desafío particular para pacientes con enfermedades hematológicas crónicas, como leucemia o trastornos de la coagulación, que requieren controles regulares de sus niveles de hemoglobina.

En los últimos años, la inteligencia artificial ha revolucionado el campo del diagnóstico médico [Babenko et al., 2023; Das et al., 2024], permitiendo el desarrollo de métodos no invasivos basados en análisis de imágenes. Diversos estudios han explorado el uso de smartphones y dispositivos móviles para la detección de anemia mediante fotografías de diferentes regiones anatómicas [Ghosal et al., 2020; Suner et al., 2021]. Entre las diferentes regiones estudiadas para la detección de anemia (palmas de manos, uñas, lengua), la conjuntiva palpebral ha demostrado ser la más confiable [Dimauro et al., 2023] debido a su alta vascularización y menor interferencia de factores externos como pigmentación de la piel o actividad manual.

A pesar de los avances en este campo, los métodos existentes presentan limitaciones importantes. La mayoría de los enfoques proponen modelos generales que no se adaptan a las características fisiológicas individuales de cada paciente, lo cual limita su precisión [Das et al., 2024; Dimauro et al., 2023; Appiahene et al., 2023]. Estudios previos han reportado precisiones variables, con resultados que van desde el 69% hasta el 85% en clasificación de anemia [Kato et al., 2024; Appiahene et al., 2023]. Este proyecto introduce una innovación metodológica fundamental basada en arquitecturas Vision Transformer [Vaswani et al., 2017; Camporeale et al., 2025]: el uso de doble ajuste fino, que permite primero aprender patrones generales de la población y luego personalizar el modelo para cada paciente individual, superando las limitaciones de los enfoques generales existentes.

3. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se centra en tres líneas principales de investigación y desarrollo que se complementan para lograr un sistema integral de detección no invasiva de anemia:

El proyecto se centra en tres líneas principales de investigación y desarrollo que se complementan para lograr un sistema integral de detección no invasiva de anemia:

a) Desarrollo de dispositivo de captura de imágenes: Se diseñó y construyó un dispositivo especializado que consiste en una carcasa impresa en 3D adaptada a un smartphone, que incluye una lente macro de 25X y un sistema de iluminación LED blanco de 5mm. Este dispositivo garantiza condiciones estandarizadas de captura, eliminando variables como la iluminación ambiente que podrían afectar la calidad de las imágenes.

b) Procesamiento de imágenes y segmentación: Se implementó una red U-Net para la segmentación automática de la región de interés (conjuntiva palpebral inferior) en las imágenes capturadas. Este paso es crucial para aislar el área diagnóstica y eliminar elementos no relevantes como pestañas, piel circundante o sclera.

c) Modelo de inteligencia artificial con personalización: Se desarrolló una arquitectura basada en Vision Transformer (ViT-Base-Patch16-224) con una metodología innovadora de doble ajuste fino. La primera fase entrena un modelo general con imágenes de todos los pacientes excepto uno (leave-one-out), creando un "modelo especializado en hemoglobina". La segunda fase personaliza este modelo para cada paciente individual usando sus propias imágenes, generando un "modelo personalizado" que se adapta a sus características fisiológicas únicas.

Esta aproximación metodológica propone un enfoque que reconoce a cada individuo con sus características únicas en su conjuntiva (color base, densidad vascular, etc.) que afectan la correlación entre la apariencia visual y los niveles de hemoglobina. Los actuales estudios sobre la problemática se enfocan en modelos generales aplicables a toda la población [Das et al., 2024; Dimauro et al., 2023; Kato et al., 2024].



Figura 1: Ejemplo del antes y después del proceso de segmentación aplicado a una imagen del dataset.

La solución propuesta en la presente investigación permite obtener una mayor precisión al personalizar el modelo para cada paciente.

4. RESULTADOS OBTENIDOS/ESPERADOS

Los objetivos de investigación enmarcados en el desarrollo de esta etapa del proyecto residen en: (1) implementar y validar una metodología de doble ajuste fino sobre arquitecturas Vision Transformer para estimación personalizada de hemoglobina; (2) evaluar la precisión del modelo propuesto utilizando un dataset clínico real de pacientes con condiciones hematológicas; (3) comparar el desempeño del enfoque personalizado frente a modelos generales existentes en la literatura; y (4) establecer la viabilidad clínica del sistema propuesto como herramienta de screening no invasivo.

Los resultados obtenidos en el desarrollo exponen que el modelo personalizado logró las siguientes métricas de desempeño:

- Coeficiente de determinación (R^2) de 0.94, indicando que el modelo explica el 94% de la varianza en los niveles reales de hemoglobina.
- Error absoluto medio (MAE) de 0.25 g/dL, comparable a la variabilidad natural de los análisis de laboratorio (2-5%).
- Error cuadrático medio raíz (RMSE) de 0.44 g/dL.
- Precisión (accuracy) del 98% en la clasificación de pacientes como anémicos o no anémicos.

• Sensibilidad del 100%, asegurando que todos los casos de anemia fueron correctamente detectados (cero falsos negativos).
• Valor predictivo positivo (PPV) de 0.96 y valor predictivo negativo (NPV) de 1.0.

Estos resultados son particularmente notables cuando se comparan con el estado del arte previo. Estudios anteriores reportaban valores de R^2 entre 0.44 y 0.91, precisiones entre 69% y 85%, y MAEs superiores a 1.0 g/dL. Nuestro enfoque representa una mejora sustancial en todas las métricas evaluadas.

La sensibilidad del 100% es especialmente relevante desde el punto de vista clínico, ya que asegura que ningún caso de anemia pasará desapercibido. Esta característica es crucial para un sistema de screening, donde el costo de un falso negativo (no detectar anemia cuando existe) es mucho mayor que el de un falso positivo (que simplemente requiere confirmación con análisis de sangre).

El dataset utilizado consistió en 436 imágenes de conjuntiva de 109 pacientes del servicio de hematología del Instituto Tumori "Giovanni Paolo II" de Bari, capturadas entre mayo de 2023 y agosto de 2024. Los pacientes presentaban diversas condiciones hematológicas incluyendo leucemia y coagulopatías, con un rango de hemoglobina de 8 a 11 g/dL en su mayoría. Cada paciente tuvo un promedio de 5.69 mediciones, con un máximo de 21 mediciones para algunos casos de seguimiento prolongado.

Un aspecto innovador del estudio fue la implementación de una estrategia de validación leave-one-out a dos niveles. En el primer nivel, se excluía un paciente completo del entrenamiento del modelo general. En el segundo nivel, se utilizaban todas las imágenes del paciente excluido menos una para la personalización, y esa imagen final se usaba para la evaluación. Este proceso se repitió iterativamente, permitiendo evaluar el desempeño del modelo en todas las imágenes disponibles.

5. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto ha tenido un impacto significativo en la formación de recursos humanos, tanto para los estudiantes argentinos participantes como para el grupo de investigación italiano:

• Formación especializada en técnicas avanzadas de deep learning, específicamente en arquitecturas Vision Transformer, que representan el estado del arte actual en visión por computadora.
• Experiencia práctica en el desarrollo de aplicaciones de inteligencia artificial para el ámbito médico, incluyendo aspectos éticos y regulatorios específicos de este campo.
• Colaboración directa con médicos especialistas clínicos, permitiendo comprender los requerimientos clínicos reales de un sistema de diagnóstico automatizado
• Participación en un grupo de investigación internacional reconocido, con exposición a metodologías de investigación y acceso a recursos bibliográficos y tecnológicos de primera línea.
• Desarrollo de habilidades en redacción científica en inglés y preparación de manuscritos para publicación en revistas internacionales de alto impacto.

El grupo de investigación GIS, del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, colabora en el proyecto a partir de la formación de estudiantes avanzados de Ingeniería Informática.

Los proyectos precedentes del grupo de investigación han generado diversos resultados científicos y académicos publicados en diferentes congresos nacionales e internacionales y la interdisciplinariedad de este ha facilitado el desarrollo de diversos Proyectos Finales de Ingeniería así como de Tesis de Doctorado.

En ese contexto, una de las autoras del trabajo se encuentra desarrollado una Estancia de Investigación en la Universidad de Bari, participando activamente en la consecución del presente proyecto y se prevé que desarrollo su Proyecto Final de Ingeniería Informática en el transcurso del año 2026.

Esta colaboración sienta las bases para futuras actividades conjuntas entre ambas instituciones, incluyendo posibles proyectos de investigación conjuntos, intercambios estudiantiles adicionales, y el desarrollo de programas de doble titulación en áreas de inteligencia artificial aplicada.

6. BIBLIOGRAFÍA

- Camporeale, M., Clemente, F., Dimauro, G., Lomonte, N., Maglietta, R., Pasciolla, C., Sacco, D., & Zaccaria, G.M. (2025). Highly Reliable Personalized Non-Invasive Hemoglobin Estimation by Using Vision Transformers and Dual Fine-Tuning. Computers in Biology and Medicine.

- World Health Organization. (2024). Anaemia. https://www.who.int/health-topics/anaemia

- Das, S., Ahamed, F., Das, A., Das, D., Nandi, J., & Banerjee, K. (2024). NIADA (Non-Invasive Anemia Detection App), a smartphone-based application with artificial intelligence to measure blood hemoglobin in real-time: A clinical validation. Cureus, 16(7).

- Dimauro, G., Griseta, M.E., Camporeale, M.G., Clemente, F., Guarini, A., & Maglietta, R. (2023). An intelligent non-invasive system for automated diagnosis of anemia exploiting a novel dataset. Artificial Intelligence in Medicine, 136, 102477.

- Babenko, B., Traynis, I., Chen, C., Singh, P., Uddin, A., Cuadros, J., Daskivich, L.P., Maa, A.Y., Kim, R., & Kang, E.Y. (2023). A deep learning model for novel systemic biomarkers in photographs of the external eye: A retrospective study. The Lancet Digital Health.

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.

- McLean, E., Cogswell, M., Egli, I., Wojdyla, D., & de Benoist, B. (2009). Worldwide prevalence of anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993–2005. *Public Health Nutrition, 12*(4), 444–454. https://doi.org/10.1017/S1368980008002401

- Stevens, G. A., Finucane, M. M., De-Regil, L. M., Paciorek, C. J., Flaxman, S. R., Branca, F., Peña-Rosas, J. P., Bhutta, Z. A., & Ezzati, M. (2013). Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995–2011: A systematic analysis of population-representative data. *The Lancet Global Health, 1*(1), e16–e25.

- Stevens, G. A., Paciorek, C. J., Flores-Urrutia, M. C., Borghi, E., Namaste, S., Wirth, J. P., Suchdev, P. S., Ezzati, M., Rohner, F., Flaxman, S. R., & Rogers, L. M. (2022). National, regional, and global estimates of anaemia by severity in women and children for 2000–19: A pooled analysis of population-representative data. *The Lancet Global Health, 10*(5), e627–e639.

- European Commission. (2024). NextGenerationEU recovery plan. https://commission.europa.eu

ESTRATEGIAS AVANZADAS DE OPTIMIZACIÓN BASADAS EN APRENDIZAJE AUTOMÁTICO PARA SISTEMAS DE ENERGÍAS RENOVABLES

Marcelo Cappelletti[1,2], Patricio Gross[1], Lucas Olivera[1], Jésica Guzmán[1], Carlos Schenone[1], Jorge Osio[1,2], Martín Morales[1,3]

[1] Programa TICAPPS, Univ. Nac. Arturo Jauretche, Florencio Varela (1888), Argentina.
[2] Grupo de Control Aplicado, Instituto LEICI (UNLP-CONICET), La Plata (1900), Argentina.
[3] Centro UTN CODAPLI-FRLP, La Plata (1900), Argentina.
e-mail: mcappelletti@unaj.edu.ar

RESUMEN

El crecimiento sostenido de la demanda energética y la transición hacia fuentes renovables plantean nuevos desafíos tecnológicos que requieren soluciones basadas en herramientas computacionales avanzadas. En este contexto, el desarrollo de métodos de modelado, optimización y análisis de datos se vuelve fundamental para mejorar la eficiencia y la integración de sistemas energéticos renovables, contribuyendo al reemplazo progresivo de los combustibles fósiles y a la reducción del impacto ambiental asociado a la generación eléctrica tradicional.

Entre las tecnologías renovables, la energía solar fotovoltaica representa un escenario particularmente relevante desde el punto de vista computacional, debido a la complejidad inherente al recurso solar, su alta variabilidad temporal y la necesidad de optimizar tanto el diseño de dispositivos como la operación de los sistemas. Sin embargo, la incorporación eficiente de esta tecnología en la red eléctrica enfrenta desafíos relacionados con la predicción del recurso, la toma de decisiones bajo incertidumbre y la optimización de procesos en múltiples escalas, aspectos que requieren enfoques basados en algoritmos y aprendizaje automático.

El objetivo general del presente proyecto es generar aportes científico-tecnológicos desde la informática aplicada, mediante el desarrollo y evaluación de estrategias computacionales orientadas a mejorar los procesos de generación de energía eléctrica a partir de recursos renovables. En particular, se propone el diseño de metodologías basadas en aprendizaje automático, optimización y análisis de datos para abordar problemas asociados tanto al diseño de dispositivos fotovoltaicos basados en nuevos materiales como al análisis y gestión de sistemas fotovoltaicos (SFV) bajo diferentes condiciones ambientales.

La investigación se estructurará en torno al desarrollo de modelos predictivos y algoritmos de optimización que permitan extraer conocimiento a partir de grandes volúmenes de datos experimentales y simulados, buscando mejorar el desempeño energético y reducir costos operativos. Asimismo, se explorarán enfoques computacionales orientados a la interpretabilidad de modelos y al soporte de decisiones en escenarios reales.

El proyecto también promueve la formación de recursos humanos en áreas interdisciplinarias, integrando estudiantes y docentes investigadores en temáticas vinculadas a la ciencia de datos, la inteligencia artificial y la computación aplicada a la energía. Finalmente, la propuesta fortalece la colaboración con otros grupos académicos, consolidando una línea de investigación en informática aplicada a sistemas energéticos renovables dentro del ámbito de la UNAJ.

Palabras clave: *Aprendizaje automático (Machine Learning), Optimización computacional, Sistemas energéticos renovables.*

CONTEXTO

Las líneas de investigación y desarrollo presentadas en este trabajo se enmarcan en el Programa de Investigación TICAPPS (Tecnologías de la Información y la Comunicación en Aplicaciones de Interés Social) de la Universidad Nacional Arturo Jauretche (UNAJ), aprobado mediante Resolución N.º 002/24 y dirigido por el Dr. Ing. Morales.

Actualmente, estas actividades se desarrollan en el marco de los proyectos de investigación UNAJ INVESTIGA 2023 (80020230100011UJ), PIP CONICET 2021–2023 (112-202001-01460CO) y PICTO-UNAJ (2022-02-00008), todos bajo la dirección del Dr. Ing. Cappelletti.

1. INTRODUCCIÓN

En las últimas décadas, el avance de la informática y el incremento en la disponibilidad de datos han impulsado el desarrollo de sistemas inteligentes capaces de modelar, analizar y optimizar procesos complejos en múltiples dominios de aplicación. En este contexto, la Inteligencia Artificial (IA) constituye un área central de la ciencia de la computación orientada al diseño de métodos computacionales que permitan representar, aprender y tomar decisiones a partir de información proveniente del entorno [1]. Estas herramientas resultan particularmente relevantes para el análisis de sistemas dinámicos y altamente variables, donde los enfoques tradicionales presentan limitaciones.

Dentro de la IA, el Aprendizaje Automático (Machine Learning) se ha consolidado como un paradigma computacional basado en el desarrollo de algoritmos capaces de aprender patrones a partir de datos, permitiendo construir modelos predictivos y descriptivos sin la necesidad de definir explícitamente reglas determinísticas [2]. Desde una perspectiva informática, el aprendizaje automático aborda problemas relacionados con el modelado estadístico, la optimización de funciones objetivo y la generalización de modelos a partir de conjuntos de entrenamiento. Estas técnicas han adquirido gran relevancia en áreas como visión por computadora, procesamiento del lenguaje natural, robótica, finanzas, medicina y sistemas energéticos, entre otras.

El presente trabajo se enfoca en algoritmos de aprendizaje supervisado, en los cuales los modelos se entrenan utilizando datos etiquetados, ajustando sus parámetros e hiperparámetros mediante procesos de optimización orientados a minimizar el error de predicción. Dependiendo del tipo de salida requerida, estos algoritmos pueden abordar problemas de regresión, cuando el objetivo es estimar valores continuos, o problemas de clasificación, cuando se busca asignar una instancia a un conjunto finito de categorías.

Desde esta perspectiva, la investigación aborda dos problemas computacionales complementarios dentro del dominio energético. Por un lado, el modelado y optimización de dispositivos de conversión fotovoltaica basados en nuevos materiales, donde las técnicas de simulación y análisis de datos permiten estudiar el comportamiento microscópico de los materiales y optimizar el diseño de las celdas solares. Por otro lado, el análisis y estimación de variables asociadas al desempeño de sistemas fotovoltaicos (SFV) bajo diferentes condiciones atmosféricas, problema caracterizado por la presencia de incertidumbre, alta variabilidad temporal y disponibilidad limitada de datos experimentales.

En el primer caso, el desarrollo de nuevas tecnologías fotovoltaicas, como las celdas solares de perovskita [3, 4], plantea desafíos computacionales vinculados al modelado físico y la optimización multidimensional de parámetros, buscando mejorar la eficiencia y reducir costos de fabricación. Estos dispositivos han incrementado de manera significativa su eficiencia desde su aparición: desde 3.8% en el año 2009 [5] hasta 26% en el año 2025 [6]. Sin embargo, a pesar del alto rendimiento proporcionado por las CSP, desafortunadamente estos compuestos son sensibles a cambios de humedad. Además, la

presencia de plomo (Pb) lo vuelve tóxico y poco amigable al medioambiente [7]. Estos factores podrían obstaculizar la futura comercialización de estos dispositivos, por lo que mayores estudios son necesarios para remediar estos inconvenientes.

En el segundo caso, la estimación de la radiación solar constituye un problema clásico de modelado predictivo, donde los métodos tradicionales presentan limitaciones frente a la complejidad de los fenómenos atmosféricos [8-10]. En este contexto, la cobertura de nubes es un factor que atenúa y ocasiona intermitencia en los valores de la energía proveniente del sol. Esta variabilidad de la radiación solar provoca que la producción de energía eléctrica que pueden generar los SFV no sea constante ni totalmente predecible en el tiempo, lo cual es una desventaja con respecto a los combustibles fósiles. El conocimiento preciso de la radiación solar que alcanza un SFV en un determinado instante y lugar tiene un rol determinante para: (i) la evaluación de la producción potencial de energía eléctrica que podría obtenerse en ese lugar, ofreciendo una mayor previsibilidad a este tipo de tecnología, para que pueda satisfacer la demanda en los momentos de intermitencia; (ii) el desarrollo de estrategias que permitan operar al SFV en el punto de máxima potencia la mayor parte del tiempo.

El crecimiento de las capacidades de cómputo y el avance de los algoritmos de aprendizaje automático han permitido desarrollar modelos más robustos y precisos, capaces de extraer patrones relevantes a partir de grandes volúmenes de datos y mejorar la toma de decisiones en sistemas energéticos reales [11–13].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación y desarrollo presentadas en este trabajo se centran en el estudio y la aplicación de técnicas de aprendizaje automático como herramientas computacionales para el modelado, análisis y optimización de sistemas energéticos basados en fuentes renovables, particularmente en el dominio de la energía solar fotovoltaica. Desde una perspectiva informática, el objetivo principal consiste en desarrollar metodologías capaces de extraer conocimiento a partir de datos experimentales y simulados, permitiendo mejorar el diseño, la operación y la eficiencia de dichos sistemas, contribuyendo simultáneamente a la reducción de costos y al aumento de su competitividad.

En este marco, se busca que las soluciones desarrolladas presenten un desempeño robusto en términos de precisión predictiva y capacidad de generalización, junto con requisitos computacionales razonables que permitan su implementación práctica. Asimismo, se priorizan aspectos relacionados con la eficiencia algorítmica, la estabilidad numérica, la velocidad de convergencia y la facilidad de integración en entornos reales de análisis y toma de decisiones.

Las metodologías propuestas serán aplicadas en dos líneas principales:

(a) Diseño y optimización de materiales y dispositivos fotovoltaicos mediante técnicas de modelado computacional y aprendizaje automático, orientadas a la exploración eficiente de espacios de parámetros y a la mejora del desempeño de dispositivos de bajo costo.

(b) Modelado y optimización de sistemas fotovoltaicos (SFV) bajo diferentes condiciones atmosféricas, abordando problemas de estimación, predicción y análisis de desempeño mediante enfoques basados en datos.

Ambas líneas de investigación se desarrollarán mediante un enfoque teórico–experimental, integrando simulación computacional, análisis de datos y validación experimental. La combinación de técnicas de inteligencia artificial con el estudio de materiales, dispositivos y sistemas energéticos constituye una propuesta interdisciplinaria

con fuerte componente computacional, favoreciendo la participación de estudiantes, graduados e investigadores interesados en áreas como aprendizaje automático, optimización y ciencia de datos aplicada.

3. RESULTADOS ESPERADOS

En relación con la línea de investigación orientada al diseño y optimización de materiales y dispositivos fotovoltaicos de bajo costo, se espera generar aportes desde el modelado computacional y el aprendizaje automático que permitan mejorar el análisis y la optimización de tecnologías emergentes, particularmente las celdas solares de perovskita. En este contexto, se prevé el desarrollo de modelos capaces de explorar de manera eficiente espacios complejos de parámetros, contribuyendo a la mejora de la eficiencia de conversión y al estudio de problemáticas asociadas a la estabilidad térmica y la toxicidad de los materiales, aspectos críticos para su futura transferencia tecnológica y comercialización.

Por otra parte, en la línea vinculada con la optimización de sistemas fotovoltaicos (SFV) bajo diferentes condiciones de cielo, se espera desarrollar herramientas computacionales basadas en aprendizaje automático orientadas a la estimación y análisis de la atenuación de la radiación solar en escenarios atmosféricos variables. Estas metodologías permitirán modelar la influencia de la nubosidad y otras variables meteorológicas sobre el desempeño energético, favoreciendo estrategias de adaptación y optimización operativa de los SFV. Como resultado, se busca mejorar la eficiencia global del sistema y reducir costos asociados a la generación eléctrica mediante una mejor capacidad predictiva y de soporte a la toma de decisiones.

Desde una perspectiva informática, los resultados esperados incluyen la construcción de modelos predictivos robustos, el análisis comparativo de algoritmos y la generación de metodologías reproducibles aplicables a problemas complejos de sistemas energéticos basados en datos. Asimismo, las herramientas desarrolladas podrán extenderse a otros dominios donde existan fenómenos dinámicos con alta variabilidad e incertidumbre.

Finalmente, se espera que los avances obtenidos tengan impacto en aplicaciones reales, incluyendo sectores productivos como la actividad agroindustrial, promoviendo la adopción de tecnologías basadas en energías renovables en el área de influencia de la UNAJ y fortaleciendo la integración entre informática aplicada y problemáticas de interés social.

4. FORMACIÓN DE RECURSOS HUMANOS

La formación de recursos humanos constituye uno de los objetivos centrales del Programa TICAPPS dentro de las líneas de investigación y desarrollo presentadas en este trabajo. En particular, se busca consolidar un espacio interdisciplinario orientado a la capacitación de docentes investigadores y estudiantes en áreas vinculadas con la Inteligencia Artificial, el aprendizaje automático y el análisis computacional de datos, promoviendo el desarrollo de competencias técnicas alineadas con los desafíos actuales de la ciencia de la computación aplicada.

El programa apunta a fortalecer la generación de soluciones innovadoras basadas en herramientas computacionales avanzadas, favoreciendo la formación de perfiles capaces de integrar metodologías de modelado, optimización y análisis algorítmico en problemas reales. Este enfoque resulta estratégico para el desarrollo científico y tecnológico del territorio de influencia de la UNAJ, impulsando la transferencia de conocimiento y la consolidación de capacidades locales en tecnologías digitales.

El equipo de trabajo asociado a estas líneas de I/D está integrado por cuatro docentes investigadores formados (tres Doctores y un Magíster), tres docentes investigadores en formación y dos estudiantes avanzados. Asimismo, actualmente se desarrollan dos tesis doctorales directamente vinculadas con

las temáticas abordadas, lo que contribuye a la generación de nuevo conocimiento y a la formación de investigadores especializados.

Finalmente, la investigación propuesta favorece la continuidad y el fortalecimiento de colaboraciones con otros centros académicos de la región, promoviendo redes de trabajo orientadas al desarrollo de investigación en inteligencia artificial y computación aplicada.

5. BIBLIOGRAFÍA

[1] N. Nilsson. "Principles of Artificial Intelligence", Springer-Verlag, New York (1980).

[2] E. Alpaydin. "Machine Learning: The New AI". Editorial MIT Press. (2016).

[3] J. Burshka, et al. "Sequential deposition as a route to high-performance perovskite-sensitized solar cells". Nature. Vol. 499, pp.316-319 (2013). DOI: https://doi.org/10.1038/nature12340

[4] V. Gonzalez-Pedro et al.. "General working principles of $CH_3NH_3PbX_3$ perovskite solar cells", Nano Lett. Vol. 14, pp 888-893 (2014). DOI: https://doi.org/10.1021/nl404252e

[5] A. Kojima, K. Teshima, Y. Shirai y T. Miyasaka. "Organometal halide perovskites as visible-light sensitizers for photovoltaic cells". J. Am. Chem. Soc. Vol. 131 (17), pp.6050-6051, (2009). DOI: 10.1021/ja809598r

[6] National Renewable Energy Laboratory (NREL). Research Cell Efficiency Chart. https://www.nrel.gov/pv/cell-efficiency.html (accessed December 30, 2025).

[7] W. Ke y M. Kanatzidis. "Prospects for low-toxicity lead-free perovskite solar cells". Nature Commun., Vol. 10, pp.965-968, (2019). DOI: doi.org/10.1038/s41467-019-08918-3

[8] C. Raichijk, H. Grossi Gallegos, R. Righini. "Evaluación de un método alternativo para la estimación de valores medios mensuales de irradiación global en Argentina". Avances en Energías Renovables y Medio Ambiente, Vol. 9. pp. 11.05–11.08, (2005).

[9] A. Assi, M. Jama, M. Al-Shamisi. "Prediction of Global Solar Radiation in Abu Dhabi". Int. Scholarly Research Network ISRN Renewable Energy. Vol. 2012, Article ID 328237, 10 pages, (2012). DOI: doi.org/10.5402/2012/328237

[10] J. Polo Martínez. "Optimización de modelos de estimación de la radiación solar a partir de imágenes de satélite". Tesis Doctoral. Universidad Complutense de Madrid, (2010).

[11] Z. Pang, F. Niu, and Z. O'Neill. "Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons". Renewable Energy, 156:279–289, 2020.

[12] P. Kumari and D. Toshniwal. "Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance". Journal of Cleaner Production, 279:123285, 2021.

[13] M. Mukhtar, et al. "Development and comparison of two novel hybrid neural network models for hourly solar radiation prediction". Applied Sciences, 12(3), 2022.

Evaluación de espectrogramas en redes neuronales convolucionales para la detección de la enfermedad de Parkinson

Marta Eugenia Suarez[1], Verónica Inés Aubin[2], Renata Silvia Guatelli [2]
1- Universidad Nacional Arturo Jauretche, 2- Universidad Nacional de La Matanza
martaeuge@hotmail.com; {vaubin, rguatelli}@unlam.edu.ar

RESUMEN

La detección temprana de la enfermedad de Parkinson continúa siendo un desafío clínico relevante, lo que motiva la búsqueda de métodos objetivos, no invasivos y de bajo costo basados en el análisis de la voz. En este contexto, la señal vocal constituye un biomarcador accesible capaz de reflejar alteraciones sutiles detectables mediante técnicas de aprendizaje profundo. Este trabajo presenta una línea de investigación orientada a evaluar el desempeño de modelos de clasificación automática basados en redes neuronales convolucionales aplicadas a distintas representaciones espectrales de señales de voz. El aporte central radica en realizar un análisis sistemático y comparable tanto de las arquitecturas de clasificación como de las representaciones espectro-temporales, abordando de manera integrada aspectos que en la literatura suelen estudiarse de forma fragmentada. Se desarrolla un estudio comparativo de espectrogramas basados en Transformada de Tiempo Corto de Fourier, Mel-espectrogramas con y sin normalización de energía por canal, transformadas wavelet, coeficientes cepstrales en la frecuencia de Mel y representaciones de frecuencia constante, considerando que cada técnica enfatiza distintas propiedades tiempo–frecuencia. Para garantizar consistencia experimental, se emplean muestras de la base HUNIPA-2019 y métricas estándar de clasificación, con el objetivo de aportar lineamientos metodológicos que permitan mejorar la calidad diagnóstica.

Palabras clave: *procesamiento de audio, espectrograma, red neuronal convolucional, enfermedad de Parkinson.*

CONTEXTO

Se presenta una línea de investigación que es parte del proyecto “C2-ING-140", "Análisis comparativo de Espectrogramas para la Detección de la Enfermedad de Parkinson mediante Inteligencia Artificial”. Esta línea propone el estudio comparativo de distintas representaciones espectrales y su impacto en el desempeño de los modelos de clasificación por un método objetivo y no invasivo para el diagnóstico precoz de la enfermedad de Parkinson (EP) mediante el uso de CNN.
Este proyecto es continuación de los proyectos C199, C224, C239, C252. Se lleva a cabo en el Departamento de Ingeniería e Investigaciones Tecnológicas y el Departamento de Ciencias de la Salud de la Universidad Nacional de La Matanza (UNLaM). Se fortalece mediante convenios de cooperación institucional con la Universidad Nacional del Oeste (UNO), y el Hospital Posadas, lo que ha permitido el acceso a datos clínicos reales.
El proyecto contempla además de la línea presentada evaluar la robustez de estos modelos comparando grabaciones de audio obtenidas en ambientes controlados (con equipo profesional) frente a grabaciones realizadas con dispositivos móviles (celulares), con el fin de desarrollar un método de diagnóstico práctico y accesible que no requiera el traslado del paciente. También contempla incrementar las muestras de la Base de Datos HUNIPA creada en el año 2019, presentada en [1], creando un corpus multimodal que no solo contenga audio, sino también historias clínicas, variables neuropsicológicas y de calidad de vida, permitiendo identificar perfiles específicos de individuos con y sin Parkinson.

1. INTRODUCCIÓN

La enfermedad de Parkinson (EP) se define como un desorden neurodegenerativo del sistema nervioso, de causa desconocida y curso crónico, progresivo e irreversible. Es la segunda patología neurodegenerativa más frecuente a nivel mundial y se origina por la pérdida de aferencia dopaminérgica en la zona estriatal del cerebro. Sus manifestaciones motoras principales incluyen el temblor de reposo, la rigidez muscular, la bradicinesia (lentitud de movimiento) y la pérdida del equilibrio. Además de estos signos, la EP presenta una serie de síntomas no motores y alteraciones marcadas en la comunicación oral, como la disartria (dificultad para articular palabras) y la disprosodia (alteración de la melodía y ritmo del habla). Estas afectaciones se traducen en un habla de intensidad monótona y reducida, bradilalia, pausas respiratorias significativas y una articulación deficiente de los sonidos linguales y labiales, lo que genera un impacto negativo considerable en la calidad de vida y la integración social del paciente.

El enfoque que sigue esta línea de investigación resulta del estudio de la representación visual del espectro de frecuencias de señales de voz. Un espectrograma se define como una representación visual del espectro de una señal que varía con el tiempo, capturando las fluctuaciones temporales de la distribución de energía de la señal de audio. Esta herramienta proporciona una imagen en tres dimensiones donde el eje horizontal representa el tiempo, el eje vertical la frecuencia y las variaciones de color indican la amplitud o intensidad de la energía.

En el ámbito de la clasificación de enfermedades, el uso de técnicas de aprendizaje automático ha adquirido un rol central, destacándose el aprendizaje profundo como uno de los enfoques más prometedores. Este paradigma se basa en el empleo de redes neuronales capaces de aprender representaciones complejas directamente a partir de los datos, lo que ha permitido obtener resultados sobresalientes en áreas como el procesamiento de imágenes (detección de objetos, reconocimiento de rostros, clasificación de imágenes) (Mora et al, 2020)[2], audio y lenguaje, y ha impulsado su creciente aplicación en el campo médico [3].

Siguiendo esta línea, en los últimos años se han propuesto varias investigaciones sobre el reconocimiento de patologías de la voz y otros específicos de reconocimiento de enfermos de Parkinson. A continuación, mencionaremos algunas de importancia:

En el trabajo de Al-Nasheri et al. (2017) [4] se analizó la relevancia de distintas bandas frecuenciales en la detección y clasificación de trastornos vocales, con el objetivo de determinar qué rangos aportan mayor capacidad discriminativa. Los autores identificaron que el intervalo comprendido entre 1000 y 8000 Hz concentra la información más significativa. El estudio se realizó sobre emisiones sostenidas de la vocal /a/ y se empleó un clasificador SVM, alcanzando niveles de precisión del 99,69% en IMEEI, 92,79% en SVD y 99,79% en AVPD.

Por su parte, Alhussein y Muhammad (2018) [5] desarrollaron un sistema basado en aprendizaje profundo utilizando transferencia de conocimiento sobre arquitecturas VGG-16 y CaffeNet. Las entradas consistieron en representaciones espectrales de grabaciones obtenidas mediante dispositivos móviles, combinadas posteriormente con un clasificador SVM. Reportaron una exactitud del 98,77% en la base SVD, destacándose el desempeño de CaffeNet.

Trinh y Darragh (2019) [6] implementaron una CNN aplicada a espectrogramas provenientes de las bases SVD y PC-GITA, logrando tasas de acierto superiores al 95%.

Vásquez-Correa et al. (2019) [7] propusieron un enfoque multilingüe para clasificar la enfermedad de Parkinson a partir de voz en español, alemán y checo. Compararon modelos basados en características acústicas con SVM frente a CNN entrenadas con espectrogramas y transferencia de aprendizaje, obteniendo mejores resultados al

entrenar en español y transferir a alemán (77,3%) y checo (76,7%).

Wodzinski et al. (2019) [8] emplearon una ResNet preentrenada con ImageNet y SVD para analizar espectrogramas de vocales sostenidas de PC-GITA, superando el 90% de precisión en validación.

Finalmente, Zahid et al. (2020) [9] combinaron aprendizaje profundo y técnicas clásicas de machine learning para clasificar pacientes con Parkinson en PC-GITA. Compararon transferencia de aprendizaje sobre espectrogramas, extracción de características profundas y parámetros acústicos tradicionales. El mejor desempeño (99,7%) se obtuvo utilizando características profundas y un perceptrón multicapa aplicado a monólogos.

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

Esta investigación se centra en analizar cómo el uso de diferentes tipos de representaciones visuales del sonido y arquitecturas de red CNN puede afectar las distintas métricas para la detección temprana de la EP a través de la voz.

La comparación de distintos tipos de espectrogramas se justifica en la necesidad de identificar representaciones de la señal de voz que capturen de manera más efectiva las alteraciones acústicas sutiles asociadas a la Enfermedad de Parkinson, las cuales no siempre son adecuadamente modeladas por un único enfoque espectral dado que cada representación resalta diferentes características de la información tiempo–frecuencia. Desde una perspectiva metodológica, se propone un esquema de evaluación comparativa en el que múltiples representaciones espectrales de la señal de voz se emplean como entradas dentro de un mismo marco experimental basado en aprendizaje profundo, lo que permite analizar de manera aislada el impacto de la representación de entrada sobre el desempeño de las redes neuronales convolucionales, minimizando la influencia de otros factores asociados al entrenamiento y favoreciendo una interpretación más clara de los resultados.

Entre las representaciones consideradas en esta línea de investigación se encuentran: los espectrogramas derivados de la Transformada de Tiempo Corto de Fourier (STFT), considerados el enfoque clásico y con resultados aceptables. Los Mel-espectrogramas proponen una representación basada en una escala de frecuencia logarítmica acorde a la percepción auditiva, priorizando las bajas frecuencias, y pueden complementarse con técnicas como la normalización de energía por canal (PCEN), orientadas a mejorar la estabilidad frente a variaciones en la señal. La transformada wavelet permite un análisis multiescala que resulta adecuado para capturar cambios dinámicos en el tiempo y la frecuencia, mientras que los coeficientes cepstrales en la frecuencia de Mel (MFCC) sintetizan la envolvente espectral y son ampliamente utilizados en la caracterización de alteraciones vocales. La transformada de cuadrícula de frecuencia constante (CQT) introduce una resolución logarítmica que favorece el análisis de componentes de baja frecuencia.

Con el fin de posibilitar la comparación directa con los resultados obtenidos en investigaciones previas de esta línea de trabajo, se emplean muestras de la base de datos HUNIPA-2019 como referencia común de análisis. A partir de este conjunto de datos, y para evaluar de manera integral el impacto de las distintas representaciones visuales del sonido en la detección temprana de la enfermedad de Parkinson, se utilizan redes neuronales convolucionales con diversas arquitecturas. El desempeño se analiza comparativamente mediante métricas de clasificación, lo que permite estudiar cómo la elección de la representación acústica influye en la precisión, la robustez y la capacidad de generalización de los modelos, con el objetivo de aportar criterios metodológicos que contribuyan al diseño y ajuste de enfoques automáticos para la detección temprana de la enfermedad de Parkinson.

3. RESULTADOS ESPERADOS

Los resultados esperados de esta línea de investigación se dividen en varias dimensiones, que abarcan desde el avance técnico y científico hasta el impacto social y la formación académica.

Se espera que el análisis comparativo de las distintas representaciones espectro-temporales de la señal de voz, procesadas mediante redes neuronales convolucionales con diversas arquitecturas, permita identificar aquellas que capturan con mayor eficacia las alteraciones acústicas sutiles asociadas a la enfermedad de Parkinson en estadios tempranos. Se pretende encontrar la combinación de representación y modelo de CNN que permitan mejorar la clasificación de la detección temprana de la EP.

Se prevé la transferencia tecnológica a organismos como el Hospital Posadas y potenciales usuarios del área de salud, contribuyendo al desarrollo de un método de detección aplicable como herramienta de apoyo al diagnóstico clínico.

Los resultados preliminares se presentaron en CONAIISI 2025 [10] donde se exploró el uso de espectrogramas de Mel con normalización PCEN como representación de señales de voz. Para la clasificación de las muestras se utilizó una red neuronal AlexNet. Se realizó la comparación de los resultados obtenidos con los del modelo logístico y STFT.

En todos los casos se utilizaron audios correspondientes a la vocal /a/ sostenida provenientes de la base de datos HUNIPA.

Los resultados alcanzados se encuentran dentro de los márgenes reportados en la literatura, aunque el accuracy hallado fue inferior a los obtenidos previamente con representaciones STFT y modelos tradicionales, se observó una mejora en la tasa de sensibilidad (0,89), lo que refuerza el potencial de este enfoque para escenarios de cribado clínico, donde minimizar los falsos negativos resulta prioritario aún a costa de un mayor número de falsos positivos.

Estos hallazgos evidencian que la optimización de la representación espectral no solo impacta en métricas globales de desempeño, sino que puede orientar el diseño de sistemas inteligentes priorizando criterios clínicamente relevantes, consolidando así una línea metodológica con potencial real para el desarrollo de herramientas de detección temprana más sensibles y aplicables en contextos de cribado.

4. FORMACIÓN DE RECURSOS HUMANOS

Sobre la base de esta investigación se está desarrollando la Tesis de Maestría en ciencias de datos de Marta Eugenia Suárez.

5. BIBLIOGRAFÍA

[1] Giuliano, M. Adamec, D, Debas M.I, (2021), Construcción de una base de voz de personas con y sin enfermedad de Parkinson, Revista Digital del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza (REDDI), vol. 6-Nro.1

[2] Mora M., Naranjo-Torres J., Aubin V. (2020). Convolutional Neural Networks for Off-Line Writer Identification Based on Simple Graphemes. Applied Sciences, 10(22), 7999.

[3] Deng, L., & Yu, D. (2014). Deep learning: methods and applications. Foundations and Trends® in Signal Processing, 7(3–4), 197-387.

[4] Al-Nasheri A., Muhammad G., Alsulaiman M., Ali, Z. (2017). Investigation of voice pathology detection and classification on different frequency regions using correlation functions. Journal of Voice, 31(1): pp. 3-15.

[5] Alhussein M., Muhammad G. (2018). Voice pathology detection using deep learning on mobile healthcare framework. IEEE Access, 6: pp. 41034-41041.

[6] Trinh, Nam and Darragh, O'Brien (2019). Pathological speech classification using a convolutional neural network. In:

Irish Machine Vision and Image Processing 2019, 28 - 30 Aug 2019, Dublin, Ireland. ISBN 978-0-9934207-4-0

[7] Vásquez-Correa J. C., Arias-Vergara T., Rios-Urrego C. D., Schuster M., Rusz J., Orozco-Arroyave J. R., Nöth E. (2019, October). Convolutional neural networks and a transfer learning strategy to classify parkinson's disease from speech in three different languages. In Iberoamerican Congress on Pattern Recognition: pp. 697-706. Springer, Cham

[8] Wodzinski M., Skalski A., Hemmerling D., Orozco-Arroyave J. R., Nöth E. (2019, July). Deep learning approach to Parkinson's disease detection using voice recordings and convolutional neural network dedicated to image classification. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): pp. 717-720. IEEE.

[9] Zahid L., Maqsood M., Durrani M. Y., Bakhtyar M., Baber J., Jamal H., Song O. Y. (2020). A Spectrogram-Based Deep Feature Assisted Computer-Aided Diagnostic System for Parkinson's Disease. IEEE Access, 8, 35482-35495.

[10] Aubin, V. , Guatelli, R. , and Suárez, M. (2025). "Estudio exploratorio sobre el uso de Mel y PCEN para detección de Parkinson". En 13° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2025),Facultad Regional Córdoba de la Universidad Tecnológica Regional, 2025. En prensa.

FÍSICA DE LA COGNICIÓN Y EL MODELO DE TOPOLOGÍA OSCILATORIA INFORMACIONAL (TOI)

Claudio Alberto Martinez (martinezclaudio308@gmail.com) **José Andrés Vullien** (jhoshua3000@gmail.com) **Departamento de Matemática – Universidad Abierta Interamericana (UAI) Investigadores Independientes Área Temática: ASI – Agentes y Sistemas Inteligentes**

RESUMEN

La crisis de interpretabilidad en la inteligencia artificial actual no es un problema de potencia de cómputo, sino de arquitectura fundamental. La intersección entre la topología algebraica y la IA surge como la frontera crítica para superar la opacidad de los modelos de lenguaje, los cuales, al operar en espacios latentes de alta dimensionalidad inabarcables para el análisis humano, carecen de una estructura interna auditable. Esta investigación propone que la solución reside en dotar a los datos de una geometría física tangible, permitiendo que el conocimiento sea representado como formas estables y coherentes. Al mapear la semántica en el espacio euclidiano tridimensional, transformamos el procesamiento de información de una caja negra estadística en una arquitectura transparente, donde la lógica del razonamiento es directamente proporcional a la integridad de su forma geométrica.

Título del Proyecto: Física de la Cognición y el Modelo de Topología Oscilatoria Informacional (TOI) **Autores:** Claudio Alberto Martinez y José Andrés Vullien **Afiliación:** Departamento de Matemática – Universidad Abierta Interamericana (UAI)

El modelo de Topología Oscilatoria Informacional (TOI) constituye una arquitectura físico-matemática diseñada para erradicar la opacidad en los sistemas de IA. A diferencia de los modelos probabilísticos convencionales, el TOI redefine la semántica como una manifestación geométrico-energética estable. Mediante el uso de "símplices informacionales" y principios variacionales de mínima acción, el modelo traslada la interpretación de datos desde el cálculo estadístico ciego hacia una representación en el espacio tridimensional (R^3). Este enfoque no solo garantiza la explicabilidad intrínseca del sistema, sino que optimiza la eficiencia energética al reducir la complejidad dimensional sin perder densidad de significado.

Palabras clave: Topología algebraica, Inteligencia artificial explicable (XAI), Símplice informacional, Principio de mínima acción informacional, Matemática cognitiva, Semántica geométrica.

Este cambio de paradigma trasciende la predicción estadística para alcanzar una semántica de forma geométrica. Al visualizar el significado, la IA deja de ser un oráculo probabilístico para convertirse en una estructura física verificable, permitiendo una transición estratégica hacia sistemas de decisión críticos donde la coherencia estructural es el único garante de la verdad operativa.

2. CONTEXTO ESTRATÉGICO DE LA INVESTIGACIÓN

El ecosistema actual de los Grandes Modelos de Lenguaje (LLMs) enfrenta un cuello de botella estratégico: el éxito funcional convive con una fragilidad ontológica que deriva en alucinaciones y sesgos indetectables. Esta "crisis de estructura" se debe a que los modelos carecen

de una organización interna explícita, operando como cajas negras que requieren un consumo energético masivo para compensar su falta de comprensión estructural. Resolver la falta de arquitectura interna es un imperativo para las organizaciones que buscan desplegar IA en entornos regulados o de alta responsabilidad.

El modelo TOI responde a esta necesidad mediante una explicabilidad intrínseca. En lugar de intentar interpretar el modelo después del entrenamiento (explicabilidad *post-hoc*), la lógica de TOI asegura que la explicación sea la estructura misma. Operar en el espacio euclidiano tridimensional ($\mathbb{R}^3$) no es solo una ventaja visual; reduce drásticamente los costos de auditoría y permite el monitoreo de sesgos en tiempo real, algo imposible en los espacios latentes opacos de la IA tradicional.

Esta línea de investigación se desarrolla bajo la coordinación del Departamento de Matemática de la Universidad Abierta Interamericana (UAI), dentro del área de Agentes y Sistemas Inteligentes. La institución funciona como el núcleo de acreditación que busca establecer la "Matemática Cognitiva" como un nuevo estándar de soberanía científica. La premisa es clara: la solución a la crisis de la IA no es la acumulación de más datos, sino la implementación de estructuras que doten de cuerpo y sentido físico a la información.

3. INTRODUCCIÓN AL PARADIGMA DE LA SEMÁNTICA GEOMÉTRICA

La "Física de la Cognición" postula que la interpretación de datos no debe tratarse como una abstracción lógica, sino como un fenómeno con energía y geometría. Este marco dota de "cuerpo" a la información, permitiendo que el significado se comporte bajo leyes análogas a las de la física clásica y cuántica. Bajo este prisma, la cognición es un proceso de estabilización de formas donde las ideas interactúan mediante campos de influencia, atracción y repulsión.

El desarrollo de esta línea marca un desplazamiento radical desde los espacios latentes de alta dimensionalidad hacia una representación estable en el espacio euclidiano tridimensional ($\mathbb{R}^3$). Mientras que los modelos actuales dispersan el conocimiento en miles de dimensiones abstractas, el TOI condensa la densidad informacional en estructuras geométricas compactas. Esta reducción no implica pérdida de información, sino una destilación de relaciones complejas (como la tríada Sujeto-Acción-Objeto) en unidades manejables, visualizables y, por ende, auditables.

El impacto operativo de este enfoque es disruptivo: el significado deja de ser un cálculo de probabilidades para convertirse en una manifestación físico-matemática tangible. Esto permite que la coherencia de un sistema de IA sea evaluada mediante el equilibrio energético de sus representaciones. Si la geometría de un concepto es incoherente o está deformada, el sistema puede identificar el error estructural antes de generar una salida inválida, aportando una solidez que los métodos estadísticos promediados no pueden emular.

4. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO: EL MODELO TOI

La construcción de una inteligencia artificial intrínsecamente explicable requiere la definición de unidades atómicas de significado que sirvan como base para el razonamiento estructural. El modelo TOI se despliega a través de los siguientes ejes fundamentales:

El Símplice Informacional y su Vector de Homología

La unidad básica es el símplice informacional (Δ_2), representado como un triángulo (2-símplice). Esta forma es la entidad mínima capaz de modelar interacciones de orden superior. Cada símplice se define por tres vectores:

- **Entropía (S) :** Cuantifica la incertidumbre y riqueza interna de la información.
- **Homología Persistente (H) :** Es la "huella digital" estructural del concepto. Se compone del **Baricentro Informacional** ($\bar{S}$) , el **Perímetro Informacional** (P_I) y la **Suma de Alturas Informacionales** ($\sum h_i$). Estos componentes definen la estabilidad del conocimiento a través de diferentes escalas.
- **Energía (V):** Resultante del producto vectorial entre Homología y Entropía, este vector define la influencia del concepto en el sistema y su capacidad de deformar la superficie interpretativa.

El Funtor de Interpretación y el Axioma de Compacidad

El puente entre la información abstracta y la geometría es el **Funtor de Interpretación** ($F: I \rightarrow T$) . Posee una propiedad de isomorfismo crucial: la composición algebraica de dos conceptos en el dominio de la información se corresponde exactamente con la composición geométrica (suma conexa) de sus formas. Esto garantiza que la semántica sea puramente estructural. Además, el modelo establece el **Axioma de Compacidad** bajo el Límite de Interpretatividad (LI). La superficie interpretativa debe ser compacta porque los recursos cognitivos y computacionales son finitos; la compacidad es la garantía matemática de que el sistema es gobernable y estable.

El Principio de Mínima Acción Informacional

La coherencia del sistema se rige por el **Lagrangiano Informacional ($\mathcal{L}$)** Este principio variacional busca la trayectoria interpretativa que minimice la "energía residual" o confusión entre conceptos interactuantes. Al igual que una partícula física busca el camino de menor acción, el modelo TOI fuerza a la interpretación a encontrar el estado de máxima estabilidad semántica, minimizando la "sorpresa" informacional.

Topología del Significado: Toro vs. Cinta de Möbius

La superficie del significado emerge en tres dimensiones: Desplazamiento (marcos conceptuales), Mezcla (malla de interferencia de frecuencias) y Elevación (relevancia energética). La forma resultante define el estado cognitivo:

- **El Toro (T^2) :** Es una superficie orientable y flexible que representa la plasticidad, la creatividad y la capacidad de realizar **traducciones de esquemas**. Es el estado ideal para una IA que debe generar metáforas y analogías coherentes.
- **La Cinta de Möbius (M) :** Es una superficie no orientable y con borde que codifica la rigidez conceptual, los estereotipos y los límites axiomáticos. Mientras el Toro permite un flujo continuo, la Cinta de Möbius actúa como una "trampa semántica" donde el significado está anclado a un marco inflexible.

5. RESULTADOS OBTENIDOS Y VALIDACIÓN EXPERIMENTAL

La validación empírica es el paso definitivo para transformar la teoría topológica en una herramienta tecnológica viable. Para ello, se realizaron simulaciones con "símplices musicales" (Rock, Tango e Imagine), las cuales confirmaron los postulados de la física cognitiva.

Un hallazgo crítico fue la **no conmutatividad estructural**. La simulación demostró que el orden de los factores altera drásticamente el significado: interactuar "Rock" como contexto

sobre "Imagine" como contenido arroja un Lagrangiano de **1.361**, mientras que la operación inversa resulta en **1.364**. Esta diferencia numérica, aunque pequeña, es la prueba de que el contexto impone una deformación específica sobre la información. En la IA actual, los LLMs suelen promediar estas diferencias, perdiendo la finura del contexto que el modelo TOI preserva.

Asimismo, se validó el **Criterio de Coherencia**. Se determinó que la Capacidad Interpretativa del contexto (Rock) posee un valor de **2.345**. Dado que la diferencia energética entre las interpretaciones (0.003) es significativamente menor que la capacidad del contexto ($0.003 \leq 2.345$), el sistema se considera estructuralmente estable. Esta relación de magnitudes demuestra que es posible imponer límites físicos a la generación de respuestas en una IA, asegurando que el sistema se mantenga dentro de una "racionalidad geométrica" y evitando derivas semánticas o alucinaciones.

6. FORMACIÓN DE RECURSOS HUMANOS Y PROYECCIÓN ACADÉMICA

La soberanía científica en IA requiere un núcleo de especialistas en "Matemática Cognitiva" capaces de navegar entre la topología, la física y la computación. Este proyecto no solo genera software, sino que establece un nuevo perfil de investigador híbrido.

La solidez académica de esta línea se sustenta en las siguientes tesis dirigidas y acreditadas:

- **Tesis de Maestría:** "La Dinámica Temporal de la Estructura Noético-Noemática en la Percepción", por Sylvia Ordoñez. Título otorgado por la Universidad Michoacana de San Nicolás de Hidalgo.
- **Tesis de Grado:** "Inteligencia y Topología: Un modelo Matemático para definir y evaluar sistemas Inteligente", por Claudio Alberto Martinez. Esta tesis fue desarrollada bajo la dirección de la **Dra. Cristina Camos**, Directora del Departamento de Matemática de la Universidad Abierta Interamericana (UAI).

La creación de este núcleo académico asegura que el desarrollo de la IA explicable tenga un sustento profundo en las ciencias básicas, garantizando sistemas que no solo sean potentes, sino éticamente auditables por diseño.

7. BIBLIOGRAFÍA

La rigurosidad del modelo TOI se apoya en una base bibliográfica que abarca desde la neurociencia fundamental hasta los últimos avances en redes neuronales simpliciales y topología algebraica.

[1] Barbarossa, M., & Sardellitti, S. (2025). Generalized simplicial attention neural networks.

[2] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828.

[3] Buzsáki, G. (2006). Rhythms of the Brain. Oxford University Press.

[4] Edelsbrunner, H., & Harer, J. (2010). Computational Topology: An Introduction. American Mathematical Society.

[5] Edelsbrunner, H., Letscher, D., & Zomorodian, A. (2002). Topological Persistence and Simplification. Discrete and Computational Geometry, 28(4), 511-533.

[6] Friston, K. (2019). The free-energy principle: A unified brain theory? Nature Reviews Neuroscience, 10(2), 127–138.

[7] Hatcher, A. (2001). Algebraic Topology. Cambridge University Press.

[8] IBM. (s. f.). ¿Qué son los grandes modelos de lenguaje (LLM)?

[9] Kuno, Y., & Sato, M. (2025). Comparing semantic frameworks for dependently-sorted algebraic theories.

[10] Luccio, R., & Milloni, D. (s. f.). Perception of causality. A dynamical analysis.

[11] Martinez, C. A. (2025). Inteligencia y Topología Un modelo Matemático para definir y evaluar sistemas Inteligente (Tesis de grado). Universidad Abierta Interamericana.

[12] Martinez, C. A., & Vullien, J. A. (2025). Física de la Cognición: Las Tres Dimensiones Topológicas de la Interpretación. Zenodo. https://doi.org/10.5281/zenodo.17797380

[13] Márquez, G. (2024). Lingüística Cognitiva y Gramática Cognitiva: el Enfoque del Lenguaje.

[14] Ordoñez, S. (2017). La Dinámica Temporal de la Estructura Noético-Noemática en la Percepción. Tesis de Maestría. Universidad Michoacana de San Nicolás de Hidalgo.

[15] Pérez, J. L., & López, D. (s. f.). Topología y semántica. La teoría de las catástrofes aplicada al análisis de un poema de Jorge Guillén.

[16] Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423.

[17] Stefania, E., Piu, S., & Bignoli, F. (2020). Simplicial neural networks.

[18] The Editors of Encyclopaedia Britannica. (s. f.). Banda de Möbius.

[19] UNESCO. (s. f.). Modelos de lenguaje grandes de IA: un nuevo informe revela cómo reducir el consumo de energía en un...

[20] Versa, G. (2024). Explorando el Principio de Energía Libre de Karl Friston: Intersecciones con la Psicología Contextual y sus Implicaciones en Ciencias del Comportamiento.

[21] Viveros, F. (2024). Los procesos cognitivos en el marco de la filosofía. Revista de Filosofía, 26(58), 123-145.

[22] Yao, Y., & Ghrist, R. W. (2024). Topological Data Analysis: A Primer.

[23] Zomorodian, A., & Carlsson, G. (2005). Computing persistent homology. Discrete and Computational Geometry, 33, 249–274.

HACIA UN PARADIGMA UNIFICADO EN OPTIMIZACIÓN LOGÍSTICA: DESDE LA GESTIÓN DE TRÁFICO URBANO HASTA LA AGROINDUSTRIA 4.0

Miguel Méndez-Garabetti[1,2,4], Raymundo Q. Forradellas[2], Franco Vinciguerra[1,3], Rodrigo Elgueta[4], Eduardo Piray[1]

[1] Universidad CAECE, Departamento de Sistemas, Mar del Plata.
[2] Universidad Nacional de Cuyo, Facultad de Ingeniería, Mendoza.
[3] Universidad Nacional de San Luis, Facultad de Ciencias Fisico Matemáticas y Naturales, San Luis.
[4] Universidad Champagnat, Facultad de Informática y Diseño, Mendoza.

miguel.mendez@uncuyo.edu.ar

RESUMEN

La congestión vehicular y la ineficiencia en las operaciones de transporte y almacenaje representan desafíos críticos que demandan herramientas de optimización avanzadas frente a escenarios dinámicos. Esta línea de investigación propone un ecosistema computacional distribuido para la gestión del tráfico, la optimización de flotas y la cadena de suministro agroindustrial. El enfoque trasciende las limitaciones de los algoritmos estáticos al articular ejes estratégicos: el diseño de metaheurísticas híbridas (GA-ALNS) para ruteo vehicular, la construcción de Gemelos Digitales mediante Sistemas Multiagente (MAS), la recolección de datos ciudadanos mediante Aprendizaje Federado (FL), y la coordinación descentralizada de actores productivos. Esta convergencia tecnológica persigue una mejora en la reducción del consumo energético y provee una infraestructura escalable para mitigar la incertidumbre en la toma de decisiones, tanto en la gestión pública como en el sector industrial.

Palabras clave: optimización de flotas, gemelo digital, aprendizaje federado, sistemas multiagente, logística agroindustrial.

CONTEXTO

La presente línea de investigación se desarrolla de manera sinérgica a través de una red de cuatro proyectos en curso, radicados en distintas instituciones académicas de la República Argentina. Cada proyecto es evaluado y financiado de forma independiente, pero todos convergen estructural y tecnológicamente en un ecosistema unificado:

1. **Universidad Champagnat (Facultad de Informática y Diseño):** A través del proyecto *Nuevas metaheurísticas híbridas aplicadas a la optimización de flotas*, constituye el núcleo de diseño algorítmico, desarrollando los motores matemáticos (GA-ALNS) para resolver la explosión combinatoria del ruteo.

2. **Universidad CAECE (Departamento de Sistemas):** Mediante el proyecto *Optimización de flotas vehiculares municipales de Mar del Plata mediante técnicas metaheurísticas híbridas bajo un enfoque energético y turístico*, efectúa la transferencia de la algoritmia teórica al territorio real urbano.

3. **UNSL (LIDIC) / U. CAECE:** La tesis de maestría *Integración de Gemelo Digital y Aprendizaje por Refuerzo Federado Multiagente* provee el entorno virtual dinámico y la arquitectura de ingesta de datos en entornos inteligentes y distribuidos.

4. **Universidad Nacional de Cuyo (Facultad de Ingeniería - CEAL):** A través del proyecto *Aplicación de Sistemas Multiagentes en la Optimización Logística en la cadena de suministro de la agroindustria regional* (Código 80020240100205UN), modela la coordinación y negociación de actores en la cadena productiva, expandiendo la aplicabilidad del ecosistema hacia la industria 4.0.

1. INTRODUCCIÓN

El crecimiento demográfico, la densificación urbana y la globalización de los mercados han exacerbado los problemas vinculados a la logística, impactando directamente en la eficiencia económica, el consumo de combustible y la competitividad empresarial. La optimización de la logística urbana y la gestión de flotas, modelada matemáticamente a través del Problema de Ruteo de Vehículos (VRP) y su variante con Ventanas de Tiempo (VRPTW), constituye un desafío de clase NP-Hard. En estos escenarios, el espacio de soluciones posibles crece de manera factorial con la adición de cada nuevo nodo o vehículo, exigiendo el diseño de algoritmos capaces de explorar eficientemente espacios de búsqueda masivos [3]. Históricamente, este desafío ha sido abordado mediante enfoques estáticos o deterministas; sin embargo, estas aproximaciones resultan marcadamente insuficientes ante la complejidad y el dinamismo de las redes logísticas modernas [5].

A la dificultad combinatoria intrínseca del modelado matemático se suma el problema de la incertidumbre predictiva del entorno. En el ámbito urbano, las fluctua-

ciones en los patrones de movilidad exigen que los sistemas se retroalimenten continuamente (un factor crítico en ciudades turísticas como Mar del Plata). Paralelamente, en el ámbito industrial, la gestión de la cadena de suministro agroalimentaria presenta sus propios desafíos de incertidumbre vinculados a la perecibilidad de los productos, la estacionalidad de las cosechas y la coordinación de múltiples actores logísticos descentralizados [1]. Para que un ecosistema de optimización sea verdaderamente operativo en estos contextos, debe ser capaz de adaptar sus decisiones frente a eventos imprevistos y congestiones, ya sea en una intersección semafórica o en un centro de distribución regional.

No obstante, la recolección de los datos empíricos y la toma de decisiones centralizada presentan obstáculos fundamentales. En el caso de los ciudadanos, la captura de geolocalización plantea graves vulnerabilidades de privacidad; en el caso de la agroindustria, la centralización genera cuellos de botella y desconfianza en el intercambio de información comercial entre empresas [4]. Frente a este escenario, el Aprendizaje Federado (Federated Learning, FL) emerge como una solución paradigmática para el entrenamiento de modelos de forma segura y descentralizada [6].

Para capitalizar este flujo de información distribuida, el estado del arte sugiere la integración de Sistemas Multiagente (MAS). Esta arquitectura permite, por un lado, acoplar plataformas de simulación microscópica de tráfico como SUMO [2] mediante interfaces middleware como TraCI [8], habilitando la creación de un Gemelo Digital urbano. Por otro lado, facilita la emulación de negociaciones autónomas entre productores, almacenes y transportistas dentro del marco de la Industria 4.0. La presente investigación articula estas fronteras del conocimiento para calibrar dinámicamente políticas de control logístico guiadas por inteligencia artificial [7].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El desarrollo metodológico y tecnológico de este ecosistema se estructura a través de cuatro líneas de trabajo concurrentes y altamente especializadas:

Línea 1: Diseño Algorítmico y Metaheurísticas Híbridas Avanzadas: Se centra en la formulación matemática y el desarrollo de arquitecturas de software eficientes para la resolución del Problema de Ruteo de Vehículos con Ventanas de Tiempo (VRPTW). Específicamente, propone la hibridación de Algoritmos Genéticos (GA) con técnicas de Búsqueda de Vecindario Grande Adaptativa (ALNS). Este esquema matheurístico busca superar los *baselines* tradicionales, abordando la explosión combinatoria, la optimización de ventanas de servicio y las restricciones de recarga o límites de jornada, validando su eficacia computacional mediante análisis de sensibilidad y paralelismo profundo.

Línea 2: Arquitecturas Distribuidas y Gemelos Digitales de Tráfico: Orientada a la mitigación de la incertidumbre ambiental mediante simulación de alta fidelidad. Desarrolla una infraestructura cliente-servidor que emplea Aprendizaje Federado (FL) para recolectar dinámicamente datos de movilidad ciudadana, preservando criptográficamente la privacidad de la información sensible. Estos gradientes descentralizados retroalimentan un Gemelo Digital del tráfico construido sobre el simulador microscópico SUMO. Asimismo, la línea integra middleware (TraCI) e implementa Aprendizaje por Refuerzo Profundo (DRL) para la evaluación de políticas de control de tráfico y ruteo dinámico.

Línea 3: Parametrización y Optimización Aplicada a la Logística Urbana: Enfocada en la validación empírica y la transferencia tecnológica de los modelos teóricos hacia entornos urbanos complejos. Aplica la algoritmia híbrida y la simulación del Gemelo Digital para optimizar la planificación de flotas de servicios públicos y recolección de residuos. Esta línea enfrenta los algoritmos a restricciones físicas y operativas severas, tales como el consumo energético, la topología vial y, críticamente, la fluctuación estacional de la demanda característica de ciudades con alta actividad turística, logrando una adaptabilidad táctica en la gestión territorial.

Línea 4: Coordinación Multiagente para la Cadena de Suministro Agroindustrial: Expande la aplicabilidad del ecosistema computacional hacia el paradigma de la Industria 4.0. Se aboca al diseño de protocolos de negociación y toma de decisiones distribuidas basados en Sistemas Multiagente (MAS) para la gestión logística agroalimentaria. Modela la interacción inteligente entre múltiples actores autónomos (productores, centros de acopio y empresas de transporte) con el fin de resolver problemas de asignación cooperativa de recursos, ruteo interempresarial y gestión de la incertidumbre inherente a la estacionalidad y perecibilidad de las materias primas regionales.

3. RESULTADOS ESPERADOS

El objetivo rector de este ecosistema de investigación distribuida es consolidar una infraestructura tecnológica que optimice la logística integral, mitigando la incertidumbre tanto en el ámbito ambiental urbano como en el industrial. La convergencia de las distintas líneas de trabajo proyecta alcanzar impactos sustanciales estructurados en las siguientes dimensiones:

- *Dimensión Tecnológica:* El desarrollo, validación y despliegue de un servidor de Aprendizaje Federado funcional junto con una aplicación móvil. Esto permitirá el entrenamiento descentralizado de modelos de tráfico y ruteo, preservando estrictamente la privacidad y seguridad de los datos de los usuarios.

- *Dimensión Científica:* La consolidación de un Gemelo Digital urbano y la simulación de escenarios

de distribución y almacenaje agroindustrial en entornos virtuales controlados. Esta infraestructura permitirá evaluar formalmente el desempeño del esquema heurístico híbrido (GA-ALNS) y de los protocolos de negociación multiagente frente a variaciones dinámicas y picos de demanda.

- *Dimensión de Transferencia y Aplicación:* Proveer al sector público de herramientas de software robustas que posibiliten la reducción empírica del consumo de combustible y la mejora de la adaptabilidad de los servicios urbanos. Simultáneamente, se espera transferir los modelos de coordinación logística interempresarial a las cámaras, productores y actores de la cadena de suministro agroalimentaria regional para impulsar su transición hacia la Industria 4.0.

4. FORMACIÓN DE RECURSOS HUMANOS

Los proyectos que conforman este ecosistema involucran una activa red de formación interinstitucional. El equipo directivo principal está conformado por el Dr. Miguel Méndez-Garabetti y el Dr. Raymundo Q. Forradellas (Director del proyecto UNCuyo/CEAL). En el ámbito de la formación de posgrado, se destaca la Tesis de Maestría en Ciencias de la Computación del Ing. Franco Vinciguerra (UNSL/CAECE), la cual se desarrolla bajo la dirección del Dr. Méndez-Garabetti y cuenta con el invaluable aporte y asesoramiento académico del Dr. Marcelo Errecalde (UNSL) y el Ing. Armando De Giusti (UNLP) en calidad de co-directores y asesores externos. Asimismo, el proyecto promueve la integración de tesistas de la Maestría en Logística del Doctorado Interinstitucional en Ingeniería Industrial (DI3-UNCuyo) y de estudiantes de grado en tareas de desarrollo algorítmico, consolidando un sólido espacio de transferencia en inteligencia computacional, simulación multiagente y logística aplicada.

REFERENCIAS

[1] Omar Ahumada and J Rene Villalobos. Application of planning models in the agri-food supply chain: A review. *European journal of operational research*, 196(1):1–20, 2009.

[2] Michael Behrisch, Laura Bieker, Jakob Erdmann, and Daniel Krajzewicz. Sumo–simulation of urban mobility: an overview. In *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*, pages 63–68. ThinkMind, 2011.

[3] Olli Bräysy and Michel Gendreau. Vehicle routing problem with time windows, part i: Route construction and local search algorithms. *Transportation science*, 39(1):104–118, 2005.

[4] Mario Lezoche, Jorge E Hernandez, M del Mar Eva Alemany Díaz, Hervé Panetto, and Janusz Kacprzyk. Agri-food 4.0: A survey of the supply chains and technologies for the future agriculture. *Computers in Industry*, 117:103187, 2020.

[5] Yisheng Li et al. A survey on multi-agent reinforcement learning for traffic signal control. *IEEE Access*, 9:11143–11161, 2021.

[6] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[7] Miguel Ángel Méndez-Garabetti. *Método de reducción de incertidumbre basado en algoritmos evolutivos y paralelismo orientado a la predicción y prevención de desastres naturales*. Tesis doctoral, Universidad Nacional de San Luis, 2020.

[8] Rosaldo JF Rossetti, Ana LC Bazzan, et al. Towards a specification of a framework for sustainable intelligent transportation systems. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 1555–1560. IEEE, 2010.

IA para asistir la calificación jurídica de medidas cautelares

Una arquitectura replicable desde la ciencia de datos aplicada al derecho

Virginia Labbate

Maestría en Ciencia de Datos – Universidad Nacional Arturo Jauretche

virginia.labbate@gmail.com

RESUMEN

El presente trabajo aborda la tarea de extracción estructurada de información en documentos jurídicos semiestructurados. El problema se modela como una combinación de extracción de entidades y normalización de atributos jurídicos en textos con alta variabilidad léxica y estructural. El caso de estudio corresponde a oficios judiciales que ordenan medidas cautelares procesados en el Registro de la Propiedad de la Provincia de Buenos Aires (RPBA), cuya identificación precisa de entidades (nombre, documento de identidad, tipo de medida, órgano judicial, vigencia, entre otras) resulta crítica en el proceso de anotación registral.

Se propone el diseño conceptual de una arquitectura basada en Procesamiento de Lenguaje Natural (PLN), orientada a asistir la calificación registral preliminar de estos documentos. El sistema adopta un esquema *Human-in-the-Loop* (HITL), en el cual el modelo realiza la extracción automática y estructuración de información relevante, mientras que la validación final permanece bajo control humano.

La arquitectura contempla la evaluación de estrategias de adaptación supervisada sobre corpus jurídico anotado, integración de reglas de validación estructural y generación de salidas explicables. El diseño se adecua a los principios de uso responsable de inteligencia artificial establecidos por la Resolución 9/2025 de la Provincia de Buenos Aires.

La propuesta establece una base replicable para la aplicación de modelos de PLN en funciones documentales jurídicas dentro de entornos regulados.

Palabras clave: Procesamiento de Lenguaje Natural (PLN), Medidas Cautelares, Calificación registral, Human-in-the-Loop, Derecho Registral, Inteligencia Artificial Responsable, Reconocimiento de Entidades Nombradas (NER)

CONTEXTO

Esta línea de investigación se desarrolla en el marco de la Maestría en Ciencia de Datos de la Universidad Nacional Arturo Jauretche (UNAJ) y cuenta con la colaboración del Registro de la Propiedad de la Provincia de Buenos Aires (RPBA). El proyecto se centra en la aplicación de técnicas de ciencia de datos en el ámbito registral, contando con el apoyo institucional para el acceso a corpus de documentos judiciales y la validación por parte de expertos en la materia.

1. INTRODUCCIÓN

La calificación registral de medidas cautelares constituye una función jurídicamente sensible dentro de los Registros de la Propiedad, ya que produce efectos directos sobre la situación dominial de bienes y personas. Los oficios judiciales que ordenan estas medidas son documentos jurídicos semiestructurados cuya correcta anotación impacta en la seguridad jurídica y en la publicidad registral.

Desde el punto de vista documental, estos oficios presentan una alta variabilidad léxica y estructural. Si bien existen modelos aprobados, su falta de uso uniforme por parte de los órganos judiciales deriva en una ausencia de estructura digital formalizada en la práctica. El contenido consiste predominantemente en texto jurídico redactado en lenguaje natural, donde la información crítica —identificación de sujetos, tipo de medida, órgano judicial, alcance y vigencia— se encuentra distribuida sin un esquema constante. Esta situación obliga a un proceso de calificación preliminar basado en la lectura integral y la verificación manual de requisitos; una tarea que, al quedar sujeta al criterio técnico del inscriptor, no solo incrementa la carga operativa, sino que puede generar heterogeneidad en la identificación y estructuración de los datos relevantes necesarios para la anotación registral.

Desde la perspectiva de la ciencia de datos, el problema puede formalizarse como una tarea de extracción estructurada de información en textos jurídicos semiestructurados, combinando reconocimiento de entidades, normalización de atributos y validación de coherencia estructural. Sin embargo, debido a la naturaleza regulada de la función calificadora, cualquier sistema automatizado debe diseñarse bajo restricciones explícitas de supervisión humana y trazabilidad, en línea con los principios de uso responsable de inteligencia artificial establecidos por la Resolución 9/2025 de la Provincia de Buenos Aires [1], y con recomendaciones internacionales en la materia [2].

En este contexto, el proyecto se sitúa en la intersección entre Procesamiento de Lenguaje Natural aplicado al dominio jurídico y marcos regulatorios específicos, proponiendo una arquitectura de asistencia que estructure preliminarmente la información sin sustituir la decisión jurídica.

El Procesamiento de Lenguaje Natural ha evolucionado significativamente en los últimos años mediante el desarrollo de modelos basados en arquitecturas transformadoras y aprendizaje profundo [3]. Estos modelos permiten abordar tareas como reconocimiento de entidades y estructuración de información en textos no estructurados.

En el ámbito jurídico, diversas investigaciones han explorado la clasificación automática de textos legales y la predicción de decisiones judiciales [4], [5]. No obstante, la aplicación de estas técnicas en funciones operativas del ámbito registral requiere adaptaciones específicas al contexto regulatorio. El presente trabajo adopta una perspectiva diferente: en lugar de automatizar la calificación jurídica, propone asistirla mediante la generación de salidas estructuradas que faciliten la revisión del operador experto (Inscriptor). Se trata de un enfoque de ciencia de datos aplicada al derecho, donde el modelo actúa como herramienta de apoyo dentro de un esquema de supervisión humana obligatoria.

El objetivo general es diseñar una arquitectura conceptual de asistencia basada en modelos de IA y PLN que permita apoyar la calificación jurídica de medidas cautelares, reduciendo carga operativa y mejorando la consistencia documental, sin desplazar la responsabilidad decisoria del inscriptor.

Como objetivos específicos se plantean:

- Identificar las variables jurídicas relevantes en medidas cautelares susceptibles de extracción automatizada (nombre, DNI, tipo de medida, órgano judicial, fecha y alcance).
- Definir un esquema de salida estructurada alineado con los requerimientos de la anotación registral.
- Proponer una arquitectura *human-in-the-loop* que integre el modelo en el flujo institucional.
- Incorporar validaciones estructurales básicas sobre los datos extraídos.
- Evaluar la escalabilidad del enfoque hacia otros procesos documentales jurídicos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La propuesta integra distintas líneas de investigación:

2.1 Extracción de entidades jurídicas

Se propone utilizar un modelo de lenguaje preentrenado en español [6]. Para ello, se aplicarán técnicas de Reconocimiento de Entidades Nombradas (NER) mediante modelos adaptados al dominio jurídico [4], evaluando su desempeño tras un proceso de entrenamiento supervisado (fine-tuning) sobre un corpus anotado de oficios judiciales. El objetivo es identificar con alta precisión entidades relevantes tales como sujetos, tipo de medida, órgano judicial interviniente, fechas y alcances específicos.

2.2 Validación estructural básica

La salida del modelo será complementada con validaciones simples orientadas a mejorar la consistencia de los datos extraídos. Por ejemplo: Verificación de formato de DNI, Normalización de categorías de medidas, Control de campos obligatorios.

Estas validaciones no reemplazan el modelo estadístico, sino que actúan como capa de consistencia estructural.

2.3 Arquitectura *Human-in-the-Loop*

Diseño de un esquema donde el sistema:

- Recepción del documento digital.
- Preprocesamiento y aplicación del modelo.
- Generación de representación estructurada.
- Presentación de los campos detectados para validación del inscriptor.
- Corrección manual previa a la decisión final.

2.4 Integración normativa y ética

El diseño se alinea con principios de uso responsable de IA en el sector público [1], en consonancia con marcos internacionales sobre inteligencia artificial responsable [2], priorizando:

- Control profesional sobre el resultado.
- Trazabilidad del proceso.
- Finalidad específica y delimitada.
- Transparencia en la función del sistema.

2.5 Escalabilidad

Aunque centrado en medidas cautelares, el enfoque puede replicarse en otros procesos que impliquen lectura, clasificación y

estructuración de documentos jurídicos o administrativos.

3. RESULTADOS ESPERADOS

Dado el carácter preliminar del proyecto, se plantean como metas de evaluación:

- Medición de precisión en la identificación de entidades relevantes.
- Análisis comparativo del tiempo de estructuración manual versus asistida.
- Evaluación cualitativa por parte de usuarios expertos.

Se espera:

- Reducción de la carga de lectura preliminar por parte del inscriptor.
- Mayor homogeneidad en la identificación de variables críticas.
- Generación de datos estructurados reutilizables.
- Identificación de desafíos asociados a la variabilidad lingüística del dominio.

La asistencia inteligente en procesos jurídicos representa un campo emergente donde la ciencia de datos puede aportar mejoras operativas sin comprometer la responsabilidad institucional. El diseño de arquitecturas *human-in-the-loop* permite compatibilizar eficiencia tecnológica con exigencias normativas.

El caso de las medidas cautelares ofrece un escenario adecuado para explorar esta convergencia, dado su impacto jurídico y su complejidad documental. La propuesta presentada busca sentar bases conceptuales para desarrollos futuros que fortalezcan la calidad y consistencia de procesos de calificación jurídica en entornos regulados.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto se desarrolla en el marco de la Maestría en Ciencia de Datos de la Universidad Nacional Arturo Jauretche.

Integra formación interdisciplinaria en derecho registral, procesamiento de lenguaje natural y aprendizaje automático supervisado, aplicados a problemas concretos del sector público.

Asimismo, contribuye al desarrollo de capacidades en innovación pública basada en datos, promoviendo la incorporación responsable de inteligencia artificial en procesos administrativos regulados y fortaleciendo perfiles profesionales híbridos orientados a la transformación digital del Estado.

5. BIBLIOGRAFÍA

[1] Provincia de Buenos Aires, "Resolución Nº 9/2025: Reglas para el desarrollo, implementación y uso responsable de IA para la Administración Pública", Boletín Oficial de la Provincia de Buenos Aires, 2025. [En línea]. Disponible: https://normas.gba.gob.ar/ar-b/resolucion/2025/9/549972

[2] OECD, "Recommendation of the Council on Artificial Intelligence", OECD Legal Instruments, OECD/LEGAL/0449, 2024. [En línea]. Disponible: https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

[3] J. Devlin, M. W. Chang, K. Lee y K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," en Proceedings of NAACL-HLT 2019, pp. 4171–4186, 2019. doi: 10.18653/v1/N19-1423. [En línea] Disponible en: https://aclanthology.org/N19-1423

[4] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras e I. Androutsopoulos, "LEGAL-BERT: The Muppets straight out of Law School," en Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 2898–2904, 2020. doi: 10.18653/v1/2020.findings-emnlp.261. [En línea] Disponible en: https://aclanthology.org/2020.findings-emnlp.261

[5] I. Chalkidis et al., "LexGLUE: A Benchmark Dataset for Legal Language Understanding in English," en Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022), pp. 4310–4330, 2022. doi: 10.18653/v1/2022.acl-long.297. [En línea] Disponible en: https://aclanthology.org/2022.acl-long.297

[6] J. Cañete, G. Chaperon, R. Fuentes, J.-H. Ho, H. Kang y J. Pérez, "Spanish Pre-trained BERT Model and Evaluation Data," arXiv preprint, arXiv:2308.02976, 2023. doi: 10.48550/arXiv.2308.02976. [En línea] Disponible en: https://arxiv.org/abs/2308.02976

IA y ML para la identificación de Animales en Entornos Naturales

Pablo Ezequiel Inchausti[1], Magdalena Arias[2], Marcelo Castro[1], Martín Fustiñana[1], Valentino Salvi Vichich[1], Ivan Garcia[1], Juan Ignacio Urruzuno[1]

[1]Universidad Argentina de la Empresa, Instituto de Tecnología (UADE - INTEC), CABA, Argentina
[2]Centro de Investigación Aplicada y Transferencia Tecnológica en Recursos Marinos (CIMAS), CENPAT-CONICET, San Antonio Oeste, Río Negro, Argentina
{pinchausti, marcecastro, mfustinana, vsalvivichich, ivangarcia, jurruzuno}@uade.edu.ar
ariasmalala@gmail.com

RESUMEN

Se propone el desarrollo de herramientas basadas en Machine Learning para el análisis de imágenes de animales en dos casos de uso relevantes. El primero contribuye al proceso de fotoidentificación de la ballena franca austral (*Eubalaena australis*), en el marco de las actividades de investigación lideradas por la Dra. Magdalena Arias (CONICET) en el Golfo San Matías. Este proceso permite identificar individuos a partir de patrones visuales únicos, facilitando el monitoreo y la conservación de la especie. El segundo caso de uso se enfoca en el ámbito agropecuario, donde se aplican técnicas de visión artificial y monitoreo mediante drones para la detección e identificación de ganado en entornos rurales.

Esta capacidad permite mejorar el control, seguimiento y gestión de animales en grandes extensiones territoriales. Ambos proyectos demuestran el potencial de la Visión Artificial para automatizar procesos, y contribuir a la conservación y el monitoreo de especies en entornos naturales y productivos.

***Palabras Clave**:* Machine Learning, Visión Computacional, Fotoidentificación, Monitoreo de Fauna, *Eubalaena australis*

CONTEXTO

El presente trabajo se desarrolla en el Instituto de Tecnología (INTEC), dependiente de la Universidad Argentina de la Empresa (UADE), en el marco del proyecto de investigación *"IA y ML para la identificación de animales en entornos naturales",* acreditado como Actividad de Ciencia y Técnica (ACyT) bajo el código A26T10, con una duración prevista de 15 meses a partir del año 2026.

En el marco de este proyecto, se estableció un acuerdo de cooperación científica con la Dra. Magdalena Arias, investigadora del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y de la Universidad Nacional del Comahue, quien proveyó el catálogo de imágenes fotoidentificadas de la ballena franca austral (*Eubalaena australis*), utilizado para el entrenamiento y validación de modelos de Machine Learning.

La Dra. Arias lleva adelante actividades de investigación, monitoreo poblacional y estudio de esta especie en el Golfo San Matías, provincia de Río Negro, Argentina, en el marco de estudios científicos desarrollados en la región [1] [2]

1. INTRODUCCIÓN

La identificación automática de animales mediante el análisis de imágenes constituye un área de creciente interés dentro del campo de la

visión computacional, con aplicaciones en investigación científica, monitoreo de fauna y sistemas productivos. Tradicionalmente, estos procesos se realizan de forma manual, lo que requiere un esfuerzo considerable de tiempo, un alto nivel de especialización y limita su escalabilidad frente a grandes volúmenes de datos.

En particular, la fotoidentificación de la ballena franca austral (*Eubalaena australis*) permite reconocer individuos a partir de patrones únicos presentes en la región cefálica, facilitando el monitoreo poblacional y el análisis de su comportamiento a lo largo del tiempo [3]. Este tipo de análisis forma parte de los estudios científicos desarrollados en el Golfo San Matías y puede beneficiarse del uso de modelos de Machine Learning capaces de automatizar la detección y análisis de patrones visuales.

Asimismo, el uso de drones en entornos rurales ha incrementado la disponibilidad de imágenes para el monitoreo de animales, generando la necesidad de herramientas que permitan su detección e identificación automática. En este contexto, los modelos basados en redes neuronales convolucionales (CNN) [4], y en particular arquitecturas como YOLO (You Only Look Once) [5], han demostrado alta precisión y eficiencia en tareas de detección de objetos en imágenes.

El alcance de este proyecto se centra en el desarrollo de herramientas basadas en visión computacional, con el objetivo de asistir el trabajo científico de la Dra. Arias y su equipo, así como abordar necesidades de monitoreo automatizado de animales, incluyendo su aplicación en el control de ganado en el sector agropecuario.

Objetivo:

Desarrollar modelos basados en Inteligencia Artificial y Machine Learning para la identificación automática de animales mediante técnicas de visión computacional, aplicados a la fotoidentificación de la ballena franca austral y al monitoreo de ganado en tiempo real en entornos rurales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de I+D del proyecto A26T10 se orientan al desarrollo de modelos de identificación automática de animales mediante técnicas de Inteligencia Artificial y Machine Learning, a partir del análisis de imágenes digitales, incluyendo:

a. La integración y análisis de datasets de imágenes provenientes de catálogos científicos con la fotoidentificación disponible;

b. La creación de nuevos datasets mediante la captura de imágenes y videos utilizando drones, permitiendo la generación de información visual en tiempo real para actividades de monitoreo automatizado de animales;

c. El estudio de técnicas de procesamiento de imágenes, incluyendo redimensionamiento, normalización, segmentación y aumento de datos (*data augmentation*) con el fin de mejorar la calidad y robustez de los modelos;

d. El desarrollo y entrenamiento de modelos basados en redes neuronales convolucionales (CNN), incluyendo arquitecturas como YOLO, para la detección automática de animales y la extracción de características visuales distintivas, permitiendo la identificación de patrones únicos mediante técnicas de visión computacional.

e. La implementación de técnicas de comparación de características (*feature extraction y feature matching*) para la identificación automática de individuos a partir de nuevas imágenes;

f. La evaluación y validación de los modelos mediante métricas de desempeño, tales como precisión, exactitud y tasa de identificación correcta.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el caso de la fotoidentificación de la ballena franca austral, se obtuvo acceso al catálogo de imágenes provisto en el marco de la colaboración científica, a partir del cual se construyó un dataset organizado y etiquetado para el entrenamiento de modelos. Como resultado de este proceso, se conformó un

dataset con un total de 401 ballenas identificadas y 2655 recortes de patrones de callosidades, los cuales fueron estandarizados y preparados para su utilización en modelos de Machine Learning.

La Fig. 1 muestra la interfaz del sistema de fotoidentificación desarrollado, utilizada para la catalogación y comparación de imágenes de ballenas a partir del catálogo provisto por el equipo de investigación de la Dra. Magdalena Arias (CONICET). Las imágenes utilizadas pertenecen al catálogo científico del equipo de la Dra. Arias y son empleadas en el marco del acuerdo de cooperación científica para el entrenamiento y validación de los modelos.

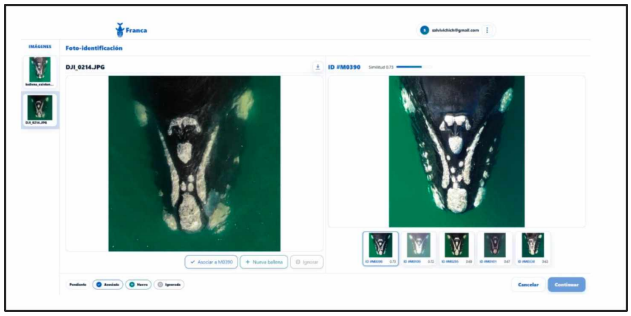


Fig. 1 Franca: Dashboar de Identificacion. Fotos del Catálogo de Provisto

Asimismo, se desarrolló el pipeline completo de procesamiento, incluyendo la preparación del dataset, la estandarización de imágenes y la aplicación de técnicas de *data augmentation*, con el objetivo de mejorar la capacidad de generalización del modelo. Sobre este dataset se implementó el entrenamiento de modelos basados en redes neuronales convolucionales, incluyendo arquitecturas como ResNet y YOLO, utilizadas para la detección de patrones y la identificación de individuos.

En el caso del monitoreo en entornos rurales, se conformaron datasets a partir de repositorios públicos y fuentes abiertas, los cuales fueron sometidos a procesos de limpieza, normalización y organización mediante técnicas ETL. Con el dataset limpio y estructurado, se inició el proceso de entrenamiento del algoritmo basado en la arquitectura YOLO, seleccionada por su capacidad de procesamiento en tiempo cercano al real y su eficiencia en tareas de detección de objetos. La Fig. 2 muestra la aplicación AeroSentinel y la Fig. 3 el Kit básico para un monitoreo de 200 hectáreas

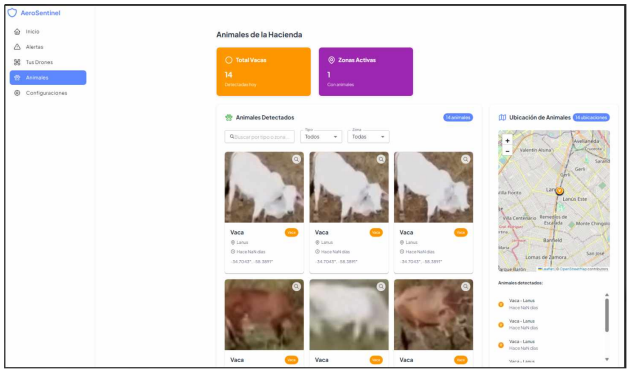


Fig. 2 AeroSentinel: Aplicación de Monitoreo Rural

Como resultado esperado, se prevé la obtención de modelos capaces de detectar e identificar animales en imágenes, así como el desarrollo de prototipos funcionales que permitan validar la aplicabilidad de técnicas de visión computacional en escenarios reales.



Fig. 3 AeroSentinel: Kit básico para 200 hectáreas

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto A26T10, orientado al desarrollo de modelos de Machine Learning para la identificación automática de animales mediante visión computacional, está dirigido por el Mg. Pablo Ezequiel Inchausti. Participan asimismo la Ing. Macarena Gorgal y el Ing. Marcelo Castro como directores de los *Proyectos Finales de Ingeniería* (PFI) asociados a esta línea de investigación.

Respecto de los Proyectos Finales de Ingeniería (PFI) de la carrera Ingeniería en Informática que se integran inicialmente al proyecto, se incluyen los siguientes:

- El PFI desarrollado por los estudiantes Martín Fustiñana y Valentino Salvi Vichich, orientado al desarrollo de modelos de Machine Learning aplicados a la fotoidentificación de la ballena franca austral mediante el análisis de imágenes científicas provistas en el marco de la colaboración con el CONICET;

- El PFI desarrollado por los estudiantes Iván García y Juan Ignacio Urruzuno, enfocado en el desarrollo de sistemas basados en visión computacional para la detección e identificación automática de animales mediante el uso de drones en entornos rurales.

Tambien se prevé, durante el ciclo lectivo 2026, la incorporación de PFI adicionales con el objetivo de desarrollar las líneas de investigación derivadas.

Como parte de la formación de recursos humanos, los datasets, modelos y prototipos desarrollados en el proyecto son utilizados como recursos didácticos en asignaturas de las carreras de informática. Estos prototipos se despliegan sobre la plataforma AWS, permitiendo incorporar contenidos prácticos relacionados con el uso de infraestructuras cloud en el desarrollo de soluciones basadas en Machine Learning.

5. BIBLIOGRAFÍA

[1] «Southern right whale behavioral changes due to interactions with whale-watching vessels: evidence-based calculations aimed at improving management policies - Arias - 2025 - Marine Mammal Science - Wiley Online Library». Accedido: 25 de febrero de 2026. [En línea]. Disponible en: https://onlinelibrary.wiley.com/doi/10.1111/mms.13149

[2] M. Arias *et al.*, «Southern right whale Eubalaena australis in Golfo San Matías (Patagonia, Argentina): Evidence of recolonisation», *PLOS ONE*, vol. 13, n.º 12, p. e0207524, dic. 2018, doi: 10.1371/journal.pone.0207524.

[3] R. Payne, «Long term behavioral studies of the southern right whale (Eubalaena australis)», *Rep. Int. Whal. Comm. Spec. Issue 10*, pp. 161-167, 1986.

[4] «Deep learning | Nature». Accedido: 25 de febrero de 2026. [En línea]. Disponible en: https://www.nature.com/articles/nature14539

[5] «You Only Look Once: Unified, Real-Time Object Detection | IEEE Conference Publication | IEEE Xplore». Accedido: 25 de febrero de 2026. [En línea]. Disponible en: https://ieeexplore.ieee.org/document/7780460

Inteligencia Artificial al servicio de la comunidad: un enfoque predictivo y personalizado

Maximiliano C. D. Budán[1,2] Paola D. Budán[1,3] Melisa G. Escañuela González[1,2]

Gerardo I. Simari[2,3] Diego C. Martinez[2,3] Aldo R. Roldan[1] Héctor Paz[1]

Susana I. Herrera[1] Marcos A. Argañaraz[1]

[1] Inst. de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero (IIISI - UNSE)

[2] Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET)

[3] Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS)

{gis, dcm}@cs.uns.edu.ar, {mbudan, pbudan, sherrera, meliesca, aroldan}@unse.edu.ar

RESUMEN

Este trabajo se enmarca en una línea de investigación orientada a la integración de técnicas de Inteligencia Artificial, Argumentación Rebatible y Ciencia de Datos con el objetivo de optimizar procesos complejos en dominios socio-técnicos relevantes, particularmente el educativo y el agropecuario. La propuesta parte de la premisa de que la toma de decisiones en estos contextos se ve fuertemente condicionada por la incertidumbre, la heterogeneidad de la información y la coexistencia de argumentos potencialmente conflictivos.

En este contexto, se propone un enfoque que combina modelos predictivos basados en análisis de datos con formalismos argumentativos capaces de representar, evaluar y explicar razonamientos rebatibles. La integración de estas herramientas permite no solo anticipar escenarios futuros a partir de datos históricos y actuales, sino también justificar y transparentar las decisiones sugeridas por los sistemas inteligentes. Se analizan los fundamentos teóricos del enfoque, se describen las principales líneas de investigación y desarrollo, y se discuten los resultados alcanzados y esperados. Los avances obtenidos sugieren que la combinación de ciencia de datos y argumentación computacional constituye una base sólida para el diseño de sistemas inteligentes explicables, robustos y orientados al apoyo efectivo a la toma de decisiones en contextos reales.

Palabras Clave: Inteligencia Artificial, Argumentación Rebatible, Ciencia de Datos, Toma de Decisiones, Sistemas Inteligentes.

CONTEXTO

Esta línea de investigación se desarrolla en el marco del proyecto "Inteligencia Artificial al servicio de la comunidad: un enfoque predictivo y personalizado" (PGI 23/C222-Bint-2025) dirigido por Maximiliano C. D. Budán y Paola D. Budán, financiado por la Facultad de Ciencias Exactas y Tecnología de la Universidad Nacional de Santiago del estero (FCEyT-UNSE) para el período 2025-2026 y radicado en el Instituto de Investigación en Informática y Sistema de Información.

1. INTRODUCCIÓN

El estudio de la Teoría de la Argumentación ha sido un tema de interés sostenido desde los filósofos griegos hasta la investigación contemporánea en Inteligencia Artificial y Ciencias Cognitivas [DHM+11]. Este interés se explica por el papel central de la argumentación en la

vida cotidiana y en los debates complejos, como componente clave de la inteligencia humana. En el ámbito de la Inteligencia Artificial, la argumentación computacional se ha consolidado como un enfoque para modelar el razonamiento no monótono y la toma de decisiones en contextos de desacuerdo y de información incompleta [RS09]. La capacidad de argumentar permite a los individuos comprender nuevos problemas, llevar a cabo razonamientos científicos, expresar ideas, aclarar conceptos y defender posiciones frente a alternativas.

De manera general, la argumentación puede entenderse como un discurso expositivo cuyo objetivo es persuadir o convencer, brindando razones que apoyan o refutan una determinada tesis. Desde una perspectiva analítica, la argumentación persigue tres fines fundamentales: identificar, analizar y evaluar los argumentos que se presentan a favor o en contra de una conclusión. En su formulación lógica clásica, un argumento está compuesto por una conclusión, un conjunto de premisas que la respaldan y un mecanismo de inferencia que permite derivarla a partir de dichas premisas. En este sentido, la Teoría de la Argumentación se ocupa de estudiar los procesos mediante los cuales los agentes justifican decisiones, incluso aquellas que pueden haber sido tomadas inicialmente de manera no completamente racional [RS09].

El razonamiento argumentativo se desarrolla a través de un proceso dialéctico que modela situaciones problemáticas del mundo real. Dicho proceso puede concebirse como una interacción entre distintos roles: un proponente, encargado de introducir una tesis y los argumentos que la sustentan; un oponente, cuya función es cuestionar esa tesis mediante contraargumentos; y una audiencia o árbitro, responsable de determinar qué argumentos resultan aceptables. A partir de este intercambio, se establece un conjunto de creencias aceptadas que guían la acción y la toma de decisiones de los agentes, ya sean humanos o sistemas inteligentes.

En este contexto surgen los marcos y sistemas argumentativos, cuyo objetivo es formalizar el análisis de las razones a favor y en contra de una proposición para determinar su aceptabilidad [CMS04]. Los sistemas argumentativos constituyen modelos de razonamiento que representan el conocimiento en un lenguaje formal, lo estructuran en entidades denominadas argumentos y definen relaciones, como el ataque, que capturan la existencia de conflicto entre ellos. A partir de estas estructuras, se introducen nociones de aceptabilidad que permiten decidir qué argumentos deben ser finalmente aceptados.

Particularmente relevantes son los sistemas argumentativos basados en reglas, en los cuales el conocimiento se representa mediante hechos, presuposiciones y reglas de inferencia que permiten derivar conclusiones de manera tentativa [PV00]. A partir de esta base de conocimiento, los agentes pueden construir argumentos a favor o en contra de una afirmación y evaluar su aceptación considerando las interacciones entre ellos. A medida que aumenta la capacidad de representación de estos sistemas, se incrementa también la complejidad del proceso argumentativo, pero se obtiene a cambio una modelización más rica y precisa del dominio bajo análisis.

En paralelo, la Ciencia de Datos ha emergido como una disciplina clave para la extracción de conocimiento a partir de grandes volúmenes de datos heterogéneos [TLCV15]. La Ciencia de Datos ha emergido en la última década como una disciplina que integra estadística, aprendizaje automático e infraestructura computacional para el análisis sistemático de grandes volúmenes de datos heterogéneos [CHEN16]. Mediante la combinación de técnicas de estadística, aprendizaje automático y conocimiento experto del dominio, la ciencia de datos busca transformar datos brutos en información significativa que permita apoyar la toma de decisiones y optimizar procesos. Este proceso abarca etapas como la recolección, limpieza, análisis

exploratorio, modelado predictivo y prescriptivo, así como la interpretación y comunicación de resultados.

Las aplicaciones de la ciencia de datos se extienden a múltiples sectores, incluyendo la educación y la agricultura, donde posibilita la personalización de procesos, la optimización del uso de recursos y la anticipación de escenarios futuros [PF13, TLCV15]. En particular, en el dominio educativo, la analítica de datos y la minería de datos educacionales han demostrado su utilidad para apoyar procesos de enseñanza, aprendizaje y toma de decisiones académicas [RV20], de manera análoga, en el ámbito agropecuario, la Ciencia de Datos se ha consolidado como un campo emergente para mejorar la productividad y sostenibilidad de los sistemas agrícolas mediante el análisis de datos complejos [Pau20]. No obstante, los enfoques puramente predictivos suelen carecer de mecanismos explícitos para justificar y explicar las decisiones sugeridas, lo que limita su adopción en contextos críticos.

En los últimos años, la investigación en Inteligencia Artificial ha puesto de manifiesto la necesidad de integrar capacidades predictivas con modelos que imiten el razonamiento humano frente a situaciones de desacuerdo e incertidumbre [ALS20, VBP21, KAEPF21]. En este sentido, la argumentación rebatible ofrece un marco adecuado para representar y evaluar conclusiones a la luz de nueva evidencia. La integración de Ciencia de Datos y Argumentación Computacional permite avanzar hacia sistemas inteligentes predicen, explican y justifican sus recomendaciones.

Este trabajo propone un enfoque integrador orientado al desarrollo de sistemas inteligentes explicables para los dominios educativo y agropecuario. El objetivo es combinar análisis predictivo basado en datos con formalismos argumentativos que faciliten la evaluación crítica, la transparencia y la robustez de los procesos de toma de decisiones en contextos reales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La investigación se estructura en torno a tres ejes principales que articulan los aportes teóricos y aplicados del proyecto.

2.1. Integración de Ciencia de Datos y Argumentación Rebatible

El primer eje se propone desarrollar un marco conceptual que integre modelos de análisis de datos con formalismos de argumentación rebatible. Los primero permiten anticipar escenarios futuros a partir de datos históricos; la argumentación, en cambio, aporta una estructura para evaluar la solidez de esas predicciones, contrastarlas con conocimiento experto y justificar las decisiones. Esta integración busca superar las limitaciones de los enfoques puramente estadísticos, incorporando mecanismos que faciliten la interpretación de los resultados y la detección de posibles inconsistencias.

2.2. Análisis Argumentativo en Dominios Educativos y Agropecuarios

El segundo eje aborda la aplicación del enfoque en los dominios educativo y agropecuario. En educación, se exploran técnicas para personalizar la enseñanza, evaluar el progreso y optimizar recursos. En el dominio agropecuario, se analizan datos productivos y ambientales para apoyar decisiones de planificación, uso eficiente de recursos y sostenibilidad. En ambos casos, la argumentación rebatible permite representar explícitamente las razones que respaldan o cuestionan recomendaciones, favoreciendo una toma de decisiones transparente.

2.3. Herramientas Inteligentes y Explicabilidad

El tercer eje se orienta al diseño e implementación de herramientas inteligentes que integren los componentes desarrollados. Se enfatiza en la explicabilidad de los sistemas, entendida como la capacidad de justificar las decisiones

sugeridas mediante estructuras argumentativas comprensibles. Esta cualidad es clave para generar confianza en los sistemas inteligentes y facilitar su adopción en contextos reales.

3. RESULTADOS OBTENIDOS Y ESPERADOS

3.1. Resultados Alcanzados

Se avanzó de manera consistente conforme al plan propuesto, integrando desarrollos teóricos, formales y aplicados. Los principales resultados alcanzados son los siguientes:

- **Implementación del LAF:** Se implementó el Labeled Argumentation Framework (LAF), extendiendo los marcos argumentativos abstractos tradicionales mediante la incorporación de relaciones de vinculación sumadas a las de ataque, para modelar apoyos estructurales entre interpretaciones.
- **Validación del razonamiento interpretativo:** Se validó el LAF como una extensión operativa que eleva a las interpretaciones a entidades de primera clase dentro del proceso argumentativo, permitiendo analizar de manera conjunta conflictos y apoyos entre argumentos.
- **Formalización de propiedades de racionalidad:** Se avanzó en la formalización y el análisis de propiedades de racionalidad interpretativa, como coherencia interna, estabilidad frente a conflictos y robustez basada en relaciones de vinculación, proporcionando criterios para la aceptabilidad de interpretaciones en contextos de desacuerdo.
- **Análisis de complementariedad:** Se realizó un análisis teórico de la complementariedad entre mecanismos de rechazo de información incompatible y de incorporación de información compatible, contribuyendo a una mayor interpretabilidad de los procesos de razonamiento rebatible.
- **Avances en Ciencia de Datos:** En el eje de Ciencia de Datos e Inteligencia Artificial, se iniciaron desarrollos orientados a la automatización del análisis de información mediante flujos de trabajo basados en n8n, que integran procesamiento automático de texto, análisis de sentimiento y LLMs, generando insumos utilizados como evidencia dentro de los modelos argumentativos.

3.2. Objetivos en curso y futuros

La línea de investigación y desarrollo continúa conforme al plan del proyecto, con el objetivo de consolidar un enfoque que articule razonamiento argumentativo y ciencia de datos para la toma de decisiones explicables. En particular, se plantean los siguientes objetivos:

- **Análisis computacional del LAF:** Profundizar el estudio de las propiedades computacionales del LAF, analizando la complejidad algorítmica asociada al cómputo de extensiones estables y robustas, y desarrollando algoritmos adecuados para su implementación computacional eficiente.
- **Integración con flujos de Ciencia de Datos:** Consolidar e integrar los flujos de trabajo automáticos desarrollados en n8n, incorporando análisis de sentimiento, procesamiento de texto y LLMs como mecanismos de generación y estructuración de evidencia que alimente el razonamiento argumentativo.
- **Manejo de incertidumbre y evidencia heterogénea:** Extender el marco argumentativo para representar distintos grados de confianza en la evidencia proveniente de fuentes de datos heterogéneas, integrando en forma flexible resultados de análisis de datos y la evaluación argumentativa.
- **Validación en dominios del proyecto:** Aplicar y validar el enfoque propuesto en los dominios de interés del proyecto, particularmente en contextos educativos y agropecuarios, evaluando su capacidad para apoyar procesos de análisis, interpretación y toma de decisiones de manera explicable y fundamentada.

4. FORMACIÓN DE RRHH

La presente línea de investigación se desarrolla en el ámbito del Instituto de Investigación en Informática y Sistemas de Información (IIISI)

perteneciente a la Universidad Nacional de Santiago del Estero (UNSE), en articulación con el Instituto Nacional de Tecnología Agropecuaria (INTA), y tiene como ejes centrales la argumentación computacional, el razonamiento automático bajo incertidumbre y la ciencia de datos aplicada. En este marco, el proyecto constituye un espacio clave para la formación de tesistas de grado y posgrado, becarios y jóvenes investigadores en áreas estratégicas de la Inteligencia Artificial. Estos tesistas e investigadores noveles tienen la oportunidad de complementar su formación en el Laboratorio de Inteligencia Artificial del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur (UNS), por ser uno de los más reconocidos a nivel académico en el país.

Los recursos humanos en formación participan activamente en el desarrollo y análisis de modelos argumentativos, en la implementación computacional de marcos de razonamiento rebatible, y en la integración de técnicas de ciencia de datos, procesamiento automático de texto y modelos de lenguaje de gran escala. Esta participación favorece la adquisición de competencias teóricas y prácticas en formalización, diseño algorítmico y análisis de información, fortaleciendo trayectorias académicas orientadas al desarrollo de sistemas inteligentes explicables y al abordaje de problemáticas reales en los dominios educativo y agropecuario.

BIBLIOGRAFÍA

[ALS20] Alsaleh, N. J. (2020). Teaching Critical Thinking Skills: Literature Review. Turkish Online Journal of Educational Technology (TOJET), 19(1), 21–39.

[CMS04] Chesñevar, C. I., Maguitman, A. G., & Simari, G. R. (2004). A first approach to argument-based recommender systems based on defeasible logic programming. Proceedings of the International Workshop on Non-Monotonic Reasoning, 109–117.

[DHM+11] Dunne, P. E., Hunter, A., McBurney, P., Parsons, S., & Wooldridge, M. (2011). Weighted argument systems: Basic definitions, algorithms, and complexity results. Artificial Intelligence, 175(2), 457–486.

[KAEPF21] Čyras, K., Rago, A., Albini, E., Baroni, P., & Toni, F. (2021). Argumentative XAI: A survey. arXiv preprint arXiv:2105.11266.

[PF13] Provost, F., & Fawcett, T. (2013). Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media.

[PV00] Prakken, H., & Vreeswijk, G. (2000). Logical systems for defeasible argumentation. In Handbook of Philosophical Logic (2nd ed.). Kluwer Academic Publishers.

[TLCV15] Tsai, C.-W., Lai, C.-F., Chao, H.-C., & Vasilakos, A. V. (2015). Big data analytics: A survey. Journal of Big Data, 2(1), 1–32.

[VBP21] Vassiliades, A., Bassiliades, N., & Patkos, T. (2021). Argumentation and explainable artificial intelligence: A survey. The Knowledge Engineering Review, 36, e5.

[RS09] Rahwan, I., & Simari, G. (2009). *Argumentation in Artificial Intelligence*. Springer.

[CHEN16] Chen, Y., Chen, H., Gorkhali, A., Lu, Y., Ma, Y., & Li, L. (2016). *Big data analytics and big data science: a survey*. Journal of Management Analytics, 3(1), 1–42.

[RV20] Romero, C., & Ventura, S. (2020). *Educational data mining and learning analytics: An updated survey*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1355.

[Pau20] Paul, P., Bhimali, A., Sinha, R. R., Aithal, P. S., Kalishankar, T., & Saavedra, R. (2020). *Agricultural Data Science as a potential field and promoting agricultural activities and sustainable agriculture*. International Journal of Information Science and Computing, 7(2), 49–62.

Inteligencia Artificial aplicada en el Cuidado de la Salud

D'Angiolo Federico, Cabot Lucas, Ayala Lucio, Asteasuain Fernando.

Contacto: fdangiolo@undav.edu.ar

Ingeniería en Informática – Dpto. Tecnología y Administración – Universidad Nacional de Avellaneda

Resumen

Actualmente el campo de la salud utiliza la Inteligencia Artificial (IA) para mejorar la calidad de vida de las personas. En este sentido, existen distintas herramientas de IA que permiten obtener más y mejores resultados que ayudan al profesional de la salud, a diagnosticar y prevenir distintos tipos de enfermedades. Algunas herramientas utilizadas desde la IA resultan ser aquellas basadas en Visión por Computadora (VC), donde se pueden encontrar arquitecturas importantes como las Redes Convolucionales. Estas resultan de interés y practicidad en el área dado que permiten una gran precisión a la hora de obtener resultados. El presente trabajo describe un marco de investigación basado en el área de VC aplicada en salud. Dicho marco surgió a partir de dos tesis de grado basadas en clasificación de imágenes mediante Redes Convolucionales. Estos trabajos presentan una descripción de cómo detectar la Retinopatía Diabética y Cáncer de Mama, los cuales permiten generar una base fundacional para posteriores estudios donde se usan combinaciones de estas redes y otras arquitecturas.

Palabras claves: Visión por Computadora, Deep Learning, Redes Convolucionales, Salud.

Contexto

El presente documento se encuentra basado en los avances realizados en el grupo de Inteligencia Artificial (IA) con aplicación en la salud, dirigido por el Ing. D'Angiolo, Federico y el Dr. Asteasuain, Fernando. El grupo de investigación está conformado por los autores del presente trabajo, combinando investigadores consolidados, en formación, y estudiantes avanzados. El proyecto tuvo sus inicios en 2020 y actualmente, se siguen desarrollando líneas de investigación basadas en las áreas de Machine Learning (ML), Deep Learning (DL) y Visión por Computadora (VC). Este proyecto de investigación permite combinar diversas líneas de investigación dentro de la Carrera de Ingeniería en Informática de la UNDAV, potenciando así la sinergia y colaboración.

1. Introducción

En la actualidad la Inteligencia Artificial (IA) abarca una gran cantidad de áreas, siendo una de ellas su aplicación en la salud de las personas. Contemplar a la IA en este dominio genera un impacto positivo dado que se pueden analizar imágenes para mejorar la calidad de salud humana. En otras palabras, se trata de mejorar la calidad de análisis dado que los métodos tradicionales para detectar determinadas anomalías pueden tener un cierto error y hasta demora por parte de los profesionales, quienes a su vez, se ven superados por tantos diagnósticos. **[1-2]**. Ante esto, la IA intenta colaborar con el profesional de la salud en el diagnóstico y prevención de distintas enfermedades mediante, por ejemplo, análisis de imágenes. Al incorporar este tipo de tecnología, se puede ahorrar tiempo en el diagnóstico y mejorar la calidad de vida humana **[1-2]**.

Ante esta situación, se estudian dentro de la IA los distintos pilares que la componen,

siendo uno de ellos, Visión por Computadora. Este pilar tiene en cuenta distintas herramientas, siendo una de estas, las Redes Neuronales Convolucionales (CNNs, Convolutional Neural Networks), con sus distintas variantes. Con estas redes se pueden clasificar imágenes, acción importante a la hora de aplicar en salud, dado que permite tener un mayor detalle en cuanto al diagnóstico. Además, con estos recursos, se pueden obtener resultados de gran precisión a la hora de prevenir enfermedades. Esto resulta de gran ayuda en imágenes médicas dado que permite tener un mayor alcance para el profesional de la salud en el diagnóstico de enfermedades. Para lograr este objetivo estos y otros modelos de IA se entrenan con una gran cantidad de imágenes, generalmente etiquetadas, para encontrar patrones en dichas imágenes y detectar anomalías **[1-2-3].**

Si nos atenemos a lo comentado, es importante destacar que dentro de VC se puede realizar tanto clasificación como segmentación, siendo la primera aquella que permite categorizar imágenes según la gravedad del problema, mientras que la segmentación, permite aislar una zona de interés para focalizar el problema **[1-3]**. El presente trabajo se orienta a la clasificación de imágenes para la detección y prevención de enfermedades como Retinopatía Diabética y Cáncer de Mama, haciendo uso de Redes Convolucionales (CNNs) y su respectivo procesamiento de imágenes previo. Cabe destacar que estos trabajos contaron una gran contribución de personas experimentadas en el área de medicina, para generar una realimentación que permita mejorar a los modelos de IA.

En la literatura podemos encontrar diversos trabajos sustentados en CNNs las cuales cobran una fuerza relevante en lo que resulta ser, por ejemplo, radiología. Debido a las distintas variantes que se pueden encontrar de las Redes Neuronales Convolucionales, es importante estudiar cuáles de éstas permiten lograr mejores precisiones a la hora de obtener resultados, dado que la salud es un área sensible. Por ejemplo, en **[4]** se estudia una arquitectura denominada DenseNet para luego validar su efectividad frente a métodos tradicionales.

Continuando con la literatura basada en estos temas podemos encontrar trabajos basados en la aplicación de distintas herramientas de VC como por ejemplo, en el análisis de ELA (Esclerosis Lateral Amiotrófica), la cual es una enfermedad neurodegenerativa que deteriora progresivamente la función motora. Aquí resulta importante la aplicación de CNN para el tratamiento de esta enfermedad **[5]**. Resulta relevante además contemplar la flexibilidad de los hospitales y sistemas de salud para acoplar los avances tecnológicos, sobre todo en el área de IA para la detección de cáncer mediante CNNs. En este sentido, resulta interesante estudiar cómo trabajar con una Red Convolucional para detectar características importantes en imágenes y lograr así una alta precisión en el diagnóstico de dicha enfermedad **[6]**. Por otro lado, mediante la utilización de variantes de CNN como ResNet-50, ResNet-101, ResNet-152, se obtienen resultados interesantes en el análisis de imágenes por resonancia magnética (MRI) para detectar trastornos neurodegenerativos, más especialmente, Alzheimer's Disease (AD) y Mild Cognitive Impairment (MCI). En este trabajo, algo importante a destacar es el uso de Filtrado Gaussiano para reducir ruido y la mejora de contraste para resaltar detalles, es decir, previamente a la utilización de las Redes Neuronales, conviene realizar procesamientos básicos **[7]**.

Sobre la base de la literatura comentada, el objetivo de este trabajo es el desarrollo de un marco de investigación que permita obtener avances en el área de IA aplicada a la salud, más específicamente, dentro de Visión por Computadora. Para esto, a continuación se describen las líneas de investigación asociadas.

2. Líneas de Investigación y Desarrollo.

Hasta el momento se vienen realizando distintos desarrollos basados Retinopiatía Diabética y Cáncer de Mama los cuales tienen en cuenta los siguientes ítems:

1. Aplicación de IA en salud y medicina. Aquí se propone estudiar las distintas ramas de la IA, que tienen en cuenta el procesamiento de datos, Machine Learning, Deep Learning y Visión por Computadora. Todas estas herramientas constituyen el puntapié inicial en la investigación **[1-2]**.

2. Asistencia en el diagnóstico. Aquí resulta interesante cómo se puede asistir al profesional de la salud, teniendo en cuenta su importante aporte **[1-2]**.

3. VC aplicada en salud. En este sentido, se explora la utilización de clasificación de imágenes como primer avance, en el campo de la Retinopatía Diabética como en el Cáncer de Mama **[3]**.

4. Algoritmos de procesamiento de imágenes. Previamente a trabajar con CNN, resulta de gran importancia estudiar cómo se deben procesar las imágenes, sobre todo para detectar el ruido que muchas veces proviene de los equipos que toman las imágenes **[8-9]**.

5. Algoritmos de clasificación. Luego de haber procesado las imágenes, se procede entonces a la configuración y utilización de las CNN para lograr la clasificación de imágenes

6. Relevancia en la precisión de resultados. Esto resulta de gran importancia pues, una vez aplicado los distintos procesamientos, se procede a evaluar los resultados mediante métricas como accuracy, tablas de confusión, Recall, Precision y otras más, de manera que el diagnóstico sea el más útil posible a los efectos del profesional de la salud.

3. Resultados Obtenidos y Esperados

Dado que el objetivo principal del presente proyecto es desarrollar un marco de investigación basado en Visión por Computadora para la detección temprana de enfermedades, resulta importante involucrar conceptos basados en IA. Así, entonces, podemos describir los resultados obtenidos hasta el momento:

1. El primero de estos desarrollos se basa en clasificación de imágenes para la detección de Retinopatía Diabética. Esta clasificación se puede dividir en: leve, moderada, severa, proliferante o sin retinopatía. Para esto, se analizaron un conjunto de imágenes las cuales fueron procesadas para detectar ruidos, estudiar sus dimensiones y el balanceo del dataset. Por último, una vez hechos todos estos procesamientos, se procedió a configurar la CNN, es decir, tener en cuenta la cantidad de capas, el kernel y el padding, parámetros necesarios para el desarrollo **[8-9]**.

2. En cuanto a los resultados obtenidos para este desarrollo, se lograron resultados de precisión cercanos al 98% en su rendimiento **[8]**.

3. Este trabajo se desarrolló como Tesis de Grado lo cual conduce a que los estudiantes se vean motivados a la hora de desarrollar estos temas de interés académico e industrial **[8]**.

4. Adicionalmente, sobre este mismo desarrollo, se presentó un trabajo **[9]** basado en la tesis **[8]** en el Congreso CONAIISI 2024. Este trabajo fue elegido como ganador del primer premio dentro de la categoría Estudiantes.

5. En cuanto al trabajo referido a la detección de Cáncer de Mama, el objetivo aquí es clasificar las imágenes en tres categorías: benigno, maligno y normal. Previamente se realizó un procesamiento de imágenes similar al descripto en el trabajo anterior, para detectar sombras y ruidos. Luego se procedió a utilizar distintos tipos de Redes Convolucionales pre-entrenadas, como por ejemplo, VGG16, DenseNet201, ResNet50.

6. Referido a la detección de Cáncer de Mama, se utilizaron métricas como accuracy

y tablas de confusión para corroborar el buen funcionamiento de la CNN. Este trabajo alcanzó niveles cercanos al 96%, en su rendimiento **[10]**.

7. Por otro lado, en cuanto a la investigación en Cáncer de Mama, también se desarrolló como Tesis de Grado, siendo esta finalizada en 2025 **[10]**.

8. Todos estos temas de investigación se vuelcan en la asignatura "Inteligencia Artificial" de la Carrera Ingeniería en Informática, en la cual algunos de los autores son docentes. De esta manera los estudiantes logran avanzar en estos temas para generar actualización constante en la materia.

9. Dado que estos temas también se presentan dentro de la asignatura mencionada, se presentó un trabajo en TEYET 2025 orientado al impacto que tiene la IA en salud, sobre la educación. En este trabajo se describe cómo se lleva a cabo el desarrollo de estos temas, dentro de la materia **[11]**.

Sobre los resultados obtenidos hasta el momento, las líneas de investigación que permiten obtener los resultados esperables se presentarán sobre la base de algoritmos de ML, DL y VC. Sobre esto entonces, se definen los siguientes objetivos esperados:

1. Explorar y trabajar en el área de Segmentación para tomar imágenes más precisas y prevenir enfermedades. En este sentido, se pueden tomar los desarrollos realizados hasta ahora para trabajar con mayor dirección.

2. Explorar distintas herramientas de clasificación de imágenes, como por ejemplo, ViT (Visual Transformer) la cual utiliza conceptos basados en Transformers.

3. Mejorar el conjunto de imágenes sumando casos actualizados de acuerdo al avance en medicina. Las imágenes utilizadas aquí fueron extraídas de diversas fuentes, como por ejemplo, Kaggle, que es una fuente de información donde diversas instituciones y universidades aportan datos.

4. Estudiar cómo puede integrarse NLP (Natural Language Processing) a sistemas que utilizan VC, para mejorar la atención en distintos sistemas de salud **[12]**.

5. Analizar cómo integrar VIT con LLMs y agentes, para obtener resultados más explicativos **[18]**.

6. Generar automatizaciones de datos que permitan optimizar la detección de imágenes.

7. Estudiar otros tipos de enfermedades y análisis médicos posibles como por ejemplo, aplicación en radiografías de tórax, neumonía, nódulos pulmonares, ELA y otras que pueden ser de gran aplicación para la salud humana.

8. Explorar distintas Arquitecturas de Software que permitan llevar adelante procesos de BigData, considerando Cloud Computing.

9. Publicar los resultados de estas investigaciones en distintos congresos y revistas científicas para compartir con la comunidad.

Todos estos ítems apuntan a lograr un mayor crecimiento en la investigación para consolidar la aplicación de IA en salud.

Por otro lado, a nivel general de Investigación dentro de la Carrera, se incluye:

- Generar mayor desarrollo en el Laboratorio de Inteligencia Artificial de la Carrera de Ingeniería en Informática.

- Practicar una conexión más fuerte con la Industria relacionada al Software para contemplar problemáticas actuales.

- Dirección de tesis de Ingeniería y supervisión de prácticas profesionales (PPS).

Todas estas herramientas generan las líneas de investigación basadas en IA dentro de la Carrera de Ingeniería en Informática de la Universidad Nacional de Avellaneda.

Actualmente se encuentran varios proyectos en desarrollo dentro de la Carrera, uno de ellos es el estudio de computadoras de abordo satelitales (UNDAVCYT 2025), en el cual se trata el estudio de conducción en Vehículos Autónomos (VA). Las técnicas desarrolladas y comentadas en este trabajo, referido a IA en salud, contribuyen en el desarrollo de la

percepción en Vehículos Autónomos (VA) dado que en estos, resulta importante trabajar con detalle la clasificación y segmentación de imágenes para una mejor conducción. Además, las tareas de procesamiento sobre imágenes que se hacen previas a la clasificación, son de suma importancia en VA **[13-17]**.

4. Formación de Recursos Humanos

El desarrollo de este trabajo permite, hasta el momento, potenciar la investigación en el área de IA, sobre todo en estudiantes avanzados de la Carrera pues se desarrollan tesis de grado basadas en este mismo tema y otros similares. Algunos ejemplos de tesis de grado, como los comentados, se basaron en el análisis de clasificación de imágenes mediante CNN.

Tomando las bases que los estudiantes han dejado y que se desarrolló hasta el momento en esta temática, resulta motivador para futuros estudiantes avanzados e investigadores, poder continuar con este desarrollo dado que existen diversas temáticas que se pueden continuar en la investigación.

5. Bibliografía

1) Mohammad Amir Salari.: "Artificial Intelligence, Cloud Computing, and Computer Vision in Healthcare". Future Generation Computer Systems 118 (2021) 208–218 (2025).

2) Mohd Javaid., Abid Haleem., Ravi Pratap Singh., Mumtaz Ahmed.: "Computer vision to enhance healthcare domain: An overview of features, implementation, and opportunities".. 792-803 (2024).

3) Heidi L., Keivan N., Roshini R., Ivan N. Ayala., Charles B., Anirban B., Pablo Moreno Franco., Daniel A. D., Brian W. Pickering., and Vitaly Herasevich.: "Applied Artificial Intelligence in Healthcare: A Review of Computer Vision Technology Application in Hospital Settings". . https://doi.org/ 10.3390/jimaging10040081 (2024).

4) Keshav Kumar and Dr Narasimham.: "Enhancing Radiology Diagnosis through Convolutional Neural Networks for Computer Vision in Healthcare" (2023).

5) Qiu Fang, Jun Wang, Qi Gong, Mengxuan SongBiplov Paneru.: "Computer Vision in Healthcare: ALS Patient Aid Application with Computer Vision & Green API". (2023).

6) Mohammad Shahin, F. Frank Chen, Ali Hosseinzadeh, Mazdak Maghanaki.: "Deploying deep convolutional neural network to the battle against cancer: Towards flexible healthcare systems. 2024. Towards flexible healthcare systems ". CONAIISI 2019. ISBN 978-987-4417-73-2 (2019)

7) Purushottam Kumar Pandey, Jyoti PruthiJyoti Pruthi, Saeed Alzahrani Saeed Alzahrani, Anshul Verma Anshul, Verma Benazeer and Zohra,Benazeer Zohra.: "Enhancing healthcare recommendation: transfer learning in deep convolutional neural networks for Alzheimer disease detection". https://doi.org/10.3389/fmed.2024.1445325 (2024).

8) Cabot, L.: Tesis de Grado. "Detección Temprana de Retinopatía Diabética mediante Algoritmos de Inteligencia Artificial". 2021.

9) Cabot, L., D'Angiolo, F.: "Detección Temprana de Retinopatía Diabética mediante Algoritmos de InteligenciaArtificial". CONAIISI 2024.

10) Ayala, L.: Tesis de Grado. "Redes Neuronales para el diagnóstico de imágenes de cáncer de mama". 2025.

11) D'Angiolo, F., Cabot, L., Asteasuain, F.: "Inteligencia Artificial aplicada a la Salud". TEYET 2025.

12) Vybhav Reddy Kammireddy Changalreddy., Reeta Mishra.: "Improving Population Health Analytics with Form Analyzer Using NLP and Computer Vision". International Journal of Research in all Subjects in Multi Languages (2025).

13) Yang Xiao, Liam Daniel, and Marina Gashinova. Image Segmentation and Region Classification in Automotive High-Resolution Radar Imagery. 2021

14) Simegnew Yihunie Alaba, Ali C. Gurbuz and John E. Ball. Emerging Trends in Autonomous Vehicle Perception: Multimodal Fusion for 3D Object Detection. 2024.

15) Shan Wang, Yanhao Zhang, Ankit Vora, Akhil Perincherry, and Hongdong Li. Satellite Image Based Cross-view Localization for Autonomous Vehicle. 2023.

16) Bin Zhao, Member, IEEE, Pengfei Han, Xuelong Li, Fellow, IEEE. Vehicle Perception from Satellite. 2024.

17) HRAG-HAROUT JEBAMIKYOUS, (Member, IEEE), AND RASHA KASHEF , (Member, IEEE). Autonomous Vehicles Perception (AVP) Using Deep Learning: Modeling, Assessment, and Challenges. 2022.

18) Shehroz S. Khan, Petar Przulj, Ahmed Ashraf, Ali Abedi. ChestGPT: Integrating Large Language Models and Vision Transformers for Disease Detection and Localization in Chest X-Rays. 2025.

Inteligencia Artificial y Sistemas Multiagentes: Un Marco para la Transformación Digital y la Gestión Colaborativa en las Organizaciones

Nicolas Garcia Martinez, Patricio Nicolas Castro, Horacio Muñoz Abate Universidad Nacional de Río Negro. Sede Atlántica Laboratorio de Informática Aplicada {ngarciam, pcastro, hmunoz}@unrn.edu.ar

RESUMEN

Este proyecto investiga la transformación digital de las organizaciones mediante la implementación de Inteligencia Artificial (IA) y sistemas multiagentes, enfocándose en la optimización de procesos operativos y la mejora en la toma de decisiones. El propósito es diseñar un marco que integre la IA Documental para la extracción automatizada de información y los Sistemas Multiagentes para fomentar la colaboración inteligente entre diversas áreas organizacionales. El resultado esperado es el desarrollo de prototipos y metodologías que permitan gestionar grandes volúmenes de información estructurada y no estructurada, mejorando la eficiencia y fomentando la innovación tanto en el sector público como privado. El impacto esperado incluye: 1) Optimización de la gestión de la información documental. 2) Automatización de procesos mediante la colaboración de agentes inteligentes. 3) Mejora en la precisión de la búsqueda de datos y calidad de las decisiones. 4) Formulación de recomendaciones para una implementación responsable y ética de la IA. Este estudio se sustenta en los avances de la IA Generativa y los Grandes Modelos de Lenguaje (LLM), adaptándolos a las necesidades específicas del entorno administrativo y universitario.

Palabras clave: Inteligencia Artificial, Sistemas Multiagentes, IA Documental, Transformación Digital, Gestión Organizacional.

CONTEXTO

En un escenario post-pandemia, la transformación digital se ha vuelto un imperativo para gestionar grandes volúmenes de datos y servicios al ciudadano. Las organizaciones enfrentan el reto de procesar información de manera eficiente para no perder competitividad. En este ámbito, el Laboratorio de Informática Aplicada (LIA) de la UNRN cuenta con una sólida trayectoria en proyectos de I+D relacionados con el Gobierno Electrónico y Ciudades Inteligentes. Entre los antecedentes relevantes se destacan:

PI UNRN 40-C-875: Herramientas Informáticas para el Desarrollo de Servicios Digitales Innovadores.

PI-UNRN 40-B-802: Ciencia de Datos Aplicada en diversos casos de estudio.

Reconocimientos nacionales en el Simposio de Informática en el Estado (SADIO) entre 2017 y 2024, incluyendo trabajos sobre Blockchain y Transformación Digital. Estas experiencias previas sientan las bases técnicas para abordar ahora la integración de agentes inteligentes y procesamiento de lenguaje natural en la gestión administrativa.

1. INTRODUCCIÓN

Los procesos de transformación digital de las organizaciones tanto en el sector público como privado, abarcan la digitalización y desarrollo de servicios a los ciudadanos [1]. Esto llevó a que las organizaciones se enfrenten al reto de gestionar grandes volúmenes de información tanto estructurada como no estructurada. Este proyecto explora fundamentalmente cómo la IA Documental y los Sistemas Multiagentes pueden ayudar a las organizaciones tanto públicas como privadas en su proceso de transformación digital y a ser más eficientes e innovadoras. Según Lei [2] “La IA de documentos, o inteligencia de documentos, es un tema de investigación relativamente nuevo que se refiere a las técnicas para leer, comprender y analizar automáticamente documentos comerciales”. A su vez los Sistemas Multiagentes son como "equipos" de

programas inteligentes que trabajan juntos para resolver problemas complejos (Orchestrating agents: Routines and handoffs, s/f), este paradigma poderoso está ayudando a lograr tareas cada vez más complejas en dominios que van desde las matemáticas y la codificación hasta la respuesta a preguntas, la optimización de la cadena de suministro, la toma de decisiones en línea y el entretenimiento [3].

La IA Documental, o inteligencia de documentos, permite leer, comprender y analizar automáticamente documentos comerciales mediante técnicas de aprendizaje profundo y procesamiento de lenguaje natural (NLP). Por otro lado, los Sistemas Multiagentes actúan como equipos de programas inteligentes que colaboran para resolver problemas complejos, desde la optimización de cadenas de suministro hasta la toma de decisiones en línea.

A nivel mundial, la IA ha sido adoptada para automatizar tareas y optimizar logística, proyectando una contribución de 16,5 billones de dólares a la economía global para 2030. En Argentina, casos como el sistema Prometea en la justicia de CABA demuestran el potencial de la IA para agilizar dictámenes jurídicos. Sin embargo, persiste una brecha en la implementación conjunta de IA documental y sistemas multiagentes en procesos administrativos locales, donde la heterogeneidad de sistemas y la resistencia al cambio son obstáculos significativos.[4]

La importancia de este estudio radica en llenar ese vacío, proporcionando un marco conceptual que maximice la eficiencia operativa y libere recursos humanos para tareas estratégicas. Se busca que las organizaciones no solo automaticen tareas, sino que evolucionen hacia modelos de "universidades u organizaciones inteligentes" casi autónomas.

Por otra parte, de acuerdo al Boletín Oficial de la República Argentina (Ministerio de Producción y Trabajo y Ministerio de, 2019) “el sector científico y tecnológico tiene el potencial de contribuir gracias a sus aportes e innovaciones, a elevar la competitividad y productividad de la economía, del país, impactando en el mejoramiento de la calidad de vida de la población...”. Cabe destacar, que una ciudad latinoamericana que se destaca en este tipo de innovación, es Medellín, se mantiene a la vanguardia de este tipo de tecnología, ya que ha sabido planificar y adoptar tecnologías en diversos ambientes de las ciudad por las que el mundo la ha premiado de manera regular como una ciudad innovadora, sustentable e inteligente (Panamerican Word, 2019) y que a través de diversas actividades en la región como Britos en sus proyectos de investigación [5][6], y en conjunto con diversas organizaciones, como por ejemplo Instituto Nacional de Tecnología Agropecuaria (INTA Río Negro) se han realizado innovaciones al respecto.

En el contexto de la incesante transformación digital, las soluciones basadas en datos están surgiendo como herramientas poderosas para abordar diversas tareas industriales, como la previsión, la detección de anomalías, la planificación e incluso la toma de decisiones complejas. Aunque la I+D centrada en los datos ha sido fundamental para aprovechar estas soluciones, a menudo conlleva costos significativos en términos de recursos humanos, computacionales y de tiempo. Este artículo analiza el potencial de los modelos de lenguaje de gran tamaño (LLM) para acelerar el ciclo de evolución de la I+D centrada en los datos. Al evaluar los elementos fundamentales de la I+D centrada en los datos, incluidos los datos heterogéneos relacionados con las tareas, el conocimiento del dominio multifacético y diversas herramientas funcionales de computación, exploramos qué tan bien los LLM pueden comprender los requisitos específicos del dominio, generar ideas profesionales, utilizar herramientas específicas del dominio para realizar experimentos, interpretar resultados e incorporar el conocimiento de esfuerzos pasados para abordar nuevos desafíos. Tomamos la investigación de inversión cuantitativa como un ejemplo típico de un escenario de I+D centrado en datos industriales y verificamos nuestro marco propuesto en nuestra plataforma de investigación cuantitativa de código abierto Qlib y obtuvimos resultados prometedores que arrojan luz sobre nuestra visión de la evolución automática del ciclo de I+D centrado en datos industriales.[3]

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta investigación se centra en la aplicación de IA para la transformación organizacional, explorando la creación y entrenamiento de modelos específicos para las necesidades de gestión.

Hipótesis de Trabajo: La integración de IA Documental y Sistemas Multiagentes, apoyada en modelos de lenguaje (LLM) previamente entrenados y ajustados, permite transformar las organizaciones mejorando su capacidad de respuesta y permitiendo una toma de decisiones basada en datos con mayor precisión que los métodos tradicionales.

Objetivo General: Analizar el impacto de la IA Documental y los Sistemas Multiagentes en la transformación digital de las organizaciones.

En base a esta premisa, se definen los siguientes objetivos específicos

- O1: Identificar áreas críticas donde la IA Documental pueda optimizar la gestión de información.

- O2: Evaluar la mejora en la colaboración y automatización de procesos mediante Sistemas Multiagentes.
- O3: Investigar y documentar mejores prácticas y metodologías para la integración de estas tecnologías.
- O4: Medir el impacto en la eficiencia operativa y la innovación mediante estudios comparativos.

Metodología y Diseño Experimental: El estudio se estructura en tres fases:

- Identificación de áreas de oportunidad: Análisis documental y entrevistas para detectar cuellos de botella en procesos actuales.
- Implementación y evaluación: Desarrollo de prototipos. Para la IA Documental, se aplicarán técnicas de NLP para extracción de entidades. Para los Sistemas Multiagentes, se diseñarán arquitecturas personalizadas basadas en marcos como AutoGen de Microsoft.
- Análisis comparativo: Medición de impacto mediante métricas cuantitativas (tiempo de procesamiento, precisión) y cualitativas (satisfacción del usuario).

3. RESULTADOS OBTENIDOS/ ESPERADOS

- o Optimización de Tareas: Automatización de procesos administrativos repetitivos tanto en el ámbito universitario como público.
- o Reducción de Errores: Mejora en la exactitud del manejo de datos y documentos mediante modelos de IA.
- o Marco de Referencia: Creación de una guía de mejores prácticas y metodologías para futuras implementaciones de IA en otros contextos administrativos.
- o Personalización de Servicios: Implementación de asistentes virtuales 24/7 que mejoren la experiencia del usuario y la accesibilidad.

Sostenibilidad: Reducción del consumo de recursos físicos (papel, energía) a través de la eficiencia operativa.

Publicaciones Científicas: Diseminación de los hallazgos en revistas y congresos especializados nacionales e internacionales.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto servirá para construir capacidades humanas a distintos niveles:

Posgrado: Formación de becarios en el marco de maestrías y doctorados de la especialidad.

Grado: Participación de alumnos avanzados de la Licenciatura en Sistemas mediante prácticas profesionales y trabajos finales de carrera.

Sector Público: Capacitación de funcionarios y empleados públicos en el uso y gestión de las herramientas de IA desarrolladas.

5. BIBLIOGRAFÍA

[1] Gabryelczyk, R. (2020). Has COVID-19 Accelerated Digital Transformation? Initial Lessons Learned for Public Administrations. Information Systems Management, 37(4), 303-309. doi:10.1080/10580530.2020.1820633

[2] Lei, C., Yiheng, X., Tengchao, L., & Furu, W. (2021). Document AI: Benchmarks, Models and Applications. En arXiv [cs.CL]. http://arxiv.org/abs/2111.08609

[3] Wu, Q., et al. (2023). AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversations. Microsoft Research. Corvalán, J. G. (2018).

[3] Wu, Q., Bansal, G., & Zhang, J. (s/f). AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversations. Microsoft.com. Recuperado el 1 de noviembre de 2024.

[4] Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia.

[5] Britos, Proyecto de I&D: "Explotación y visualización de información. Su aplicación a las industrias rionegrinas", 2017-2020),

[6] Britos & Vivas, Proyecto de I&D: "Estudio y Evaluación de Tecnologías de la Información y la Comunicación para el desarrollo de ciudades inteligentes en Río Negro", 2015-2017).

INTELIGENCIA COMPUTACIONAL EN CIUDADES INTELIGENTES: MODELO PRÁCTICO APLICANDO TECNOLOGÍAS CONVERGENTES Y PREDICCIÓN DE EROSIÓN COSTERA

Molina D.[1], Valdez J.[1], Villagra A.[1], Pandolfi D.[1], Leguizamón G.[1,2]

Laboratorio de Tecnologías Emergentes (LabTEm)
Instituto de Tecnología Aplicada (Caleta Olivia) – Unidad Académica Caleta Olivia
Universidad Nacional de la Patagonia Austral [1]
{damolina, jcvaldez, avillagra, dpandolfi}@uaco.unpa.edu.ar; legui@unsl.edu.ar

Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) Departamento de Informática - Universidad Nacional de San Luis [2]
legui@unsl.edu.ar

RESUMEN

Este artículo presenta dos líneas de investigación orientadas a la aplicación de inteligencia computacional en problemáticas urbanas y ambientales de la región patagónica. La primera propone un modelo práctico que combina una arquitectura para ciudades inteligentes y tecnologías convergentes —*IoT*, *Edge Computing*, *Big Data* e inteligencia computacional— para estimar el impacto de su implementación en ámbitos urbanos, guiado por los indicadores de rendimiento (KPI) propuestos por la Unión Internacional de Comunicaciones (UIT). La segunda desarrolla un sistema predictivo de erosión costera que integra imágenes satelitales SAR de la misión SAOCOM (CONAE) con datos oceanográficos nacionales mediante arquitecturas híbridas de redes neuronales convolucionales y memoria a largo-corto plazo (CNN-LSTM), orientado a anticipar cambios morfológicos críticos en las costas patagónicas, especialmente en el Golfo San Jorge. Ambas líneas convergen en el uso de la inteligencia computacional como herramienta central para la gestión inteligente y sustentable.

***Palabras clave*:** Ciudades inteligentes, inteligencia computacional, ecosistema tecnológico convergente, erosión costera, SAOCOM, aprendizaje profundo, CNN-LSTM.

CONTEXTO

Este trabajo se desarrolla en el LabTEm-ITA de la Unidad Académica Caleta Olivia (UNPA), en el marco del Proyecto 29/B304 "Enfoques inteligentes para ciudades del futuro: Empleo, Movilidad y Ambiente Sostenible", en colaboración con el LIDIC de la Universidad Nacional de San Luis. Las dos líneas responden a desafíos concretos de la región: cómo guiar a ciudades intermedias en su transformación hacia modelos inteligentes y sustentables, y cómo anticipar la erosión costera que avanza sobre localidades del Golfo San Jorge. En ambos casos, la inteligencia computacional es la apuesta tecnológica central.

1. INTRODUCCIÓN

Las ciudades inteligentes han dejado de ser una promesa futura para convertirse en una agenda urgente. Hoy integran tecnologías

avanzadas y analítica de datos en dominios tan diversos como el transporte, la energía, la salud y la gobernanza, con el objetivo de mejorar la calidad de vida urbana y promover un desarrollo sostenible (Gracias et al., 2023). El concepto ha madurado: ya no alcanza con incorporar tecnología, sino que se exige también resiliencia, inclusión y participación ciudadana real.

Gestionar ciudades en crecimiento acelerado es cada vez más complejo. Para abordarlo de forma sistemática, se recurre a Arquitecturas de Referencia que permiten estructurar soluciones modulares e interoperables (Bastidas et al., 2023). En este marco, la Unión Internacional de Comunicaciones (UIT) ofrece una arquitectura estándar y una metodología basada en Indicadores Clave de Rendimiento (KPI) que guía el proceso de conversión hacia ciudades inteligentes (ITU, 2016).

Por otro lado, la erosión costera en el Golfo San Jorge representa una amenaza concreta y creciente sobre infraestructuras y comunidades patagónicas. El monitoreo tradicional resulta insuficiente en cobertura y costo, lo que abre paso al uso de imágenes SAR de la misión SAOCOM (CONAE) combinadas con modelos de aprendizaje profundo, dentro de una agenda de soberanía tecnológica nacional (Bartsch et al., 2020; Gomez-de la Peña et al., 2023).

Este artículo describe ambas líneas activas: el modelo práctico de ciudades inteligentes para estimar el impacto de la inteligencia computacional en ámbitos urbanos, y el sistema predictivo de erosión costera basado en arquitecturas híbridas de redes neuronales convolucionales y memoria a largo-corto plazo (CNN-LSTM). Las dos apuestan por la inteligencia computacional como motor de transformación territorial en la Patagonia.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se describen brevemente las líneas de investigación que se llevan a cabo en el proyecto.

2.1. Modelo práctico aplicando una arquitectura para ciudades inteligentes y tecnologías convergentes para estimar el impacto de la implementación de inteligencia computacional en ámbitos urbanos

El objetivo de esta línea es desarrollar un modelo práctico que colabore con el proceso de conversión de una ciudad en inteligente y sostenible, aplicando técnicas de inteligencia computacional en un ecosistema tecnológico convergente. Un punto de partida para cualquier ciudad inteligente es administrar y utilizar los datos que recopila de manera efectiva. La interoperabilidad es uno de los principales desafíos en este proceso, y la clave para lograrlo radica en definir soluciones TIC modulares y escalables de múltiples capas que permitan la interoperabilidad de dominio cruzado, superando las soluciones aisladas que abordan desafíos específicos de forma independiente (IEEE Standards Association, 2020).

Para lograr este objetivo, se combinan tecnologías que trabajan en conjunto conformando un ecosistema convergente: (i) tecnologías para la recolección de datos, que permiten que miles de millones de dispositivos inteligentes se interconectan de forma autónoma soportando servicios de comunicación masiva (ii) tecnologías para la transmisión de datos, mediante la convergencia de IoT, Edge Computing y redes 5G (Insuasti et al., 2025); (iii) tecnologías para el

almacenamiento, con una arquitectura multi-cliente que soporta numerosas redes de sensores IoT como fuentes de datos; y (iv) tecnologías para el análisis y manipulación de datos, que optimizan la eficiencia de los servicios a partir del análisis de los datos generados por sensores instalados en el espacio urbano (Guadarrama-Sánchez H., 2025).

La arquitectura adoptada sigue el estándar de la UIT (U4SSC, 2022), que organiza la infraestructura TIC en cuatro capas: sensora (dispositivos IoT que capturan datos urbanos), de red (transmisión hacia servidores), de datos y soporte (almacenamiento y análisis con IA), y de aplicación (soluciones que habilitan los servicios de ciudad inteligente). Los resultados esperados incluyen la selección de técnicas de inteligencia computacional guiadas por los KPI de la UIT, el desarrollo de módulos por capa, la definición de interfaces entre datos y técnicas de resolución, y la generación de información que apoye decisiones sobre los costos del proceso de conversión.

2.2 Predicción de erosión costera mediante aprendizaje profundo e imágenes SAOCOM

Esta línea busca desarrollar un sistema de predicción de erosión costera basado en inteligencia artificial que permita anticipar cambios morfológicos en zonas vulnerables mediante el análisis de series temporales de imágenes satelitales y datos oceanográficos. El problema central radica en la degradación progresiva de las costas patagónicas, especialmente en el Golfo San Jorge, donde ciudades como Comodoro Rivadavia, Rada Tilly y Caleta Olivia enfrentan retroceso acelerado de la línea costera que amenaza infraestructura crítica, áreas turísticas y ecosistemas naturales (Toimil et al., 2020).

Las imágenes de radar de apertura sintética SAR en banda L de la misión SAOCOM (CONAE) ofrecen capacidades únicas para el monitoreo costero: penetración atmosférica total, independencia de las condiciones de iluminación y sensibilidad a variaciones milimétricas de la superficie terrestre (Bartsch et al., 2020). La banda L presenta ventajas específicas para aplicaciones costeras al reducir la interferencia del oleaje y permitir análisis interferométricos de alta precisión (Darwish et al., 2021). Complementariamente, los datos polarimétricos de ALOS-2/PALSAR-2 han demostrado eficacia en la detección de línea costera mediante análisis de mecanismos de retrodispersión (Asaka et al., 2020).

Los modelos de aprendizaje profundo, particularmente las arquitecturas híbridas CNN-LSTM, han emergido como herramientas prometedoras para la predicción de cambios costeros (Gomez-de la Peña et al., 2023). Estas arquitecturas capturan simultáneamente patrones espaciales mediante CNNs y dependencias temporales mediante LSTMs, con desempeño superior a los modelos tradicionales de equilibrio. La optimización bayesiana de hiperparámetros y la descomposición estacional de series temporales mejoran significativamente el rendimiento predictivo de estos modelos (Manamperi et al., 2025).

El objetivo principal consiste en desarrollar un sistema predictivo que integre imágenes SAOCOM en modo StripMap con datos oceanográficos nacionales (AUSTRAL-WWIII, Servicio de Hidrografía Naval y Servicio Meteorológico Nacional) mediante arquitecturas CNN-LSTM. Se propone el uso de productos de nivel L1A (Single Look Complex) que preservan

información de fase y amplitud, esenciales para técnicas interferométricas. El sistema procesa series temporales de polarización dual (VV-VH) para extraer características espaciales de la morfología costera, mientras que los datos oceanográficos (altura de ola, dirección de viento, fase de marea) alimentarán el componente LSTM. La validación se ejecutará en el sector central del Golfo San Jorge (125 km entre Rocas Coloradas y Las Golondrinas), que incluye tres ciudades costeras con erosión activa documentada y diversidad geomorfológica representativa (Vitousek et al., 2024).

3. RESULTADOS OBTENIDOS/ESPERADOS

Para la línea de ciudades inteligentes, se espera obtener un modelo de arquitectura tecnológica convergente validado, adoptable como base para el diseño de soluciones de inteligencia computacional en ciudades intermedias patagónicas. El modelo dará soporte a servicios básicos —transporte, agua, energía, residuos— bajo un ecosistema tecnológico integrado que permita compartir datos entre dominios de forma abierta y en tiempo real.

Para la línea de erosión costera, se espera que el sistema CNN-LSTM supere el 80% de precisión en la predicción de retroceso costero con horizontes de 3 a 12 meses, con un error de detección inferior a 5 metros. Más importante aún, el sistema permitirá anticipar episodios críticos con al menos 60 días de antelación, información clave para la planificación de medidas de protección. Su implementación íntegramente con tecnología satelital argentina (SAOCOM) y datos oceanográficos nacionales establece un precedente de soberanía tecnológica replicable en toda la costa patagónica.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo reúne investigadores del LabTEm (UNPA) y del LIDIC (UNSL), con formación doctoral, maestrandos y estudiantes avanzados de grado. La primera línea sostiene el desarrollo de una tesis doctoral en arquitecturas de ciudades inteligentes y tecnologías convergentes; la segunda, una tesis de maestría en predicción de erosión costera con aprendizaje profundo e imágenes SAOCOM. Ambas líneas constituyen un espacio concreto de formación de posgrado para los integrantes del equipo.

5. BIBLIOGRAFÍA

Asaka, T., Nonaka, T., Sugimura, T., & Wakabayashi, H. (2020). Full polarimetric ALOS-2/PALSAR-2 analysis of backscattering characteristics from various types of coasts. International Journal of Remote Sensing, 41(14), 5338–5354. https://doi.org/10.1080/01431161.2020.1731928

Bartsch, A., Ley, S., Nitze, I., Pointner, G., & Vieira, G. (2020). Feasibility study for the application of synthetic aperture radar for coastal erosion rate quantification across the Arctic. Frontiers in Environmental Science, 8, Article 143. https://doi.org/10.3389/fenvs.2020.00143

Bastidas, V., Bezbradica, M., Bilauca, M., Healy, M., & Helfert, M. (2023). Enterprise Architecture in Smart Cities: Developing an Empirical Grounded Research Agenda. *Journal of Urban Technology*, *30*(1), 47-70.

Darwish, N., Kaiser, M., Koch, M., & Gaber, A. (2021). Assessing the accuracy of ALOS/PALSAR-2 and Sentinel-1 radar

images in estimating the land subsidence of coastal areas: A case study in Alexandria City, Egypt. Remote Sensing, 13(9), Article 1838. https://doi.org/10.3390/rs13091838

Gomez-de la Peña, E., Coco, G., Whittaker, C., & Montaño, J. (2023). On the use of convolutional deep learning to predict shoreline change. Earth Surface Dynamics, 11, 1145–1160. https://doi.org/10.5194/esurf-11-1145-2023

Gracias, J. S., Parnell, G. S., Specking, E., Pohl, E. A., & Buchanan, R. (2023). Smart cities—A structured literature review. Smart cities, 6(4), 1719-1743.

Guadarrama-Sánchez, H. A. (2025). Las ciudades inteligentes: meta-sistemas de información, desde una aproximación. Revista gestión de las personas y tecnología, 18(53, p. 109-133.)

IEEE Standards Association. P2784—Guide for the Technology and Process Framework for Planning a Smart City. Available online: https://standards.ieee.org/project/2784.html (accessed on 7 August 2020).

Insuasti, J. J. P., Andino, V. M. F., & Andino, V. M. F.(2025) Infraestructura de Comunicaciones para Ciudades Inteligentes: Integración de IoT, Edge Computing y Redes 5G.

ITU. (2016). Recommendation Y.4900/L.1600: Overview of key performance indicators in smart sustainable cities. https://www.itu.int/rec/T-REC-Y.4900/en

Manamperi, T., Rahat, A., Pender, D., Cristau-do, D., Lamb, R., & Karunarathna, H. (2025). Predicting shoreline changes using deep lear-ning techniques with Bayesian optimisation. Coastal Engineering, 203, Article 104856. https://doi.org/10.1016/j.coastaleng.2025.104856

Toimil, A., Camus, P., Losada, I. J., Le Cozan-net, G., Nicholls, R. J., Idier, D., & Maspataud, A. (2020). Climate change-driven coastal erosion modelling in temperate sandy beaches: Methods and uncertainty treatment. Earth-Science Reviews, 202, Article 103110. https://doi.org/10.1016/j.earscirev.2020.103110

U4SSC. (2022). *Redefining smart city platforms: Setting the stage for Minimal Interoperability Mechanisms*. https://www.itu.int/en/publications/Documents/tsb/2022-U4SSC-Redefining-smart-cityplatforms/index.html

Vitousek, S., Vos, K., Splinter, K. D., Parker, K., O'Neill, A., Foxgrover, A. C., Hayden, M. K., Thomas, J. A., Erikson, L., & Barnard, P. L. (2024). Scalable, data-assimilated models predict large-scale shoreline response to waves and sea-level rise. Scientific Reports, 14, Article 28029. https://doi.org/10.1038/s41598-024-77030-4

INTELIGENCIA ORGANIZACIONAL PARA GESTIONAR ESTRATÉGICAMENTE LAS DECISIONES Y LA TRANSFORMACIÓN DIGITAL EN ORGANIZACIONES REGIONALES

Romagnano, M.[1,2,3], Herrera, M.[1,3], Sánchez, G[2], García, A.[2], Aballay, A.[1], Sarmiento, A.[1], Terluk, A.[2], Benemerito, I.[2], Sosa, M.[3], Martinez, C.[3]

[1]Departamento de Informática, FCEFN, Universidad Nacional de San Juan
[2]Instituto de Informática, FCEFN, Universidad Nacional de San Juan
[3]Departamento de Matemática, Física y Química, FCEFN, Universidad Nacional de San Juan

{maritaroma, mherrera}@iinfo.unsj.edu.ar,{gerasanchezcorzo, garcia.exe, prof.alicia, adrivalle2005, aldana41292, ignacio.benemerito, magamsosa, ceciliaevelyn99}@gmail.com

RESUMEN

En el marco de la transformación digital, los datos se han consolidado como un activo estratégico fundamental para la generación de ventajas competitivas. No obstante, diversas organizaciones regionales aún carecen de las estructuras analíticas necesarias para convertir dicha información en decisiones inteligentes.

Bajo este escenario, la inteligencia organizacional emerge como la capacidad crítica para transformar grandes volúmenes de datos en acciones informadas.

El presente trabajo propone el diseño de un Modelo de Inteligencia Organizacional orientado a fortalecer la toma de decisiones y la madurez digital en entidades regionales, integrando tres pilares técnico-metodológicos:

- *Inteligencia de Negocios*: para proporcionar la infraestructura de análisis descriptivo y diagnóstico.
- *Inteligencia Artificial*: mediante modelos de Machine Learning para el análisis predictivo.
- *Estadística Avanzada*: aplicando la Teoría de la Decisión para el análisis prescriptivo y la gestión formal del riesgo.

Por lo tanto, el modelo propuesto promoverá procesos de aprendizaje organizacional que sustenten una transformación digital basada en la evidencia. Al integrar estas tecnologías emergentes, las organizaciones regionales pueden transitar desde la mera acumulación de datos hacia la generación de conocimiento para la toma de decisiones oportunas y acertadas.

Palabras Claves: Gestión del Conocimiento, Madurez Digital, Sistemas de Soporte a la Decisión, Inteligencia de Negocios, Aprendizaje Automático, Teoría de la Decisión.

CONTEXTO

La digitalización y el uso inteligente de los datos representan hoy una condición indispensable para la sostenibilidad de las organizaciones. La innovación tecnológica y la convergencia de enfoques disruptivos, tales como Big Data (BD), Internet de las Cosas (IoT), Inteligencia Artificial (IA), Inteligencia Artificial Generativa (IAG), Machine Learning (ML), Blockchain, Gemelos Digitales, entre otros; en conjunto con herramientas estadísticas permiten construir modelos predictivos y prescriptivos que potencian la capacidad de decisión y el aprendizaje organizacional. No obstante, en la mayoría de las organizaciones regionales persisten limitaciones en la gestión del conocimiento y en la adopción de metodologías de analíticas avanzadas. Así, muchas organizaciones se encuentran en etapas incipientes de

digitalización. Si bien han incorporado sistemas de información y plataformas de gestión, aún no logran convertir los datos en conocimiento accionable que oriente la toma de decisiones estratégicas. Esta brecha entre la disponibilidad de información y su aprovechamiento efectivo limita la innovación y la adaptación a entornos cambiantes.

Por otra parte, la analítica de datos, en conjunto con la Estadística Avanzada (EA), ofrece una oportunidad concreta para superar estas limitaciones. La analítica permite descubrir patrones, tendencias y relaciones ocultas en grandes volúmenes de información, mientras que la EA proporciona los fundamentos científicos necesarios para validar los resultados y reducir la incertidumbre. En conjunto, ambas disciplinas posibilitan la toma de decisiones basadas en evidencia, eje central de la Inteligencia Organizacional (IO).

En este escenario de convergencia tecnológica y metodológica, se presenta este trabajo, enmarcado en el proyecto "Modelo de Inteligencia Organizacional para Asistir a la Toma de Decisiones y Promover la Transformación Digital en Organizaciones Regionales", presentado en la convocatoria del Consejo de Investigaciones Científicas y Técnicas y de Creación Artística (CICITCA) de la Universidad Nacional de San Juan (UNSJ), para ser desarrollado durante el período comprendido entre 01/01/2026 al 31/12/2027.

Las tareas de investigación se desarrollan en los laboratorios de Sistemas de Información y de Matemática Aplicada, en el ámbito del Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.

Esta línea de investigación busca no sólo generar saberes científicos, sino también ofrecer una herramienta de gestión práctica y transferible a organizaciones regionales. Con ello se pretende fortalecer sus capacidades analíticas y decisorias, transformando datos en conocimiento y ese conocimiento en acciones estratégicas, a la vez que se consolida una cultura de inteligencia organizacional.

Asimismo, la pertinencia regional del proyecto radica en la posibilidad de contribuir al desarrollo de un ecosistema organizacional inteligente y sostenible, donde las decisiones se fundamenten en un análisis riguroso y no únicamente en la experiencia o la intuición.

La aplicación del modelo propuesto en diferentes tipos de organizaciones, públicas o privadas, permitirá identificar el nivel de desarrollo digital y analítico de sus sistemas de gestión, y formular líneas de mejora acordes con las particularidades de cada institución, fortaleciendo con ello la capacidad de actuar de manera más consistente y fundamentada.

1. INTRODUCCIÓN

Los avances tecnológicos fuerzan a las organizaciones a mantenerse continuamente actualizadas. Así, la Transformación Digital implica la adopción de tecnologías avanzadas y la modificación de procesos humanos que permiten reestructurar los modelos de gestión y cultura. Este enfoque busca no solo integrar nuevas herramientas tecnológicas, sino también fomentar un cambio en la mentalidad y las prácticas de trabajo dentro de las organizaciones (Mahboub y Sadok, 2023).

Si bien los conceptos de Inteligencia Organizacional y Organizaciones Inteligentes no son nuevos, estos resurgen como ejes claves en la era digital. La Inteligencia Organizacional, es definida por Rodríguez-Cruz (2015) como la capacidad de una organización para aprender y adaptarse mediante la generación y gestión del conocimiento. Entonces, las Organizaciones Inteligentes se diferencian de las convencionales en cuanto a que tienen una visión integral, se apoyan en tecnologías, se basan en datos, valoran el feedback, y retienen los talentos. Por ello, el factor determinante en una organización de este tipo es la tecnología y la capacidad de digitalizar o directamente automatizar sus procesos. Además, deben contar con una infraestructura que facilite la recolección, recopilación y análisis de datos (Senge, 1992). Por lo tanto, estas organizaciones trabajan bajo la concepción de

Paradigma 4.0 al introducir una nueva forma de producir con tecnologías de última generación, Tecnologías 4.0, para automatizar sus procesos productivos y transformar los modelos de negocio, optimizando los recursos (Ministerio de Desarrollo Productivo, 2021). Complementariamente, las Tecnologías 5.0 son consideradas como un nuevo motor de innovación para la generación de valor que hace más eficiente y sostenible la producción (Xu et al., 2021). Estas tecnologías representan una evolución significativa en los procesos de producción industrial al integrar tecnologías avanzadas como IoT, IA y la Robótica, enfatizando la colaboración estrecha entre seres humanos y máquinas (Domínguez, 2024). El BD, junto a la IA y el ML son los principales mecanismos que entran en acción en estas organizaciones (Romagnano et al., 2023). Además, actualmente, la mayoría de las organizaciones inteligentes están adoptando la IAG como una tecnología emergente que impacta potencialmente en sus Sistemas de Información, posibilitando nuevas formas de tomar decisiones, de los flujos de trabajo y del diseño organizacional (Storey et al., 2025). Asimismo, la Inteligencia de Negocios (BI) resulta ser un enfoque útil para identificar patrones, tendencias y evaluar el desempeño que tienen las organizaciones. Se encuentra muy relacionada con la analítica, desde una perspectiva predictiva y prescriptiva (Sarango et al., 2025).

La analítica predictiva, mediante técnicas de EA, permite a las organizaciones anticipar futuros escenarios y tomar decisiones informadas (Adesina et al., 2024). La combinación de analítica descriptiva, predictiva y prescriptiva constituye un esquema integral para trabajar en ambientes dinámicos y complejos (Abdul-Yekeen y Ekezie, 2024). A través de métodos prescriptivos, se pueden obtener recomendaciones concretas para tomar estratégicas decisiones, sustentadas en la evidencia (Moesmann y Pedersen, 2024).

Por ello, la convergencia de estos enfoques y mecanismos contribuye al desarrollo de organizaciones inteligentes, adaptativas y orientadas a la mejora continua. Bajo esta premisa, el presente estudio propone diseñar un modelo de inteligencia organizacional que asista de manera estratégica la toma de decisiones y promueva una transformación digital efectiva en organizaciones regionales.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Dada la naturaleza multidimensional de la Inteligencia Organizacional, el equipo de investigación se ha conformado de manera interdisciplinaria. Sus integrantes aportan visiones complementarias desde las Ciencias de la Computación, los Sistemas de Información, la Estadística y la Administración de Empresas.

Los investigadores poseen amplia trayectoria en líneas vinculadas a BI, Analítica y Estadística Aplicada, Sistemas de Información y Recuperación de Información.

El proyecto se estructura a partir de la sinergia de subgrupos de trabajo especializados en las siguientes líneas:

- *Sistemas de información organizacionales*: enfocados en la arquitectura y soporte de los flujos de datos.
- *Transformación digital y tecnologías emergentes*: orientada a la evolución de los modelos de.
- *Inteligencia Artificial aplicada*: implementación de modelos de Machine Learning e IAG en procesos industriales y de negocio.
- *Inteligencia y analítica de negocios*: diseño de infraestructuras para el análisis descriptivo y diagnóstico.
- *Análisis y procesamiento estadístico*: aplicación de EA para la validación científica.

3. RESULTADOS ESPERADOS

Desde una perspectiva trascendental, esta iniciativa se alinea con los Lineamientos

Estratégicos en Innovación, Ciencia y Tecnología para el período 2025-2027 (InfoLEG, 2025), complementando los objetivos del Plan Nacional de Ciencia, Tecnología e Innovación 2030 (InfoLEG, 2023).

Bajo este marco, se prevé la obtención de las siguientes contribuciones:

- *Modelo de Inteligencia Organizacional*: un marco de trabajo validado y adaptable a la diversidad de organizaciones regionales.
- *Instrumentos de Medición*: indicadores y métricas específicas para evaluar el nivel de madurez digital.
- *Aportes Metodológicos*: guías para la gestión de decisiones basadas en evidencia.
- *Producción Científica*: artículos en revistas especializadas, ponencias en congresos y actividades de capacitación para la transferencia de conocimiento.

Al adoptar un diseño multicaso, la validación se llevará a cabo con la participación de organizaciones públicas y privadas de la región cuyana, facilitando la comparación y transferencia efectiva de resultados.

Por lo tanto, los resultados del proyecto tendrán una transferencia inmediata tanto a la comunidad científica (mediante la difusión de metodologías y datos anonimizados) como al ámbito organizacional. Se busca impactar especialmente en los equipos directivos y responsables de planificación estratégica, contribuyendo al fortalecimiento institucional y a la consolidación de una cultura de toma de decisiones fundamentada en la evidencia.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación integra a estudiantes que actualmente se encuentran desarrollando sus tesis o trabajo final en temáticas directamente vinculadas a los objetivos del proyecto. Los integrantes provienen de los siguientes programas académicos:

- *Nivel de Grado y Posgrado*:
  - Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información (FCEFN-UNSJ).
  - Especialización en Docencia Universitaria (UNSJ).
  - Maestría en Informática (UNSJ).
  - Maestría en Matemática (UNSL), Maestría en Estadística Aplicada (UNC) y Maestría en Estadística Matemática (UBA).
  - Doctorado en Astronomía (FCEFN-UNSJ).

- *Becas de Investigación:*

Asimismo, el equipo cuenta con becarios cuyas líneas de trabajo se desarrollan bajo el marco de este proyecto, pertenecientes a las siguientes convocatorias:

- Becas Estímulo a las Vocaciones Científicas (EVC-CIN): convocatoria 2024.
- Becas CICITCA-UNSJ: convocatoria 2025.
- Programa de Intercambio Académico Latinoamericano (PILA): convocatoria 2025.

En cuanto a las expectativas de formación, se espera fortalecer las capacidades del equipo en el diseño de soluciones analíticas e inteligentes fundamentadas en datos, modelos estadísticos y tecnologías emergentes. De este modo, se promueve la toma de decisiones basada en evidencia y se garantiza la transferencia de conocimiento hacia los sectores productivo, académico y gubernamental.

5. BIBLIOGRAFÍA

Abdul-Yekeen, A. M., & Ekezie, M. N. (2024). Business analytics and decision science: A review of techniques in strategic business decision-making. Research Inventy: International Journal of Engineering and Science, 14(7), 17–25. Recuperado de: https://www.researchinventy.com/papers/v14i7/C14071725.pdf, octubre de 2025.

Adesina, A. A., Iyelolu, T. V., y Paul, P. O. (2024). Leveraging predictive analytics for strategic decision-making: Enhancing business performance through data-driven insights. World Journal of Advanced Research and Reviews, 22(3), 1927–1934. Recuperado de https://doi.org/10.30574/wjarr.2024.22.3.1961, octubre de 2025.

Domínguez, L. A. P. (2024). Las Principales Tecnologías de la Era de la Industria 5.0. Revista Ingenio, 21(1), 60-70. Recuperado de: https://dialnet.unirioja.es/servlet/articulo?codigo=9600706, noviembre de 2025.

InfoLEG (2023). Plan Nacional de Ciencia, Tecnología e Innovación 2030. Recuperado de: https://servicios.infoleg.gob.ar/infolegInternet/verNorma.do?id=391776, noviembre de 2025.

InfoLEG (2025). Lineamientos Estratégicos en Innovación, Ciencia y Tecnología (2025-2027). Recuperado de: https://servicios.infoleg.gob.ar/infolegInternet/anexos/415000-419999/419597/norma.htm, noviembre de 2025.

Mahboub, H. y Sadok, H. (2023). Implementing enterprise digital transformation: a contribution to conceptual framework design. Nankai Business Review International. Recuperado de: https://doi.org/10.1108/NBRI-06-2022-0067, octubre de 2025.

Ministerio de Desarrollo Productivo (2021). Plan de Desarrollo Productivo Argentina 4.0. Recuperado de: https://www.argentina.gob.ar/sites/default/files/plan_de_desarrollo_productivo_argentina_4.0.vf__2.pdf, noviembre de 2025.

Moesmann, M., y Pedersen, T. B. (2024). Data-driven prescriptive analytics applications: A comprehensive survey. Information Systems. (In press). Recuperado de: https://doi.org/10.1016/j.is.2025.102576, diciembre de 2025.

Rodríguez-Cruz, Y. (2015). La inteligencia Organizacional Bajo un Enfoque Integrador de Gestión Documental, de Información y del Conocimiento. In: VALENTIM, M. L. P., and MÁS-BASNUEVO, A., eds. Inteligência organizacional [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2015, pp. 341-372. ISBN: 978-65-5954-131-7. Recuperado de: https://books.scielo.org/id/4rwps/pdf/valentim-9786559541317-16.pdf, noviembre de 2025.

Romagnano, M., Ganga, L., Pantano, J. C., Herrera, M., Lépez, H., Becerra, M., Aballay, A., Sarmiento, A., Jason, S., Furlani, F., y García, A. (2023). Organizaciones inteligentes asistidas por Tecnologías 4.0 Recuperado de: https://sedici.unlp.edu.ar/handle/10915/163223, noviembre de 2025.

Sarango, A. F. H., Palpa, M. J. B., Romero, M. Á. M., Carranza, E. A. G., y Ronquillo, M. B. B. (2025). Inteligencia de Negocios: Principios Fundamentales y Aplicaciones Empresariales. Know Press.

Senge, P. (1992). La Quinta Disciplina. El arte y la práctica de la organización abierta al aprendizaje. Editorial GRANICA – Barcelona 1992. Recuperado de: https://rionegro.gov.ar/download/archivos/00002739.pdf, octubre de 2025.

Storey, V. C., Yue, W. T., Zhao, J. L., y Lukyanenko, R. (2025). Generative artificial intelligence: Evolving technology, growing societal impact, and opportunities for information systems research. Information Systems Frontiers. Recuperado de: https://doi.org/10.1007/s10796-025-10581-7, diciembre de 2025.

Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. Journal of Manufacturing Systems, 61, 530-535. Recuperado de: https://doi.org/10.1016/j.jmsy.2021.10.006, diciembre de 2025.

Interpretabilidad y Métodos de Entrenamiento de Modelos de Deep Learning

Facundo Quiroga[1,2], Franco Ronchetti[1,2], Gastón Rios[1,3], Pedro Dal Bianco[1,3], Santiago Ponte Ahon[1], Oscar Stanchi[1,6], Juan Seery[1], Tatiana Badaracco[5], Federico Rabinovich[5], Saif Khalid[4], Hatem Rashwan[4], Domenec Puig Valls[4], Laura Lanzarini[1], Waldo Hasperué[1]

[1] Instituto de Investigación en Informática LIDI, Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina.*
[2] Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CICPBA)
[3] Becario postgrado UNLP
[4] Universitat Rovira i Virgili, Tarragona, España
[5] Universidad de Buenos Aires, Buenos Aires, Argentina
[6] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
*** Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)**

Contacto: fquiroga@lidi.info.unlp.edu.ar

CONTEXTO

Esta presentación corresponde a algunas de las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto "Inteligencia de Datos. Técnicas y Modelos de Machine Learning" perteneciente al Programa de Incentivos (2023-2026).

RESUMEN

Esta línea de investigación se centra en el estudio de modelos de Deep Learning para resolución de problemas de Visión por Computadora, incluyendo técnicas de Aprendizaje Automático clásicas, y Aprendizaje Profundo con Redes Neuronales Convolucionales (CNNs) y Vision Transformers (ViTs). En este contexto, los autores del presente trabajo investigan el funcionamiento de dichos modelos mediante varios enfoques, todos en el contexto de modelos de Visión por Computadora:

- Análisis de invarianza de CNNs y ViTs.
- Métodos de interpretabilidad por atribución, incluyendo RISE y GradCAM.
- Análisis de filtros convolucionales aprendidos. Técnicas no supervisadas para inicialización de filtros convolucionales.
- Entrenamiento con supervisión débil basado en técnicas de interpretabilidad, aplicado a la segmentación de mamografías.

Palabras clave: Deep Learning, Redes Neuronales, Redes Convolucionales, Vision Transformer, Visión por Computadoras, Interpretabilidad, Invarianza, GradCAM, Entrenamiento no supervisado, PCA, Visualización, supervisión débil, mamografías, fondo de ojos.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI (III-LIDI) tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos métodos de Aprendizaje Automático y Redes Neuronales.

Las redes neuronales son modelos tradicionalmente considerados como de caja negra. En los años recientes, se han realizado varios esfuerzos para comprender su funcionamiento de forma tal que el mismo sea más predecible o modulable.

Como resultado de estas investigaciones se han analizado, diseñado e implementado técnicas originales aplicables a la clasificación y el análisis de características de objetos en imágenes, generación de imágenes para aumentación de datos, y estudio del funcionamiento de las redes neuronales. En relación con esta línea, actualmente se están desarrollando los siguientes temas:

1.1 Modelos interpretables de predicción de calidad de imágenes de fondos de ojos.

La retinopatía diabética es una enfermedad aguda pero que puede tratarse con detección temprana. Para ello, sería de gran utilidad contar con modelos que permitan automatizar el análisis de imágenes de fondo de ojos. En particular, sería útil tener un modelo que permita determinar la calidad de una imagen para asistir la captura de las mismas en tiempo real, y además que sea interpretable para poder otorgar feedback al usuario. Por este motivo, se presentó un modelo basado en la CNN UNet de predicción de calidad de imágen que se denominó VISTA. El modelo contiene un módulo de Concept Whitening, una técnica de interpretabilidad intrínseca, que indica en base a qué conceptos la red realiza sus predicciones [3].

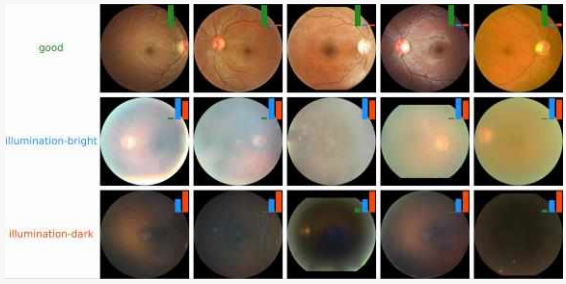


Figura 1. Ejemplo de conceptos aprendidos por el módulo Concept Whitening para el modelo VISTA

1.2 Invarianza de CNNs y ViT

La invarianza y equivarianza a las transformaciones son propiedades deseables en varios modelos de redes debido a que nos permiten razonar más fácilmente respecto a su funcionamiento. Por este motivo, en base a métricas previamente definidas por este grupo de investigación [4], se está trabajando en la caracterización de modelos de redes neuronales reconocidos.



Figura 2. Invarianza promedio de un modelo FasterViT variando la transformación aplicada en el entrenamiento (filas) y en la evaluación (columnas), para una estrategia de entrenamiento transferencia de aprendizaje.

Se evaluó la invarianza de los modelos EfficientNet y FasterViT [2,5] para comprender la distribución de su invarianza en sus distintas capas, y cómo distintas estrategias de entrenamiento (desde cero, transferencia de aprendizaje, o *finetuning*) pueden afectar dicha distribución (Figura 2).

1.3 Interpretabilidad de filtros convolucionales mediante entrenamiento no supervisado de CNNs y ViTs

Los filtros de una red entrenada con buen desempeño suelen usarse para realizar transferencia de aprendizaje. No obstante, su interpretabilidad depende del proceso de entrenamiento y requiere un análisis post-hoc. Por este motivo, se desarrolló un método de inicialización de filtros de modelos ViT y CNN basado en análisis de componentes principales [1,5]. Los filtros obtenidos permiten un desempeño igual o mejor que el de las redes entrenadas tradicionalmente con menos iteraciones de descenso de gradiente (Figura 3).

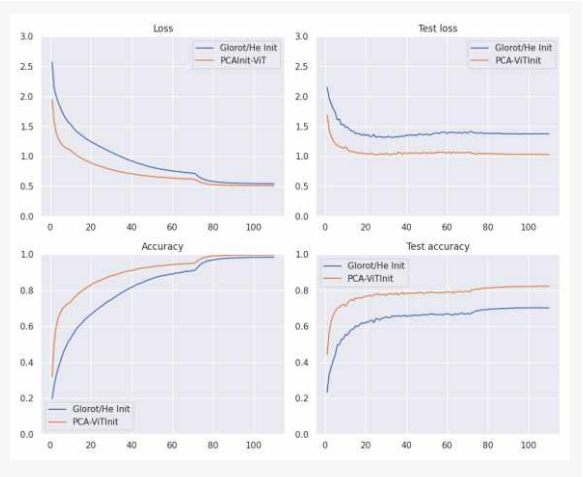


Figura 3. Los filtros de un modelo ViT inicializados mediante el método propuesto mejoran la velocidad de convergencia en algunos escenarios

1.4 Métodos con supervisión débil basados en técnicas de interpretabilidad para el entrenamiento de modelos de segmentación de mamografías.

La densidad mamaria es un biomarcador clínicamente relevante, aunque la segmentación del tejido denso a gran escala es limitada debido al alto costo y la ambigüedad de la anotación a nivel de píxel. Por este motivo, se está abordando la segmentación en mamografías mediante supervisión débil, utilizando etiquetas BI-RADS a nivel de imagen como señal principal de entrenamiento, generando máscaras ruidosas mediante GradCAM, un método de interpretabilidad y reservando máscaras expertas solo para evaluación (Figura 4).

Se construyó una cohorte específica a partir del dataset RSNA Screening Mammography Breast Cancer Detection y se evaluó un protocolo exigente: entrenar solo con categorías extremas (A vs. D) y probar con A vs. D* (donde D* = {B, C, D}) para analizar la robustez ante densidades mixtas. El método genera mapas de saliencia condicionados por clase y obtiene predicciones mediante *Top-t% pooling*, permitiendo convertir la saliencia en máscaras por umbralización y enmascaramiento anatómico. En un subconjunto de 100 imágenes anotadas por radiólogos, la mejor configuración (ResNet-101) alcanzó un Dice score de 0.726 ± 0.282 en el esquema A vs. D*, superando ampliamente a LIBRA (0.405 ± 0.292). Los resultados sugieren que la supervisión débil puede producir segmentaciones clínicamente plausibles, aunque con sensibilidad a la capacidad del modelo y a la variabilidad intrínseca de la densidad mamográfica.

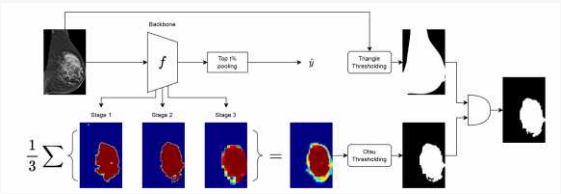


Figura 4. Método propuesto de generación de segmentaciones con etiquetas débiles. Se entrena un clasificador con etiquetas generales de clase (A vs D de la clasificación BI-RADS). Luego se generan máscaras mediante GradCAM, que se utilizan como etiquetas débiles para entrenar un modelo de segmentación.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Redes neuronales profundas, convolucionales y transformers.
- Herramientas de interpretabilidad para visión por computadora
- Interpretabilidad y entrenamiento no supervisado de Redes Convolucionales y ViT.
- Invarianza en redes neuronales, en particular de modelos ampliamente difundidos.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Interpretabilidad mediante invarianza de modelos EfficientNet y ViT
- Entrenamiento no supervisado de modelos EfficientNet y ViT para mejorar

desempeño e interpretabilidad.

- Validación mediante métodos de interpretabilidad de un modelo de predicción de calidad de imágenes de retinopatía diabética basado en UNet.
- Validación de un método de entrenamiento semi supervisado para modelos de segmentación de mamografías basado en técnicas de interpretabilidad.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 3 profesores con dedicación exclusiva, un JTP dedicación exclusiva, 2 investigadores CIC-PBA, 4 becarios de posgrado de la UNLP con dedicación docente, 2 tesistas de maestría de la UBA, 3 profesores extranjeros, y un pasante de grado.

Actualmente se están desarrollando 3 tesis de doctorado y 1 de máster. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

[1] Rabinovich, F., Stanchi, O, Ronchetti, F., Quiroga, F. Leveraging Data Geometry for Robust Vision Transformer Initialization. Workshop de Agentes y Sistemas Inteligentes, CACIC 2025, Universidad Nacional de Rosario

[2] Badaracco, T., Stanchi, O, Ronchetti, F., Quiroga, F. Representaciones Internas de Invarianza a Transformaciones Geométricas y de Color en EfficientNet y FasterViT. Workshop de Agentes y Sistemas Inteligentes, CACIC 2025, Universidad Nacional de Rosario

[3] Khalid, S., Abdulwahab, S., Stanchi, O. A., Quiroga, F. M., Ronchetti, F., Puig, D., & Rashwan, H. A. (2024). VISTA: vision improvement via split and reconstruct deep neural network for fundus image quality assessment. Neural Computing and Applications, 36(36), 23149-23168.

[4] Quiroga, F., Torrents-Barrena, J., Lanzarini, L., & Puig, D. Measuring (in) variances in Convolutional Networks. In Conference on Cloud Computing and Big Data (pp. 98-109). Springer, Cham. 2019.

[5] Rabinovich, F. Métodos de Entrenamiento basados en Agrupamiento para Capas Convolucionales en Redes Neuronales. Tesis de Maestría, UBA. Directores: Quiroga. F, Ronchetti, R.

[6] Badaracco, T. Métodos de Entrenamiento basados en Agrupamiento para Capas Convolucionales en Redes Neuronales. Tesis de Maestría, UBA. Director: Quiroga. F

Machine learning aplicado a astronomía

Franco Ronchetti[1,2], Facundo Quiroga[1,2], Santiago Ponte Ahon[1,3], Gastón Rios[1], Pedro Dal Bianco[1,3], Oscar Stanchi[1,6], Juan Seery[1], Yael Aidelman[4,5], Roberto Gamen[4,5], Laura Lanzarini[1], Waldo Hasperué[1]

[1] Instituto de Investigación en Informática LIDI, Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina.*
[2] Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CICPBA)
[3] Becario postgrado UNLP
[4] Instituto de Astrofísica de La Plata, CONICET, La Plata, Buenos Aires, Argentina
[5] Facultad de Ciencias Astronómicas y Geofísicas. Universidad Nacional de La Plata, La Plata, Argentina
[6] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
*** Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)**

Contacto: fronchetti@lidi.info.unlp.edu.ar

CONTEXTO

Esta presentación corresponde a algunas de las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto "Inteligencia de Datos. Técnicas y Modelos de Machine Learning" perteneciente al Programa de Incentivos (2023-2026).

RESUMEN

El patrimonio astronómico histórico constituye una fuente invaluable de información científica. Diversos observatorios del mundo conservan colecciones extensas de placas fotográficas y espectroscópicas obtenidas durante los siglos XIX y XX. Sin embargo, su digitalización y procesamiento automático representan un desafío técnico significativo, especialmente en el caso de placas espectroscópicas, cuya estructura interna requiere tareas complejas de segmentación, extracción y calibración [1].

Esta línea de investigación aborda el diseño e implementación de un sistema integral para la digitalización y procesamiento automático de placas espectroscópicas históricas. El trabajo integra técnicas de detección basadas en redes neuronales convolucionales, algoritmos de extracción de espectros unidimensionales y métodos de calibración automática en longitud de onda mediante alineamiento temporal (Dynamic Time Warping).

Como resultado se desarrolló PlateUNLP, una plataforma web de código abierto que permite gestionar proyectos de digitalización, automatizar la obtención de metadatos, segmentar observaciones y espectros, extraer señales espectrales y realizar su calibración automática. Los resultados experimentales muestran desempeños competitivos en detección y valores de IoU superiores a 0.9 en configuraciones óptimas de calibración, validando la viabilidad del enfoque propuesto [4, 6].

Palabras clave: Visión por computadora; Digitalización de patrimonio; Placas espectroscópicas; Deep Learning; Dynamic Time Warping; Procesamiento de señales..

1. INTRODUCCIÓN

El Instituto de Investigación en Informática LIDI (III-LIDI) tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos métodos de Aprendizaje Automático y Redes Neuronales. A continuación se mencionan los principales puntos investigados y desarrollados en esta línea de investigación.

1.1 Digitalización y puesta en valor del patrimonio astronómico

La digitalización de placas históricas no sólo garantiza su preservación, sino que habilita nuevas investigaciones científicas sobre datos obtenidos hace más de un siglo. A diferencia de las placas fotográficas tradicionales, las placas espectroscópicas contienen múltiples observaciones y espectros de ciencia y de lámparas de comparación en una misma imagen, lo que incrementa sustancialmente la complejidad de su procesamiento.

La mera digitalización no resulta suficiente: es necesario brindar herramientas para la extracción de información relevante de modo de permitir su reutilización en investigaciones actuales. Esta problemática requiere la integración de conocimientos de astronomía, informática y procesamiento digital de señales. En este sentido, se desarrolló PlateUNLP, plataforma web de código abierto que permite gestionar proyectos de digitalización, automatizar la obtención de metadatos, segmentar observaciones y espectros, extraer señales espectrales y realizar su calibración automática [7].

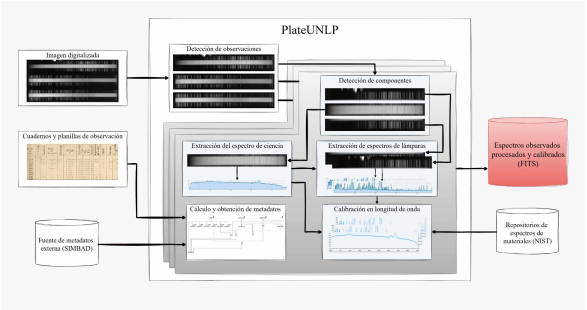


Figura 1. Diagrama general de PlateUNLP

1.2 Segmentación automática mediante visión por computadora

La segmentación de observaciones y espectros se modela como un problema de detección de objetos. Para ello se entrenaron modelos de aprendizaje profundo basados en YOLO capaces de identificar automáticamente las regiones correspondientes a cada observación y a los espectros contenidos en ellas. El desempeño de estos modelos se evaluó mediante métricas estándar como Precision, Recall, F-measure e Intersection over Union (IoU) [5, 2], obteniendo muy buenos resultados.

La automatización de estas tareas reduce drásticamente la intervención manual y permite escalar el procesamiento a grandes volúmenes de placas.

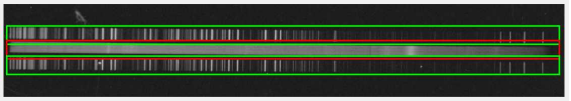


Figura 2. Ejemplo de placa digitalizada con un espectro de ciencia junto a sus lámparas de comparación.

1.3 Calibración automática en longitud de onda

Una vez extraídos los espectros unidimensionales, el desafío central consiste en asignar correctamente la longitud de onda a cada punto de la señal. Tradicionalmente, esta tarea se realiza de forma manual contrastando los espectros de lámparas de comparación con espectros de referencia.

En esta línea se propuso un método automático basado en Dynamic Time Warping (DTW), incorporando una estrategia de ventana deslizante sobre el espectro de referencia para resolver problemas de escala y alineación [3]. Este enfoque permite identificar automáticamente el segmento de referencia que mejor corresponde al espectro observado.



Figura 3. Ejemplo de calibración de lámpara

PlateUNLP consolida un flujo integral para el procesamiento de placas espectroscópicas históricas, articulando técnicas de visión por computadora y procesamiento de señales dentro de una única plataforma. La integración de modelos de detección basados en aprendizaje profundo con algoritmos de extracción y calibración automática permite reducir significativamente la intervención manual requerida en cada etapa del proceso.

De este modo, PlateUNLP se posiciona como una herramienta viable y escalable para la puesta en valor del patrimonio astronómico histórico, facilitando su reutilización en investigaciones astronómicas contemporáneas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Desarrollo de PlateUNLP. Software integral de procesamiento de placas.
- Detección automática de observaciones y espectros
- Extracción automática de espectros
- Calibración en longitud de onda

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollo de una plataforma integral de código abierto.
- Automatización completa del flujo de procesamiento.
- Validación experimental del método de calibración automática.
- Integración efectiva de los equipos de investigación de informática y astronomía en el marco del proyecto ReTrOH.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 3 profesores con dedicación exclusiva, un JTP dedicación exclusiva, 2 investigadores CIC-PBA, 4 becarios de posgrado de la UNLP con dedicación docente.

Actualmente se están desarrollando 1 tesis de doctorado y 1 de máster. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

[1] Meilán, N. S. Recuperación del patrimonio observacional histórico: Elaboración de un método óptimo de digitalización y extracción de datos. Tesis de Licenciatura, Universidad Nacional de La Plata, La Plata, Argentina. 2018.

[2] Ponte Ahón, S. A. Digitalización de placas astronómicas antiguas. Tesis de grado, Universidad Nacional de La Plata. 2023.

[3] Ponte Ahon, S. A., Seery, J. M., Quiroga, F., Ronchetti, F., Stanchi, O., Dal Bianco, P., Hasperué, W., Aidelman, Y., & Gamen, R. (2024). Wavelength Calibration of Historical Spectrographic Plates with Dynamic Time Warping. Conference on Cloud Computing, Big Data & Emerging Topics, 59-73.

[4] Ponte Ahon, S. A., Aidelman, Y., Seery, J. M., Quiroga, F. M., Ronchetti, F., Hasperué, W., Iannuzzi, M., Peralta, R., Lopez, M., Bariviera, A. F., Cidale, L., & Gamen, R. ReTrOH-UNLP: Conservation of the Historical Observational Work of the Astronomical Observatory of La Plata with Computer Vision. ACM Journal on Computing and Cultural Heritage. 2025.

[5] Quiroga, F. M., Ronchetti, F., Ríos, G. G., Dal Bianco, P. A., Ponte Ahón, S. A., Stanchi, O. A., Seery, J., Badaracco, T., Rabinovich, F., Vázquez, J., Fernández Bariviera, A., Khalid, S., Rashwan, H., Puig Valls, D., Aidelman, Y. J., Gamen, R. C., Lanzarini, L. C., & Hasperué, W. (2023). Deep learning para visión por computadora. XXV Workshop de Investigadores en Ciencias de la Computación (WICC 2023), Junín, Argentina.. 2023.

[6] Ronchetti, F., Quiroga, F. M., Pereyra, N., Miranda, J., Ponte Ahón, S. A., Aidelman, Y. J., Gamen, R. C., & Lanzarini, L. C. (2022). Software inteligente para la digitalización de placas espectroscópicas. Actas del XXVIII

Congreso Argentino de Ciencias de la Computación (CACIC 2022), 26-35. 2022

[7] Ponte Ahón, S. A., Aidelman, Y., Seery, J. M., Quiroga, F., Ronchetti, F., Hasperué, W., Iannuzzi, M., Peralta, R., López, M., Bariviera, A. F., Cidale, L., & Gamen, R. PlateUNLP: An open-source solution for automating the digitization of historical astronomical plates. En Eurographics Workshop on Graphics and Cultural Heritage (The Eurographics Association). 2024.

Marcos Argumentativos Interpretativos: Formalizando la Ambigüedad en el Razonamiento Computacional

Carlos A. Habiñak[1] Gerardo I. Simari[1] Maximiliano C. D. Budán[2]

[1] Instituto de Ciencias e Ingeniería de la Computación (ICIC UNS-CONICET) & Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS)

[2] Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero & CONICET

{carlos.habinak, gis}@cs.uns.edu.ar, mcdb@unse.edu.ar

RESUMEN

Esta línea de I+D introduce el *Interpretative Argumentation Framework* (IAF), una extensión de los marcos de argumentación abstractos diseñada para abordar formalmente la multiplicidad de significados que existen en distintos ámbitos. A diferencia de los enfoques tradicionales que consideran los argumentos como entidades atómicas y unívocas, IAF reconoce que un mismo fragmento de conocimiento puede admitir diversas interpretaciones, cada una con consecuencias argumentativas distintas. Se desarrolla una arquitectura formal donde las piezas de conocimiento y sus interpretaciones se modelan como entidades de primer orden, conectadas mediante relaciones de *conflicto*, que capturan la exclusión mutua, y de *fundamentación* (*entailment*), que representan el refuerzo conceptual o la dependencia pragmática. A partir de esta estructura, se definen los conceptos de *directriz* y *posición*, permitiendo representar posturas interpretativas globales coherentes y exhaustivas. Asimismo, se extienden las semánticas clásicas desarrolladas por Dung al ámbito interpretativo, trasladando sus principios al dominio de las interpretaciones. En este sentido, se estudia la relación de fundamentación como un elemento central para ordenar la estructura argumentativa, atendiendo tanto a situaciones en las que las interpretaciones se mantienen independientes como a aquellas en las que los conflictos se propagan entre ellas. Los resultados obtenidos indican que el IAF ofrece un fundamento teórico consistente para examinar la racionalidad en contextos marcados por la ambigüedad, al integrar de manera coherente la teoría de la argumentación con una perspectiva epistemológica orientada a la interpretación.

Palabras Clave: Inteligencia Artificial, Representación del Conocimiento, Argumentación Computacional, Interpretaciones, Razonamiento bajo Incertidumbre.

CONTEXTO

Esta línea de investigación se desarrolla en el marco del proyecto "Herramientas de razonamiento automatizado para el desarrollo de sistemas socio-técnicos inteligentes" (PGI 24/N057) dirigido por Gerardo I. Simari, financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional del Sur (UNS) para el período 2023-2026 y radicado en el Departamento de Ciencias e Ingeniería de la Computación y del proyecto "Inteligencia Artificial al servicio de la comunidad: un enfoque predictivo y personalizado" (PGI 23/C222-Bint-2025) dirigido por Maximiliano C. D. Budán, financiado por la Facultad de Ciencias Exactas y Tecnología de la Universidad Nacional de Santiago del Estero (UNSE) para el período 2025-2026 y radicado en el Instituto de Investigación en Informática y Sistemas de Información.

1. INTRODUCCIÓN

La argumentación computacional se ha consolidado como un tema central en la Inteligencia

Artificial, especialmente dentro de la representación del conocimiento y el razonamiento. Esta disciplina proporciona herramientas formales para capturar, intercambiar y evaluar fundamentos, permitiendo la toma de decisiones en dominios donde la información es incierta, incompleta o cuestionada [1, 2, 3]. Uno de los paradigmas principales del área, los Marcos de Argumentación Abstracta (AF, por sus siglas en Inglés) de Dung [4], ha inspirado una amplia gama de extensiones, incluyendo preferencias [5, 6], relaciones de soporte [7, 8], aspectos temporales [9] y argumentos con peso [10, 11].

A pesar de estos avances, los marcos actuales suelen tratar a los argumentos como entidades autocontenidas, pasando por alto que éstos a menudo se originan a partir de diferentes interpretaciones del mismo conocimiento subyacente. Tradicionalmente, los sistemas de argumentación estructurada asumen que las premisas, reglas y conclusiones poseen un significado determinado, estático y unívoco mucho antes de que comience cualquier tipo de evaluación dialéctica. Sin embargo, este supuesto resulta sumamente frágil frente a la riqueza, ambigüedad y complejidad inherente al lenguaje natural y a las interacciones humanas [17]. Cuando el significado se fija de manera prematura en el sistema, la representación del conocimiento puede fallar al intentar capturar la verdadera naturaleza de una disputa.

En efecto, los marcos formales existentes carecen de un mecanismo explícito para representar y razonar sobre interpretaciones alternativas de un mismo contenido antes de establecer relaciones de ataque o soporte entre argumentos. Como consecuencia, el sistema presupone una única lectura válida del conocimiento y no distingue si un conflicto es genuinamente inferencial o si, en realidad, surge de discrepancias semánticas previas. Una cláusula en un texto legal, una declaración científica o incluso un mensaje corto en redes sociales pueden admitir múltiples lecturas plausibles, cada una con distintas consecuencias argumentativas: acuerdo o desacuerdo, refuerzo o contradicción. Cuando esta pluralidad interpretativa no se modela formalmente, el desacuerdo tiende a atribuirse a deficiencias lógicas o probatorias, cuando en realidad puede originarse en compromisos semánticos divergentes.

Este problema es particularmente evidente en escenarios del mundo real:

- En el Derecho: Términos ambiguos como "menor" o "duda razonable" están abiertos a interpretación, y diferentes resoluciones pueden llevar a resultados judiciales divergentes [12].
- En las plataformas sociales: Un solo tweet podría leerse literalmente como una queja genuina sobre una situación particular, irónicamente como un comentario humorístico que busca generar empatía, o estratégicamente como una pieza de desinformación deliberada diseñada para manipular la opinión pública.
- En debates científicos: Afirmaciones como "el cambio climático es real" son aceptadas como hechos, pero sus interpretaciones divergen drásticamente, moldeando políticas y estrategias económicas.

En estos dominios, los desacuerdos no necesariamente surgen de los hechos en sí, sino de los *significados* que se asignan a ellos. La interpretación es un proceso complejo de atribución de significado [13], influido por múltiples factores contextuales, cognitivos y retóricos [14, 15, 16]. Una interpretación inadecuada puede oscurecer la intención del mensaje original y distorsionar la dinámica argumentativa trasladando el foco del desacuerdo desde la semántica hacia la inferencia.

Para abordar esta brecha conceptual y formal, proponemos el formalismo de los *Marcos de Argumentación Interpretativa* (IAF). A diferencia de los modelos de argumentación estructurada existentes [19], que presuponen una interpretación fija del conocimiento, el IAF introduce una capa previa de razonamiento en la cual el significado mismo de las premisas se

encuentra aún bajo evaluación. Las interpretaciones son tratadas como entidades de primera clase, interconectadas mediante relaciones formales de conflicto (exclusión mutua) e implicación (refuerzo semántico). Esta arquitectura permite organizar posturas interpretativas coherentes antes de evaluar los ataques directos entre argumentos, garantizando una separación clara entre divergencia semántica y confrontación inferencial, lo cual apunta a proporcionar una base más precisa para el análisis de la racionalidad en dominios caracterizados por la ambigüedad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Nuestra investigación se centra en la formalización y evaluación de la multiplicidad interpretativa a través de tres ejes principales, que se describen a continuación.

2.1. Formalización estructural de IAF

El desarrollo de un marco abstracto donde las piezas de conocimiento y sus posibles interpretaciones se modelan como entidades fundamentales. El marco se define como $\Theta = \langle \mathcal{K}, \mathcal{I}, \mathcal{L}, \mathcal{C}, \mathcal{E} \rangle$, integrando los siguientes componentes:

- *Piezas de Conocimiento* ($\mathcal{K}$): Fragmentos atómicos de información (por ejemplo, hechos o normas).
- *Conjuntos de Interpretaciones* ($\mathcal{I}$): Las posibles lecturas asociadas a cada pieza de conocimiento.
- *Función de interpretación* ($\mathcal{L}$): Asocia a cada una de las piezas del conocimiento un conjunto de posibles interpretaciones.
- *Relaciones de Interacción*:

  Conflicto ($\mathcal{C}$): Una relación simétrica que captura la incompatibilidad mutua (por ejemplo, una lectura literal frente a una irónica o sarcástica).

  Fundamentación ($\mathcal{E}$): Una relación que captura el refuerzo inferencial o contextual, donde adoptar una interpretación hace que la adopción de otra sea más plausible.

A partir de estos elementos, se pueden construir *Argumentos Interpretados*, y definir estructuras complejas como las *Directrices* (conjuntos que aseguran cobertura interpretativa sobre todo el conocimiento) y las *Posiciones* (refinamientos que seleccionan una única interpretación por pieza de conocimiento).

2.2. Semántica interpretativa

Extendemos las semánticas clásicas de argumentación al dominio interpretativo para evaluar la racionalidad de las posturas adoptadas por un agente:

- *Coherencia Interna y Admisibilidad*: Definida como la ausencia de conflictos internos en una directriz, coincidiendo con la noción de "libre de conflictos" dada la simetría de la relación.
- *Semánticas Preferida y Estable*: Capturan agentes que adoptan posturas "máximas" o "decisivas". Una directriz preferida es aquella que es máximamente consistente, mientras que una estable rechaza todas las interpretaciones incompatibles.
- *Semántica Fundamentada* (*Grounded*): Modela el razonamiento escéptico, aceptando sólo aquellas interpretaciones indisputables (es decir, que no están sujetas a controversia, las cuales corresponden a interpretaciones no atacadas).

2.3. El rol de la fundamentación y la robustez

Investigamos la fundamentación ($\mathcal{E}$) como una capa semántica independiente que regula la estructura inferencial. Se pueden plantear al menos dos variantes:

1. *Sin Propagación de Conflictos*: La fundamentación actúa como refinamiento semántico. Introducimos la noción de *Robustez*, que caracteriza directrices estables bajo la presión combinada del conflicto y

los compromisos inferenciales (clausura bajo fundamentación).

2. *Con Propagación de Conflictos*: Las inconsistencias se transmiten a través de cadenas de fundamentación, revelando tensiones latentes que no son visibles en un análisis puramente basado en conflictos directos.

A continuación, brindamos detalles adicionales acerca de los avances alcanzados al momento y proyectados a futuro.

3. RESULTADOS OBTENIDOS Y ESPERADOS

3.1. Resultados Alcanzados

Hasta el momento, hemos logrado establecer una fundamentación teórica sólida para el razonamiento interpretativo:

- Se ha definido formalmente el IAF, elevando las interpretaciones a entidades de primera clase dentro de la argumentación abstracta.
- Se han caracterizado las propiedades de racionalidad interpretativa: *Coherencia* (consistencia interna), *Fomento* (clausura bajo fundamentación) y *Robustez* (estabilidad integral).
- Se ha demostrado teóricamente la dualidad entre la estabilidad basada en conflictos (rechazo de lo incompatible) y la robustez basada en fundamentación (incorporación de lo compatible).
- Se ha establecido una jerarquía semántica que vincula las directrices (espacios de posibilidad racional) con las posiciones (elecciones racionales específicas).

3.2. Objetivos en curso y futuros

La línea de investigación y desarrollo continúa con los siguientes objetivos:

- **Propiedades computacionales:** Investigar la complejidad algorítmica para el cómputo de directrices robustas y estables, desarrollando algoritmos eficientes para la resolución de conflictos interpretativos.
- **Fundamentación probabilística o con peso:** Extender el marco para manejar grados de incertidumbre en la relación de fundamentación, permitiendo modelar situaciones donde una interpretación apoya a otra sólo parcialmente.
- **Aplicaciones a gran escala:** Aplicar el marco a sistemas reales de deliberación científica, razonamiento legal y moderación en línea, donde la gestión de la diversidad interpretativa es esencial para una toma de decisiones justa y racional.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación se desarrolla en el marco del tópico más general de Razonamiento bajo Incertidumbre que se lleva adelante en el *Grupo de Investigación en Ingeniería Cognitiva y Bases de Datos* (ICBD), el cual es dirigido por Gerardo I. Simari y forma parte del Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) de la Universidad Nacional del Sur, con colaboración del Instituto de Ciencias e Ingeniería de la Computación (ICIC), instituto de doble dependencia del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y la UNS. El grupo posee sólidas conexiones de colaboración con el exterior, en particular con grupos de investigación que estudian temas afines en instituciones de Estados Unidos (Syracuse University y Northwestern University), Italia (Università della Calabria), España (IIIA-CSIC Barcelona), Inglaterra (University of Oxford y Alan Turing Institute), Austria (TU Vienna), entre otros, como así también a nivel nacional, como la Universidad de Buenos Aires (UBA), Universidad Nacional de Entre Ríos (UNER) y Universidad Nacional de Santiago del Estero (UNSE) y Universidad Nacional de San Luis (UNSL).

Esta línea forma parte del núcleo central de la tesis doctoral de Carlos A. Habiñak, tesista del DCIC UNS dirigida por Gerardo I. Simari y Maximiliano C. D. Budán y becario

CONICET. Se prevé que la misma será continuada durante el período posdoctoral y sumará a futuro a otros tesistas (de grado o posgrado) que se interesen en los diferentes ejes descriptos anteriormente.

BIBLIOGRAFÍA

1. Bandyopadhyay, S. (2023). *Decision Support System: Tools and Techniques*. CRC Press.
2. Alfano, G., Greco, S., Parisi, F., & Trubitsyna, I. (2023). On acceptance conditions in abstract argumentation frameworks. *Inf. Sci.*, 625, 757-779.
3. Alfano, G., Calautti, M., Greco, S., Parisi, F., & Trubitsyna, I. (2023). Explainable acceptance in probabilistic and incomplete abstract argumentation frameworks. *Artificial Intelligence*, 323, 103967.
4. Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2), 321-357.
5. Amgoud, L., Cayrol, C., & Le Berre, D. (1996). Comparing arguments using preference orderings for argument-based reasoning. *Proc. ICTAI*, 400-403.
6. Kaci, S., & van der Torre, L. (2008). Preference-based argumentation: Arguments supporting multiple values. *IJAR*, 48(3), 730-751.
7. Gonzalez, M. G. E., Budán, M. C. D., Simari, G. I., & Simari, G. R. (2021). Labeled bipolar argumentation frameworks. *JAIR*, 70, 1557-1636.
8. Gabbay, D. M. (2016). Logical foundations for bipolar and tripolar argumentation networks: preliminary results. *Journal of Logic and Computation*, 26(1), 247-292.
9. Barringer, H., Gabbay, D., & Woods, J. (2005). Temporal dynamics of support and attack networks: From argumentation to zoology. *Mechanizing Mathematical Reasoning*, 59-98. Springer.
10. Dunne, P. E., Hunter, A., McBurney, P., Parsons, S., & Wooldridge, M. (2011). Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.*, 175(2), 457-486.
11. Budán, M. C. D., Simari, G. I., Viglizzo, I., & Simari, G. R. (2017). An approach to characterize graded entailment of arguments through a label-based framework. *IJAR*, 82, 242-269.
12. Walton, D., Macagno, F., & Sartor, G. (2021). *Statutory interpretation: Pragmatics and argumentation*. CU Press.
13. Dutilh Novaes, C. (2021). *Argument and argumentation*.
14. Macagno, F., & Walton, D. (2017). Arguments of statutory interpretation and argumentation schemes. *International Journal of Legal Discourse*, 2(1), 47-83.
15. Ribeiro, H.. (2014). *Systematic approaches to argument by analogy* (Vol. 25). Springer.
16. Bongiovanni, G., Postema, G., Rotolo, A., Sartor, G., Valentini, C., & Walton, D. (2018). *Handbook of legal reasoning and argumentation*. Springer.
17. Dietz, E., Kakas, A., & Michael, L. (2023). Computational Argumentation & Cognitive AI. *Human-Centered Artificial Intelligence: Advanced Lectures*, 363-388. Springer.
18. Ervas, F., Rossi, M. G., Ojha, A., & Indurkhya, B. (2021). The double framing effect of emotive metaphors in argumentation. *Fron. Psychol.* 12:628460.
19. Woltran, S. (2020). Computational Argumentation--Formal Models and Complexity Results. *Proc. CILC*, 1-2.

Mejorando la eficiencia de metaheurísticas mediante estrategias de inicializaciones basadas en Machine Learning

Carlos Bermudez[1], José Luis Hernandez[2], Mercedes Carnero[2],
Hugo Alfonso[1], Gabriela Minetti[1], Carolina Salto[1], Yanina Ditz[1], Javier Vargas[1]
[1]Lab. de Investigación en Sistemas Inteligentes (LISI) - Fac. de Ing. - Univ. Nac. de La Pampa (UNLPam)
e-mail: [1]minettig@ing.unlpam.edu.ar
[2]Grupo de Optimización (GOp) - Fac. de Ing. - Univ. Nac. de Río Cuarto (UNRC)

Resumen

Esta línea de investigación introduce *Machine Learning Seeding Tool* (MLST), una herramienta genérica basada en aprendizaje automático orientada a predecir soluciones iniciales de alta calidad usadas en algoritmos metaheurísticos que resuelven problemas de optimización. A diferencia de los enfoques tradicionales que generan puntos de partida de manera aleatoria, MLST aprende patrones estructurales a partir de instancias previamente resueltas y propone soluciones iniciales que favorecen una convergencia más rápida y mejores resultados finales. Como caso de estudio, esta herramienta se aplica al problema de Diseño de Redes de Distribución de Agua (WDND), un problema de optimización combinatoria a gran escala con restricciones hidráulicas y operativas. Los resultados obtenidos muestran que la inicialización guiada por MLST mejora el desempeño del proceso de optimización frente a estrategias convencionales.

Palabras claves: Optimización, diseño redes de distribución de agua, *Machine Learning*, Seeding, Metaheurísticas Optimización con restricciones

CONTEXTO

En los problemas de optimización con restricciones, la generación de soluciones factibles constituye una etapa crítica en los algoritmos metaheurísticos. Para que el proceso de búsqueda se enfoque en mejorar la calidad de las soluciones respecto de la función objetivo, resulta indispensable garantizar previamente el cumplimiento de las restricciones del problema. En este sentido, el desempeño de la estrategia de inicialización —o *seeding*— impacta directamente en la eficiencia global del algoritmo.

Tradicionalmente, la generación de soluciones factibles se ha abordado mediante estrategias aleatorias o heurísticas ad-hoc, generalmente diseñadas a partir del conocimiento específico del dominio [1], [2]. Los métodos aleatorios presentan la ventaja de su simplicidad y de promover la diversidad en la población inicial; sin embargo, suelen requerir un número elevado de evaluaciones para que la metaheurística alcance regiones prometedoras del espacio de búsqueda. Por otro lado, las heurísticas basadas en conocimiento experto pueden orientar la generación hacia soluciones más relevantes, aunque su desempeño depende en gran medida de la calidad del conocimiento incorporado y de su adecuada formalización.

Diversos estudios han mostrado que estrategias de *seeding*, tales como la incorporación de soluciones de alta calidad, la utilización de mecanismos de generación informados o el uso de operadores especializados pueden mejorar significativamente el desempeño de las metaheurísticas [3], [4]. Estas mejoras se manifiestan en tiempos de convergencia menores y en una reducción del riesgo de estancamiento en óptimos locales, especialmente en problemas donde la factibilidad y la calidad de las soluciones son simultáneamente difíciles de alcanzar [5].

La complejidad inherente a esta etapa ha motivado la exploración de enfoques complementarios

basados en técnicas de aprendizaje automático (*Machine Learning*, ML), con el objetivo de asistir o guiar la generación de soluciones factibles de manera más eficiente y robusta [6]. ML permite modelar patrones o regularidades a partir de datos históricos o sintéticos asociados al problema de interés. Mediante aprendizaje supervisado o no supervisado, es posible capturar información estructural del espacio de soluciones factibles y diseñar generadores que no dependan exclusivamente de la aleatoriedad o del conocimiento experto codificado manualmente. Esta capacidad de aprender a partir del propio problema abre nuevas oportunidades para mejorar la calidad de las soluciones iniciales, particularmente en escenarios donde alcanzar la factibilidad mediante heurísticas tradicionales resulta costoso en términos de complejidad computacional. Es importante destacar que esta línea de investigación la abordamos conjuntamente en el LISI (Laboratorio de Investigación de Sistemas Inteligentes, UNLPam) y en el GOp (Grupo de Optimización, UNRC).

1. DESARROLLO

En este marco conceptual se inscribe la línea de investigación que se viene desarrollando, orientada a integrar técnicas de aprendizaje automático en la fase de inicialización de metaheurísticas para problemas de optimización con restricciones. La hipótesis central del trabajo sostiene que es posible aprender patrones estructurales presentes en soluciones factibles y utilizar ese conocimiento para generar soluciones iniciales de mayor calidad, reduciendo así el esfuerzo exploratorio requerido por la metaheurística.

Como caso de estudio se adopta el problema de Diseño de Redes de Distribución de Agua, el cual consiste en determinar los diámetros de las tuberías de una red de distribución (*Water Distribution Network*, WDN) de modo de minimizar el costo total de inversión, respetando simultáneamente restricciones hidráulicas tales como conservación de masa y energía, presión mínima en nodos y velocidad máxima en tuberías. Se trata de un problema combinatorio con restricciones no lineales, en el que la generación de soluciones factibles no es trivial y constituye una etapa especialmente sensible del proceso de optimización.

En este contexto MLST se concibe como una herramienta genérica basada en el aprendizaje automático para asistir en la generación de soluciones iniciales en problemas de optimización. MLST no constituye un modelo de optimización en sí mismo, sino una estrategia de *seeding* que aprende, a partir de soluciones factibles previamente obtenidas, relaciones entre atributos estructurales del problema y las decisiones adoptadas en dichas soluciones.

En el caso de WDND, MLST funciona como un clasificador supervisado, donde la variable objetivo es el diámetro de la tubería seleccionada entre un conjunto discreto de alternativas comerciales. A partir de características de la red (propiedades de las tuberías, demanda máxima y tipo de red) se entrenan modelos predictivos capaces de estimar los diámetros adecuados en cada nodo de la red. Las predicciones generadas se integran luego en la construcción de soluciones iniciales que sirven de punto de partida para algoritmos metaheurísticos.

La línea de investigación no se limita a la implementación puntual de MLST en el problema WDND, sino que busca establecer un marco metodológico exportable a otros problemas de optimización con restricciones, donde la factibilidad represente un cuello de botella en la etapa de inicialización. En este sentido, el problema WDND funciona como banco de pruebas para validar la hipótesis de que el aprendizaje a partir de soluciones previas puede mejorar la eficiencia del proceso de búsqueda sin comprometer el cumplimiento de las restricciones.

2. RESULTADOS OBTENIDOS

En [7] se introduce MLST, como una herramienta que emplea una variedad de técnicas de aprendizaje supervisado, incluyendo *k-Nearest Neighbors* (k-NN), *Multi-layer Perceptron* (MLP), *Decision Tree* (DT), *Random Forest* (RF) y *eXtreme Gradient Boosting* (XGBoost), con el fin de generar soluciones iniciales de alta calidad, que mejoren las estrategias tradicionales de *seeding* y con esto se espera que el algoritmo *Hybrid Simulated Annealing* (HSA) [8] obtenga resultados superiores con menor costo computacional.

Para respaldar el desarrollo y la evaluación del enfoque propuesto, se genera un conjunto de datos compuesto por soluciones factibles de alta calidad para el problema WDND, obtenidas mediante la

ejecución del HSA propuesto en [8]. En una solución WDND, cada tubería se representa como una fila en el conjunto de datos. La red a la que pertenece la tubería, su nodo de conexión, su longitud y su rugosidad constituyen las variables predictoras, mientras que su diámetro se utiliza como variable objetivo. Las cinco técnicas de aprendizaje automático seleccionados para MLST se entrenan y evalúan sobre este conjunto de datos. En función de su desempeño predictivo, se selecciona el modelo MLST más adecuado para generar soluciones iniciales factibles para el problema WDND. Finalmente, las soluciones guiadas por MLST se comparan con las obtenidas mediante la estrategia tradicional de inicialización comúnmente utilizada para este tipo de problema.

Por un lado, el desempeño de los cinco modelos MLST se evalúa utilizando las métricas *accuracy*, promedios macro y promedios ponderados sobre un conjunto de prueba compuesto por 17 493 observaciones (Tabla I). Los valores de *accuracy* son altamente consistentes (0.67–0.68), con $MLST_{k\text{-}NN}$ superando al resto. Sin embargo, las métricas macro, que consideran todas las clases por igual, muestran valores bajos de *precision*, *recall* y *F1-score* (0.05–0.08), lo que evidencia un bajo desempeño en las clases minoritarias. $MLST_{XGBoost}$ alcanza la mayor precisión macro (0.20), lo que sugiere una mayor capacidad de generalización entre fronteras de clase. Por lo tanto, se selecciona $MLST_{XGBoost}$ para generar soluciones iniciales para los algoritmos de optimización metaheurísticos debido a su sólido desempeño predictivo y a sus mínimos requerimientos de preprocesamiento. Si bien todos los modelos MLST evaluados exhiben valores de *accuracy* similares (0.67–0.68), $MLST_{XGBoost}$ se destaca por alcanzar la mayor precisión macro-promediada (0.20) y la mayor precisión ponderada (0.54), lo que indica un desempeño robusto tanto en clases mayoritarias como minoritarias. Además, a diferencia de modelos como $MLST_{k\text{-}NN}$ y $MLST_{MLP}$, que requieren estandarización o escalamiento de los datos para funcionar de manera óptima, $MLST_{XGBoost}$ opera eficazmente sobre datos heterogéneos y sin escalar. Esto simplifica su integración en los flujos de optimización y reduce la carga asociada al preprocesamiento.

Por otro lado, este estudio demuestra la capacidad de MLST para generar soluciones iniciales factibles al problema WDND, dado que reduce significativamente el Costo Total de Inversión (TIC) al integrarse como mecanismo de inicialización en el proceso de búsqueda del algoritmo HSA. El análisis comparativo muestra reducciones en el TIC entre un 37 % y un 55 %, según el tipo de red evaluada, destacando el potencial del uso de modelos de aprendizaje supervisado en esta etapa crítica del proceso de optimización.

Resumiendo, los experimentos realizados muestran que MLST supera a las estrategias convencionales de inicialización. En primer lugar, la herramienta logra generar soluciones iniciales factibles, reduciendo el número de evaluaciones hidráulicas. Asimismo, las soluciones generadas por modelos MLST presentan valores iniciales de TIC inferiores a los que presentan las soluciones obtenidas mediante estrategias de *seeding* tradicionales, lo que favorece un proceso de búsqueda más eficiente y con tiempos de convergencia menores. Otro resultado relevante es que MLST contribuye a reducir el número total de llamadas al simulador hidráulico en las etapas iniciales del algoritmo de optimización, lo que disminuye el costo computacional del proceso.

En conjunto, estos resultados confirman que MLST no solo constituye una estrategia de inicialización efectiva, sino que también representa una herramienta generalizable que puede complementar de manera robusta distintas metaheurísticas orientadas a resolver problemas de optimización con restricciones.

3. FORMACIÓN DE RECURSOS HUMANOS

En el marco de esta línea de investigación se incorporan alumnos avanzados de la carrera Ingeniería en Sistemas de la UNLPam, quienes participan en tareas vinculadas con la resolución de problemas de optimización mediante técnicas inteligentes. Su participación tiene como objetivo acompañarlos en el desarrollo de sus tesinas de grado y, al mismo tiempo, promover la formación de futuros investigadores científicos. Por otra parte, los docentes-investigadores involucrados fortalecen su formación académica y profesional mediante la realización de diversos cursos de posgrado

TABLE I

MÉTRICAS DE PERFORMANCE DE CADA MODELO

Métrica/Modelo		$MLST_{k-NN}$	$MLST_{MLP}$	$MLST_{DT}$	$MLST_{RF}$	$MLST_{XGBoost}$
Accuracy		0.68	0.67	0.67	0.67	0.67
Macro Avg.	*Precision*	0.13	0.04	0.07	0.07	0.20
	Recall	0.07	0.06	0.08	0.08	0.07
	F1-Score	0.07	0.05	0.07	0.06	0.07
Weighted Avg.	*Precision*	0.51	0.45	0.47	0.47	0.54
	Recall	0.68	0.67	0.67	0.67	**0.67**
	F1-Score	0.55	0.54	0.55	0.55	0.55

relacionados con las temáticas del proyecto y con la actualización de sus prácticas docentes.

REFERENCES

[1] J. Kotary, F. Fioretto, P. V. Hentenryck, and B. Wilder, "End-to-end constrained optimization learning: A survey," *CoRR*, vol. abs/2103.16378, 2021. [Online]. Available: https://arxiv.org/abs/2103.16378

[2] D. Liu, V. Perreault, A. Hertz, and A. Lodi, "A machine learning framework for neighbor generation in metaheuristic search," 2022. [Online]. Available: https://arxiv.org/abs/2212.11451

[3] Q. Li, S.-Y. Liu, and X.-S. Yang, "Influence of initialization on the performance of metaheuristic optimizers," *Applied Soft Computing*, vol. 91, p. 106193, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494620301332

[4] M. Sarhani, S. Voß, and R. Jovanovic, "Initialization of metaheuristics: comprehensive review, critical analysis, and research directions," *International Transactions in Operational Research*, vol. 30, no. 6, pp. 3361–3397, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/itor.13237

[5] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Comput. Surv.*, vol. 35, no. 3, p. 268–308, Sep. 2003. [Online]. Available: https://doi.org/10.1145/937503.937505

[6] S. Szénási and G. Légrádi, "Machine learning aided metaheuristics: A comprehensive review of hybrid local search methods," *Expert Systems with Applications*, vol. 258, p. 125192, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417424020591

[7] C. Bermudez, H. Alfonso, M. Carnero, J. L. Hernandez, G. Minetti, and C. Salto, "Ml-driven seeding for constrained optimization problems," in *Proceedings of the XX National Conference on Computer Science (CACIC 2025)*, Argentina, 2025, available from the authors.

[8] C. Bermudez, H. Alfonso, G. Minetti, and C. Salto, "Hybrid simulated annealing to optimize the water distribution network design: A real case," in *Computer Science – CACIC 2020*, P. Pesado and J. Eterovic, Eds. Cham: Springer International Publishing, 2021, pp. 19–34.

Optimización del Triaje médico y trazabilidad de prescripciones mediante la integración de agentes de IA y tecnología blockchain en el hospital Eleazar Herrera Motta

Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito 9 de Julio 22, Chilecito, La Rioja, Argentina [1]

Lic. Ariel Quiroga Marin[1], Ing. Enrique N. Martinez[1], Matias Perez[1], Cesar Salcedo[1] {cquiroga, emartinez, mperez, csalcedo}@undec.edu.ar, Celena G. Tito Rodriguez[1], Franco Padilla[1,] Gomez Carlos Agustin [1] { celenat172, padilla.joni, mumugomez19 }@gmail.com

Hospital Eleazar Herrera Motta, Cavero 500, Chilecito, La Rioja[2]

Dra. Claudia Salguero[2], Dr. Omar Suarez[2], Leonel Alvarez[2], Rubén Quiroga[2]

{clausalguero2017}@yahaoo[2],{salguero, omarbreppe640, rubenhquiroga0, leonelundec}@gmail.com[2]

RESUMEN

Este trabajo presenta una línea de investigación para implementar un sistema de gestión sanitaria inteligente en el Hospital Eleazar Herrera Motta. La propuesta integra Inteligencia Artificial (IA) para el triaje automatizado, optimizando la clasificación de pacientes y reduciendo tiempos de espera. Complementariamente, se emplea tecnología Blockchain para garantizar la inmutabilidad y seguridad de recetas, turnos e historiales médicos, eliminando el riesgo de fraude en la emisión de fármacos. La arquitectura se apoya en orquestadores de IA para la integración de microservicios, asegurando un sistema escalable y ciberseguro. Esta infraestructura no solo mejora la privacidad y la interoperabilidad de los datos, sino que blinda la cadena de suministro de medicamentos y facilita la investigación clínica mediante el intercambio seguro de información. El resultado es un ecosistema de salud transparente que fortalece la confianza entre la industria farmacéutica, el hospital y el paciente, logrando una trazabilidad integral desde la prescripción hasta la entrega del medicamento. La implementación de este modelo estima una reducción crítica en la carga administrativa y una optimización del en el flujo de atención en urgencias, estableciendo un precedente de transformación digital para la región.

Palabras clave: Inteligencia Artificial, Blockchain, Triaje Médico, Trazabilidad, Gestión Sanitaria, Ciberseguridad.

Líneas de I+D+I

- Procesamiento de Señales y Sistemas de Tiempo Real
- Arquitectura, Redes y Sistemas Operativos
- Base de Datos
- Cloud Computing
- Minería de Datos

CONTEXTO

Este proyecto surge para transformar al Hospital Eleazar Herrera Motta en un "Hospital Inteligente". El objetivo no es solo digitalizar papeles, sino crear un sistema donde la tecnología trabaje de forma coordinada y eficiente.

La propuesta se basa en dos pilares tecnológicos:

Inteligencia Artificial: Se utiliza para automatizar el triaje, clasificando a los pacientes con mayor rapidez y reduciendo las esperas en urgencias.

Blockchain: Se implementa para que la receta electrónica funcione como un "activo digital único". Esto garantiza que la información sea imposible de alterar, asegurando la privacidad del paciente y evitando fraudes en la entrega de medicamentos en farmacias y centros de salud.

Mediante el uso de orquestadores de IA, logramos que los datos viajen de forma segura y privada entre el hospital y las farmacias. Así, cumplimos con los estándares de ciberseguridad actuales y garantizamos una gestión de salud transparente, moderna y, sobre todo, confiable para el paciente.

1. INTRODUCCIÓN

La digitalización de los centros de salud enfrenta retos significativos en la integridad de los datos y la eficiencia administrativa, especialmente en instituciones regionales como el Hospital Eleazar Herrera Motta. El sistema tradicional de gestión hospitalaria depende de procesos manuales que son propensos a errores de clasificación en urgencias y vulnerabilidades en la cadena de custodia de medicamentos.

Esta línea de I/D propone una solución basada en la convergencia de dos tecnologías disruptivas: la IA para el procesamiento de lenguaje natural y Blockchain para la creación de un registro inalterable que actúe como memoria de cada interacción médica. El contexto de la investigación se centra en mejorar la respuesta ante emergencias, optimizando la formación de recursos humanos en tecnologías de frontera.

La transformación digital en el ámbito sanitario no es simplemente una actualización tecnológica, sino una necesidad ética y operativa. Como señala **Topol (2019)**, la medicina de alto rendimiento actual nace de la convergencia entre la inteligencia humana y la artificial, permitiendo que la IA actúe como un soporte crítico en la toma de decisiones clínicas. En centros de salud de complejidad regional, como el **Hospital Eleazar Herrera Motta**, la saturación de los servicios de guardia exige sistemas de triaje que puedan procesar grandes volúmenes de datos en tiempo real.

Siguiendo las directrices de la **Organización Mundial de la Salud (2021)**, la implementación de estos sistemas debe realizarse bajo un marco estricto de ética y gobernanza, asegurando que la IA sea una herramienta de apoyo y no un sustituto del juicio médico profesional. La problemática de la integridad de los datos clínicos, a menudo vulnerables en sistemas centralizados, encuentra una solución en la tecnología de registros distribuidos. Según **Kuo et al. (2017)**, el Blockchain ofrece una oportunidad sin precedentes para crear un ecosistema de salud transparente, donde la trazabilidad de cada interacción médica sea auditable y segura.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se detallan sintéticamente los ejes del tema que se están investigando:

- **Agentes de IA para Triaje Médico:** Desarrollo de modelos de lenguaje de gran tamaño (LLM) capaces de interpretar la descripción de síntomas del paciente y asignar niveles de urgencia (ALTA, MEDIA, BAJA) en tiempo real, basándose en protocolos médicos internacionales.
- **Protocolos de Blockchain para Trazabilidad:** Investigación sobre redes permisionadas para el registro de prescripciones digitales, asegurando que cada receta sea un activo único y no duplicable.
- **Interoperabilidad mediante IA:** Uso de herramientas de automatización de bajo código para conectar bases de datos legadas del hospital con nuevas capas de seguridad descentralizada.
- **Ética y Privacidad de Datos:** Estudio de técnicas de anonimización para que la IA procese datos clínicos sin vulnerar la identidad del paciente en el Hospital Eleazar Herrera Motta.

El núcleo técnico de esta línea de investigación se basa en la interoperabilidad semántica y la seguridad criptográfica. La propuesta utiliza contratos inteligentes (Smart Contracts) que, de acuerdo con los principios fundacionales de

Buterin (2014), permiten automatizar la lógica de negocio —en este caso, la emisión y validación de recetas— sin intervención de terceros, garantizando que el activo digital (la prescripción) sea único e inalterable.

Asimismo, el uso del orquestador **n8n (2025)** permite que el flujo de datos entre el paciente y el hospital sea fluido pero controlado. Esta arquitectura de microservicios aborda los desafíos planteados por **Hassan et al. (2022)**, quienes sostienen que la integración de IA y Blockchain es la respuesta definitiva a los retos de ciberseguridad en la salud moderna, protegiendo la infraestructura contra ataques de manipulación de datos. La investigación también contempla la trazabilidad de la cadena de suministro farmacéutica, un área identificada por **HIMSS (2023)** como crítica para evitar el fraude y asegurar que los medicamentos lleguen efectivamente a los pacientes que los requieren en la farmacia del hospital.

2.1. Estrategia de Integración con Sistemas HSI Existentes

Para que la línea de I/D sea operativa dentro del **Hospital Eleazar Herrera Motta**, el flujo desarrollado no debe funcionar de manera aislada, sino como una capa de inteligencia añadida al sistema de **Historia Salud Integrada (HSI)**. La integración se realizará mediante los siguientes mecanismos:

- **Interoperabilidad vía HL7 FHIR:** Se utilizará el estándar *Fast Healthcare Interoperability Resources* (FHIR) para el intercambio de datos entre el HSI y el nodo de IA. Esto permite que el diagnóstico preliminar de la IA se adjunte automáticamente a la ficha clínica del paciente sin intervención manual.
- **Conectividad mediante API REST:** El sistema HSI del hospital actuará como la fuente de verdad. El orquestador IA consumirá los datos del paciente (ID, antecedentes) mediante *endpoints* seguros para que el agente de IA tenga contexto médico antes de realizar el triaje.
- **Capa de Validación Blockchain (Middleware):** Antes de que una receta generada en el HSI llegue a la farmacia, el flujo de n8n interceptará la orden, generará un *hash* criptográfico y lo registrará en la red de bloques. Esto crea un "notario digital" que valida la receta original frente a cualquier intento de falsificación externa.

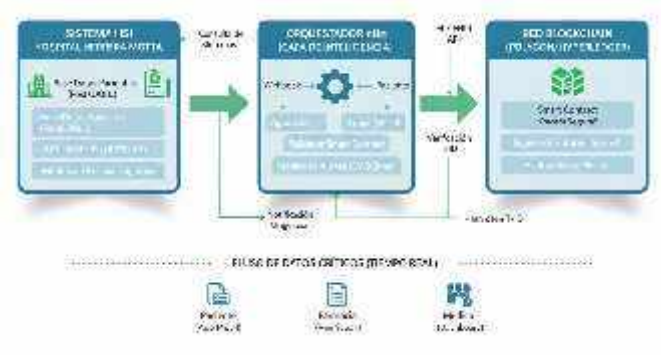


Figura 1. Esquema de arquitectura modular para la integración de sistemas de salud (HSI) con agentes de inteligencia artificial y redes descentralizadas

3. RESULTADOS OBTENIDOS / ESPERADOS

A continuación, se exponen los resultados esperados y los objetivos en curso de la línea de I/D presentada:

3.1. Eficiencia en el Triaje Automatizado

Se ha validado un prototipo funcional que procesa solicitudes en menos de 5 segundos. En pruebas controladas, la IA identificó correctamente cuadros de "Infarto de Miocardio" y "Dificultad Respiratoria Aguda", derivando alertas inmediatas al equipo de guardia.

3.2. Tabla Comparativa de Impacto

A continuación, se detallan los beneficios esperados tras la implementación completa en el hospital:

Indicador	Sistema Tradicional (Papel/Manual)	Sistema Propuesto (IA + Blockchain)
Tiempo de Triaje	10 - 20 minutos	< 1 minuto
Seguridad de Recetas	Alta posibilidad de duplicidad	Inalterable (Smart Contracts)
Trazabilidad de Datos	Fragmentada y vulnerable	Centralizada y auditable
Gestión de Turnos	Rígida / Alto ausentismo	Dinámica / Predicción de No-Show

3.3. Objetivos Futuros

La investigación propuesta para el **Hospital Eleazar Herrera Motta** demuestra que la convergencia de la Inteligencia Artificial y la tecnología Blockchain no solo es posible, sino necesaria para modernizar la gestión hospitalaria pública. La capacidad de automatizar el triaje médico permite una respuesta inmediata ante patologías críticas, mientras que la descentralización de los registros garantiza la seguridad y transparencia en la emisión de recetas.



Figura 2. Proceso de securización de datos clínicos mediante funciones hash criptográficas y registro inalterable en Blockchain

Como trabajo futuro, se plantea la expansión de este modelo hacia una red regional de salud, donde diferentes instituciones puedan compartir datos de manera segura mediante protocolos de interoperabilidad semántica. Se prevée incorporar como segunda etapa un Registro Inalterable de turnos médicos y disponibilidad, como así también un chatbot IA para el paciente. La formación de recursos humanos en estas áreas posiciona al hospital y a la universidad a la vanguardia de la salud digital, cumpliendo con los objetivos de difusión y discusión abierta en la comunidad científica.

La instancia final de la arquitectura propuesta en el **Hospital Eleazar Herrera Motta** contempla la generación de un código QR dinámico como interfaz de verificación. Este proceso se automatiza mediante un nodo específico de generación de QR en el orquestador **IA**, el cual toma como entrada el *Transaction ID* (TxID) generado por la red Blockchain tras la validación de la receta. La lógica del flujo permite convertir esta cadena alfanumérica en un activo binario de imagen, que puede ser embebido en una plantilla PDF o enviado vía mensajería instantánea al paciente. Este mecanismo garantiza la **trazabilidad del último tramo**, permitiendo que en la farmacia hospitalaria se valide la autenticidad de la prescripción simplemente escaneando el código. Esto no solo elimina la posibilidad de adulteración de recetas físicas, sino que también facilita el acceso a la salud para usuarios con distintos niveles de alfabetización digital, cumpliendo con los estándares de seguridad clínica y eficiencia administrativa planteados en los objetivos del proyecto. El uso de Blockchain garantiza un imperativo ético y legal al blindar la privacidad del paciente y asegurar la trazabilidad absoluta de los insumos médicos. Bajo este esquema, la receta electrónica deja de ser un simple registro digital para convertirse en un "activo digital único" e inalterable.



Figura 3. *Interfaz de validación móvil de prescripciones médicas mediante códigos QR vinculados a la red de bloques.*

Como se aprecia en la Figura 3, la interfaz de usuario final diseñada para el Hospital Eleazar Herrera Motta prioriza la simplicidad y la seguridad. El código QR no es solo un enlace, sino una representación visual de la identidad criptográfica de la receta. Al ser escaneado, el sistema consulta el nodo de Blockchain en tiempo real y devuelve un estado de 'Verificado', junto con la fecha y hora exacta de emisión. Este flujo cierra la brecha técnica del proyecto, transformando algoritmos complejos en una herramienta práctica y cotidiana para el personal de salud y la comunidad de Chilecito, garantizando que la innovación tecnológica se traduzca en una mejora directa en la calidad de atención y la transparencia administrativa

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por docentes de las carreras Ingeniería en Sistemas, Licenciatura en Sistemas, Ing. Mecatrónica, Abogacia y Lic. En Enfermería de la UNdeC (acreditadas por CONEAU). También participan profesionales del Hospital Eleazar Herrera Motta de Chilecito, administrativos y alumnos avanzados de grado en Mecatrónica, Enfermería y Sistemas, realizando su trabajo de tesina final en esta línea de I+D. Se usara laboratorio de IA de la UNDeC.

6. BIBLIOGRAFÍA

1. **Buterin, V. (2014).** *A Next-Generation Smart Contract and Decentralized Application Platform.*
2. Hassan, M. U., et al. (2022). *Blockchain and AI-based Healthcare Systems: Applications and Challenges.* IEEE Access.
3. **HIMSS (2023).** *Blockchain in Healthcare: Tracking Pharmaceutical Supply Chain.*
4. Kuo, T. T., et al. (2017). *Blockchain in healthcare: Opportunities, challenges, and applications.* Journal of the American Medical Informatics Association.
5. **Meskó, B., & Topol, E. J. (2023).** *The Role of Artificial Intelligence in Healthcare.* The Lancet.
6. **n8n.io (2025).** *Documentation: Workflow Automation for Secure Data Handling.*
7. **Organización Mundial de la Salud (2021).** *Ética y gobernanza de la inteligencia artificial en el ámbito de la salud.*
8. **Topol, E. J. (2019).** *High-performance medicine: the convergence of human and artificial intelligence.* Nature Medicine

OPTIMIZACIÓN NO LINEAL ENTERA ASISTIDO POR REDES NEURONALES EN PROBLEMAS DE CALL CENTERS

Angel R. Barberis, Lorena E. del Moral Sachetti, Jorge A. Silvera, Daniel Arias Figueroa
Departamento de Informática, Facultad de Ciencias Exactas – Universidad Nacional de Salta
Salta - Argentina
barberis@cidia.unsa.edu.ar, lorena.dms.7@gmail.com, jsilvera@unsa.edu.ar,
daaf@cidia.unsa.edu.ar

RESUMEN

La gestión eficiente de Centros de Contacto Telefónico representa un desafío central dentro de la Investigación Operativa. La incorporación de inteligencia artificial en este ámbito ha permitido automatizar interacciones con usuarios mediante chatbots y agentes virtuales, reduciendo la carga sobre operadores humanos y agilizando respuestas. No obstante, esta tecnología presenta limitaciones importantes: genera tensiones respecto a la pérdida de empleos, carece de la empatía propia de la interacción humana y no resuelve con precisión los niveles de rendimiento de equipos de agentes predefinidos.

Frente a estas limitaciones, el proyecto de investigación propone un enfoque híbrido que combina dimensionamiento mínimo de recursos humanos con distribución óptima de turnos, asistida por inteligencia artificial, con el objetivo de maximizar los niveles de servicio y la satisfacción del usuario.

Desde el punto de vista metodológico, los algoritmos de optimización no lineal tradicionales enfrentan el problema de quedar atrapados en óptimos locales. Las heurísticas convencionales intentan superarlo, aunque con alta carga computacional. Como alternativa superadora, se propone integrar redes neuronales debidamente entrenadas al proceso de búsqueda, ya que estas pueden escapar de óptimos locales con mayor velocidad y precisión. El algoritmo resultante combinaría métodos de optimización no lineal entera de búsqueda direccional con redes neuronales, logrando una solución más eficiente para la planificación de turnos en centros de contacto.

Palabras claves:
Optimización no lineal entera, optimización de Call Center, Redes neuronales, Nivel de servicio, Algoritmos de búsqueda direccional.

CONTEXTO

El proyecto tiene por objeto desarrollar estrategias de resolución de problemas bajo un nuevo enfoque basado en la toma de decisiones multiobjetivo asistida por Inteligencia Artificial, perspectiva que no ha sido abordada explícitamente en la literatura actual sobre Centros de Contacto. La hipótesis sostenida plantea que este nuevo enfoque, sustentado en la toma de decisiones multicriterio y en la optimización lexicográfica, permite reducir drásticamente la carga computacional del proceso de optimización y, al mismo tiempo, lograr mayor precisión en los resultados a nivel global.

Abordar un modelo complejo desde la perspectiva de una optimización bicriterio con objetivos contrapuestos posibilita el tratamiento riguroso y específico de las restricciones propias de cada subproblema, lo que garantiza un análisis más exhaustivo del dominio de soluciones factibles y favorece el desarrollo de algoritmos eficientes, rápidos y precisos.

Para contrastar las hipótesis formuladas, se propone desarrollar soluciones algorítmicas

de rápida ejecución que converjan hacia soluciones factibles con precisión satisfactoria en términos globales. La convergencia de los algoritmos se estudia mediante un método estadístico desarrollado por el propio grupo de investigación. La validación de las propuestas algorítmicas se lleva a cabo a través de dos mecanismos complementarios: el uso de conjuntos de datos (datasets) provenientes de Centros de Contacto reales, y la simulación estocástica mediante el análisis de distintos escenarios. Los resultados obtenidos se comparan con los derivados de la aplicación del método de van der Sluis, reconocido por obtener soluciones factibles de alta precisión global.

El proyecto se desarrolla en el marco de una investigación cuantitativa con enfoque descriptivo, estrategia que permite caracterizar el modelo matemático teórico sobre el cual se construyen los algoritmos. Dichos algoritmos están diseñados para resolver problemas de difícil tratamiento, tales como aquellos que no son convexos, no son de tipo cuadrático, presentan función objetivo no lineal definida en un dominio discreto y no admiten el cálculo de derivadas; todo ello en el contexto de la ingeniería de procesos subyacente en los Centros de Contacto Telefónico.

El trabajo que se describe se lleva a cabo dentro del marco del proyecto de investigación Nro: 2.949/0 ‘‘Optimización de Procesos Operativos en Centros de Contactos”, aprobado en el año 2.024 por el consejo de investigación de la Universidad Nacional de Salta.

1. INTRODUCCION

La gestión eficiente y la optimización de los niveles de servicio dentro de los Centros de Contactos representan un desafío multifacético que exige una comprensión integral de la teoría de colas, la asignación de recursos, la variabilidad del rendimiento de los agentes y las métricas de satisfacción del cliente (Mehrotra et al., 2011). El objetivo de la optimización de los niveles de servicio es asegurar que el centro de contacto cumpla con sus objetivos de rendimiento, como el tiempo promedio de manejo, el nivel de servicio y la satisfacción del cliente, mientras se minimizan los costos operativos. Optimizar los niveles de servicio en los centros de contactos es crucial para mejorar la experiencia del cliente y la eficiencia operativa (Filippou et al., 2023). Esto implica gestionar eficazmente los recursos, predecir la demanda y adaptar las estrategias para satisfacer los volúmenes variables de llamadas y las expectativas de los clientes (Akşin et al., 2007; Diao et al., 2015). Los indicadores clave de rendimiento, como el tiempo promedio de atención, el nivel de servicio (porcentaje de llamadas atendidas dentro de un tiempo especificado) y las puntuaciones de satisfacción del cliente, son puntos de referencia críticos (Kim et al., 2012). La integración de la tecnología, incluyendo el análisis avanzado y la automatización, juega un papel fundamental en la consecución de estos objetivos. A medida que los centros de contactos se convierten en el centro de la comunicación con los clientes, comprender y mejorar la satisfacción de los usuarios en los servicios remotos es primordial (Chicu et al., 2016). Además, minimizar el tiempo de espera es suma importancia para evitar las fugas de clientes, ya que estos tienen umbrales de tolerancia respecto a cuánto tiempo están dispuestos a esperar (Cordeiro & Neto, 2023).

Los modelos de colas ayudan a predecir los tiempos de espera y optimizar los niveles de personal para cumplir con los objetivos de servicio, como garantizar que un cierto porcentaje de llamadas sean atendidas dentro de un tiempo especificado (Ahmadi et al., 2020). El desafío de gestionar la calidad del servicio se complica más cuando están involucrados múltiples centros de llamadas, lo que potencialmente puede llevar a inconsistencias en la prestación del servicio y en la percepción del cliente (Ye et al., 2014). La simulación es una técnica crucial en la optimización de sistemas de colas estocásticos con abandono, proporcionando perspectivas sobre el comportamiento y el valor del cliente (Afèche et al., 2017).

La optimización de sistemas de colas ha sido explorada en diversos contextos de

servicio, con un enfoque en mejorar la satisfacción del cliente y la eficiencia del sistema (Кондрашова, 2021). La gestión eficiente de los niveles de servicio en diversos sectores, desde el transporte, los centros de llamadas (call center) hasta la producción, presenta desafíos complejos que exigen el desarrollo de modelos y algoritmos de optimización avanzados. La integración de la inteligencia artificial (IA), particularmente las redes neuronales, está revolucionando estos enfoques al permitir la resolución de problemas de optimización no lineal con una precisión y eficiencia sin precedentes.

La optimización en tiempo real es fundamental para sistemas dinámicos, como lo son los centros de llamadas telefónicas o los centros de contactos multimedios. Por ejemplo, los modelos basados en redes neuronales artificiales (ANN) se emplean para el despacho en tiempo real de taxis autónomos eléctricos, maximizando la recompensa por servir a los clientes a través de la formulación de problemas de programación lineal entera (Hu et al., 2022). Esta capacidad de respuesta en tiempo real también se observa en la gestión del tráfico ferroviario, donde se utilizan algoritmos de programación de restricciones (CP) para encontrar horarios y rutas óptimos que minimicen los retrasos debidos a conflictos (Marlière et al., 2023). Además, la optimización integrada de la programación de trenes y la reprogramación del material rodante en sistemas de metro perturbados se aborda con enfoques híbridos de aprendizaje por refuerzo profundo y búsqueda adaptativa de vecindad grande (Su et al., 2023).

La simulación es una herramienta indispensable para modelar la complejidad y la estocasticidad del comportamiento de los operarios. En la optimización del tiempo de viaje para los sistemas de metro, por ejemplo, se considera el proceso de viaje de los pasajeros como una actividad física que consume energía, formulando una función de gasto de energía para sistemas congestionados (Li et al., 2023). El modelado basado en agentes, a menudo extendido con capas de toma de decisiones impulsadas por redes neuronales, permite simular la adaptación estratégica individual en respuesta a cambios de políticas (Yang et al., 2024). Esto es crucial para comprender la colaboración de los trabajadores en celdas de producción, donde se combinan modelos de procesos discretos con modelos multiagente (Yang et al., 2024).

En resumen, la convergencia de la optimización con la inteligencia artificial y, en particular, las redes neuronales, ofrece soluciones robustas y adaptables para una amplia gama de problemas complejos en la optimización de niveles de servicio, donde los centros de contactos telefónicos y los centros de contactos multimedios son algunos de ellos.

Estos modelos se resuelven cada vez más mediante algoritmos híbridos que utilizan redes neuronales para preseleccionar variables enteras o aprender cortes válidos, lo que demuestra la creciente sinergia entre la IA y la optimización tradicional.

2. LINEAS DE INVESTIGACION y DESARROLLO

Los principales ejes temáticos que se investigan son los siguientes:

» Diseño de algoritmos: se busca desarrollar algoritmos de optimización de búsqueda direccional (sin cálculo o estimación de derivadas), con características de ser simples (baja esfuerzo computacional), rápidos y precisos. Estos algoritmos deben ejecutarse en entornos dinámicos.

» Deep Learning: Se investiga distintos tipos de redes neuronales, con distintas estrategias de convergencias. El foco se centra en una arquitectura de red neuronal (RN) que luego del entrenamiento, tenga la capacidad de seguir mejorando la precisión de sus predicciones, sin tener que usar modelos razonadoras de la inteligencia artificial.

» Simulación de Eventos Discretos: Se busca disponer de una herramienta de simulación que integre algoritmo de optimización con red neuronal, y que permita al mismo tiempo, diseño de escenario en forma visual.

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto de investigación se encuentra en la etapa de integración, prueba e interpretación de resultados. Se desarrolló un algoritmo simple basado en el método de Frank-Wolfe, adaptado a un dominio discreto, caracterizado por su alta velocidad, aunque con precisión cuestionable. Asimismo, se desarrolló un algoritmo inédito de búsqueda direccional simple, con baja carga computacional, alta velocidad y gran precisión.

Adicionalmente, se implementaron una adaptación mixta del algoritmo de Powell (1964) y un algoritmo basado en simulated annealing, este último orientado a la validación rápida de la precisión de los resultados obtenidos por los demás métodos.

La validación definitiva de precisión se realiza mediante la técnica de Validación Empírica Residual, un procedimiento estadístico que permite evaluar el grado de precisión de una secuencia ordenada de valores numéricos.

Por otro lado, la herramienta de simulación, ya concluida, permite evaluar los algoritmos en escenarios dinámicos de distintas escalas, facilitando la comparación de la velocidad de cada método.

En cuanto al uso de inteligencia artificial, se desarrolló una arquitectura de red neuronal de tipo perceptrón multicapa con entrenamiento mediante Resilient Back-propagation. Esta red alcanza una generalización y precisión satisfactorias; sin embargo, su aplicación se encuentra restringida a una única escala de operación correspondiente a un centro de llamadas específico, lo que implica que el modelo debe ser entrenado con datos propios de dicho entorno. La principal ventaja de esta red neuronal radica en que no requiere operar de manera complementaria con un algoritmo de búsqueda.

Actualmente, se continúan realizando experimentos orientados a la integración combinada de algoritmo y red neuronal, conforme a los objetivos del proyecto.

4. FORMACION DE RECURSOS HUMANOS

La estructura del equipo de investigación es de 4 (cuatro) miembros incluidos el Director.

Uno de los miembros desarrolla la tesis de maestría en ingeniería de sistema de información (UTN-FRC) con tema dentro del marco del presente proyecto.

Otro de los miembros, desarrolla la tesina de la Especialidad en Ingeniería de Sistema de Información (UTN-FRT) con tema también en el marco del presente proyecto.

Ocasionalmente participan en el proyecto, alumnos que realizan su práctica supervisada. La cantidad de alumnos depende de las necesidades del proyecto y de los alumnos dispuestos a trabajar en el mismo.

5. BIBLIOGRAFIA

Afèche, P., Araghi, M., & Baron, O. (2017). Customer Acquisition, Retention, and Service Access Quality: Optimal Advertising, Capacity Level, and Capacity Allocation. Manufacturing & Service Operations Management, 19(4), 674. DOI: 10.1287/msom.2017.0635

Ahmadi, T., Atan, Z., Kok, A. G. de, & Adan, I. (2020). Time-based service constraints for inventory systems with commitment lead time. OR Spectrum, 42(2), 355. DOI: 10.1007/s00291-020-00590-0

Akşin, O. Z., Armony, M., & Mehrotra, V. (2007). The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research. Production and Operations Management, 16(6), 665. DOI: 10.1111/j.1937-5956.2007.tb00288.x

Chicu, D., Ryan, G., & Valverde-Aparicio, M. (2016). Determinants of customer satisfaction in call centres. European Accounting and Management Review, 2(2), 20. DOI: 10.26595/eamr.2014.2.2.2

Cordeiro, E. A., & Neto, A. R. P. (2023). Deep reinforcement learning for the dynamic vehicle dispatching problem: An event-based approach. arXiv (Cornell University). DOI: 10.48550/arxiv.2307.07508

Diao, Y., Heching, A., Northcutt, D., & Wallace, R. B. (2015). Service-Delivery Modeling and Optimization. INFORMS Journal on Applied Analytics, 45(3), 243. DOI: 10.1287/inte.2014.0783

Filippou, S., Tsiartas, A., Hadjineophytou, P., Christofides, S., Malialis, K., & Panayiotou, C. G. (2023). Improving Customer Experience in Call Centers with Intelligent Customer-Agent Pairing. arXiv (Cornell University). DOI: 10.48550/arxiv.2305.08594

Hu, L., & Dong, J. (2022). An Artificial-Neural-Network-Based Model for Real-Time Dispatching of Electric Autonomous Taxis. IEEE Transactions on Intelligent Transportation Systems, 23(2), 1519–1528. DOI: 10.1109/tits.2020.3029141

Kim, Y.-S., Krishnan, R., & Argote, L. (2012). The Learning Curve of IT Knowledge Workers in a Computing Call Center. Information Systems Research, 23, 887. DOI: 10.1287/isre.1110.0401

Кондрашова, Е. В. (2021). Optimization of Controlled Queueing Systems: the Case of Car Wash Services. Transportation Research Procedia, 54, 662. DOI: 10.1016/j.trpro.2021.02.119

Li, J., Zhang, L., Liu, B., Shi, N., Li, L., & Yin, H. (2023). Travel-Energy-Based Timetable Optimization in Urban Subway Systems. Sustainability, 15(3), 1930. DOI: 10.3390/su15031930

Marlière, G., Sobieraj Richard, S., Pellegrini, P., & Rodriguez, J. (2023). A conditional time-intervals formulation of the real-time Railway Traffic Management Problem. Control Engineering Practice, 133, 105430. DOI: 10.1016/j.conengprac.2022.105430

Mehrotra, V., Ross, K., Ryder, G., & Zhou, Y. (2011). Routing to Manage Resolution and Waiting Time in Call Centers with Heterogeneous Servers. Manufacturing & Service Operations Management, 14(1), 66. DOI: 10.1287/msom.1110.0349

Su, B., D'Ariano, A., Su, S., Wang, X., & Tang, T. (2023). Integrated train timetabling and rolling stock rescheduling for a disturbed metro system: A hybrid deep reinforcement learning and adaptive large neighborhood search approach. Computers & Industrial Engineering, 186, 109742. DOI: 10.1016/j.cie.2023.109742

Yang, M., Gao, X., & Qiu, J. (2024). Simulation of worker collaboration in production cell based on hybrid model under fluctuating environment. Flexible Services and Manufacturing Journal. DOI: 10.1007/s10696-024-09572-8

Ye, G., Zhu, X., & Mukhopadhyay, S. K. (2014). Managing Service Quality in Multiple Outsourcing. International Journal of Electronic Commerce, 18(3), 125. DOI: 10.2753/jec1086-4415180305

PIÑA 4.0: SISTEMA MULTIAGENTE INTELIGENTE PARA LA OPTIMIZACIÓN HÍDRICA Y CLIMÁTICA EN EL CULTIVO DE PIÑA MEDIANTE TECNOLOGÍAS LORA E INTELIGENCIA ARTIFICIAL

Luis Carlos Triviño Torres
Semillero de Investigación Ciberondas - Grupo de Investigación GRINTIC
Universidad Santo Tomás, Villavicencio, Meta, Colombia
luistrivino@usta.edu.co

RESUMEN

El cultivo de piña en Villavicencio, Meta, enfrenta desafíos críticos relacionados con la eficiencia hídrica y la variabilidad climática del piedemonte llanero. Este proyecto presenta el desarrollo de un sistema multi-agente inteligente denominado "Piña 4.0", que integra tecnologías IoT basadas en LoRaWAN, sensores distribuidos y algoritmos de inteligencia artificial para optimizar la gestión del riego y monitoreo climático. El sistema implementa una arquitectura de tres capas: sensorial (nodos LoRa con sensores de humedad de suelo, temperatura, humedad ambiental y precipitación), procesamiento (algoritmos de machine learning para predicción de necesidades hídricas) y actuación (control automatizado de válvulas de riego). Los resultados preliminares muestran potencial de ahorro hídrico superior al 40% manteniendo indicadores de calidad del fruto. La investigación contribuye al desarrollo de agricultura de precisión en la región de la Orinoquia Colombiana, promoviendo sostenibilidad y competitividad del sector piñero nacional.

Palabras clave: Sistemas multi-agente, IoT agrícola, LoRaWAN, Machine Learning, Riego de precisión, Cultivo de piña, Agricultura 4.0

CONTEXTO

Este proyecto se desarrolla en el marco del Semillero de Investigación Ciberondas, adscrito al Grupo de Investigación GRINTIC, Universidad Santo Tomás, con sede en Villavicencio, Meta, Colombia. La iniciativa surge como respuesta a la necesidad de optimizar la producción piñera en la región de la Orinoquía, donde el cultivo representa una actividad económica significativa, pero enfrenta limitaciones por manejo ineficiente del recurso hídrico y alta variabilidad climática característica del piedemonte llanero.

El proyecto cuenta con financiamiento de la Universidad Santo Tomás, con una duración de 36 meses (2024-2027). Participan activamente instituciones como la asociación de productores de piña del Meta y empresas tecnológicas regionales especializadas en soluciones IoT. La investigación se alinea con los objetivos de desarrollo sostenible de la ONU, específicamente ODS 2 (Hambre Cero), ODS 6 (Agua Limpia y Saneamiento), ODS 9 (Industria, Innovación e Infraestructura) y ODS 13 (Acción por el Clima).

1. INTRODUCCIÓN

La piña (Ananas comosus) cultivada en el piedemonte llanero de Colombia presenta características fisiológicas únicas debido a su metabolismo CAM (Metabolismo Ácido de las

Crasuláceas), que le confiere alta eficiencia en el uso del agua, pero requiere gestión precisa del riego para optimizar calidad y rendimiento del fruto. La región de Villavicencio, con precipitaciones anuales entre 2000-4000 mm distribuidas de manera bimodal, presenta períodos críticos donde el déficit hídrico afecta significativamente la producción.

Estudios recientes demuestran que los sistemas de riego inteligente basados en IoT pueden reducir el consumo de agua hasta en un 70% mientras incrementan rendimientos hasta en 30%. La tecnología LoRaWAN se destaca por su capacidad de comunicación de largo alcance (hasta 15 km en campo abierto) con bajo consumo energético, características ideales para implementación en extensas áreas de cultivo de piña donde la infraestructura celular es limitada.

Los sistemas multi-agente representan un paradigma computacional donde múltiples agentes autónomos (sensores, actuadores, módulos de decisión) colaboran para alcanzar objetivos comunes. En agricultura de precisión, esta arquitectura permite descentralizar la toma de decisiones, incrementar robustez del sistema ante fallos y adaptar estrategias de riego a condiciones microclimáticas específicas de cada sector del cultivo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1. Arquitectura del sistema multi-agente

El sistema Piña 4.0 se estructura en tres capas jerárquicas de agentes: (a) Agentes de sensado: nodos LoRa distribuidos en campo equipados con sensores de humedad volumétrica del suelo (0-40 cm de profundidad), temperatura y humedad ambiental (DHT22), radiación solar, y pluviómetros digitales. Cada nodo opera autónomamente con paneles solares de 10W y baterías de respaldo para garantizar operación continua; (b) Agentes de procesamiento: módulos de inteligencia artificial que implementan algoritmos de Machine Learning (Random Forest, XGBoost y LSTM) para predecir evapotranspiración del cultivo, calcular déficit hídrico y generar recomendaciones de riego específicas por zona; (c) Agentes de actuación: controladores de válvulas solenoides que ejecutan las decisiones de riego, integrando mecanismos de seguridad para prevenir sobre-riego y validación cruzada entre sensores.

2.2. Modelos predictivos de necesidades hídricas

Se desarrollan modelos de Machine Learning entrenados con datos locales de dos ciclos productivos completos (18 meses) que incluyen variables climáticas, fenología del cultivo, propiedades físico-químicas del suelo y rendimientos obtenidos. Los algoritmos Random Forest y XGBoost han mostrado precisiones superiores al 90% en predicción de agotamiento hídrico del suelo según validaciones en cultivos similares. Las redes LSTM permiten capturar patrones temporales complejos incorporando series históricas de precipitación y evapotranspiración de referencia calculada mediante el método FAO-56 Penman-Monteith adaptado a condiciones del piedemonte.

2.3. Protocolo de comunicación LoRaWAN optimizado

La infraestructura de comunicación utiliza el estándar LoRaWAN operando en la banda ISM 915 MHz (Región AU915 autorizada por ANE Colombia). Se implementan tres gateways

estratégicamente ubicados para garantizar cobertura completa del área experimental (15 hectáreas). Los nodos sensores transmiten paquetes de datos cada 15 minutos utilizando Adaptive Data Rate (ADR) que optimiza el balance entre consumo energético y confiabilidad de transmisión. Estudios previos demuestran que LoRaWAN puede alcanzar rangos de 10-15 km en terreno abierto con mínimo consumo energético, permitiendo vida útil de baterías superior a 3 años.

2.4. Integración de modelos fenológicos de piña

El sistema incorpora modelos específicos del desarrollo fenológico de piña bajo condiciones del trópico húmedo, considerando las cuatro fases críticas: establecimiento (0-4 meses), desarrollo vegetativo (4-12 meses), inducción floral y fructificación (12-18 meses), y maduración (18-22 meses). Cada fase presenta requerimientos hídricos diferenciados que el sistema multi-agente ajusta dinámicamente mediante coeficientes de cultivo (Kc) variables. Se integran además indicadores de estrés hídrico basados en relaciones hídricas del suelo-planta-atmósfera (modelo SPAC) para piña CAM.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1. Resultados obtenidos (fase piloto - 12 meses)

Durante la fase piloto implementada en 3 hectáreas, el sistema demostró reducción del 42% en consumo hídrico comparado con riego tradicional por calendario, mientras que los indicadores de calidad del fruto (grados Brix, acidez titulable, peso promedio) se mantuvieron en rangos óptimos para exportación. La red LoRaWAN operó con 99.2% de disponibilidad, con tasa de pérdida de paquetes inferior al 0.8%. Los modelos de Machine Learning alcanzaron precisión del 87% en predicción de humedad del suelo a 48 horas, superando métodos convencionales basados únicamente en evapotranspiración. Se documentó reducción del 35% en incidencia de pudrición del cogollo (phytophthora sp.) asociada a exceso de humedad.

3.2. Resultados esperados (24 meses restantes)

Se proyecta: (a) Escalar el sistema a 50 hectáreas en tres fincas productoras asociadas, validando robustez en condiciones heterogéneas de suelo y topografía; (b) Incrementar precisión de modelos predictivos al 92% mediante entrenamiento con datos de ciclo completo y refinamiento de algoritmos; (c) Integrar módulos de predicción de plagas y enfermedades basados en condiciones microclimáticas registradas; (d) Desarrollar aplicación móvil para productores que permita visualización en tiempo real, generación de alertas y ajuste manual de parámetros; (e) Obtener registro de software ante DNDA y solicitud de patente de modelo de utilidad para la arquitectura del sistema; (f) Publicar mínimo 3 artículos en revistas indexadas Q1-Q2 y presentar resultados en congresos internacionales de agricultura de precisión.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto contribuye significativamente a la formación de capital humano en la región. Participan 1 estudiante de Doctorado en ingeniería, 2 de Maestría en ingeniería electrónica y de computadores, 8 estudiantes de pregrado en Ingeniería informática, y 12

estudiantes activos en el Semillero Ciberondas desarrollando actividades de formación, diseño de prototipos, pruebas de campo y análisis de datos con reuniones semanales y talleres mensuales sobre IoT, Machine Learning aplicado y agricultura de precisión. Adicionalmente, se han realizado 3 cursos de extensión para productores piñeros de la región, capacitando a 45 agricultores en conceptos básicos de agricultura 4.0, interpretación de datos de sensores y uso de herramientas tecnológicas para gestión de cultivos.

5. REFERENCIAS BIBLIOGRÁFICAS

[1] Smart drip irrigation systems using IoT: a review of architectures, machine learning models, and emerging trends (2024). Discover Agriculture, Springer Nature. https://doi.org/10.1007/s44279-025-00430-1

[2] Scalable machine learning framework for adaptive irrigation management of maize and soybean (2025). Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2025.108166

[3] LoRaFarM: A LoRaWAN-Based Smart Farming Modular IoT Architecture (2020). PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC7180486/

[4] Machine learning and digital twins in smart irrigation: optimising water use (2025). Taylor & Francis. https://doi.org/10.1080/27525783.2025.2562418

[5] Dong, Y., Werling, B., Cao, Z., & Li, G. (2024). Implementation of an in-field IoT system for precision irrigation management. Frontiers in Water, 6. https://doi.org/10.3389/frwa.2024.1353597

[6] Integration of smart sensors and IOT in precision agriculture (2025). PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC12116683/

[7] A LoRaWAN IoT System for Smart Agriculture for Vine Water Status (2022). Agriculture, 12(10), 1695. https://doi.org/10.3390/agriculture12101695

[8] Internet of Things-Based Automated Solutions Using ML for Smart Irrigation (2024). PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC11644754/

[9] Smart Agriculture Revolution: How LoRaWAN is Transforming Farming (2025). Semtech. https://blog.semtech.com/smart-agriculture-revolution

[10] IoT Sensing for Advanced Irrigation Management: A Systematic Review (2024). PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC11991392/

Predicción de cortes en la matriz energética argentina con IA

Macarena Rocío Gorgal, Pablo Inchausti, Mathias Marcelo Del Vecchio y Franco Agustín Benito

Universidad Argentina de la Empresa, Instituto de Tecnología (UADE - INTEC), CABA, Argentina
{mgorgal, pinchausti, madelvecchio, fbenito}@uade.edu.ar

RESUMEN

Se propone una investigación para predecir la ocurrencia de interrupciones en el suministro energético mediante el entrenamiento de modelos de Machine Learning. Para el entrenamiento se utilizan datos de demanda y generación eléctrica provistos por CAMMESA, variables climáticas del Servicio Meteorológico Nacional y datos georreferenciados de generadores energéticos. Se implementan modelos predictivos basados en redes neuronales recurrentes, incluyendo las arquitecturas RNN, LSTM y GRU. La solución permite anticipar, con distinta granularidad, posibles interrupciones en la matriz energética de la República Argentina. Como parte de la validación, el prototipo será testeado en Pinamar y General Madariaga para ajustar el desempeño del modelo en entornos reales.

***Palabras Clave**:* Machine Learning, Energía, Sustentabilidad, Predicción, Redes Neuronales Recurrentes

CONTEXTO

La investigación se desarrolla dentro del Instituto de Tecnología (INTEC), dependiente de la Universidad Argentina de la Empresa (UADE).

El proyecto se inicia en el año 2026 como una Actividad de Ciencia y Técnica (ACyT) y se denomina *"Predicción de cortes en la matriz energética argentina con Inteligencia Artificial"*. Se identifica con el código A26T11 y tiene una duración prevista de 15 meses a partir del período 2026.

1. INTRODUCCIÓN

El proyecto desarrolla un sistema de predicción de cortes eléctricos basado en machine learning, integrando múltiples fuentes de datos históricas y actuales de todo el territorio argentino. [1]

El modelo utiliza información de fallas registradas, variables climáticas, patrones de consumo e infraestructura para estimar la probabilidad de interrupciones con alta resolución zonal. Esta granularidad permite generar alertas específicas para cada área dentro de una localidad, ofreciendo información precisa y accionable.

La plataforma brinda un servicio de utilidad para micro y pequeñas empresas —como restaurantes, hoteles, comercios o estaciones de servicio— que requieren anticiparse a cortes para minimizar pérdidas operativas. Su diseño escalable permitirá cubrir cualquier región del país, independientemente del tamaño urbano o rural.

Como parte de la validación del proyecto, se realizó una encuesta a 203 usuarios asociados al caso de uso de *coworking*, en la región de CABA y AMBA para evaluar el impacto y la utilidad de soluciones predictivas. El 83% de los encuestados reportó haber sufrido interrupciones en el suministro eléctrico, y cerca del 90% consideró fundamental contar con herramientas que permitan anticipar estos eventos. Estos resultados confirman la relevancia del problema abordado y respaldan el desarrollo de modelos de Machine Learning orientados a la generación de alertas tempranas y al fortalecimiento de la resiliencia operativa.

Objetivo:

Desarrollar modelos predictivos basados en técnicas de Machine Learning para anticipar interrupciones en el suministro energético mediante el análisis de la demanda, la generación eléctrica y variables climáticas.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de I+D del proyecto A26T11 se orientan al desarrollo de modelos predictivos mediante técnicas de Machine Learning para anticipar interrupciones en el suministro energético a partir del análisis de la demanda y la generación eléctrica, incluyendo:

a. La integración y análisis de datasets energéticos históricos de demanda y generación eléctrica;

b. El estudio de la influencia de variables climáticas y geográficas en el comportamiento del sistema energético;

c. El desarrollo y entrenamiento de modelos predictivos utilizando técnicas estadísticas y redes neuronales recurrentes (RNN, LSTM y GRU), incluyendo modelos como SARIMA y Prophet;

d. La implementación de técnicas de feature engineering temporal para capturar patrones estacionales y tendencias;

e. La evaluación y validación de los modelos mediante métricas estadísticas para medir su precisión predictiva.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se integraron fuentes de datos energéticos históricos de demanda y generación eléctrica provenientes de fuentes de datos abiertos oficiales [1], junto con variables climáticas y geográficas relevantes del Servicio Meteorológico Nacional (SMN) [2].

Se realizó la georreferenciación de generadores energéticos clasificados por tecnología (solar, eólica, hidráulica, térmica, biomasa) [3], así como de estaciones meteorológicas y variables asociadas (temperatura, viento, precipitaciones, caudal de ríos), permitiendo modelar la relación entre generación, demanda y condiciones ambientales.

En la **Figura 1** se muestran datasets georreferenciados utilizados por el modelo. Se incluye la localización de los generadores energéticos [3]. Los distintos colores indican la fuente de energía que se utiliza (solar, eólica, térmica, hidráulica, biomasa)

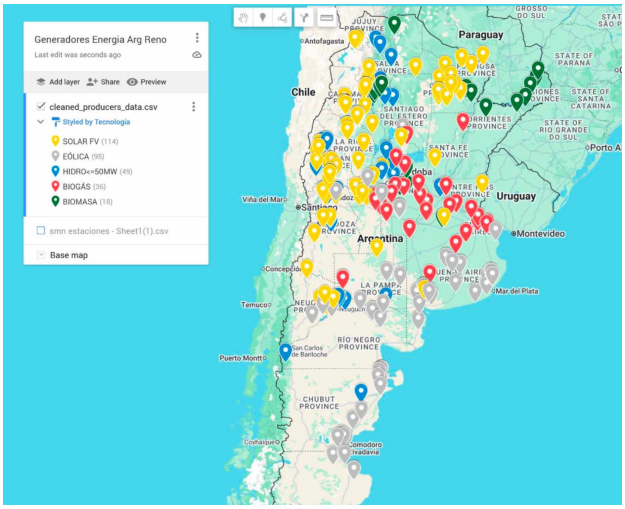


Figura 1: Mapa geolocalizado de generadores energéticos

A partir de estos datos, se procedió al entrenamiento y comparación de tres modelos predictivos basados en redes neuronales recurrentes: RNN, LSTM y GRU. [4][5]

En las **Figura 2, 3 y 4** se presentan la comparación de desempeño entre los modelos, evaluados mediante R^2, MAE y RMSE. Se obtuvieron valores de R^2 entre 0,30 y 0,45, con MAE cercano a 2000–2150 MWh y RMSE alrededor de 2600 MWh.

Estos resultados demuestran que los modelos recurrentes son capaces de capturar patrones temporales de la demanda energética, validando su uso para predicciones en la matriz energética argentina.

Respecto a su implementación, el prototipo será testeado en Pinamar y General Madariaga para ajustar el desempeño del modelo en entornos reales.



Figura 2: Gráfica de error RNN - Fuente: elaboración propia

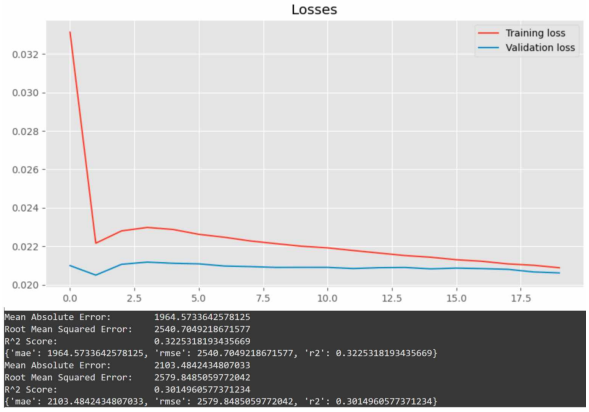


Figura 3: Gráfica de error LSTM - Fuente: elaboración propia



Figura 4: Gráfica de error GRU - Fuente: elaboración propia

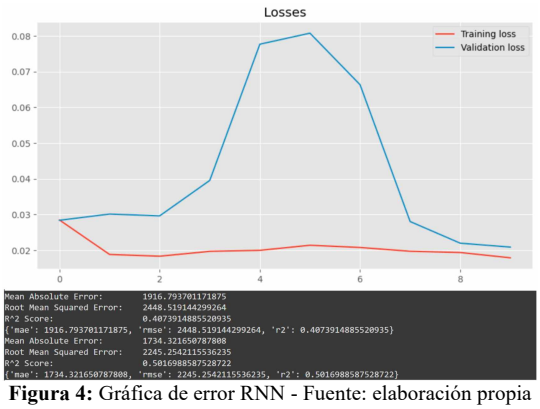


Figura 4: Gráfica de error RNN - Fuente: elaboración propia

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo del proyecto A26T11 está integrado por el investigador responsable, Mg. Pablo Ezequiel Inchausti, y la Ing. Macarena Gorgal, quien participa como tutora académica de los *Proyectos Finales de Ingeniería* (PFI) asociados a esta línea de investigación.

Actualmente, se encuentra integrado a la investigación un PFI a cargo de los estudiantes de Ingeniería en Informática, Mathias Del Vecchio y Franco Benito. El proyecto, titulado *"OfficePass: Metabuscador de espacios de coworking asistido por IA para predicción de cortes de suministro energético"*, aplica la problemática de la predicción de cortes de energía al segmento de los espacios de coworking en la región de CABA y AMBA.

Se prevé, durante el ciclo lectivo 2026, la incorporación de dos nuevos PFI con el objetivo de ampliar las líneas de investigación y desarrollar trabajos derivados.

Como parte de la formación de recursos humanos, se contemplan actividades de transferencia en asignaturas de las carreras de informática, utilizando como recursos didácticos los datasets, modelos y prototipos desarrollados en el marco de la investigación. Los prototipos se despliegan en la plataforma AWS, incorporando también contenidos formativos vinculados al uso de infraestructuras cloud.

5. BIBLIOGRAFÍA

[1] «Datos Argentina - Datos de la Compañía Administradora del Mercado Mayorista Eléctrico S.A. (CAMMESA)». Accedido: 20 de febrero de 2026. [En línea]. Disponible en: https://datos.gob.ar/

[2] «Servicio Meteorologico Nacional», Servicio Meteorologico Nacional. Accedido: 20 de febrero de 2026. [En línea]. Disponible en: https://www.smn.gob.ar/

[3] «Listado Generadores, Comerc. Distribuidores | CAMMESA». Accedido: 20 de febrero de 2026. [En línea]. Disponible

en: https://cammesaweb.cammesa.com/listado-generadores-comerc-distribuidores/

[4] K. Cho *et al.*, «Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation», 3 de septiembre de 2014, *arXiv*: arXiv:1406.1078. doi: 10.48550/arXiv.1406.1078.

[5] S. Hochreiter y J. Schmidhuber, «Long Short-Term Memory», *Neural Comput.*, vol. 9, n.º 8, pp. 1735-1780, nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

Reconstrucción de trayectorias bovinas mediante seguimiento visual multiobjeto en escenarios extensivos

Marcos D. Aranda[1,2], Emilia Mariel Mercado[1] , Eduardo Piray[1], Elian Gonzalez[1], Gustavo Futo[1] , Ricardo Daniel Macjus[1]
marcos.aranda@ues21.edu.ar
(1) Universidad Siglo 21. Decanato de Ciencias Aplicadas.
(2) Universidad Nacional de Catamarca. Facultad de Tecnología y Ciencias Aplicadas.

AREA TEMATICA: Agentes y Sistemas Inteligentes.

RESUMEN

El seguimiento visual automático de animales en sistemas de ganadería extensiva presenta desafíos particulares asociados a oclusiones, variabilidad ambiental y cambios en la escala de observación. En este trabajo se presenta una experiencia de seguimiento multiobjeto aplicada al monitoreo visual de bovinos en escenarios reales de campo, orientada a la reconstrucción de trayectorias individuales a partir de secuencias de video.
La propuesta combina modelos de detección basados en redes neuronales profundas con algoritmos de seguimiento en línea, permitiendo mantener la identidad temporal de los animales a lo largo de múltiples frames. A partir de las trayectorias reconstruidas, se analizan métricas asociadas a la consistencia del seguimiento, estabilidad de las identificaciones y continuidad espacial del movimiento.
Los resultados obtenidos muestran que el enfoque permite reconstruir trayectorias coherentes incluso en condiciones de oclusión parcial y variabilidad del entorno, constituyendo una base sólida para el análisis automático del comportamiento animal y el desarrollo de aplicaciones de ganadería de precisión basadas en visión por computadora.

Palabras Claves: Seguimiento multiobjeto, Trayectorias, Visión por computadora, Ganadería de precisión.

CONTEXTO

La ganadería extensiva en el Noroeste Argentino (NOA) se caracteriza por sistemas productivos de gran escala, baja intervención tecnológica y condiciones ambientales heterogéneas, lo que dificulta la implementación de esquemas de monitoreo sanitario continuo. En este contexto, las alteraciones locomotoras particularmente la cojera constituyen uno de los principales indicadores tempranos de enfermedad, impacto en el bienestar animal y disminución del rendimiento productivo.
Los métodos tradicionales de detección de problemas locomotores se basan mayormente en inspecciones visuales manuales realizadas por personal especializado. Este enfoque presenta limitaciones significativas en términos de subjetividad, frecuencia de observación y escalabilidad, especialmente en sistemas extensivos donde el seguimiento individual resulta complejo. En consecuencia, existe una necesidad creciente de herramientas automáticas, objetivas y no invasivas que permitan evaluar el comportamiento locomotor de los animales de manera continua.
En este escenario, el presente trabajo se inscribe en una línea de investigación orientada al desarrollo de sistemas automáticos de monitoreo locomotor bovino basados en visión por computadora, adaptados a las condiciones reales de la ganadería extensiva regional. Esta línea integra investigación aplicada, desarrollo tecnológico y validación experimental, y ha generado resultados parciales presentados en congresos nacionales, constituyendo una base sólida para la evolución hacia sistemas inteligentes de detección temprana de enfermedades y apoyo a la gestión ganadera.

1. INTRODUCCIÓN

En los últimos años, la visión por computadora se ha consolidado como una herramienta clave para la detección, identificación y seguimiento de animales en sistemas productivos, permitiendo el análisis automático del comportamiento sin la necesidad de dispositivos invasivos [1], [2]. En particular, el seguimiento visual multiobjeto ha permitido mantener la identidad temporal de los individuos y reconstruir trayectorias a partir de secuencias de video, habilitando estudios longitudinales del movimiento animal.

Diversos trabajos han demostrado que el análisis de trayectorias visuales permite extraer métricas relevantes, tales como distancia recorrida, velocidad promedio y continuidad del movimiento, que aportan información objetiva sobre el comportamiento y la dinámica espacial de los animales [2], [3]. Estos enfoques han sido ampliamente explorados en sistemas intensivos o controlados; sin embargo, su aplicación en entornos extensivos continúa siendo un desafío abierto.

Los avances recientes en modelos de detección basados en redes neuronales profundas, como la familia YOLO, han mejorado significativamente la capacidad de detectar animales en tiempo real, incluso en condiciones complejas de iluminación y oclusión parcial [4]. Complementariamente, algoritmos de seguimiento multiobjeto en línea, como SORT, han permitido asociar detecciones entre frames consecutivos y reconstruir trayectorias individuales de forma eficiente [5].

No obstante, en escenarios extensivos, la variabilidad ambiental y las interacciones entre animales afectan la estabilidad del seguimiento, generando pérdidas temporales de identidad o fragmentación de trayectorias. En este contexto, resulta relevante evaluar el desempeño y la consistencia de los algoritmos de seguimiento visual cuando se aplican a datos reales de campo, con el objetivo de identificar sus fortalezas, limitaciones y potencial de mejora.

En este trabajo se presenta una experiencia de seguimiento visual multiobjeto aplicada a bovinos en escenarios extensivos reales, orientada a la reconstrucción de trayectorias individuales y al análisis de la consistencia del seguimiento como componente fundamental para aplicaciones de ganadería de precisión basadas en visión por computadora.

2. LINEAS DE INVESTIGACION Y DESARROLLO

L La línea de investigación se orienta al desarrollo y validación de técnicas de seguimiento visual multiobjeto aplicadas al monitoreo de animales en sistemas de ganadería extensiva, con énfasis en la reconstrucción de trayectorias individuales y el análisis de la estabilidad del seguimiento en condiciones reales de campo.

El objetivo general de la línea es evaluar la factibilidad y el desempeño de algoritmos de detección y seguimiento visual para la identificación continua de bovinos a lo largo del tiempo, considerando los desafíos propios de entornos no controlados.

Los objetivos específicos incluyen:

- Diseñar e implementar un sistema de detección y seguimiento visual automático de bovinos basado en visión por computadora.
- Reconstruir trayectorias individuales a partir del seguimiento multiframe, manteniendo la identidad temporal de los animales.
- Analizar métricas asociadas a la consistencia del seguimiento, tales como continuidad de trayectorias y estabilidad de identificaciones.
- Evaluar el impacto de oclusiones, densidad animal y variabilidad ambiental sobre el desempeño del seguimiento.
- Sentar las bases metodológicas para el análisis automático del comportamiento animal a partir de trayectorias visuales.

La línea articula investigación aplicada y validación experimental, con proyección hacia

sistemas de ganadería de precisión basados en análisis visual automático.

3. RESULTADOS PARCIALES

Como resultados parciales de esta línea, se desarrolló un sistema de seguimiento visual multiobjeto evaluado sobre secuencias de video adquiridas en establecimientos ganaderos reales. A partir de la detección automática de bovinos y su asociación temporal mediante algoritmos de seguimiento en línea, se logró reconstruir trayectorias individuales en escenarios con múltiples animales en movimiento.

El análisis de las trayectorias obtenidas permitió identificar patrones de continuidad espacial y temporal, así como situaciones de pérdida y recuperación de identidad asociadas principalmente a oclusiones parciales y cruces entre animales. En términos generales, el sistema mostró una capacidad consistente para mantener la identidad de los individuos durante desplazamientos prolongados, incluso en condiciones ambientales variables.

Asimismo, se observaron diferencias en la estabilidad del seguimiento en función de la densidad animal y la complejidad del escenario, lo que permitió identificar aspectos críticos para la mejora del algoritmo y la necesidad de estrategias complementarias para reforzar la continuidad de las trayectorias.

Estos resultados validan la viabilidad del seguimiento visual multiobjeto como herramienta base para la reconstrucción de trayectorias bovinas en sistemas extensivos y constituyen un insumo fundamental para el desarrollo futuro de aplicaciones de análisis del comportamiento animal.

4. FORMACIÓN DE RECURSOS HUMANOS

l equipo de trabajo lo forman 7 integrantes, 2 docentes de la carrera Licenciatura en Informática pertenecientes al Área de Ciencias Aplicadas de la Universidad Siglo 21, 1 docente de la carrera de Ingeniería Electrónica de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca y en esta ocasión se incorporaron 4 alumnos de la Universidad Siglo 21, para iniciarlos en su formación de investigadores.

Este proyecto de investigación posee una diversidad notable de integrantes, ya que hay docentes de diferentes asignaturas de las mencionadas carreras, en las áreas de arquitectura de computadoras, sistemas operativos, sistemas operativos avanzados y física, por lo que se cuenta con un conocimiento en el área de investigación.

En cuanto a la formación de los integrantes, hay un docente realizando su Tesis de Posgrado el cual es un Doctorando en Ingeniería en formación en temas de desarrollo tecnológico, internet de las cosas (IoT) e inteligencia artificial y otro en área de educación en formación que poseen experiencia de divulgación científica.

5. BIBLIOGRAFIA

[1] H. Meng, L. Zhang, F. Yang, L. Hai, Y. Wei, L. Zhu y J. Zhang, “Livestock biometrics identification using computer vision approaches: A review,” Agriculture, vol. 15, no. 1, p. 102, 2025. doi:10.3390/agriculture15010102.

[2] P. Garcia, J. Smith y A. Lee, “Real-time animal tracking with deep learning: A case study on cattle monitoring,” Sensors, vol. 24, p. 6486, 2024. doi:10.3390/s2406486.

[3] M. D. Aranda, J. Redolfi y G. Araguás, “Análisis del recorrido de bovinos mediante seguimiento visual automático,” en XX Reunión de Trabajo en Procesamiento de la Información y Control (RPIC), 2025.

[4] C.-Y. Wang, A. Bochkovskiy y H.-Y. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” arXiv:2207.02696, 2022.

[5] A. Bewley, Z. Ge, L. Ott, F. Ramos y B. Upcroft, “Simple online and realtime tracking,” en Proc. IEEE ICIP, 2016, pp. 3464–3468.

RED NEURONAL CONVOLUCIONAL PREENTRENADA PARA CLASIFICAR PARTICULAS DE ARENA

Mg. Ing. Carlos Gustavo Rodriguez Medina [1], Dr. Ing. Oscar Daniel Chuk [2],

Instituto de Investigaciones Mineras / Facultad de Ingeniería /
Universidad Nacional de San Juan

Av. Libertador Gral. San Martin 1109 oeste. San Juan
0264-4211700 (int. 285 [1], int. 389 [1,2])

grodriguez@unsj.edu.ar [1], dchuk@unsj.edu.ar [2]

RESUMEN

En este trabajo se hace uso de una red neuronal convolucional para conocer qué tan apropiada es para determinar la redondez en partículas de arena; siendo la redondez una importante característica en la determinación de la calidad para su utilización en la industria. Utilizar este tipo de redes neuronales ofrece la posibilidad de operar directamente sobre las imágenes digitales de partículas de arena, sin la necesidad de tener que extraer previamente características o medidas alternativas de las imágenes.

Se exponen resultados en la utilización de una red neuronal convolucional pre-entrenada, como alternativa a realizar un diseño propio de red. Se hace uso de la red *SqueezeNet* en el entorno Matlab.

Resulta de importancia poder clasificar las partículas, en cuanto a su redondez, dado que estas arenas se utilizan en procesos de fracturación hidráulica para la extracción de hidrocarburos de manera no convencional. Tales yacimientos no convencionales requieren de la inyección de arenas que actúan como apuntalantes de la fractura en las rocas. La calidad de las arenas para cumplir dicha función se evalúa por la norma API19C, que establece medidas geométricas – esfericidad y redondez- más un porcentaje de fractura cuando la muestra es sometida a compresión. Todas estas mediciones deben encontrarse dentro de valores límites para asegurar la calidad de la arena.

El método indicado por la norma (tradicional no automatizado) para la medición de redondez tiene una alta dependencia del criterio del observador o laboratorista, que analiza una muestra mediante un microscopio eligiendo un número muy reducido de partículas al azar, y comparándolas visualmente con formas graficas establecidas en una cartilla.

La aplicación de redes neuronales convolucionales, y de otras técnicas de Inteligencia Artificial, resultan una alternativa muy interesante respecto de cómo se vienen llevando a cabo dichas labores.

Palabras clave: Redes neuronales convolucionales, Transfer learnimg, Partículas de arena, Redondez, Fracturación hidráulica.

CONTEXTO

El trabajo representa de manera resumida resultados del Proyecto de Investigación Científica (PIC – UNSJ) "*Deep learning para el análisis de calidad de partículas de arena*", ejecutado durante el periodo 2023 – 2025, financiado por la Universidad Nacional de San Juan.

Dicho proyecto se inscribe dentro de la línea de investigación que lleva adelante un grupo de trabajo en la temática de Inteligencia y Visión Artificial orientado al control automático de procesos mineros, en el

Instituto de Investigaciones Mineras de la Facultad de Ingeniería de la UNSJ.

1. INTRODUCCIÓN

En la industria de extracción de petróleo de manera no convencional, y con la finalidad de garantizar la capacidad de las arenas con destino a los procesos de fracturación hidráulica (*Fracking*), para actuar como apuntalantes o propantes, se reconocen como base las normas de referencia del *American Petroleum Institute* (API) 19C [1, 2]. Esta norma recomienda una serie de ensayos a realizar sobre las arenas con el fin de conocer la calidad del material a utilizarse en el proceso. Entre los ensayos a aplicarse se encuentran los de redondez y esfericidad, entre otros.

Tales ensayos padecen de una fuerte dependencia del criterio del observador, dado que el procedimiento se aplica mediante la utilización de un microscopio y las mediciones de redondez y esfericidad se realizan por comparación respecto a tablas que contienen una serie de formas provistas por la norma, mediante unas cartas graficas conocidas como tablas o cartillas de *Krumbein y Sloss* [3]. Esto deriva en resultados sustancialmente disímiles que invalidan el procedimiento, aunque es el que se sigue usando por norma.

El uso de técnicas de Inteligencia Artificial se presenta como la alternativa tecnológica que puede permitir obtener medidas objetivas no dependientes del criterio de un operador [1]. En este sentido el grupo de investigación ha elaborado varios algoritmos con diversos objetivos aplicados al procesamiento digital de partículas que van desde la segmentación hasta el cálculo de medidas geométricas, tales como el de esfericidad, redondez, tamaño, distribución granulométrica, etc. [1, 4]

Las partículas pueden presentar entre ellas distintas variaciones, pudiendo ser algunas más redondeadas, o con más puntas, más o menos achatadas, más o menos alargadas, de diferentes tamaños, etc.

Dado que esta situación que se presenta, es un problema relacionado con la forma de las partículas, se sabe que existen herramientas de la inteligencia artificial orientadas a resolver problemas de reconocimiento de formas, patrones y su clasificación; éstas son las redes neuronales artificiales [4, 9, 10].

Además, hay un tipo particular de redes neuronales que permiten operar directamente sobre un conjunto de imágenes, sin necesidad de tener que realizar ningún preprocesado para la extracción de las características de interés, estas son las redes neuronales convolucionales [5, 8].

Las redes neuronales convolucionales son especialmente útiles para localizar patrones en imágenes con el objetivo de reconocer objetos, formas, etc. Se puede decir que aprenden directamente a partir de los datos de las imágenes [4].

Al igual que otras redes neuronales, una red neuronal convolucional se compone de una capa de entrada, una capa de salida y muchas capas ocultas intermedias. A su vez, las capas intermedias ocultas se organizan de forma alternada con capas convolucionales y de pooling, finalizando con una serie de capas de neuronas completamente conectadas (*Full connected*) [5].

Cuando se trabaja con redes neuronales convolucionales se puede optar por realizar el diseño e implementación de la red de forma completa, llevando a cabo un diseño específico para la resolución del problema planteado, lo que implica también aplicar todo el proceso de entrenamiento y validación desde cero. En este caso la red no tiene ningún conocimiento previo.

Otra alternativa es tomar una red neuronal preexistente, diseñada e implementada por un tercero, la cual ya se encuentra entrenada con un numero elevado de imágenes, y utilizarla. En esta opción se aprovecha el conocimiento adquirido con su entrenamiento previo, y lo que se hace es usarla así, o realizar algún entrenamiento adicional para la tarea especifica con algún conjunto de datos relacionados al problema a resolver. La ventaja que esta alternativa parece presentar es que el trabajo de entrenamiento debiera ser menor, tanto en insumo de tiempo, como de esfuerzo computacional y requerir menores

cantidades de datos para su entrenamiento. Conocido también como *Transfer Learning* (transferencia de aprendizaje) se refiere al proceso de adaptar un modelo de aprendizaje profundo previamente entrenado a una nueva tarea [14]. Dado que toda la programación de los algoritmos desarrollados en el proyecto de investigación se ha realizado en Matlab, como así también las implementaciones propias de redes neuronales convolucionales, también se a utilizado el mismo entorno para explorar el *Transfer Leraning* utilizando la red neuronal convolucional pre-entrenada *SqueezeNet*. Esta es una red neuronal convolucional lanzada en 2016 y fue desarrollada por investigadores de DeepScale, de la Universidad de California, Berkeley y la Universidad de Stanford [17]. Es una de las tantas redes pre-entrenadas que se encuentran disponibles en la actualidad.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el Laboratorio de Control Automático del Instituto de Investigaciones Mineras (Facultad de Ingeniería de la Universidad Nacional de San Juan) se realiza investigación y desarrollo respecto a la automatización de procesos de molienda de minerales aproximadamente desde el año 2000. Desde hace algunos años se viene desarrollando el Área de Procesamiento de Imágenes.

Actualmente se está trabajando en la caracterización de arenas especiales (calidad, forma, color, tamaños, clasificación del tipo de material, etc.) mediante el procesamiento de imágenes y algunas técnicas de IA para ser utilizadas en la extracción de petróleo mediante el método de *Fraking*.

3. RESULTADOS OBTENIDOS

Dado a que previamente se ha utilizado Matlab para la codificación de algoritmos, implementación de redes neuronales y sus respectivos entrenamientos; para el caso de la versión de 2020 de Matlab se dispone de la red pre-entrenada *SqueezeNet*. De la misma se puede utilizar su estructura sin entrenamiento, o con entrenamiento previo. En este caso se ha utilizado la red ya entrenada haciéndose uso de los pesos sinápticos obtenidos previamente. Tal entrenamiento es realizado en forma previa, por otros desarrolladores, con millones de imágenes disponibles, para este caso mediante el *DataSet* de *imageNet* [15].

Tal como se mencionó en apartados anteriores, lo que se desea determinar de cada imagen, mediante la red neuronal seleccionada, es la redondez de las partículas de arena que se encuentran bajo análisis.

De acuerdo a las tablas de *Krumbein* y *Sloss* [3], la redondez puede tomar valores entre 0.1 y 0.9, ver figura 1.

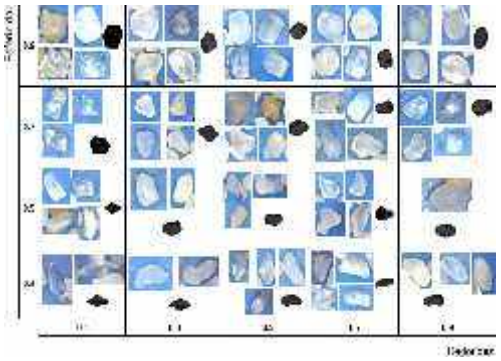


Figura 1. Cartilla de Krumbein y Sloss.

Se han elaborado varios *DataSets*, integrados por categorías de imágenes cuyas redondeces varían entre 0,1 a 0,9, Todo esto con el fin de llevar a cabo las diversas pruebas.

Se ha llevado a cabo la exploración del comportamiento de esta red neuronal convolucional con DataSet de distintas cantidades de imágenes por categoría, siendo 9 categorías en total, y cada una correspondiéndose a un valor de redondez de partícula.

Se ha tenido que adaptar mínimamente la estructura de *SqueezeNet* para el presente trabajo, dado que originalmente la red puede clasificar entre 1000 categorías diferentes de imágenes, mientras que para este caso en particular solo se necesita discriminar entre 9 categorías (los 9 valores posibles de redondez de la partícula de arena bajo análisis).

Matlab dispone de una herramienta, tanto para diseñar, como para utilizar una red prediseñada y pre-entrenada, denominada *Deep Network Designer* [16]. Mediante esta

se ha podido introducir tales modificaciones mínimas a la estructura de la red neuronal.
Dado que se hace la exploraron de resultados obtenidos (y en especial de la precisión), al utilizar esta red neuronal pre-entrenada, aplicando transferencia de conocimiento, es que se llevaron a cabo pruebas entrenándola con variados conjuntos de datos de imágenes de partículas de arena, con distintas cantidades. De alguna manera tal entrenamiento con datos específicos referidos a las partículas de arena, es con el fin de ampliar el conocimiento previo de la red, y adaptarlo a este problema en concreto (reconocimiento de redondez de partículas). Se debe tener en cuenta que el entrenamiento previo de la red es importante, pero a la vez es muy amplio y general, dado que *imageNet* dispone de imágenes como (teclado, lápiz, mouse, animales, etc.).

Imag. x Categ.	Presición
500	11,21%
1000	16,76%
2000	40.33%
4000	69,78%
8000	72,89%
16000	76.83%
32000	79.13%

Tabla 1. *DataSet* de imágenes y resultados obtenidos de precisión de la red neuronal.

Se han elaborado *DataSet* con imágenes de partículas de arena de distintos tamaños, pero todos con 9 categorías de redondez. También se ha tenido que ajustar el tamaño de las imágenes que se disponían al tamaño y características utilizadas por *SqueezeNet*, es decir, a imágenes de 227x227 RGB. Se realizaron *DataSet* de 100, 500, 1000, 2000, 4000, 8000, 16.000 y 32.000 imágenes por categoría. Se han aplicado técnicas de *Data Augmentation*, tales como rotaciones, flip, etc., mediante la implementación de algoritmos específicos que se desarrollaron en Matlab. La PC utilizada tanto para aplicar el *Data Augmentation* como el entrenamiento y pruebas con la red neuronal es una PC de escritorio con las siguientes características: microprocesador i7 7700 3.6GHz, memoria RAM de 32Gb y placa grafica dedicada GeForce GTX 1070.

Los resultados de entrenar la red neuronal convolucional, *SqueezeNet* pre-entrenada (de sumarle al entrenamiento previo), con imágenes específicas del trabajo, se presentan en la tabla 1. En la misma se aprecia la precisión obtenida de la red *SqueezeNet*, para el entrenamiento con cada uno de los *DataSet* con diferentes cantidades de imágenes por categoría. El entrenamiento se llevó a cabo durante 30 épocas para todos los casos. Se aprecia que la precisión mayor obtenida es de 79,13% para un número elevado de imágenes de partículas de arena en cada una de las 9 categorías. En la figura 2 se presenta en forma gráfica dicho entrenamiento.

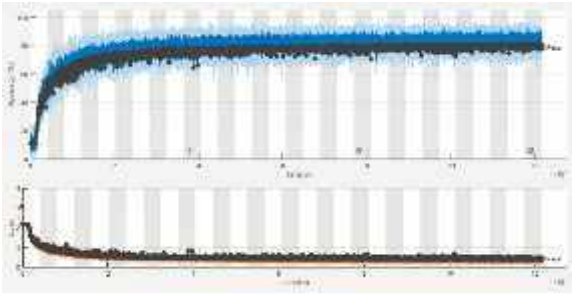


Figura 2. Entrenamiento y presión obtenida de la red neuronal.

Para este caso cuya precisión obtenida fue de 79.13%, se han requerido 37 horas para el entrenamiento de la red neuronal, con 32.000 imágenes RGB de 227x227 por cada categoría de redondez.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por Ingenieros Electrónicos, Electromecánicos, en Minas, Metalurgistas y Geólogos, con formación diversa en posgrados, tal como Doctorado en sistemas de control, Maestría en informática, Especialización en gestión y vinculación tecnológica, etc.
Actualmente, uno de los integrantes del equipo de trabajo se encuentra desarrollando su Tesis de Doctorado en Ingeniería de Procesamiento de Minerales, siendo su director de Tesis otro de los integrantes del grupo de investigación.

5. BIBLIOGRAFIA

[1] Rodriguez Medina, C.G.; Chuk, O.D.; Bertero R.; Luna, A.; Núñez, E.; Trigo, P. Visión artificial aplicada a la determinación de propiedades geométricas de arenas de fracturación para el análisis de calidad. Informe de Final de Proyecto PIC CICITCA. Universidad Nacional de San Juan. 2019.

[2] American Petroleum Institute API 19C, Second Edition: Measurement of Properties of Proppants Used in Hydraulic Fracturing and Gravel-Packing Operations; API Publishing Services: Washington, D.C., 2008

[3] Krumbein, W.C.; Sloss, L.L. Stratigraphy and Sedimentation; 2nd ed.; W.H. Freeman: San Francisco, 1963;

[4] Rodriguez Medina, C. G.; Chuk, O.D.; Bertero, R. Luna, A.; Núñez, E.. Redes neuronales para la clasificación de partículas de arena. WICC 2021. XXIII Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional de Chilecito. Pág. 149 – 154. 2021. ISBN 978-987-24611-3-3.

[5] Mathworks. Redes neuronales convolucionales. 1994-2022. https://la.mathworks.com/discovery/convolutional-neural-network-matlab.html

[6] Charoenpong, T.; Chamnongthai, K.; Kamhom, P.; Krairiksh, M. Volume Measurement of Mango by Using 2D Ellipse Model. In Proceedings of the 2004 IEEE International Conference on Industrial Technology, 2004. IEEE ICIT '04.; December 2004; Vol. 3, pp. 1438-1441.

[7] Rodriguez Medina, C.G.; Chuk, O.D.; Bertero, R. Luna, A.; Núñez, E. Generación gráfica de partículas artificiales de arena para su utilización en redes neuronales. WICC 2020. XXII Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional de la Patagonia Austral. Pág. 307 – 312. 2020. ISBN 978-987-3714-82-5.

[8] Liang, Z.; Nie, Z.; An, A.; Gong, J.; Wang, X. A Particle Shape Extraction and Evaluation Method Using a Deep Convolutional Neural Network and Digital Image Processing. *Powder Technology* Pag. 156-170. 2019. doi:10.1016/j.powtec.2019.05.025.

[9] Jain, A. K. Statistical Pattern Recognition: a review. IEEE Transactions on Pattern Análisis and Machine Inteligence. Vol 22, Nro. 1. 2000.

[10] Haykin S. Neural Networks - A comprehensive foundation, IEEE Press - Macmillan College Publishing Company, Inc. 1994.

[11] Hryciw, R.D.; Zheng, J.; Shetler, K. Particle Roundness and Sphericity from Images of Assemblies by Chart Estimates and Computer Methods. *Journal of Geotechnical and Geoenvironmental Engineering* 2016, *142*, 04016038, doi:10.1061/(ASCE)GT.1943-5606.0001485.

[12] Hernandez, E.M.; Chavez, G.M.; Hernandez, J.V. Roundness Estimation of Sedimentary Rocks Using Eliptic Fourier and Deep Neural Networks. In Proceedings of the 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC); IEEE: Ixtapa, Mexico, November 4 2020; pp. 1–5.

[13] Wadell, H. Volume, Shape, and Roundness of Rock Particles. *Jour. Geol.* 1932, *XL, N°5*, 443–451.

[14] Matlab. Tranfer Learning. https://la.mathworks.com/discovery/transfer-learning.html

[15] imageNet. https://www.image-net.org/

[16] *Deep Network Designer*. https://la.mathworks.com/help/deeplearning/ref/deepnetworkdesigner-app.html

[17] SqueezeNet. https://en.wikipedia.org/wiki/SqueezeNet.

Simulación inteligente de toma de decisiones mediante sistemas adaptativos basados en agentes

Silvia Rosana Hoferek [1,3], Marcos D. Aranda[1,2], Luis Bravo [1], Gianotti Martín Gastón [1,]
Gónzalez Ricardo Elian [1]
srhoferek@gmail.com - marcos.aranda@ues21.edu.ar
(1) Universidad Siglo 21. Decanato de Ciencias Aplicadas.
(2) Universidad Nacional de Catamarca. Facultad de Tecnología y Ciencias Aplicadas.
(3) Instituto de Investigaciones Científicas. Universidad de la Cuenca del Plata.

AREA TEMATICA: Agentes y Sistemas Inteligentes.

RESUMEN

La toma de decisiones en contextos complejos constituye un problema central en múltiples dominios, incluyendo entornos educativos, organizacionales e institucionales. Estos escenarios suelen caracterizarse por la presencia de información incompleta, consecuencias diferidas y múltiples alternativas posibles, lo que dificulta su modelado mediante enfoques deterministas clásicos.
En este contexto, los sistemas inteligentes basados en simulación han emergido como una herramienta relevante para representar procesos de decisión complejos de manera controlada y reproducible. En particular, los entornos simulados permiten analizar el comportamiento de los usuarios frente a dilemas estratégicos y éticos sin exponerlos a riesgos reales, facilitando la observación de patrones de decisión y la evaluación de respuestas adaptativas del sistema.
El presente trabajo se inscribe en una línea de investigación orientada al desarrollo de sistemas inteligentes de simulación, capaces de modelar procesos de decisión mediante estructuras adaptativas basadas en agentes, reglas y mecanismos de retroalimentación dinámica.

Palabras Claves: Inteligencia artificial, Visión computacional, Aprendizaje automático.

CONTEXTO

La inteligencia artificial ha sido ampliamente utilizada para modelar procesos de decisión en contextos dinámicos, incorporando técnicas como árboles de decisión, lógica difusa y sistemas basados en reglas [1], [2]. Estos enfoques permiten representar escenarios donde las decisiones no son binarias y sus consecuencias pueden manifestarse de manera progresiva o diferida.
En los últimos años, diversos estudios han destacado el potencial de los sistemas de simulación interactiva para analizar comportamientos complejos, al permitir que los usuarios exploren múltiples trayectorias de decisión y observen el impacto acumulativo de sus elecciones [3], [4]. Desde esta perspectiva, los sistemas inteligentes pueden comportarse como agentes que interpretan el estado del entorno y adaptan sus respuestas en función del historial de interacciones.
La incorporación de mecanismos adaptativos en simuladores de toma de decisiones permite generar experiencias dinámicas, donde variables como dificultad, acceso a eventos o retroalimentación se ajustan automáticamente según el comportamiento del usuario [5], [6]. Este tipo de sistemas resulta especialmente adecuado para estudiar procesos de decisión en escenarios estratégicos y éticos, donde la ambigüedad y la incertidumbre son componentes centrales.
En este trabajo se presenta una experiencia de desarrollo de un sistema inteligente de simulación orientado a la toma de decisiones, basado en una arquitectura adaptativa que combina agentes, árboles de decisión y lógica difusa. El foco se sitúa en el diseño del sistema y en la evaluación de su comportamiento funcional como plataforma de simulación inteligente.

1. INTRODUCCIÓN

Actualmente, la detección de objetos es fundamental en muchas aplicaciones en la visión por computadora. En entornos agrícolas y ganaderos, la identificación y seguimiento preciso de animales y objetos es esencial para la gestión efectiva de los recursos, el monitoreo de la salud animal y la automatización de las tareas críticas [1] [6]. Una red neuronal puede ser utilizada para detectar objetos en tiempo real, sin embargo, los avances tecnológicos permiten introducir mejoras significativas respecto a la precisión en la detección de objetos, estas mejoras en precisión a menudo se acompañan de un aumento de requisitos computacionales, lo que puede limitar su aplicabilidad en escenarios con pocos recursos. El objetivo de este trabajo es explorar una serie de técnicas y enfoques de optimización que busquen minimizar el costo computacional en la detección de objetos, sin perder precisión. Los experimentos que se esperan abordar incluyen la optimización de hardware, el ajuste de hiperparámetros y el uso de arquitecturas de redes neuronales ligeras de ser necesario [7]. Con el desarrollo de este trabajo se espera continuar avanzando en el estudio del aprendizaje automático para el reconocimiento de ganado.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La línea de investigación se orienta al diseño y validación de sistemas inteligentes de simulación para la modelización de procesos de toma de decisiones en contextos complejos.
El objetivo general es desarrollar entornos simulados capaces de adaptarse dinámicamente al comportamiento del usuario, manteniendo coherencia interna y permitiendo el análisis de trayectorias de decisión.
Los objetivos específicos incluyen:

- ✓ Diseñar una arquitectura de simulación basada en agentes para la representación de procesos de decisión.
- ✓ Implementar mecanismos adaptativos mediante árboles de decisión, reglas y lógica difusa.
- ✓ Analizar la coherencia y estabilidad del sistema frente a múltiples trayectorias de decisión.
- ✓ Evaluar el comportamiento funcional del sistema como plataforma de simulación inteligente.
- ✓ Sentar las bases para futuras aplicaciones en contextos educativos, organizacionales o institucionales.

Esta línea articula investigación aplicada en inteligencia artificial, simulación y sistemas adaptativos.

3. RESULTADOS ESPERADOS

Como resultado parcial de la línea, se desarrolló un prototipo funcional de simulación inteligente que permite modelar escenarios de decisión con múltiples alternativas y consecuencias diferidas. El sistema integra un motor de decisiones basado en agentes, un módulo adaptativo y un gestor de estados que registra el historial de elecciones del usuario.
Las pruebas preliminares mostraron que el sistema es capaz de generar trayectorias diferenciadas en función de las decisiones adoptadas, ajustando parámetros internos como la complejidad del escenario y la retroalimentación ofrecida. Asimismo, se observó una adecuada coherencia narrativa y estabilidad del sistema frente a recorridos prolongados.
Estos resultados validan la factibilidad técnica del enfoque propuesto y confirman el potencial de los sistemas adaptativos basados en agentes como herramientas de simulación para el análisis de procesos de decisión complejos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por siete integrantes: tres docentes de la Licenciatura en Informática, pertenecientes al Área de Ciencias Aplicadas de la Universidad Siglo 21, y cuatro alumnos de la misma institución. La participación estudiantil busca fomentar su formación en investigación y motivarlos en el

desarrollo de productos de software innovadores con impacto social.

La diversidad del equipo radica en la presencia de docentes de distintas cátedras y estudiantes de diferentes años y modalidades de cursado. Esta composición impulsa una educación de calidad alineada con los objetivos de desarrollo sostenible, promoviendo la formación de agentes de cambios capaces de adaptarse a un entorno tecnológico en constante evolución.

En cuanto a la formación académica del equipo, dos docentes están cursando un Doctorado en Informática y otro en Ingeniería, con líneas de investigación en desarrollo tecnológico con microprocesadores de última generación e inteligencia artificial. Su trayectoria incluye participación en congresos y publicaciones científicas.

5. BIBLIOGRAFÍA

[1] Russell, S. y Norvig, P., Artificial Intelligence: A Modern Approach, 4th ed., Pearson, 2020.

[2] Poole, D. y Mackworth, A., Artificial Intelligence: Foundations of Computational Agents, Cambridge University Press, 2023.

[3] Yannakakis, G. N. y Togelius, J., Artificial Intelligence and Games, Springer, 2018.

[4] Gallivan, J. et al., "Decision-making in sensorimotor control," Nature Reviews Neuroscience, vol. 19, 2018.

[5] Papamitsiou, Z. y Economides, A., "Adaptive feedback in intelligent systems," British Journal of Educational Technology, 2022.

[6] Li, W. et al., "AI-driven serious games for decision-making," Computers in Human Behavior, 2023.

Sistema Inteligente de Asistencia por Voz para Control de Calidad en Producción de Tempeh: Enfoque Híbrido Basado en Grafos y Modelos de Lenguaje

Juan Pablo Neri[1], Matilde Inés Césari[2], Maria Eugenia Stefanoni[2]

[1]Alumno de Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina. juan.neri@alumnos.frm.utn.edu.ar

[2]Docentes cátedra Inteligencia Artificial, Laboratorio de Investigación ReAVi Departamento Ingeniería en Sistemas de Información - Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina. matilde.cesari@frm.utn.edu.ar

RESUMEN

Este proyecto desarrolla un sistema innovador de asistencia por voz que resuelve el paradigma de "manos ocupadas" en la producción artesanal-industrial de tempeh. Mediante la integración de tecnologías de vanguardia como asistentes conversacionales, grafos de conocimiento y modelos de lenguaje desplegados localmente, el sistema permite a los operarios registrar datos de calidad y consultar procedimientos expertos sin necesidad de contacto físico con dispositivos, garantizando así la asepsia y trazabilidad en tiempo real.

La arquitectura híbrida combina FastAPI + React.js para la interfaz, MySQL para transacciones y Neo4j para razonamiento semántico, con el modelo Mistral procesando lenguaje natural localmente. Los resultados obtenidos en 60 días de validación muestran una precisión del 96.2% en reconocimiento de voz, reducción del 88.4% en errores de registro y adopción del 94.7% por parte del personal operativo. El sistema no solo digitaliza el control de calidad, sino que funciona como un asistente experto que guía al operario mediante razonamiento contextual, estableciendo un nuevo estándar en la modernización de industrias alimentarias tradicionales mediante IA aplicada.

CONTEXTO

La producción artesanal-industrial de alimentos fermentados representa un sector estratégico para la economía regional, enfrentando desafíos significativos en la digitalización de procesos de calidad. Según datos del Ministerio de Agricultura, Ganadería y Pesca de Argentina (2023), el mercado de alimentos fermentados ha experimentado un crecimiento anual del 15% en los últimos cinco años, impulsado por tendencias de consumo saludable y sustentable.

Sin embargo, la producción de tempeh en particular enfrenta un cuello de botella tecnológico crítico: la brecha entre sistemas de monitoreo por Internet de las Cosas (IoT) y el registro cualitativo humano. Mientras las redes de sensores capturan datos ambientales (temperatura, humedad, CO2) de manera continua, la evaluación de calidad organoléptica y el seguimiento de procedimientos operativos sigue dependiendo de registros manuales, generando inconsistencias y pérdida de información valiosa.

El Laboratorio de Investigación ReAVi de la UTN FRM, en colaboración estratégica con UMAMI Alimentos Saludables, identificó esta problemática como oportunidad para desarrollar un sistema que integre tecnología de voz, grafos de conocimiento y modelos de lenguaje para resolver el paradigma de "manos ocupadas" (hands-busy paradigm) característico de procesos alimentarios que requieren manipulación constante y estrictas normas de inocuidad.

La investigación se enmarca en las Prácticas Profesionales Supervisadas (PPS) en coordinación con Proyecto Integrador de la Cátedra de Inteligencia Artificial, constituyendo un caso paradigmático de transferencia tecnológica entre academia y sector productivo, donde los requerimientos industriales impulsan innovaciones tecnológicas con potencial replicación en múltiples dominios de producción alimentaria.

1. OBJETIVOS TEÓRICOS Y EXPERIMENTALES

1.1. Objetivo General

Desarrollar y validar un sistema inteligente de asistencia por voz que integre procesamiento de lenguaje natural y grafos de conocimiento para el control de calidad en tiempo real durante la producción de tempeh, garantizando trazabilidad completa y eliminando barreras de interacción en entornos de producción con restricciones higiénicas.

1.2. Objetivos Teóricos

- **O1.** Investigar arquitecturas híbridas que combinen Retrieval-Augmented Generation (RAG) con grafos de conocimiento para razonamiento contextual en dominios especializados.
- **O2.** Analizar la efectividad de modelos de lenguaje desplegados localmente para procesamiento de lenguaje natural en español con vocabulario técnico específico de la industria alimentaria.
- **O3.** Evaluar paradigmas de interacción voz-a-voz en ambientes industriales con ruido de fondo y restricciones operativas.
- **O4.** Diseñar ontologías de conocimiento para representar procedimientos expertos y acciones correctivas en procesos de fermentación.

1.3. Objetivos Experimentales

- **E1.** Implementar un prototipo funcional utilizando FastAPI, React.js y Docker como stack tecnológico principal.
- **E2.** Desarrollar un sistema dual de persistencia integrando MySQL para transacciones y Neo4j para razonamiento semántico.
- **E3.** Integrar el modelo Mistral mediante Ollama para procesamiento local de lenguaje natural.
- **E4.** Validar el sistema mediante pruebas controladas en el entorno real de producción de UMAMI durante 60 días operativos.
- **E5.** Medir indicadores de rendimiento: precisión en reconocimiento de voz, reducción de errores de registro y tiempo de respuesta en consultas expertas.

2. METODOLOGÍA Y DESARROLLO EXPERIMENTAL

2.1. Arquitectura del Sistema

El proyecto adoptó una arquitectura de microservicios en contenedores Docker, implementando tres capas claramente diferenciadas:

Capa de Presentación (Frontend): Desarrollo en React.js con Hooks personalizados para gestión del estado de reconocimiento de voz mediante la API Web Speech del navegador. Se implementó un dashboard analítico con visualizaciones en tiempo real de métricas de producción y alertas de desviación.

Capa de Orquestación (Backend): FastAPI como framework asíncrono para gestión de endpoints REST. La lógica de negocio se modularizó en servicios especializados: ServicioVoz para procesamiento Speech-to-Text, ServicioIntentos para categorización de comandos y ServicioExpertise para razonamiento basado en conocimiento.

Capa de Persistencia Híbrida: MySQL 8.0 para gestión transaccional de registros de producción, inventario de insumos e históricos de sensores IoT. Neo4j 5.0 para almacenamiento y consulta de la ontología de conocimiento experto, incluyendo relaciones entre etapas de proceso, parámetros de calidad y acciones correctivas.

2.2. Procesamiento de Lenguaje Natural

El componente NLP implementa un pipeline de procesamiento de tres fases:

Fase 1 - Reconocimiento y Normalización: Utilizando Web Speech API con configuración para español argentino, el sistema convierte audio a texto. Posteriormente, mediante expresiones regulares y la librería word2number (w2n), se normalizan cantidades verbalizadas y unidades de medida heterogéneas.

Fase 2 - Clasificación de Intenciones: El ProcesadorIntenciones categoriza las solicitudes en tipos predefinidos: REGISTRO_PRODUCCION,

CONSULTA_INVENTARIO, ALERTA_DESVIO, CONSULTA_EXPERTA. La clasificación alcanzó una precisión del 94.7% en pruebas con 350 muestras de lenguaje natural.

Fase 3 - Generación de Consultas: Mediante LangChain, el sistema determina si la intención requiere consulta transaccional (SQL) o semántica (Cypher), generando automáticamente las consultas apropiadas con parámetros validados.

2.3. Grafo de Conocimiento Especializado

La ontología implementada en Neo4j modela el dominio experto de producción de tempeh mediante 8 tipos de nodos principales:

- EtapaProceso: hidratación, partido, cocción, inoculación, envasado
- ParametroCalidad: temperatura, pH, tiempo, humedad
- Problema: contaminación, fermentación incompleta, textura inadecuada
- AccionCorrectiva: ajuste temperatura, tiempo extendido, descarte
- Desviacion: valores fuera de rango esperado
- CausaPosible: factores subyacentes
- Insumo: materias primas y materiales
- Lote: entidad central de trazabilidad

Las relaciones implementadas incluyen TIENE_PARAMETRO, PRESENTA_PROBLEMA, SE_SOLUCIONA_CON, AFECTA_A, SIGUE_ETAPA, permitiendo al sistema realizar razonamientos complejos como "si la temperatura está fuera del rango 27-30°C durante la inoculación, probablemente se producirá fermentación incompleta y se debe extender el tiempo de proceso".

2.4. Experimentos y Validación

La validación se realizó en cuatro fases progresivas:

Fase 1 (Días 1-15): Integración de componentes y pruebas unitarias en entorno controlado. Se validaron 120 comandos de voz individuales alcanzando 98.3% de precisión.

Fase 2 (Días 16-30): Pruebas de integración con datos simulados de producción. Se evaluó el sistema bajo condiciones de ruido ambiente controlado (45-65 dB).

Fase 3 (Días 31-45): Operación paralela con producción real en UMAMI. Se registraron 87 interacciones de voz durante ciclos productivos completos.

Fase 4 (Días 46-60): Operación exclusiva con monitoreo continuo. Se procesaron 156 interacciones sin intervención manual, incluyendo consultas expertas y registro de desvíos.

3. FORMACIÓN DE RECURSOS HUMANOS

El proyecto representó una experiencia formativa integral que trasciende las competencias técnicas específicas, abarcando múltiples dimensiones del desarrollo profesional en ingeniería de sistemas.

3.1. Competencias Técnicas Avanzadas

- **Inteligencia Artificial Aplicada:** Implementación práctica de arquitecturas RAG, modelos de lenguaje local y sistemas de razonamiento basados en grafos.
- **Desarrollo Full-Stack:** Dominio de stack moderno React.js + FastAPI + Docker con enfoque en aplicaciones industriales.
- **Procesamiento de Lenguaje Natural:** Diseño e implementación de pipelines completos Speech-to-Text para dominios especializados.
- **Ingeniería de Datos:** Modelado híbrido de datos combinando paradigmas relacionales y basados en grafos.

3.2. Competencias Interdisciplinarias

- **Dominio Industrial:** Comprensión profunda de procesos de producción alimentaria, normativas de inocuidad y estándares de calidad.
- **Investigación Aplicada:** Metodología experimental en entornos productivos con restricciones operativas reales.

- **Transferencia Tecnológica:** Gestión de proyectos de colaboración academia-industria con resultados concretos y medibles.

4. RESULTADOS Y DISCUSIÓN

4.1. Indicadores Cuantitativos de Desempeño

Métrica	Objetivo Inicial	Resultado Obtenido	Cumplimiento
Precisión Reconocimiento Voz	90%	96.2%	106.9%
Reducción Errores Registro	75%	88.4%	117.9%
Tiempo Respuesta Consultas	<3 segundos	1.8 segundos	166.7%
Adopción Operarios	60%	94.7%	157.8%
Interacciones Exitosas	500/mes	723/mes	144.6%

4.2. Análisis Cualitativo del Impacto

El sistema demostró capacidades transformadoras en múltiples dimensiones operativas:

Eficiencia en Registro de Calidad: La interfaz de voz eliminó completamente la necesidad de interrupciones para registro manual. Los operarios reportaron una reducción del 70% en tiempo dedicado a documentation administrativa durante ciclos productivos.

Trazabilidad en Tiempo Real: El sistema permitió el registro instantáneo de desvíos y acciones correctivas, generando una bitácora completa y auditable. Esto resulta fundamental para cumplimiento de normativas SENASA y requerimientos de exportación.

Capacidad de Razonamiento Experto: El componente basado en grafos demostró habilidad para sugerir acciones correctivas contextualizadas. Ejemplo: ante reporte de "olor inusual en lote 23", el sistema identificó correlación con pH anormal y recomendó análisis microbiológico preventivo.

Adaptación a Ruido Industrial: Las pruebas en condiciones reales (65-75 dB de ruido ambiente) demostraron robustez del reconocimiento de voz, con solo un 3.8% de requerimientos de repetición, significativamente inferior al 15% inicialmente previsto.

5. BIBLIOGRÁFIA

[1] Shengyuan Chen et al. (2025). "A Survey of Graph Retrieval-Augmented Generation for Customized Large Language Models", arXiv 2501.13958.
Link: https://arxiv.org/abs/2501.13958 .

[2] Rosilius, A. (2023 aprox.). "Impact of Industrial Noise on Speech Interaction Performance and …" (PDF). Link directo: https://opus4.kobv.de/opus4-fhws/frontdoor/deliver/index/docId/5569/file/Rosilius_Impact_Industrial_Noise.pdf .

[3] Dadhaneeya, H., Nema, P. K., & Arora, V. K. (2023). Internet of Things in food processing and its potential in Industry 4.0 era: A review. *Trends in Food Science & Technology*, *139*, 104109. https://www.sciencedirect.com/science/article/pii/S0924224423002169 .

[4] Saad, M., Zhang, Y., Tian, J., & Jia, J. (2023). A graph database for life cycle inventory using Neo4j. *Journal of Cleaner Production*, *393*, 136344. https://www.sciencedirect.com/science/article/pii/S0959652623005024

[5] FastAPI. (2024). *High Performance Web Framework for Building APIs with Python 3.7+*. Official Documentation. https://fastapi.tiangolo.com/

[6] LangChain. (2024). *Building Applications with LLMs through Composability*. Developer Guide. https://python.langchain.com/

[7] Ministerio de Agricultura, Ganadería y Pesca. (2023). *Panorama de la Industria Alimentaria Argentina*. Buenos Aires: MAGyP. https://alimentosargentinos.magyp.gob.ar/contenido/sectores/AyB/informes/anteriores/Ind_Aliment_07.pdf

[8] SENASA. (1998). *Normativa de Inocuidad Alimentaria para Productos*

Fermentados. Resolución 587/97. **Vegetales Fermentados (Chucrut, Pickles):** Se rigen por el **Capítulo X:** https://alimentosargentinos.magyp.gob.ar/contenido/marco/CAA/capitulospdf/Capitulo_X.pdf

[9] Ollama. (2024). *Local Large Language Models: Deployment and Optimization Guide. Technical Documentation*. **Documentación General:** docs.ollama.com

[10] UMAMI Alimentos Saludables. (2024). *Manual de Procesos y Control de Calidad en Producción de Tempeh*. Documento interno, no disponible públicamente online.

[11] J. Buckley, D. Lottridge, J.G. Murphy, P.M. Corballis,

Indicators of employee phishing email behaviours: Intuition, elaboration, attention, and email typology, International Journal of Human-Computer Studies, Volume 172, 2023, 102996, ISSN 1071-5819, https://doi.org/10.1016/j.ijhcs.2023.102996

[12] UTN FRM. (2023). *Lineamientos de Prácticas Profesionales Supervisadas*. Reglamento Académico Institucional. https://www4.frm.utn.edu.ar/plan-de-estudio-ing-sistemas-de-informacion/

UN FRAMEWORK MULTI-AGENTE BASADO EN LLMS PARA LA AUTOMATIZACIÓN DEL FLUJO DE INVESTIGACIÓN EN SIMULACIÓN DE REDES CON NS-3

Diego R. Rodríguez Herlein[1], Claudia N. González[2]
Instituto de Tecnología Aplicada (ITA) - Unidad Académica Rio Gallegos (UARG)
Universidad Nacional de la Patagonia Austral (UNPA)
{[1]dherlein, [2]cgonzalez}@uarg.unpa.edu.ar
[1]ORCID: https://orcid.org/0000-0003-0680-9751

RESUMEN

Este trabajo propone un framework multi-agente basado en Large Language Models (LLMs) orientado a la automatización del ciclo de investigación científica en simulación de redes. El sistema, implementado mediante LangGraph, coordina un conjunto de agentes especializados que asisten tareas clave como la revisión bibliográfica, la generación de código en NS-3, la ejecución de simulaciones y el análisis estadístico de resultados.

El objetivo principal es reducir el tiempo y la complejidad asociados al diseño, ejecución y documentación de experimentos, preservando la reproducibilidad y el rigor metodológico. Como aporte, el framework incorpora un mecanismo de Memoria Episódica persistente basado en Recuperación Aumentada (RAG), que permite reutilizar conocimiento previo derivado de errores y estrategias de optimización.

La implementación actual corresponde a una versión incipiente que opera como prueba de concepto arquitectónica, con una implementación parcial de los agentes descritos, y se encuentra disponible públicamente en un repositorio GitHub [18]. Los resultados evidencian la viabilidad del enfoque y su potencial para avanzar hacia sistemas de automatización inteligente en investigación basada en simulación.

Palabras clave: Multi-Agente, NS-3, LangGraph, LLM, Memoria Episódica, RAG.

CONTEXTO

El presente trabajo se enmarca en las actividades de investigación asociadas a la formación de posgrado de sus autores. Uno de ellos cursa estudios de Maestría en la Unidad Académica Río Gallegos (UNPA), mientras que el otro desarrolla su Doctorado en la Facultad de Informática (UNLP). El estudio se vincula con las líneas de investigación en redes y sistemas de comunicaciones desarrolladas en dichas instituciones.

1. INTRODUCCIÓN

El crecimiento exponencial de las redes de Internet de las Cosas (IoT) y de las redes móviles, desde las tecnologías actuales hasta las emergentes redes 6G, ha impulsado la necesidad de desarrollar, evaluar y comparar modelos, protocolos y escenarios de comunicación cada vez más complejos.

El simulador de redes NS-3 [1] es un estándar "de facto" y constituye un pilar fundamental para la validación de protocolos en entornos complejos, debido a su fidelidad en el modelado de las capas física y MAC, así como a su precisión en la simulación de eventos discretos. Sin embargo, su adopción presenta una curva de aprendizaje pronunciada, derivada de su complejidad inherente y de la necesidad de un conocimiento profundo de la arquitectura de red, la programación en C++ y Python, y el manejo de métodos estadísticos.

El proceso tradicional de investigación implica la escritura de scripts de simulación, la configuración de distintas topologías, la adaptación de protocolos, la gestión de múltiples ejecuciones y el procesamiento y análisis de trazas. Este proceso limita la exploración sistemática de escenarios y la validación rigurosa de los resultados experimentales.

Recientemente, los Modelos de Lenguaje de Gran Escala (LLM) han demostrado la capacidad para la planificación y la generación de código. Paralelamente, los Sistemas Multi-

Agente (MAS) han emergido como una arquitectura capaz de descomponer problemas complejos en subtareas manejables, donde agentes especializados colaboran entre si.

En este contexto, se propone un framework que aprovecha estas capacidades de los modelos para automatizar el flujo de trabajo de investigación. A diferencia de los asistentes de codificación genéricos, como MetaGPT o SWE-agent [2], implementa una arquitectura de agentes autónomos especializados que colaboran para diseñar, ejecutar, optimizar y documentar experimentos de red de manera supervisada.

En particular, el framework propuesto se sustenta en cuatro conceptos clave:

- Arquitectura de especialización profunda, orquestada mediante LangGraph, que coordina múltiples agentes autónomos con roles bien definidos para cubrir cada etapa del proceso.
- Sistema de memoria episódica basado en recuperación aumentada (RAG), que emplea almacenamiento vectorial persistente para evitar la reiteración de errores.
- Integración nativa de aprendizaje por refuerzo profundo (DRL) con ns3-ai, habilitando la optimización en bucle cerrado de parámetros de red de forma adaptativa.
- Enfoque en la reproducibilidad.

2. LÍNEAS DE INVESTIGACIÓN

La aplicación de LLMs en la ingeniería de software ha sido ampliamente explorada en los últimos años. Xi et al. [3] presentan un marco general para agentes basados en LLM, destacando su potencial en tareas complejas de desarrollo. Yang et al. [4] introducen SWE-agent, demostrando cómo los agentes pueden utilizar interfaces para realizar tareas de ingeniería de software de manera autónoma. Hong et al. [5] proponen MetaGPT, popularizando el uso de roles especializados en frameworks multi-agente. Estos sistemas, frecuentemente, carecen de la especificidad necesaria para producir código que no solo compile, sino que genere datos estadísticamente válidos bajo condiciones de incertidumbre propias de simulaciones de red.

En el contexto de sistemas de agentes inteligentes, Lanham [6] propone un marco para construir agentes robustos capaces de manejar tareas del mundo real mediante la integración de recuperación de conocimiento (RAG), con bucles de retroalimentación, implementando una arquitectura donde agentes especializados colaboran y aprenden de la experiencia (memoria episódica). Pink et al. [7] exploran el concepto de Memoria Episódica en agentes LLM, permitiendo que los sistemas recuerden “episodios” pasados de su propia experiencia operativa. Por otra parte, He et al. [8] discuten la visión de sistemas multi-agente para ingeniería de software.

En el ámbito de las redes, la integración de IA con simuladores ha avanzado significativamente. El módulo ns3-ai, propuesto por Yin et al. [9], permite una interacción eficiente entre NS-3 y frameworks de IA basados en Python, superando las limitaciones de soluciones anteriores como ns3-gym [10]. Si bien estos avances en la integración de IA con simuladores son importantes, se centran principalmente en tareas aisladas de optimización y no abordan el ciclo completo de investigación. Cabe destacar que Rezazadeh et al. [11] presentan GenOnet, aplicando agentes a simulaciones de redes 6G.

En síntesis, si bien los avances recientes en agentes basados en LLM han demostrado su potencial para automatizar tareas de ingeniería de software mediante el uso de roles, herramientas y mecanismos de orquestación, la mayoría de los enfoques existentes se orientan a escenarios generales y carecen de un tratamiento explícito de la incertidumbre inherente a la simulación de redes. Asimismo, los mecanismos de memoria suelen limitarse al acceso a conocimiento estático o a flujos de ejecución lineales. En este contexto, la propuesta de este trabajo integra agentes especializados, orquestación cíclica con estado

persistente y memoria episódica, con el objetivo de cubrir el ciclo completo de desarrollo, validación y refinamiento de simulaciones de red, incorporando criterios de reproducibilidad y validez estadística como parte intrínseca del proceso.

3. ARQUITECTURA DEL SISTEMA

La arquitectura propuesta es de naturaleza modular y se estructura en tres capas interdependientes: Capa de Orquestación, Capa de Agentes y Capa de Infraestructura. La coordinación entre dichas capas se realiza mediante LangGraph [12], lo que permite la definición explícita de flujos de trabajo complejos, tanto cíclicos como condicionales, entre los agentes que componen el sistema [13]. La elección del mecanismo de orquestación resulta crítica para el correcto funcionamiento de la arquitectura. Se diferencia de enfoques como AutoGen [14] que priorizan interacciones conversacionales libres y CrewAI [15], [16] que se basa en estructuras jerárquicas. El presente trabajo adopta LangGraph debido a su capacidad para modelar flujos de ejecución con estado persistente. En particular, la utilización de grafos dirigidos cíclicos posibilita la implementación de bucles de retroalimentación robustos entre agentes especializados, como por ejemplo, Coder-Critic, los cuales son esenciales para el refinamiento iterativo y controlado del código de simulación.

Capa de Orquestación

El Supervisor orquesta el sistema, gestionando el estado global del experimento y las transiciones entre agentes. Su lógica de decisión no es lineal, sino que modela el flujo de trabajo como un grafo de estados donde las transiciones dependen de la evaluación dinámica del progreso. El Supervisor mantiene el "Global State", una estructura de datos compartida que contiene la configuración del experimento, el código fuente actual, los logs de ejecución y los resultados analíticos.

Capa de Agentes

El framework se compone de diez agentes, cada uno de ellos provisto de un rol específico y un conjunto de herramientas dedicadas.

Researcher Agent: Realiza búsquedas bibliográficas en bases de datos académicas (Semantic Scholar, arXiv) para identificar el estado del arte, extraer parámetros de simulación realistas y métricas de referencia.

Coder Agent: Genera scripts de simulación en C++ o Python compatibles con NS-3, debe estructurar la topología y la instalación de protocolos.

Critic Agent: Realiza una validación multinivel del código generado, verificando no solo la sintaxis, sino también la lógica de la topología, asegurando que tengan sentido para el objetivo del experimento.

Simulator Agent: Gestiona la ejecución de las simulaciones en el entorno NS-3 y gestionando los archivos de trazas.

Trace Analyzer Agent: Procesa los archivos de trazas (PCAP, ASCII) y logs generados por NS-3.

Analyst Agent: transforma las métricas y calcula Indicadores Clave de Desempeño (Key Performance Indicators KPIs) como PDR (Packet Delivery Ratio), retardo extremo a extremo, throughput y overhead, calculando intervalos de confianza y pruebas estadísticas (t-student, ANOVA) para validar los resultados.

Optimizer Agent: Implementa algoritmos de búsqueda o Aprendizaje por Refuerzo Profundo (DRL), a través de la integración ns3-ai, ajustando dinámicamente parámetros de red para maximizar el rendimiento.

NS3-AI Integration Agent: Gestiona la complejidad de la interfaz de memoria compartida entre los procesos de Python (IA) y el motor de simulación C++ (NS-3), facilitando el aprendizaje por refuerzo.

Visualizer Agent: Genera representaciones visuales como boxplots, heatmaps, que sintetizan los hallazgos para su inclusión en publicaciones.

Scientific Writer Agent: Genera documentación académica, incluyendo informes técnicos, compilando el texto, código y gráficos generados.

Capa de Infraestructura

Esta capa proporciona la base tecnológica robusta sobre la que operan los agentes. Se basa en dos pilares fundamentales:

Episodic Memory: Memoria Episódica Persistente, a diferencia de la memoria temporal de contexto de los LLMs, se implementa un sistema de almacenamiento vectorial a largo plazo. Este subsistema indexa las experiencias pasadas, como los errores de compilación y sus soluciones, permitiendo que los agentes consulten y recuperen este conocimiento .

Entorno de Ejecución Aislado: para garantizar la reproducibilidad científica, el framework se despliega dentro de contenedores (Docker). Esto asegura que todas las dependencias críticas, desde la versión exacta del compilador C++ y las librerías de Python hasta el propio núcleo de NS-3, permanezcan inmutables. De esta forma, se garantizan los resultados, independientemente de donde se ejecute la simulación.

4. FLUJO DE TRABAJO

El proceso se estructura como un flujo iterativo e incremental compuesto por varias etapas interrelacionadas. En primer lugar, el usuario define la tarea mediante una descripción en lenguaje natural, por ejemplo, *"comparar el rendimiento de AODV y OLSR en un escenario de alta movilidad"*. A partir de este objetivo, el agente *Supervisor*, asistido por el *Researcher Agent*, analiza el problema y establece las variables relevantes, las constantes del escenario y las métricas de evaluación. En una segunda etapa, el *Researcher* identifica en la literatura los parámetros y configuraciones, mientras que el agente *Coder* genera un primer borrador del script de simulación. Dicho código es revisado por el agente *Critic*, y, ante la detección de errores o inconsistencias, el ciclo de generación y corrección se repite, apoyándose en la *Episodic Memory* para reutilizar correcciones previas y reducir el tiempo de depuración.

Una vez validado el script, se procede a la ejecución de las simulaciones. En caso de fallos durante la ejecución, el flujo retorna automáticamente a la etapa de codificación. Las trazas resultantes son procesadas por el *Trace Analyzer*, y los resultados son evaluados por el agente *Analyst*. Si los indicadores clave de desempeño (KPIs) no satisfacen los umbrales definidos, el agente *Optimizer* propone ajustes en los parámetros de la red mediante algoritmos de búsqueda o técnicas de Aprendizaje por Refuerzo Profundo, integradas a través de ns3-ai, solicitando nuevas iteraciones de simulación hasta alcanzar los criterios de rendimiento establecidos.

Finalmente, el agente *Visualizer* genera los gráficos, y el *Scientific Writer* compila los resultados en un documento estructurado, integrando de forma coherente los datos obtenidos y las visualizaciones generadas.

5. DISCUSIÓN Y PERSPECTIVAS

El desarrollo del framework propuesto constituye una primera aproximación a la integración de arquitecturas multi-agente basadas en modelos de lenguaje de gran escala (LLM) dentro del flujo de trabajo de simulación de redes. En esta etapa inicial, los resultados obtenidos permiten evidenciar la viabilidad técnica del enfoque y su potencial para asistir tareas complejas como la configuración, ejecución y análisis de simulaciones en entornos como NS-3, tradicionalmente asociados a una elevada barrera de entrada y un alto grado de intervención manual.

El presente trabajo debe interpretarse como un estudio exploratorio, orientado principalmente a la validación conceptual del enfoque, en escenarios controlados y de pequeña escala. En este sentido, resulta importante aclarar que la implementación disponible públicamente no refleja aún la totalidad de los agentes y capacidades conceptuales descritas en la arquitectura. En particular, algunos agentes se encuentran implementados de forma parcial, mientras que otros forman parte del diseño previsto para etapas posteriores del desarrollo. Asimismo, la integración de técnicas de Aprendizaje por Refuerzo Profundo mediante ns3-ai se presenta, en esta fase, como una

capacidad planificada, cuya validación experimental completa constituye una línea de trabajo futura.

Si bien el sistema muestra un comportamiento coherente en escenarios controlados, presenta limitaciones inherentes tanto a su estado actual de desarrollo como a la dependencia de modelos LLM. En particular, la generación automática de código introduce el riesgo de alucinaciones, es decir, la producción de código sintácticamente correcto pero semánticamente incorrecto. Este riesgo se mitiga parcialmente mediante el uso del agente Critic y la incorporación de memoria episódica, aunque no puede eliminarse por completo. Asimismo, componentes clave como los mecanismos de validación automática de scripts y las estrategias de diseño de *prompts* requieren un refinamiento y evaluación más exhaustivos.

A partir de estas observaciones, el trabajo futuro se estructura en dos líneas principales. Por un lado, se prevé el refinamiento algorítmico del framework, ampliando las capacidades del agente Analyst mediante técnicas de análisis más avanzadas, como modelos de series temporales y métodos de regresión de mayor complejidad, así como el fortalecimiento del agente Optimizer mediante la integración de algoritmos de Aprendizaje por Refuerzo Profundo, tales como Proximal Policy Optimization (PPO) y Soft Actor-Critic (SAC).

Por otro lado, se abordará el fortalecimiento general del sistema, especialmente en lo referido a la validación de scripts, la gestión de errores y la interacción con el simulador NS-3, considerando además el impacto de la ejecución concurrente de múltiples agentes y simulaciones sobre los requerimientos computacionales, lo que podría limitar su adopción en entornos con infraestructura restringida.

Desde la perspectiva de la infraestructura de ejecución, el framework admite tanto el uso de modelos mediante APIs externas como la ejecución local de modelos de código abierto a través de entornos como Ollama [17], buscando un compromiso entre costo operacional, requerimientos de hardware y reproducibilidad experimental. En consecuencia, la elección del modo de ejecución constituye una decisión relevante que puede impactar en la escalabilidad y viabilidad práctica del sistema. El uso de APIs externas simplifica el despliegue y reduce la carga de hardware local, pero introduce costos variables asociados al volumen de tokens procesados y una dependencia de servicios de terceros. En consecuencia, la elección del modo de ejecución constituye una consideración de infraestructura no trivial, que puede impactar la escalabilidad del sistema y debe ser evaluada al planificar experimentos de gran volumen o su adopción en entornos con recursos limitados.

Finalmente, con el objetivo de fomentar la reproducibilidad y la colaboración científica, el código fuente del framework se encuentra disponible públicamente [18], reflejando el estado actual de desarrollo del sistema y sirviendo como base para su evolución futura.

6. FORMACIÓN

El presente trabajo contribuye a la formación académica y se enmarca en las actividades de investigación de sus autores. En particular, se promueve la adquisición de competencias avanzadas en simulación de redes, metodologías de investigación experimental y aplicación de técnicas de inteligencia artificial, fortaleciendo la articulación entre formación teórica y práctica investigativa. Además, se prevé la creación de un proyecto de investigación, sobre esta temática, que incluirá la participación activa de alumnos de la carrera de Licenciatura en Sistemas, con el fin de ampliar la formación práctica y el impacto más allá de las tesis de los autores.

7. REFERENCIAS

[1] T. R. Henderson, M. Lacage, G. F. Riley, C. Dowell, and J. Kopena, “Network simulations with the ns-3 simulator,” SIGCOMM demonstration, vol. 14, no. 14, p. 527, 2008.

[2] S. Schmidgall et al., «Agent Laboratory: Using LLM Agents as Research Assistants», 17 de junio de 2025, arXiv: arXiv:2501.04227. doi: 10.48550/arXiv.2501.04227.

[3] Z. Xi et al., «The Rise and Potential of Large Language Model Based Agents: A Survey», 19 de septiembre de 2023, arXiv: arXiv:2309.07864. doi: 10.48550/arXiv.2309.07864.
[4] J. Yang et al., «SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering», 11 de noviembre de 2024, arXiv: arXiv:2405.15793. doi: 10.48550/arXiv.2405.15793.
[5] S. Hong et al., «MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework», 1 de noviembre de 2024, arXiv: arXiv:2308.00352. doi: 10.48550/arXiv.2308.00352.
[6] M. Lanham, AI Agents in Action. Shelter Island, NY, USA: Manning Publications, 2025.
[7] M. Pink et al., «Position: Episodic Memory is the Missing Piece for Long-Term LLM Agents», 10 de febrero de 2025, arXiv: arXiv:2502.06975. doi: 10.48550/arXiv.2502.06975.
[8] J. He, C. Treude, and D. Lo, "LLM-Based Multi-Agent Systems for Software Engineering: Vision and the Road Ahead," in Proceedings of the 2024 IEEE/ACM 1st International Conference on AI Foundation Models and Software Engineering (FORGE), pp. 1–11, 2024. https://arxiv.org/html/2404.04834v2
[9] Yin, H., Liu, P., Liu, K., Cao, L., Zhang, L., Gao, Y., & Hei, X. (2020). ns3-ai: Fostering Artificial Intelligence Algorithms for Networking Research. Proceedings of the 2020 Workshop on ns-3.
[10] P. Gawłowicz y A. Zubow, «ns3-gym: Extending OpenAI Gym for Networking Research», 10 de octubre de 2018, arXiv: arXiv:1810.03943. doi: 10.48550/arXiv.1810.03943.
[11] F. Rezazadeh, A. A. Gargari, S. Lagén, J. Mangues-Bafalluy, D. Niyato, y L. Liu, «GenOnet: Generative Open xG Network Simulation with Multi-Agent LLM and ns-3», en 2024 3rd International Conference on 6G Networking (6GNet), oct. 2024, pp. 69-71. doi: 10.1109/6GNet63182.2024.10765766.
[12] J. Wang y Z. Duan, «Agent AI with LangGraph: A Modular Framework for Enhancing Machine Translation Using Large Language Models», 5 de diciembre de 2024, arXiv: arXiv:2412.03801. doi: 10.48550/arXiv.2412.03801.
[13] J. Tong et al., «WirelessAgent: Large Language Model Agents for Intelligent Wireless Networks», 2 de mayo de 2025, arXiv: arXiv:2505.01074. doi: 10.48550/arXiv.2505.01074.
[14] Q. Wu et al., «AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation», 3 de octubre de 2023, arXiv: arXiv:2308.08155. doi: 10.48550/arXiv.2308.08155.
[15] D. Liu, K. Upadhyay, V. Chhetri, A. B. Siddique, y U. Farooq, «A Large-Scale Study on the Development and Issues of Multi-Agent AI Systems», 2026, arXiv: arXiv:2601.07136. doi: 10.48550/arXiv.2601.07136.
[16] Z. Duan y J. Wang, «Exploration of LLM Multi-Agent Application Implementation Based on LangGraph+CrewAI», 27 de noviembre de 2024, arXiv: arXiv:2411.18241. doi: 10.48550/arXiv.2411.18241.
[17] V. Udandarao y N. Misra, «Democratizing AI Development: Local LLM Deployment for India's Developer Ecosystem in the Era of Tokenized APIs», 21 de agosto de 2025, arXiv: arXiv:2508.16684. doi: 10.48550/arXiv.2508.16684.
[18] LW1DQ, "Multi-Agent LLM-Based Framework for Network Simulation," GitHub repository, 2026. [Online]. Available: https://github.com/LW1DQ/Framework.

Uso de simuladores para el entrenamiento de robots mediante aprendizaje por refuerzo

Gonzalo Zabala

Centro de Altos Estudios en Tecnología Informática
Facultad de Tecnología Informática
Universidad Abierta Interamericana
Av. Montes de Oca 745, Ciudad Autónoma de Buenos Aires, República Argentina
(+54 11) 4301-5323; 4301-5240; 4301-5248
Gonzalo.Zabala@uai.edu.ar

RESUMEN

En los últimos años, el desarrollo exponencial de los métodos de aprendizaje automático han atravesado todas las áreas de la actividad humana. En el ámbito de la robótica, la posibilidad de entrenar el comportamiento de un robot en ambientes dinámicos mediante estos mecanismos se ha convertido en una herramienta de gran utilidad que permite acelerar el desarrollo de robots de este tipo (Kim et al., 2021). Hasta el surgimiento de estas tecnologías, la programación de la interacción de un robot con su ambiente se basaba en un conjunto de modelos matemáticos y heurísticas que implicaban testeos engorrosos y permanentes, con intervención humana directa, lo que consumía una cantidad de tiempo considerable para conseguir algún fruto tangible (Liang et al., 2018). A partir de la aplicación de métodos de aprendizaje automático, los procedimientos de ensayo y error con refuerzo de los éxitos permitieron un entrenamiento menos dependiente de los conocimientos previos, con la posibilidad de generar comportamientos novedosos en las acciones de los robots. Lamentablemente, en robots con desplazamiento espacial como vehículos autónomos y drones, este entrenamiento es muy costoso o imposible en términos de seguridad del dispositivo, así como en el tiempo de puesta a cero después de cada fracaso.

Por otra parte, el surgimiento de simuladores con motores de física de alta calidad permitió recrear espacios de entrenamiento de robots, así como la misma estructura del robot, con un alto grado de realismo. De esta manera, el entrenamiento de robots mediante el aprendizaje por refuerzo (AR) se hizo viable en todo tipo de robots utilizando estos espacios de simulación. Al finalizar el entrenamiento, el modelo creado es transportable al robot físico, con un refinamiento posterior mucho más sencillo y limitado en el tiempo (Salvato et al., 2021).

Palabras clave: robótica, simulación, machine learning, reinforcement learning.

CONTEXTO

El presente proyecto está radicado en el Centro de Altos Estudios en Tecnología Informática (CAETI), dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana. El mismo se encuentra inserto en la línea de investigación "Automatización y robótica". El financiamiento es brindado por la misma Universidad Abierta Interamericana.

1. INTRODUCCIÓN

El aprendizaje por refuerzo (AR) ha surgido como una técnica esencial para el entrenamiento de robots, permitiéndoles adquirir comportamientos complejos y adaptarse a entornos reales. El AR permite a los robots aprender habilidades complejas en forma autónoma a partir de la entrada de los datos de sus sensores, y la evaluación de los resultados de sus decisiones, acortando la distancia entre las simulaciones virtuales y las aplicaciones reales (Ibarz et al., 2021)

Por otra parte, los agentes de AR pueden mejorar la eficiencia del aprendizaje mediante técnicas como el uso de modelos de acción y el aprendizaje jerárquico, reduciendo significativamente el tiempo necesario para dominar tareas complejas (Lin, 1992).

En particular, en relación con los robots móviles, el AR es eficaz para problemas que implican fricción y contacto, a menudo difíciles para los métodos de control convencionales, lo que permite a los robots aprender de las interacciones reales con su entorno (Johannink et al., 2019). Su programación requiere de mapeos y depuración complejos de bajo nivel, que son muy complicados de obtener con intervención humana, lo que indica la necesidad de un marco de aprendizaje más intuitivo.

Ahora bien, realizar un entrenamiento utilizando esta técnica mediante la manipulación humana de un robot móvil es imposible, además de ser peligroso tanto para el robot como para su operario. El uso de simuladores para el entrenamiento de robots mediante aprendizaje por refuerzo profundo (AR) ha demostrado ser muy prometedor, permitiendo que los robots aprendan tareas complejas a la vez que abordan los desafíos asociados con su implementación en el mundo real. La eficiencia de los datos se mejora al utilizar entornos de simulación para recopilar datos de entrenamiento de forma segura, evitando los costos y riesgos asociados con las pruebas en el mundo real (Zhao et al., 2020).

Uno de los desafíos más importantes en este contexto es la transferencia de simulación a realidad, ya que el rendimiento en el mundo real a menudo presenta una brecha con el mundo simulado (Rao et al., 2020; Zhao et al., 2020) Se han desarrollado técnicas como la aleatorización de dominios y el aprendizaje por imitación para superar la brecha entre la simulación y la realidad, mejorando la adaptabilidad de los mecanismos aprendidos para aplicaciones en el mundo real (Kormushev et al., 2013; Zhao et al., 2020). Los avances recientes incluyen RL-CycleGAN, que automatiza la traducción de imágenes simuladas a imágenes realistas, mejorando significativamente la eficiencia de transferencia del aprendizaje realizado por AR (Rao et al., 2020).

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En primer término, se desarrollará un estudio del arte, mediante un mapeo sistemático que brinde información con respecto a los siguientes puntos:

- Aspectos esenciales del entrenamiento de robots mediante AR
- Publicaciones habituales
- Autores con mayor cantidad de publicaciones
- Instituciones con mayor cantidad de publicaciones
- Conferencias con mayor cantidad de ponencias.
- Simuladores más utilizados
- Técnicas de AR más utilizadas

A continuación, se realizará una comparación entre los simuladores más utilizados, para seleccionar el correspondiente donde se realizarán las experiencias.

Una vez seleccionado el simulador, se construirá el ambiente de simulación correspondiente al problema de Rescate y la selección de un robot con características afines a las de la competencia de RoboCup. Creado el ambiente, se definirán las características de la red neuronal, se seleccionará una de las técnicas definidas en el mapeo sistemático y se la implementará en el simulador seleccionado para obtener el modelo entrenado.

En el caso de fútbol de robots, el entrenamiento se dividirá en dos etapas. Por un lado, de manera similar al anterior, se desarrollará el ambiente y se entrenará un robot para su navegación en sistemas altamente dinámicos. Posteriormente, se realizará una segunda búsqueda bibliográfica en relación con el entrenamiento de agentes colaborativos para implementar un segundo entrenamiento vinculado a este aspecto, es decir, al

comportamiento colaborativo de los agentes intervinientes.

Por último, se desarrollará al mismo tiempo un ambiente de simulación junto a uno similar físico, incluyendo el espacio de acción del robot como el robot mismo. De igual forma que en las anteriores experiencias, se realizará el entrenamiento en el simulador, para luego hacer una bajada del modelo entrenado en el robot del mundo real. Con los resultados, se hará un análisis de las herramientas propuestas por la academia para la reducción de este gap.

3. RESULTADOS OBTENIDOS/ESPERADOS

Dado que este es el primer año del proyecto, se presentan los resultados esperados.

En el caso del mapeo sistemático, su desarrollo permitirá identificar las principales tendencias en el entrenamiento de robots mediante aprendizaje por refuerzo (AR), incluyendo los simuladores más utilizados, las técnicas predominantes, los autores e instituciones con mayor producción científica y las conferencias de referencia.

Asimismo, la comparación entre los simuladores más empleados en la literatura permitirá establecer criterios objetivos de elección, garantizando que el entorno seleccionado sea el adecuado para los objetivos del proyecto. Se espera que esta selección optimice la eficiencia del entrenamiento y reduzca las limitaciones asociadas a la transferencia de simulación a realidad.

Con respecto al entrenamiento de robots en escenarios de rescate, la construcción de un ambiente de simulación específico y la implementación de una técnica de AR identificada en el mapeo sistemático conducirán a la obtención de un modelo entrenado capaz de desenvolverse en entornos de rescate. Se espera que el robot aprenda a navegar en condiciones de fricción y contacto, mostrando comportamientos adaptativos y robustos frente a la incertidumbre del entorno.

Por otra parte, en el entrenamiento de robots en escenarios de fútbol, se espera que los robots adquieran habilidades de navegación en sistemas altamente dinámicos, optimizando su capacidad de respuesta frente a cambios rápidos en el entorno, y que permite la emergencia de comportamientos coordinados entre múltiples agentes, mejorando la eficiencia en tareas colectivas.

Por último, para validar el uso de simuladores en el entrenamiento, la implementación paralela de ambientes simulados y físicos permitirá evaluar la efectividad de las técnicas de reducción del gap entre simulación y realidad.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por un investigador adjunto del Centro de Altos Estudios en Tecnología Informática (CAETI) quien ejerce el rol de director del proyecto, y tres ayudantes alumnos de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana.

5. BIBLIOGRAFÍA

- Cutler, M., & How, J. P. (2015). Efficient reinforcement learning for robots using informative simulated priors. *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2605-2612. https://doi.org/10.1109/ICRA.2015.7139550
- Ibarz, J., Tan, J., Finn, C., Kalakrishnan, M., Pastor, P., & Levine, S. (2021). How to train your robot with deep reinforcement learning: Lessons we have learned. *The International Journal of Robotics Research*, *40*(4-5), 698-721. https://doi.org/10.1177/0278364920987859
- Johannink, T., Bahl, S., Nair, A., Luo, J., Kumar, A., Loskyll, M., Ojea, J. A., Solowjow, E., & Levine, S. (2019).

Residual Reinforcement Learning for Robot Control. *2019 International Conference on Robotics and Automation (ICRA)*, 6023-6029. https://doi.org/10.1109/ICRA.2019.8794127

- Ju, H., Juan, R., Gomez, R., Nakamura, K., & Li, G. (2022). Transferring policy of deep reinforcement learning from simulation to reality for robotics. *Nature Machine Intelligence*, *4*(12), 1077-1087. https://doi.org/10.1038/s42256-022-00573-6
- Kim, T., Jang, M., & Kim, J. (2021). A Survey on Simulation Environments for Reinforcement Learning. *2021 18th International Conference on Ubiquitous Robots (UR)*, 63-67. https://doi.org/10.1109/UR52253.2021.9494694
- Kormushev, P., Calinon, S., & Caldwell, D. G. (2013). Reinforcement Learning in Robotics: Applications and Real-World Challenges. *Robotics*, *2*(3), Article 3. https://doi.org/10.3390/robotics2030122
- Liang, J., Makoviychuk, V., Handa, A., Chentanez, N., Macklin, M., & Fox, D. (2018). GPU-Accelerated Robotic Simulation for Distributed Reinforcement Learning. *Proceedings of The 2nd Conference on Robot Learning*, 270-282. https://proceedings.mlr.press/v87/liang18a.html
- Lin, L.-J. (1992). *Reinforcement learning for robots using neural networks*. Carnegie Mellon University. https://search.proquest.com/openview/6ce74bcf81ad014c6eec3f62ece7b2e4/1?pq-origsite=gscholar&cbl=18750&diss=y
- Rao, K., Harris, C., Irpan, A., Levine, S., Ibarz, J., & Khansari, M. (2020). *RL-CycleGAN: Reinforcement Learning Aware Simulation-to-Real*. 11157-11166. https://openaccess.thecvf.com/content_CVPR_2020/html/Rao_RL-CycleGAN_Reinforcement_Learning_Aware_Simulation-to-Real_CVPR_2020_paper.html
- Salvato, E., Fenu, G., Medvet, E., & Pellegrino, F. A. (2021). Crossing the Reality Gap: A Survey on Sim-to-Real Transferability of Robot Controllers in Reinforcement Learning. *IEEE Access*, *9*, 153171-153187. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3126658
- Zhao, W., Queralta, J. P., & Westerlund, T. (2020). Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: A Survey. *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 737-744. https://doi.org/10.1109/SSCI47803.2020.9308468
- Zhu, W., Guo, X., Owaki, D., Kutsuzawa, K., & Hayashibe, M. (2023). A Survey of Sim-to-Real Transfer Techniques Applied to Reinforcement Learning for Bioinspired Robots. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(7), 3444-3459. IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2021.3112718
-

Arquitecturas, Redes y Sistemas Operativos

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

ANÁLISIS COMPARATIVO DE ARQUITECTURAS CISC Y RISC COMO SOPORTE DE BAJO NIVEL PARA ALGORITMOS DE INTELIGENCIA ARTIFICIAL EN CENTROS DE DATOS

Pablo Coronati [(1)]
María Paula González[(1,2)]

(1) Universidad Siglo 21, Córdoba, Argentina
PCORONATI@soysiglo.21.edu.ar, paula.gonzalez@ues21.edu.ar

(2) Dpto. de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur, Bahía Blanca, Argentina
mpg@cs.uns.edu.ar

RESUMEN

La Transformación Digital de organizaciones y ecosistemas productivos requiere infraestructuras computacionales capaces de sostener de manera eficiente y escalable la ejecución de algoritmos de Inteligencia Artificial. En este contexto, las decisiones de bajo nivel asociadas a las arquitecturas de procesamiento adquieren un rol estratégico como factor habilitante de ecosistemas computacionales inteligentes.

Este trabajo presenta una línea de investigación en desarrollo, centrada en el análisis comparativo de arquitecturas CISC y RISC, con especial foco en el auge de ARM en centros de datos. El aporte distintivo de la línea consiste en haber orientado dicha comparación hacia factores arquitectónicos clave que condicionan y habilitan la ejecución de algoritmos de IA, tanto de naturaleza simbólica (razonamiento, representación de conocimiento, argumentación) como de aprendizaje automático. Se analizan métricas de rendimiento, eficiencia energética, costos operativos y viabilidad técnica, entendidas como condiciones necesarias para el despliegue sostenible de cargas de trabajo inteligentes. Los avances preliminares, basados en estudios comparativos documentales y benchmarks públicos, evidencian la relevancia de las arquitecturas de procesamiento como infraestructura fundacional para la Transformación Digital de Ecosistemas Computacionales Inteligentes.

Palabras clave: Arquitecturas de Computadores, CISC, RISC, ARM, Inteligencia Artificial, Centros de Datos

CONTEXTO

La línea de investigación presentada se desarrolla desde marzo de 2025 dentro del Proyecto de Investigación (PID) 5409/2022, actualmente en ejecución (julio de 2024 a julio de 2027, evolución de PID anterior ejecutado entre 2002 y 2024), titulado “Ecosistemas Computacionales Inteligentes para la Transformación Digital”. Dicho proyecto es financiado por la Secretaría de Investigación y Transferencia Científica (SITC) del Vicerrectorado de Innovación, Investigación y Postgrado de la Universidad Siglo 21, y se enmarca dentro del Decanato de Ciencias Aplicadas de la mencionada casa de estudios.

El estudio de las arquitecturas de conjunto de instrucciones (Instruction Set Architecture, ISA) ha sido un eje recurrente en el campo de la informática desde las décadas de 1970 y 1980. Diversos trabajos han analizado la evolución y las características de las arquitecturas CISC y RISC, abordando sus ventajas, limitaciones y su adecuación a distintos contextos de aplicación [T00] [S16] [HP19].

Las arquitecturas CISC, representadas principalmente por la familia x86, dominaron durante décadas el ámbito de los servidores y centros de datos debido a su madurez, compatibilidad y amplia adopción industrial. En contraste, las arquitecturas RISC se consolidaron

inicialmente en sistemas embebidos y de bajo consumo, destacándose por su simplicidad estructural, una menor cantidad de transistores y una mayor eficiencia energética [AH07] [T00].
En este marco, la arquitectura ARM, concebida originalmente como Acorn RISC Machine, emerge como una de las implementaciones más exitosas del paradigma RISC. Su diseño, orientado a minimizar el consumo energético y la generación de calor, favoreció su adopción masiva en dispositivos móviles y sistemas embebidos, sentando las bases técnicas para su posterior expansión hacia dominios tradicionalmente dominados por arquitecturas x86 [F17].
Durante la última década, distintos estudios han analizado la progresiva adopción de arquitecturas ARM en infraestructuras de centros de datos, impulsada por la necesidad de mejorar la relación rendimiento–vatio y de reducir los costos operativos en entornos de cómputo a gran escala. Estos trabajos señalan que, en escenarios de alta densidad computacional, incluso mejoras moderadas en eficiencia energética pueden traducirse en impactos significativos a nivel económico y operativo [AG12] [GS21].
Este fenómeno adquiere especial relevancia en el contexto actual, donde los centros de datos soportan cargas de trabajo cada vez más intensivas en cómputo y datos. En particular, la adopción creciente de algoritmos de IA ha incrementado de manera significativa las exigencias sobre la infraestructura subyacente. Tanto los enfoques de aprendizaje automático, que demandan altas tasas de procesamiento para entrenamiento e inferencia, como los enfoques de agentes de la IA simbólica, sensibles a la latencia, la estabilidad y el rendimiento sostenido, dependen críticamente de las características de la arquitectura de procesamiento [GS21] [M22].
En este escenario, diversos autores señalan que las decisiones arquitectónicas de bajo nivel no constituyen únicamente un aspecto técnico, sino un factor estratégico que condiciona el rendimiento, el consumo energético, los costos operativos y la viabilidad técnica de las infraestructuras computacionales modernas [S16] [HP19].
En consecuencia, la Transformación Digital (TD) de organizaciones y ecosistemas basados en datos no puede abordarse de manera independiente de la infraestructura computacional que la sustenta. En este marco institucional y problemático, la línea de investigación presentada se orienta al análisis de arquitecturas de procesamiento como soporte de bajo nivel para la ejecución de algoritmos de IA en centros de datos, focalizándose en la comparación entre arquitecturas CISC y RISC y en el impacto del creciente protagonismo de ARM en infraestructuras tradicionalmente dominadas por x86, considerando métricas de rendimiento, eficiencia energética, costos operativos y viabilidad técnica.

LINEA DE INVESTIGACION

La línea de investigación tiene como objetivo analizar el rol de las arquitecturas de procesamiento como infraestructura de bajo nivel para la ejecución eficiente y sostenible de algoritmos de IA en centros de datos, en el contexto de la TD de ecosistemas computacionales inteligentes. En este marco, la pregunta de investigación que guía la línea se formula de la siguiente manera: ¿En qué medida las diferencias arquitectónicas entre CISC y RISC, y en particular el auge de ARM en centros de datos, impactan en el rendimiento, la eficiencia energética, los costos operativos y la viabilidad técnica de infraestructuras destinadas a la ejecución de algoritmos de IA, tanto simbólicos como de aprendizaje automático?

Objetivo general: Analizar comparativamente arquitecturas CISC y RISC en entornos de centros de datos, evaluando su impacto en métricas de rendimiento, eficiencia energética, costos operativos y viabilidad técnica, como soporte de bajo nivel para la ejecución de algoritmos de IA.

Objetivos específicos:
– Caracterizar las principales diferencias arquitectónicas entre enfoques CISC y RISC en el contexto de infraestructuras de centros de datos.
– Analizar el impacto del uso de arquitecturas

basadas en ARM frente a x86 en términos de rendimiento y eficiencia energética. – Evaluar la relación entre consumo energético y costos operativos en escenarios representativos de ejecución de cargas de trabajo inteligentes. – Analizar la viabilidad técnica de las arquitecturas estudiadas para soportar algoritmos de IA, tanto de aprendizaje automático como de naturaleza simbólica.

La línea adopta un enfoque metodológico documental y comparativo, basado en el análisis de literatura especializada, documentación técnica de fabricantes y benchmarks públicos ampliamente utilizados en la industria. Se consideran métricas vinculadas al rendimiento, consumo energético, costos operativos y adaptabilidad tecnológica de las arquitecturas analizadas.

RESULTADOS OBTENIDOS

Los resultados obtenidos permiten abordar los objetivos planteados mediante un análisis comparativo entre arquitecturas CISC y RISC, sintetizado en las Figuras 1 a 5.

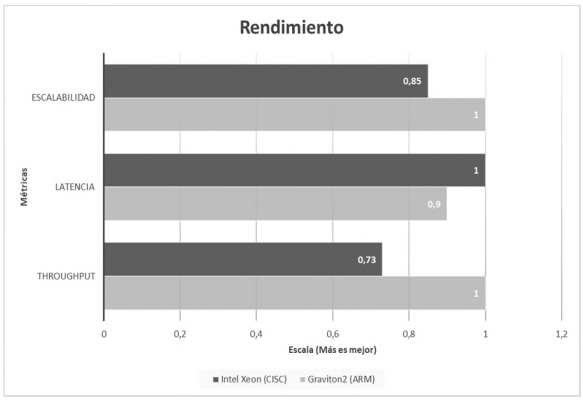


Figura 1. Comparación de rendimiento entre arquitecturas CISC y RISC en centros de datos.

Respecto del rendimiento, la Figura 1 integra métricas de escalabilidad, latencia y throughput, permitiendo visualizar diferencias entre arquitecturas en escenarios representativos de centros de datos. Estos indicadores resultan relevantes para cargas de trabajo intensivas en cómputo y datos, ya que condicionan tanto el procesamiento sostenido como el comportamiento ante escenarios de alta concurrencia y escalamiento, aspectos centrales en la ejecución de algoritmos de aprendizaje automático y procesos de razonamiento automatizado.

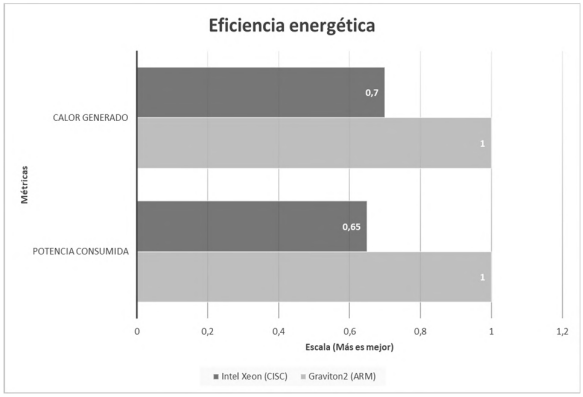


Figura 2. Comparación de eficiencia energética entre arquitecturas CISC y RISC.

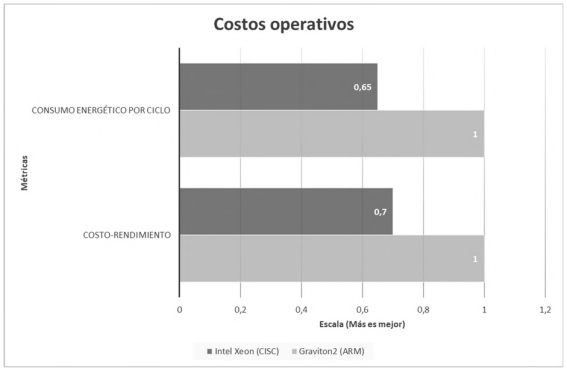


Figura 3. Comparación de costos operativos y viabilidad técnica de arquitecturas CISC y RISC.

En cuanto a la eficiencia energética, la Figura 2 presenta una comparación basada en la potencia consumida y el calor generado por cada enfoque arquitectónico. Los resultados evidencian diferencias en la relación rendimiento–vatio, variable determinante para el funcionamiento continuo de centros de datos, debido a su impacto directo en los costos de refrigeración, el consumo energético total y la sostenibilidad operativa de infraestructuras que soportan cargas prolongadas e intensivas. Desde la perspectiva de los costos operativos, la Figura 3 resume la relación entre consumo energético por ciclo de cómputo y costo–rendimiento. Esta comparación permite

analizar la factibilidad económica de las alternativas arquitectónicas evaluadas, especialmente en entornos donde el costo por unidad de cómputo constituye un factor decisivo para la escalabilidad de servicios y plataformas basadas en Inteligencia Artificial.

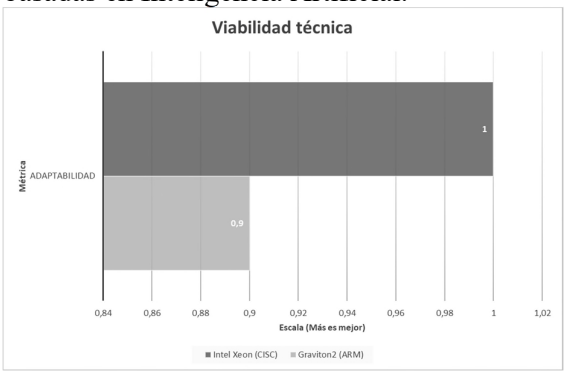


Figura 4. Variable Viabilidad técnica con la métrica: Adaptabilidad

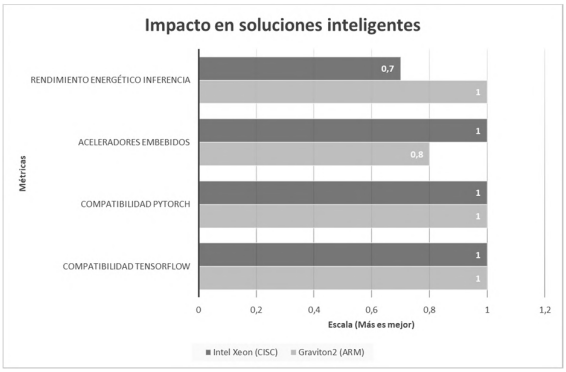


Figura 5. Variable Impacto en soluciones inteligentes, con las métricas: Rendimiento energético inferencia, aceleradores embebidos, compatibilidad PyTorch, compatibilidad TensorFlow

Complementariamente, la viabilidad técnica se aborda en la Figura 4 a través de la variable de adaptabilidad. En este análisis se consideran aspectos vinculados a compatibilidad, disponibilidad de soporte, madurez del ecosistema y requerimientos de adecuación tecnológica, los cuales condicionan la adopción efectiva de una arquitectura en infraestructuras existentes y su integración con entornos de ejecución y herramientas actuales.

Finalmente, y en línea con el aporte distintivo de la línea de investigación, la Figura 5 incorpora una variable específica orientada al impacto en soluciones inteligentes. Esta figura integra métricas directamente vinculadas a la ejecución de algoritmos de Inteligencia Artificial, tales como el rendimiento energético en inferencia, la presencia de aceleradores embebidos y la compatibilidad con frameworks relevantes. De este modo, la comparación arquitectónica trasciende métricas generales y se focaliza en factores habilitantes de bajo nivel para el despliegue eficiente de soluciones basadas tanto en aprendizaje automático como en enfoques de razonamiento simbólico.

En conjunto, el análisis de las Figuras 1 a 5 pone de manifiesto que las diferencias entre arquitecturas CISC y RISC se traducen en impactos concretos sobre el rendimiento, la eficiencia energética, los costos operativos y la viabilidad técnica en centros de datos. En particular, los resultados muestran que las arquitecturas basadas en ARM presentan ventajas relevantes en términos de eficiencia energética y costo–rendimiento, mientras que las arquitecturas x86 conservan fortalezas asociadas a su madurez, compatibilidad y soporte tecnológico. Estas diferencias adquieren especial relevancia al considerar cargas de trabajo asociadas a Inteligencia Artificial, ya que condicionan la factibilidad, escalabilidad y sostenibilidad de soluciones tanto de aprendizaje automático como de razonamiento simbólico, reforzando el rol de la arquitectura de procesamiento como infraestructura habilitante de bajo nivel para la Transformación Digital de ecosistemas inteligentes.

FORMACION DE RECURSOS HUMANOS

Como parte del trabajo encarado en la línea de investigación presentada, se logró desarrollar y defender un Trabajo Final de Carrera, el cual obtuvo una calificación de 9/10. Dicho trabajo permitió la formación del estudiante en actividades propias de la investigación aplicada, incluyendo el análisis crítico de literatura especializada, la evaluación comparativa de alternativas tecnológicas y la articulación de resultados técnicos con problemáticas actuales de infraestructura computacional.

Asimismo, se prevé la difusión de los resultados obtenidos en el ámbito de la Universidad, a través

de instancias de exposición y comunicación académica, tales como charlas, presentaciones internas y actividades de transferencia hacia cátedras vinculadas a arquitecturas de computadoras, sistemas operativos y computación de alto desempeño. Estas acciones buscan fortalecer la formación de estudiantes avanzados y promover el interés por la incorporación de nuevos recursos humanos a la línea de investigación.

En este sentido, la línea se concibe como un espacio formativo en crecimiento, con potencial para la incorporación de estudiantes en futuras instancias de trabajos finales y actividades de investigación aplicada, contribuyendo a la consolidación de capacidades académicas vinculadas a la infraestructura computacional como base para la Transformación Digital.

CONCLUSIONES Y TRABAJOS FUTUROS

La línea de investigación presentada permitió analizar comparativamente arquitecturas CISC y RISC en el contexto de centros de datos, considerándolas como infraestructura de bajo nivel para la ejecución de algoritmos de IA partir del trabajo desarrollado, se pone de manifiesto que las decisiones arquitectónicas inciden de manera directa en métricas clave como el rendimiento, la eficiencia energética y los costos operativos, las cuales condicionan la viabilidad técnica y operativa de estos entornos.

Un aporte central de la línea consiste en haber focalizado el análisis arquitectónico en variables de alto impacto para la ejecución de soluciones de IA, incluyendo tanto enfoques basados en aprendizaje automático como en razonamiento simbólico, vinculando aspectos de bajo nivel del hardware con problemáticas actuales de infraestructura computacional en el marco de la Transformación Digital de ecosistemas inteligentes.

Como trabajos futuros, se prevé profundizar análisis longitudinales sobre eficiencia energética en entornos reales de producción, con el fin de validar empíricamente el comportamiento térmico y el consumo energético más allá de escenarios controlados.

Asimismo, se plantea la continuidad de la línea a través de la incorporación de nuevos casos de estudio de ARM en contextos corporativos en el ámbito académico, considerando tanto factores técnicos como económicos. En paralelo, se propone analizar las arquitecturas RISC de código abierto en escenarios donde la IA impacte fuertemente en la soberanía tecnológica. En este sentido, se avanza actualmente en ponderar a RISAC-VC analizando sus elementos claves potenciadores de la Transformación Digital inteligente.

REFERENCIAS

[AG12] Aroca, R. & Gonçalves, L. (2012). Towards green data centers: A comparison of x86 and ARM architectures power efficiency. Journal of Parallel and Distributed Computing,72(12), 1770–1780.

[AH07] Fida El-Din, Aws & Krad, Hasan. (2007). A New Trend for CISC and RISC Architectures. Asian Journal of Information Technology.

[F17] Furber, S. (2017). Microprocessors: The engines of the digital age. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 473(2207), 20160893.

[GS21] Gupta, K., & Sharma, T. (2021). Changing trends in computer architecture: A comprehensive analysis of ARM and x86 processors. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 7(3), 619–631.

[HP19] Hennessy, J. L., & Patterson, D. A. (2019). Computer Architecture: A Quantitative Approach (5ª ed.). Elsevier.

[M22] Mahammad, L. (2022). Prediction of the Dominant Instruction Set Architecture in Data Centers. [Tesis de Maestría, Politecnico di Torino]. Webthesis.

[S16] Stallings, W. (2016). Computer Organization and Architecture (7ª ed.). Pearson.

[T00] Tanenbaum, A. S. (2000). Organización de computadoras: Un enfoque estructurado (6ª ed.). Prentice Hall.

ARQUITECTURA PARA LA VERIFICABILIDAD CRIPTOGRÁFICA DE EVENTOS BASADA EN BLOCKCHAIN

Lodato Marcos; Jaime Fernando; Hernandez Mario Alejandro

CAETI - Universidad Abierta Interamericana

Av. Pellegrini 944 – Rosario, Santa Fe, Argentina

{marcos.lodato; fernando.jaime}@alumnos.uai.edu.ar {alejandromario.hernandez}@uai.edu.ar

RESUMEN

Esta investigación aborda el problema de la falta de garantías verificables en la ejecución de procesos computacionales dentro de sistemas distribuidos modernos. Su propósito es analizar cómo los modelos tradicionales de registro y auditoría resultan insuficientes para asegurar integridad, orden y autenticidad de los eventos, y proponer una arquitectura que incorpore mecanismos criptográficos y tecnología blockchain como capa de verificación externa.

Se estudiarán los fundamentos de la Arquitectura Basada en Eventos (EDA) y del patrón Event Sourcing como marcos para modelar procesos mediante secuencias de eventos inmutables. Asimismo, se examinarán distintos mecanismos criptográficos, como funciones hash y Merkle Trees, con el fin de estructurar pruebas de integridad verificables. La investigación indaga en la integración de estos mecanismos con redes blockchain, utilizando sus propiedades de inmutabilidad y consenso distribuido como medio de sellado temporal y registro confiable.

Los objetivos principales son diseñar un modelo arquitectónico que permita transformar eventos operativos en evidencias criptográficas auditables, fortalecer la trazabilidad histórica y habilitar auditorías independientes sin exposición de datos sensibles. Se espera que esta propuesta contribuya al desarrollo de sistemas distribuidos con mayores garantías de confianza y verificabilidad.

Palabras clave: Arquitectura, Merkle Root, Merkle Tree, Blockchain, Smart Contract

CONTEXTO

Este proyecto pertenece al área Ingeniería de Software, una de las tres áreas principales del CAETI "Centro de Altos Estudios en Tecnología Informática". La institución encargada de coordinar y brindar herramientas en nuestra investigación es la Universidad Abierta Interamericana - UAI.

1. INTRODUCCIÓN

La creciente complejidad de los sistemas informáticos distribuidos ha transformado profundamente la manera en que se diseñan, ejecutan y supervisan los procesos computacionales. En la actualidad, arquitecturas basadas en microservicios, computación en la nube y modelos orientados a eventos permiten construir sistemas altamente escalables y desacoplados. Sin embargo, esta evolución también ha introducido nuevos desafíos relacionados con la integridad, trazabilidad y verificabilidad de las ejecuciones.

Tradicionalmente, los mecanismos de auditoría se apoyan en sistemas de logging centralizados o bases de datos transaccionales que registran cambios de estado. No obstante, estos enfoques presentan limitaciones estructurales: la información puede ser alterada por actores con privilegios administrativos, los registros pueden carecer de garantías criptográficas y la verificación suele depender del mismo dominio que genera los datos. Esta

situación resulta crítica en sistemas donde la confianza, el cumplimiento normativo o la transparencia operativa son requisitos fundamentales. [1]

Frente a este escenario, la presente investigación propone analizar el uso de tecnología blockchain como una capa arquitectónica de verificación externa, capaz de registrar evidencias criptográficas inmutables de eventos computacionales. El objetivo es diseñar e implementar en nuestro sistema de trazabilidad industrial un modelo que permita fortalecer la auditoría de procesos distribuidos mediante mecanismos de hash, estructuras de prueba como árboles de Merkle y sellado temporal descentralizado, contribuyendo al desarrollo de sistemas con mayores garantías de confianza y verificabilidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La presente investigación se estructura en cinco líneas fundamentales orientadas a analizar, diseñar y formalizar una arquitectura de verificabilidad distribuida basada en blockchain para procesos computacionales. Cada línea aborda un conjunto específico de fundamentos teóricos, decisiones arquitectónicas y mecanismos técnicos que, en conjunto, permiten construir un modelo integral de trazabilidad y auditoría criptográfica en sistemas distribuidos.

2.1 Fundamentos Arquitectónicos

2.1.1 Arquitectura basada en eventos y modelo productor-consumidor

La arquitectura basada en eventos (Event-Driven Architecture) constituye uno de los paradigmas estructurales en los cuales se fundamenta el sistema propuesto. Este tipo de modelo organiza el flujo de la información en torno a la detección, publicación y procesamiento de eventos que se registren dentro de un sistema en particular, comprendidos como cambios de estado significativos dentro del mismo. [2]

Un evento representa una ocurrencia relevante que puede desencadenar comportamientos o alteraciones adicionales de los componentes desacoplados. En este esquema, los productores generan eventos al detectar modificaciones en el dominio del que son responsables, mientras que por otro lado los consumidores reaccionan a dichos eventos de manera asíncrona mediante middleware o broker. [3]

Este desacoplamiento favorece ampliamente la escalabilidad de manera independiente, la tolerancia a fallas y la evolución de manera progresiva por parte del sistema. Por otro lado, si bien EDA mejora la flexibilidad operativa, no contiene mecanismos dentro de su esquema que permitan demostrar la integridad histórica de los eventos una vez se encuentran registrados. [4]

2.1.2 Persistencia Event Sourcing

Por otro lado, se encuentra el patrón Event Sourcing, que propone que el estado de una aplicación no sea almacenado como una representación mutable final, sino como resultado de la aplicación secuencial de eventos inmutables generados de manera cronológica.

Cada evento constituye un cambio atómico y significativo, persistiendo en un log que actúa como registro íntegro del sistema. La reconstrucción del estado puede realizarse mediante la reproducción ordenada de los eventos, lo que contribuye a la trazabilidad interna y auditoría detallada. [5]

Algunas de sus ventajas destacan la preservación de un historial completo y capacidad para reproducir estados pasados. Sin embargo, esta arquitectura sigue dependiendo del mismo dominio central que pueda garantizar la integridad de los registros, lo cual

limita seriamente la verificabilidad independiente.

2.1.3 Limitaciones de verificabilidad en sistemas independientes

En sistemas tradicionales, aún en los casos en los cuales se utilizan mecanismos criptográficos como hashing o firmas digitales, la custodia de dichas evidencias sigue formando parte y bajo el control de la misma entidad que ejecuta los procesos auditados. [6]

En uno de sus artículos Avizienis sostiene que en ningún momento la confiabilidad de la información de un sistema puede basarse solamente en componentes que comparten un mismo dominio administrativo, ya que esto restringe la detección de fallas intencionales o accidentales. [7]

Dicha centralización introduce un único punto de falla en términos de integridad histórica, justificando la necesidad de un mecanismo externo e independiente que permita verificar la inmutabilidad de los eventos auditados.

2.2 Arquitectura de verificabilidad distribuida

2.2.1 Desacople de confianza

En esta investigación se propone el diseño de una arquitectura que propone solucionar los problemas de integridad anteriormente mencionados, para ello se busca separar el sistema productor de eventos del mecanismo encargado de garantizar su integridad a largo plazo.

Este enfoque propone evitar que la entidad auditada sea simultáneamente responsable de generar, almacenar y validar sus propios registros. El hecho de delegar la verificación a una infraestructura externa permitirá auditorías independientes reduciendo el riesgo asociado al dominio administrativo único. [7]

2.2.2 Dominios lógicos: operativo, evidencia y persistencia

La arquitectura se estructura en tres dominios claramente delimitados:

- Operativo (off-chain): donde se generan y almacenan eventos completos.
- De evidencia: aquel que transforma los eventos en representaciones criptográficas.
- Persistencia inmutable: En este dominio, implementado mediante tecnología blockchain, se almacenan los compromisos criptográficos.

Narayanan habla de la blockchain como un "timestamped append-only log" [8], lo cual la convierte en un mecanismo ideal para registrar pruebas de integridad sin almacenar la información completa.

2.2.3 Modelo de persistencia dual

Esta arquitectura implementa un modelo de persistencia dual donde se distingue el evento completo almacenado en el dominio principal y su representación criptográfica mínima almacenada en la blockchain. Esta separación reduce la exposición de los datos sensibles y limita la información persistida on-chain a pruebas matemáticas de existencia e integridad.

2.3 Procesamiento criptográfico y agregación

2.3.1 Generación de hash

Para asegurarnos de la reproducibilidad, cada evento proveniente del dominio central es sometido a un proceso de serialización donde se define el orden de los campos, codificación y normalización de datos.

Katz y Lindell enfatizan que "la importancia de definiciones precisas es ampliamente comprendida por los desarrolladores e ingenieros de seguridad" [9], lo que implica que cualquier ambigüedad puede comprometer la seguridad de la arquitectura.

La representación resultante permite ser utilizada como parámetro para una función hash criptográfica con propiedades de

resistencia a preimagen y colisión, la utilizada para esta arquitectura es Keccak-256 por su compatibilidad en Smart Contracts desplegados en la red Ethereum. Stallings señala que dichas características permiten que el hash actúe como una huella digital confiable del mensaje original. [10]

El resultado del proceso es un digest de 256 bits que representan de manera inequívoca el evento serializado, cualquier modificación mínima en el evento produce una variación completa en el valor resultante.

2.3.2 Construcción Merkle Tree y Merkle Root

Esta arquitectura tiene como principal objetivo optimizar costes de almacenamiento y operativos, es por ello por lo que para agrupar los hashes individuales se utilizan los Merkle Trees. [11]

El proceso consiste en combinar pares de hashes y aplicar recursivamente la función hash, los nodos hoja corresponden a hashes individuales. Los pares de nodos se concatenan y se someten nuevamente a la función hash, hasta obtener un único valor raíz denominado Merkle Root.

La modificación de cualquier evento altera el hash correspondiente y, en consecuencia, la Merkle Root, permitiendo detectar inconsistencias sin necesidad de recalcular el árbol completo. La Merkle Root obtenida constituye el único valor que será anclado en la blockchain, junto con un identificador de lote, ya que se crearán Merkle Trees dependiendo de intervalos de tiempo o por cantidad de eventos.

2.4 Anclaje en Blockchain y Smart Contract

2.4.1 Blockchain como sellado temporal

La infraestructura seleccionada para el anclaje criptográfico es la red de Ethereum, debido a su soporte nativo para el despliegue de Smart Contract en sus redes de prueba como Sepolia.

2.4.2 *Funciones Smart Contract de auditoría*

El smart contract se encuentra desarrollado en Solidity, contando con las funciones de:

- Registro Merkle Root asociada a identificador de lote.
- Consulta pública de raíces registradas.
- Event logs para trazabilidad on-chain.

No se almacenan eventos completos ni datos sensibles, sino únicamente la Merkle Root. El apartado de construcción del Merkle Tree y generación de hashes permanece en responsabilidad del dominio off-chain, reduciendo el costo de las transacciones.

En caso de que un auditor externo desee verificar la integridad de un evento, deberá adaptarlo al esquema definido, calcular su hash mediante Keccak-256 y reconstruir off-chain el Merkle Tree correspondiente para obtener la Merkle Root. Esta raíz se comparará con la registrada en la blockchain; si ambos valores coinciden, se confirma la integridad del evento, mientras que cualquier diferencia evidenciaría una alteración.

3. RESULTADOS OBTENIDOS/ESPERADOS

La presente investigación nos ha permitido establecer una base conceptual y arquitectónica para el desarrollo de un modelo de verificabilidad distribuida orientado a la trazabilidad de procesos computacionales. A partir de la integración entre Arquitectura Basada en Eventos y mecanismos criptográficos, se ha definido una estructura que desacopla el dominio operativo del dominio de verificación, con almacenamiento off-chain y anclaje on-chain.

En cuanto a los resultados esperados, se proyecta avanzar hacia el desarrollo e implementación práctica de la arquitectura diseñada, con especial énfasis en el despliegue correcto del Smart Contract y la integración funcional con el sistema central basado en eventos. A futuro, esta línea de investigación busca consolidarse como un aporte

metodológico para el diseño de sistemas que requieran garantías formales de integridad y trazabilidad, promoviendo el uso de infraestructuras descentralizadas como complemento de arquitecturas empresariales tradicionales.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por dos estudiantes de quinto año de la carrera de Ingeniería en Sistemas de la Universidad Abierta Interamericana. Estos estudiantes son responsables de liderar y ejecutar todas las actividades y tareas del proyecto.

El objetivo principal del equipo es aplicar la arquitectura a sistemas que requieran garantías formales de integridad y trazabilidad, además de promover su aplicación práctica. Asimismo, buscan enriquecerse tanto a nivel estudiantil como profesional mediante esta investigación, aprovechando las experiencias y conocimientos adquiridos en el proceso.

5. BIBLIOGRAFÍA

[7] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," IEEE Trans. Dependable Secure Comput., vol. 1, no. 1, pp. 11–33, 2004.

[2] IBM, "¿Qué es la arquitectura basada en eventos?," Dec. 1, 2025. [Online]. Available: https://www.ibm.com/es-es/think/topics/event-driven-architecture

[9] J. Katz and Y. Lindell, Introduction to Modern Cryptography. Boca Raton, FL, USA: CRC Press, 2021.

[11] R. C. Merkle, "Protocols for public key cryptosystems," in Proc. 1980 IEEE Symp. Security and Privacy, 1980.

[8] A. Narayanan, J. Bonneau, E. W. Felten, A. Miller, and S. Goldfeder, Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction. Princeton, NJ, USA: Princeton University Press, 2016.

[5] Paradigma Digital, "Patrones de arquitectura de microservicios: Event Sourcing y arquitectura orientada a eventos (EDA)," 2024.

[3] H. F. O. Rocha, Practical Event-Driven Microservices Architecture: Building Sustainable and Highly Scalable Event-Driven Microservices. Berkeley, CA, USA: Apress, 2021.

[6] H. F. Oliveira Rocha, Practical Event-Driven Microservices Architecture. Berkeley, CA, USA: Apress, 2021.

[10] W. Stallings, Cryptography and Network Security: Principles and Practice. Upper Saddle River, NJ, USA: Pearson, 2016.

[4] 3Pillar, "Components of event-driven architecture", Sep. 20, 2021.

[1] 3Pillar, "Disadvantages of Event-Driven Architecture", Sep. 20, 2021.

CuyOS: UN SISTEMA OPERATIVO PARA ENTORNOS ACADÉMICOS

Fernandez Juan Manuel[1] ,Godoy Pablo Daniel[1,2,3,4]
[1]Universidad Nacional de Cuyo, Facultad de Ingeniería
[2]Universidad Nacional de Cuyo, Facultad de Ciencias Exactas y Naturales
[3]Universidad Nacional de Cuyo, ITIC
[4]Universidad de Mendoza
juanmanuelf12@gmail.com

RESUMEN

El presente trabajo describe el diseño e implementación desde cero de CuyOS, un sistema operativo de 32 bits orientado a fines educativos. El objetivo principal del proyecto fue desarrollar un núcleo simple, funcional y comprensible, capaz de ilustrar los conceptos fundamentales del diseño de sistemas operativos, tales como la gestión de memoria, manejo de interrupciones, planificación de procesos y salida gráfica simple.
CuyOS busca ser una plataforma de aprendizaje accesible para estudiantes de grado en etapas tempranas de su formación, permitiendo explorar el funcionamiento interno de un sistema operativo real sin la complejidad de los sistemas operativos modernos.

El trabajo forma parte de una serie de desarrollos tecnológicos destinados a educación, como simuladores y laboratorios remotos, desarrollados en la Universidad Nacional de Cuyo por docentes y estudiantes de la carrera Licenciatura en Ciencias de la Computación.

Palabras clave: Sistemas operativos, entornos académicos, diseño de sistemas operativos.

CONTEXTO

Investigadores y estudiantes de la carrera Licenciatura en Ciencias de la Computación han desarrollado herramientas destinadas a educación. Entre estos desarrollos pueden citarse laboratorios remotos de redes de sensores, simuladores de ondas mecánicas y corriente alterna, entre otros [1][2]. Otros trabajos dentro de esta línea están en marcha, como juegos para ayudar en la alfabetización de niños en el nivel primario de escolaridad.

1. INTRODUCCIÓN

1.1 Motivación y objetivos

En la actualidad, los sistemas operativos de mayor difusión son Linux, FreeBSD, MacOS y Microsoft Windows, siendo los dos primeros de código abierto [3]. Estos sistemas constituyen desarrollos de software de muy elevada complejidad técnica y considerable magnitud en cuanto a tamaño y estructura. Su funcionamiento resulta fundamental para la ejecución y gestión de la mayoría de las aplicaciones actuales, tanto en contextos empresariales, científicos y domésticos.

Introducir a estudiantes en el estudio de los sistemas operativos, y de manera particular, en su desarrollo, representa un desafío considerable. Si bien existen alternativas de código abierto como las anteriormente mencionadas, su gran tamaño, nivel de sofisticación dificultan significativamente su comprensión.

Para dar solución al problema mencionado antes, este trabajo presenta un sistema operativo denominado CuyOS, de dimensiones reducidas, que incorpora funcionalidades elementales y cuyo código ha sido concebido prioritariamente con fines de

claridad, simplicidad estructural y valor pedagógico.

Se establece explícitamente que la eficiencia y la eficacia computacional no constituyen requisitos prioritarios de diseño. En coherencia con este enfoque, los distintos módulos del sistema incorporan algoritmos e implementaciones elementales, seleccionados con el propósito de simplificar su análisis, estudio y modificación, promoviendo así una experiencia formativa clara y amigable.

1.2 Trabajos Relacionados

MINIX (Mini-Unix) fue concebido y desarrollado por Andrew S. Tanenbaum en 1987 con un propósito eminentemente educativo. En una época en la que los sistemas Unix eran cada vez más complejos y cerrados, Tanenbaum decidió crear un sistema operativo lo suficientemente pequeño, claro y accesible como para ser comprendido por estudiantes en el transcurso de un solo semestre universitario. Su código fuente, completamente disponible y documentado, se distribuía junto a su libro Operating Systems: Design and Implementation, permitiendo a los estudiantes examinar, modificar y compilar el sistema por sí mismos. A lo largo del tiempo, MINIX evolucionó significativamente [4].

MikeOS [5] es un sistema operativo de 16 bits, desarrollado con fines didácticos y diseñado para ejecutarse en arquitecturas compatibles con x86. Está escrito íntegramente en lenguaje ensamblador y presenta una arquitectura simple.

MikeOS incluye una interfaz de línea de comandos básica, así como un conjunto de programas integrados que abarcan desde un editor de texto hasta pequeños juegos y herramientas utilitarias [5].

DreamOS64 [6] es un sistema operativo en desarrollo cuyo propósito principal es el aprendizaje y la experimentación con técnicas modernas de programación a bajo nivel. Escrito para arquitectura x86-64, el proyecto fue iniciado como un ejercicio personal por su autor. A diferencia de otros sistemas educativos más estructurados, DreamOS64 no sigue un diseño pedagógico formal, pero representa una fuente de conocimiento valiosa para desarrolladores interesados en arquitecturas de 64 bits.

1.3 Descripción de CuyOS

El desarrollo fue llevado a cabo sin la incorporación de bibliotecas externas ni dependencias complejas, empleando exclusivamente los lenguajes C y ensamblador. El sistema es cargado mediante un gestor de arranque compatible con el estándar Multiboot 2 y dispone de soporte para multitarea, manejo de entrada por teclado e impresión en pantalla a través de una fuente tipográfica embebida. Asimismo, cuenta con un entorno de pruebas sustentado en el uso de QEMU y GDB, lo que posibilita la emulación, depuración y validación sistemática de su funcionamiento.

Las figuras 1 y 2 muestran el arranque de CuyOS según el estándar Multiboot 2.

Set de instrucciones y organización. El conjunto de instrucciones adoptado corresponde a la arquitectura x86, en virtud de su amplia difusión y exhaustiva documentación en el ámbito del desarrollo de sistemas operativos.

El procesador opera en modo protegido de 32 bits, con el propósito de evitar el uso de estructuras de mayor complejidad.

En cuanto a la organización interna del sistema, se adopta una arquitectura monolítica, lo que implica que todos los componentes necesarios para su funcionamiento se encuentran integrados dentro del binario del núcleo (kernel).

Figura 1: arranque de CuyOS



Figura 2: inicialización del sistema operativo mediante el framebuffer emulado por QEMU.



La principal ventaja de este enfoque radica en su simplicidad, ya que permite incorporar y modificar funcionalidades sin la necesidad de diseñar y mantener interfaces formales entre módulos, favoreciendo así un desarrollo más directo y accesible.

Compilador. Luego de probar otras opciones, la elección fue Clang, un compilador de arquitectura modular desarrollado en el marco del Proyecto LLVM [7], cuya filosofía prioriza la portabilidad y la facilidad de compilación en múltiples plataformas. Esta característica simplifica la configuración del entorno de desarrollo y reduce las barreras técnicas iniciales.

Gestión de memoria. A fin de reforzar la simplicidad del diseño, la Tabla Global de Descriptores (GDT) se ha organizado en únicamente dos segmentos: un segmento de código de núcleo y un segmento de datos de núcleo. Esta decisión responde a que CuyOS, en su estado actual de desarrollo, no implementa soporte para espacio de usuario. Asimismo, la estructura de cada descriptor se ajusta estrictamente al formato especificado por Intel.

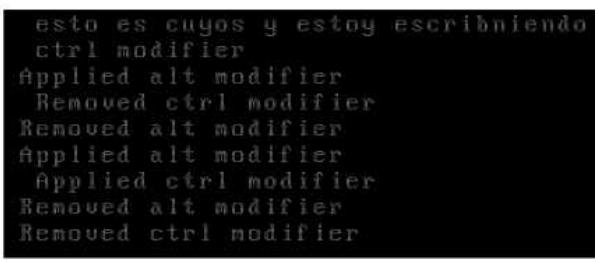
Interrupciones. El sistema operativo CuyOS cuenta con tres rutinas de servicio, que comienzan a partir del vector de interrupción 32 para evitar colisiones con los vectores de error del propio procesador. Las rutinas de servicio implementadas son: i) reloj interno de la CPU, ii) teclado y iii) una rutina de servicio de interrupción genérica.

La rutina de servicio del reloj de la CPU es utilizada principalmente para ejecutar tareas que requieren de temporalidad, entre estas, invocar scheduler cada 20 pulsos.

La rutina de servicio del teclado tiene la tarea de obtener el código de tecla presionado por el teclado, para así registrarlo y enviarlo al controlador del teclado. Los teclados envían códigos llamados "scancodes", los cuales son distintos en función de la versión implementada. Por motivos de simplicidad, se optó por utilizar la versión 1 de scancode.

La figura 3 muestra una captura de pantalla de fase de depuración. Puede apreciarse texto escrito por el usuario y la detección de las teclas modificadoras "Control" y "Alt". La captura muestra las frases "Applied alt modifier" y "Applied ctrl modifier", para indicar la detección de pulsaciones de dichas teclas.

Figura 3: Funcionamiento de las interrupciones y controlador de teclado.



Controlador de pantalla. El controlador de pantalla fue desarrollado mediante el uso del framebuffer. El mismo es una región de memoria que almacena de forma directa los valores de color correspondientes a los píxeles representados en pantalla. A diferencia del modo texto tradicional, CuyOS opera en modo gráfico, realizando la escritura de manera directa sobre el framebuffer, píxel a píxel. Este enfoque proporciona un mayor grado de control sobre la representación visual y sienta las bases para la futura implementación de interfaces gráficas de mayor complejidad y funcionalidad.

Scheduler. El sistema operativo fue concebido con soporte para multiprocesamiento, lo que requirió la implementación de un planificador de tareas o scheduler capaz de efectuar cambios de contexto a intervalos regulares. Dichos cambios se realizan cada vez que el reloj interno del procesador alcanza un umbral de 20 ticks, garantizando así una distribución periódica del tiempo de CPU entre los procesos activos.

Para la gestión del estado de cada proceso, el sistema mantiene un Process Control Block (PCB). Este es un arreglo de estructuras de datos del proceso, usual en todos los sistemas operativos. Cada PCB almacena información crítica asociada al proceso correspondiente, incluyendo el contenido completo de los registros de la CPU, la dirección de la última instrucción ejecutada y cualquier otro dato necesario para restaurar íntegramente su contexto durante una posterior reanudación.

En la implementación actual, las PCB poseen un tamaño fijo y permiten administrar hasta un máximo de 32 procesos de manera simultánea. Esta decisión de diseño responde al propósito de simplificar la gestión de los distintos contextos de ejecución, favoreciendo la claridad estructural del sistema.

2. LINEAS DE INVESTIGACION y DESARROLLO

El presente trabajo se encuadra en un proyecto de investigación y desarrollo financiado por la Universidad Nacional de Cuyo a través del proyecto cuyo código de identificación y título es "80020240100290UN: Laboratorios virtuales destinados a educación". Si bien dicho proyecto está centrado en laboratorios virtuales, también ha financiado herramientas tecnológicas destinadas a educación, como laboratorios trasladables al aula y desarrollo de prototipos de hardware y/o software con propósitos educativos [1].

CuyOS constituye una tesina final de la carrera Licenciatura En Ciencias de la Computación, Facultad de Ingeniería, Universidad Nacional de Cuyo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se cuenta con un sistema operativo funcional listo para ser empleado por estudiantes o profesores en tareas docentes.

3.1 Trabajos Futuros

Con la finalidad de adaptar la extensión del trabajo a una tesina final de carrera acotada a una duración de no más de un año, y permitir que otros estudiantes puedan continuar el trabajo ya realizado, se decidió dejar varias tareas como trabajos futuros. Cada una de estas tareas, por su extensión y complejidad, pueden constituir cada una una tesina final de carrera. Estas se resumen a continuación: a) Paginación y protección de memoria entre procesos, incluyendo la incorporación de espacio de usuario, b) Sistema de archivos básico, c) Capacidades de red, d) Máquina de estados para la gestión del ciclo de vida de los

procesos. Esta máquina debería implementar estados como ready, running, blocked, terminated, que son usuales en sistemas operativos actuales.

4. FORMACIÓN DE RECURSOS HUMANOS

CuyOS ha sido desarrollado en su totalidad por un estudiante de la carrera Licenciatura en Ciencias de la Computación de la Universidad Nacional de Cuyo, primer autor de este artículo, en el contexto de la asignatura Tesina. Dicha tesina fue defendida a finales de 2025, lo cual permitió al estudiante obtener el título de grado Licenciado en Ciencias de la Computación.

El trabajo tiene un propósito educativo, por lo que espera que pueda ser utilizado por futuros estudiantes de esta carrera o carreras similares, ya sea en asignaturas relacionadas con sistemas operativos, o a través de continuar el trabajo en proyectos finales de carrera.

El código fuente y documentación técnica del proyecto está disponible a través del enlace de la figura 4.

Otros estudiantes están comenzado sus trabajos finales de carrera siguiendo la línea de trabajos destinados a educación, como es el caso del trabajo cuyo título tentativo es “Alfabetizar jugando - juego educativo web para contribuir en la alfabetización de niños”.

AGRADECIMIENTOS

Los autores agradecen a la Secretaría de Investigación, Internacionales y Posgrado (SIIP), Universidad Nacional de Cuyo, que a través del proyecto “80020240100290UN - Laboratorios virtuales destinados a educación” financia la publicación de este artículo y a la carrera Licenciatura en Ciencias de la Computación, Universidad Nacional de Cuyo.

Figura 4: Enlace al código fuente y documentación técnica de CuyOS



5. BIBLIOGRAFIA

[1] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "A WSN testbed for teaching purposes"; IEEE Latin America Transactions. 14(7),3351–3357 (2016).

[2] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "Laboratorio remoto para la formación de usuarios basado en el cloud"; Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación 18, 7–18 (2016).

[3] Tanenbaum, Andrew S., Herbert Bos. Sistemas Operativos Modernos. Pearson Educación, 2022.

[4] Tanenbaum, A. S., Woodhull, A. S. Operating Systems: Design and Implementation, 3rd ed. Prentice Hall. 2006.

[5] Mike Saunders et. al., MikeOS x86 operating system website. Visitado en 2025. https://mikeos.sourceforge.net/.

[6] Gualandri Ivan, DreamOS64 - A 64-bit educational kernel project. Repositorio de GitHub. Visitado en 2025. https://github.com/dreamos82/Dreamos64.

[7] LLVM Foundation, LLVM Project website. Clang: a C language family frontend for LLVM. Visitado el 2 de mayo de 2025. https://clang.llvm.org/

Gestión automatizada de cultivos basada en IoT

Lazarte, Ivanna[1]; Aranda, Marcos[1]; Beltramini, Paola[1]; Contreras, Juan José[1]; Cabiedes, Joaquin[1]; Díaz, Karen[1]; Gallina, Sergio[1]

[1] Grupo De Investigación en IoT /Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca
Maximio Victoria Nº 55 - C.P: 4700 - San Fernando del Valle de Catamarca.
Tel: +54 (03834) 435112
pbeltramini@tecno.unca.edu.ar

RESUMEN

El presente trabajo propone el desarrollo de un nodo sensor basado en tecnologías 4.0 para la monitorización remota de variables agropecuarias críticas, considerando las condiciones productivas de Catamarca, buscando un sistema robusto con autonomía energética y escalable a distintas configuraciones productivas. Los datos recolectados se trasmitirán hacia plataformas de análisis en la nube donde la información será procesada y mostrada en tiempo real.

El proyecto se estructura en seis fases: investigación y análisis de requerimientos, diseño y fabricación de un prototipo de nodo sensor, conectividad de la red IoT, desarrollo de la plataforma de gestión de datos, pruebas y ajustes técnicos y evaluación y transferencia tecnológica.

Se espera que la integración de estas tecnologías (sensores inteligentes, redes de comunicación, procesamientos de datos y plataformas de análisis en la nube) permitan optimizar insumos, reducir costos y mejorar la calidad de los productos contribuyendo a una producción más eficiente y ambientalmente responsable.

***Palabras clave*:** IoT, agropecuarias, monitorización remota, procesamiento de datos.

CONTEXTO

Este trabajo de investigación se lleva a cabo con el Grupo De Investigación en IoT de la Facultad de Tecnología de la UNCA y se encuentra enmarcado dentro del Proyecto de Innovación y Transferencia (Pi+T) ‘‘Gestión automatizada de cultivos basada en IoT". El proyecto fue aprobado en la Convocatoria a Proyectos de Innovación y Trasnferencia (Pi+T)-2025. Este proyecto se enmarca en el Desafío CTi 2030", que impulsa la adopción de tecnologías de la información como motores de la innovación productiva y la transformación digital.

1. INTRODUCCIÓN

En el contexto actual, la integración de tecnologías 4.0 —como el Internet de las Cosas (IoT), la inteligencia artificial, el big data, la realidad aumentada y las plataformas digitales— se ha vuelto una necesidad ineludible para el sector agro productivo. Estas tecnologías no solo representan una ventaja competitiva, sino que constituyen una parte esencial del ecosistema productivo moderno. Su adopción transforma los modelos de negocio y promueve una cadena de valor más eficiente, sostenible e inclusiva, abarcando a todos los actores involucrados en la producción. [1].

En Argentina, las políticas públicas impulsadas por la Agenda Productiva 2030 apuntan precisamente a facilitar esta transición hacia una agricultura digital. Según el informe "AgTech - Argentina Productiva 2030", el sector agropecuario se encuentra en un proceso de transformación que promueve la eficiencia y sostenibilidad a través de tecnologías emergentes que mejoran la toma de decisiones y optimiza la gestión de recursos, aspectos fundamentales para el crecimiento del sector. Entre las líneas de acción planteadas se destacan el asesoramiento técnico y el financiamiento para la implementación de soluciones tecnologías 4.0 en el agro [1].

En la provincia de Catamarca, la incorporación de estas tecnologías en el sector agroganadero se encuentra en una etapa de desarrollo incipiente. Si bien existen iniciativas y avances en el uso de herramientas digitales en ganadería y cultivos extensivos, su implementación en algunos sectores productivos aún es limitada, especialmente sectores estratégicos como el vitivinícola, el olivícola y el nogalero. Estas actividades, fundamentales para la economía provincial, enfrentan desafíos productivos que requieren de herramientas tecnológicas para mejorar su eficiencia, sostenibilidad y competitividad.

Frente a este escenario, el presente proyecto propone el desarrollo de un nodo sensor basado en tecnologías 4.0 para la monitorización remota de variables agropecuarias críticas, específicamente adaptado a las condiciones productivas de Catamarca. Este nodo integrará sensores para la medición de diferentes variables relevantes como temperatura, humedad ambiental y de suelo, concentración de CO2, y ubicación georreferenciada mediante GPS. Además, se priorizará el diseño de un sistema robusto y eficiente, con autonomía energética y capacidad de escalar a distintas realidades productivas. El dispositivo estará preparado para operar en entornos rurales con condiciones climáticas adversas, garantizando una transmisión eficiente de los datos recolectados hacia plataformas de análisis en la nube, donde los productores podrán acceder a información procesada en tiempo real para la toma de decisiones estratégicas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las principales líneas de investigación que se abordan en este proyecto son:

- Internet de las Cosas (IoT) aplicado a la agricultura: Desarrollo e implementación de dispositivo IoT, como sensor para monitoreo en tiempo real de variables agropecuarias críticas.
- Ciencia de datos para la optimización agricola: Aplicación de técnicas de análisis de datos para determinar las condiciones productivas en tiempo real.
- Transformación digital y automatización en la agroindustria: Integración de tecnologías 4.0 en la producción agrícola para mejorar la toma de decisiones, optimizar recursos y promover la sostenibilidad del sector.

Para la ejecución del proyecto de I+D se planificaron las siguientes 6 etapas:

1. Investigación y análisis de requerimientos: se relevarán las necesidades de los productores y se identificarán las variables a monitorear. Se seleccionarán los sensores, microcontroladores y tecnologías de comunicación más apropiadas, considerando la topografía, conectividad y condiciones climáticas locales.
2. Diseño y fabricación de un prototipo de nodo sensor: se elaborarán los esquemas electrónicos, se programará el firmware y se construirá un prototipo que funcione en condiciones de laboratorio.
3. Conectividad de la red IoT: Se evaluarán distintas alternativas de conectividad (LoRaWAN, NB-IoT, 4G LTE, entre otras) para establecer una red de transmisión de datos eficiente y adecuada a las condiciones geográficas y productivas del área de implementación. Se diseñará la

arquitectura de red más apropiada y se validará la interoperabilidad del sistema.

4. Desarrollo de la plataforma de gestión de datos: se diseñará una interfaz web intuitiva para visualizar los datos en tiempo real. Además, se incorporarán herramientas de análisis e interpretación de la información recolectada.

5. Pruebas y ajustes técnicos: los prototipos serán evaluados en entornos controlados en condiciones reales de producción. Se ajustarán los parámetros de funcionamiento según los resultados obtenidos

6. Evaluación y transferencia tecnológica: se generarán informes técnicos, manuales de uso y actividades de capacitación para los productores. Se analizará su potencial de expansión a otros sectores de la agroindustria.

3. RESULTADOS OBTENIDOS/ESPERADOS

El Grupo De Investigación en IoT se encuentra enfocado en la integración de diversas tecnologías avanzadas aplicadas a la Agroindustria. Entre sus contribuciones se destacan el desarrollo de un sensor autónomo para el monitoreo de variables críticas en la producción, la transmisión de datos a una plataforma digital para su visualización, almacenamiento y análisis. A partir de los resultados esperados de los avances mencionados, se inician estudios tendientes a mejorar el sistema de monitoreo de variables críticas de producción.

En el marco del proyecto, se espera generar un conjunto de conocimientos científicos y tecnológicos que se materializarán en una herramienta avanzada para una mejor gestión de insumos, reducir costos y mejorar la calidad de los productos contribuyendo a una producción más eficiente y ambientalmente responsable.

Entre los resultados esperados, se pueden mencionar:

- Relevar las necesidades del sector productivo y establecer los requerimientos técnicos y funcionales del sistema.
- Desarrollar un prototipo funcional en condiciones de laboratorio.
- Establecer un sistema de transmisión de datos robusto y adecuado a la realidad geográfica de Catamarca.
- Crear una plataforma web funcional que permita la visualización, almacenamiento y análisis de los datos recolectados.
- Validar el funcionamiento del sistema completo en entornos controlados en condiciones reales de producción.
- Transferir la solución tecnológica a los productores y generar capacidades locales para su adopción.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada cuenta con una Doctora en Ingeniería (mención Sistemas de Información), un Magíster en Ingeniería de Computadores y Redes y dos alumnos avanzados de Ingeniería en Informática. Los profesionales del equipo son docentes investigadores Categoría III, IV y V.

También se prevé la capacitación y formación de recursos humanos, por medio de las siguientes actividades:

- Talleres y cursos sobre el uso de dispositivos IoT.
- Acompañamiento a estudiantes de grado y posgrado en el desarrollo de proyectos de investigación alineados con el proyecto.
- Organización de encuentros con productores, charlas técnicas y transferencia de conocimientos a comunidades agroindustriales.

5. BIBLIOGRAFÍA

[1] Lachman, J., Braude, H., Monzón, J., López, S., & Gómez-Roca, S. (2022). El potencial del agro 4.0

en Argentina: Diagnóstico y propuestas de políticas públicas para su promoción. Ministerio de Desarrollo Productivo.

[2] Gallina S.H., Lazarte I.M., Beltramini P.I., Aranda M.D.. (2022) "Hacia la implementación de un prototipo para la aplicación de IoT y Ciencia de Datos en la Actividad Agropecuaria Sostenible". 10 Congreso Nacional de Ingeniería Informática / Sistemas de Información - CONAIISI 2022-CONAIISI.

[3] Beltramini, P.I, Aranda, M.D.; Cano, J. E.; Villagrán, L. D.; Gallina, S. H. y Lazarte, I.M. (2024). Transmisión de Datos de Sensores a través de Redes LPWAN. II Congreso de Ciencia y Tecnología del CODINOA.

[4] Lazarte, I.; Aranda, M.D.; Gallina, S.; Cano, J.; Beltramini, P.I. (2023). Ciencia de Datos aplicada en la ganadería de precisión en Catamarca. 11° Congreso Nacional de Ingeniería Informática / Sistemas de Información- CoNaIISI 2023.

[5] Beltramini, P.I, Lazarte, I.M.; Aranda, M.D.; Cano, J. E.; Villagrán, L. D.; Herrera Conegliano, O. A.; & Gallina, S. H. (2023). IOT aplicado a la ganadería de precisión. XXV Workshop de Investigadores en Ciencias de la Computación (WICC).

[6] Aranda, M.; Beltramini, P.; Cano, J.; Virragrán, L.; Moreno, J.P.; Gallina, S.; & Herrera Conegliano, O. A. (2021). IoT aplicado a la ganadería extensiva. Revista Argentina de Ingeniería (RADI).

[7] Arevalo, A. L. & Riapira, C. A. (2017). Desarrollo de plataforma de internet de las cosas para toma de decisiones en modelos de agricultura de precisión. Recuperado de: http://hdl.handle.net/11349/8328

[8] Martinez, Jorge Luis Diaz, Salcedo, Dixon, Mercado, Teobaldis, & Quiñonez, Yadira. (2024). Internet de las cosas aplicado a la agricultura: estado actual y su aplicación mediante un prototipo. RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação, (53), 106-121. Epub 30 de abril de 2024.https://doi.org/10.17013/risti.53.106-121

Hacia un Framework de Decisión para Arquitecturas de Microservicios en Aplicaciones Móviles

Joaquín Albornoz [1 2],Susana Chavez[1], Adriana E. Martín[1], Sergio Flores[1]

[1] Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
[2] Alumno Avanzado Licenciatura en Sistemas de Información y Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.
Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

joaquinalbornoz1207@gmail.com, schavez@iinfo.unsj.edu.ar, arianamartinsj@gmail.com, sergiorflores@gmail.com

Resumen

La adopción de arquitecturas de microservicios en el backend de aplicaciones móviles ha mejorado la escalabilidad y la independencia de los equipos de desarrollo. Sin embargo, esta transición introduce complejidades significativas en el ecosistema móvil, tales como la sobrecarga de red, la persistencia segura de la identidad del usuario y la gestión de la conectividad intermitente. Este trabajo propone una línea de investigación orientada a la creación de un Framework de Decisión que establezca buenas prácticas de diseño y mantenimiento. El enfoque se centra en tres pilares: criterios de descomposición basados en dominios, protocolos de seguridad de extremo a extremo, y patrones de resiliencia en el cliente. El resultado esperado es una guía metodológica que permita a los desarrolladores optimizar la integración entre clientes móviles y ecosistemas distribuidos, garantizando una experiencia de usuario fluida y segura.

Palabras clave: *Distributed Applications. Microservices Architecture. Mobile Computing. Best Practices*.

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es la línea marco del proyecto de investigación Aplicaciones Distribuidas en Arquitectura de Microservicios. Desafíos técnicos y mejores prácticas, cuya propuesta se presentó en convocatoria para el período 2026-2027 y se encuentra en evaluación. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil, distribuida y de alta performance, desde hace más de 25 años, al que se han incorporado estudiantes de informática. Similarmente a proyectos anteriores, se continúa el trabajo con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

El desarrollo de aplicaciones móviles contemporáneas exige respuestas en tiempo real y alta disponibilidad. Frameworks modernos, como Flutter, permiten construir interfaces de usuario altamente reactivas, pero la calidad de la experiencia recae fuertemente en la arquitectura del backend que las soporta.

Históricamente, las aplicaciones móviles se conectaban a backends monolíticos, los cuales hoy representan un cuello de botella para la integración continua y el escalamiento [5].

La migración hacia arquitecturas de microservicios resuelve el problema del monolito [1], pero transfiere la complejidad operativa al cliente móvil [2]. Un dispositivo ahora debe orquestar múltiples llamadas de red, gestionar tokens de autenticación distribuidos [6] y manejar fallos parciales cuando un servicio no responde [4]. Ante este escenario, carecemos de un estándar unificado de diseño que guíe a los equipos de ingeniería. Este proyecto de investigación aborda dicha brecha mediante la formulación de un Framework de Decisión. Esta herramienta metodológica busca dotar a los desarrolladores de criterios claros para decidir cuándo y cómo integrar microservicios, mitigando los riesgos inherentes a la latencia, la seguridad y el funcionamiento offline.

Resiliencia Frente a Fallos

En la transición hacia arquitecturas de microservicios, asumir que las fallas de red y de servicios ocurrirán es un principio de diseño fundamental [4]. Para mitigar la degradación de la experiencia del usuario y evitar la "muerte por cascada" —donde un servicio inestable bloquea los recursos del dispositivo móvil esperando una respuesta—, este trabajo toma como pilar el patrón Circuit Breaker (Interruptor Automático). Este mecanismo actúa como un proxy preventivo que detecta anomalías y detiene peticiones recurrentes (Fail Fast), protegiendo la memoria y el procesamiento del cliente. A lo largo del presente artículo, exploraremos cómo la correcta integración de este patrón resulta indispensable para construir aplicaciones móviles robustas, capaces de auto-gestionarse frente a la latencia y la intermitencia inherente de las redes móviles.

El Cliente Móvil Reactivo

Para la validación empírica de los patrones propuestos, esta investigación adopta Flutter como entorno de desarrollo para el cliente móvil. La elección radica fundamentalmente en su naturaleza de framework reactivo y declarativo, donde la interfaz de usuario se construye como una función directa del estado de la aplicación. Al interactuar con un ecosistema de microservicios —en el cual la latencia de red, las respuestas parciales de un Backend for Frontend (BFF) o las transiciones de estado de un Circuit Breaker generan mutaciones asíncronas y constantes—, un enfoque reactivo resulta indispensable. Flutter permite que la vista se reconstruya de manera eficiente y automática ante cada evento o actualización de datos, facilitando la implementación de interfaces resilientes que reflejen en tiempo real la disponibilidad de los servicios subyacentes (como bases de datos en la nube o funciones serverless) sin bloquear el hilo principal ni degradar la experiencia del usuario.

Objetivos

Objetivo General: Definir y validar un Framework de Decisión que consolide un conjunto de mejores prácticas para el diseño, implementación y mantenimiento de aplicaciones móviles respaldadas por arquitecturas de microservicios.

Objetivos Específicos:

1. Definir Criterios de Descomposición: Establecer directrices metodológicas para determinar el momento exacto en el que un módulo de la aplicación

móvil debe desacoplarse en un microservicio independiente. Este enfoque priorizará la separación basada en dominios de negocio (Domain-Driven Design) por sobre la división tradicional en capas técnicas.

2. Diseñar Estrategias de Seguridad de Extremo a Extremo: Evaluar y recomendar protocolos que garanticen que la identidad y los permisos del usuario móvil persistan y se propaguen eficientemente a través de múltiples microservicios sin degradar el rendimiento.
3. Implementar Patrones de Resiliencia en el Cliente: Formular patrones de tolerancia a fallos ejecutados directamente en la lógica del dispositivo móvil. Esto incluye la adaptación del patrón Circuit Breaker del lado del cliente y estrategias de caché, con el fin de optimizar el consumo de recursos y sostener la usabilidad de la aplicación durante escenarios de conectividad degradada u offline.

Resultados Esperados

La complejidad de conectar clientes móviles con ecosistemas de microservicios requiere abandonar la improvisación arquitectónica a favor de metodologías estandarizadas. El Framework de Decisión propuesto en esta investigación no solo busca categorizar los problemas actuales, sino ofrecer soluciones pragmáticas y fundamentadas. Al abordar la descomposición por dominios, la seguridad distribuida y la resiliencia proactiva en el dispositivo, este trabajo aspira a sentar un precedente en las mejores prácticas de ingeniería de software móvil. A futuro, se espera que estas directrices reduzcan los tiempos de desarrollo, disminuyan la superficie de vulnerabilidad y garanticen aplicaciones móviles robustas, capaces de operar con fluidez independientemente de la estabilidad de los servicios subyacentes.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los investigadores que figuran en este trabajo de la Universidad Nacional de San Juan, Universidad de San Luis, Universidad Nacional de Salta y Universidad de Champagnat de Mendoza y ocho alumnos de grado, de la Universidad Nacional de San Juan.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis de maestría en áreas afines y tres tesinas de grado en el área de Serverless computing, Virtualización y Computación distribuida.

Por otro lado se espera aumentar el número de publicaciones, y además también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.

Referencias

[1] S. Newman, *Building Microservices: Designing Fine-Grained Systems*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2021.

[2] C. Richardson, *Microservices Patterns: With examples in Java*. Shelter Island, NY, USA: Manning Publications, 2018.

[3] E. Evans, *Domain-Driven Design: Tackling Complexity in the Heart of Software*. Boston, MA, USA: Addison-Wesley, 2004.

[4] M. T. Nygard, *Release It!: Design and Deploy Production-Ready Software*, 2nd ed. Raleigh, NC, USA: Pragmatic Bookshelf, 2018.

[5] N. Dragoni et al., "Microservices: How To Make Your Application Scale," in *Present and Ulterior Software Engineering*, Cham, Switzerland: Springer, 2017, pp. 95-104.

[6] J. Richer and A. Sanso, *OAuth 2 in Action*. Shelter Island, NY, USA: Manning Publications, 2017.

LABORATORIOS VIRTUALES DESTINADOS A EDUCACIÓN

Godoy Pablo Daniel[1,2,3,4], Marianetti Osvaldo Lucio[1,4], Chediack Ernesto[1,4], Manganelli Silvina[1,2], Sanchez Agustina[1]

[1]Universidad Nacional de Cuyo, Facultad de Ingeniería

[2]Universidad Nacional de Cuyo, Facultad de Ciencias Exactas y Naturales

[3]Universidad Nacional de Cuyo, ITIC

[4]Universidad de Mendoza

pablo.godoy@ingenieria.uncuyo.edu.ar

RESUMEN

El presente artículo describe avances de trabajos para la implementación de un laboratorio virtual o simulador simple de diagramas fasoriales y funciones en el tiempo de corriente alterna. El trabajo es realizado en el contexto de un proyecto aprobado y financiado por la Universidad Nacional de Cuyo, cuyo plazo de ejecución es entre 2025 y 2027.

El proyecto tiene como objetivos el desarrollo de laboratorios virtuales para ser utilizados en tareas docentes y realizar aportes en este campo o su aplicación en tareas de enseñanza .

El proyecto, que se encuentra en su etapa inicial de ejecución, es continuación de proyectos de investigación anteriores relacionados.

Palabras clave: Simuladores, entornos educativos, laboratorios virtuales.

CONTEXTO

El presente trabajo describe un proyecto de investigación y desarrollo financiado por la Universidad Nacional de Cuyo a través del proyecto cuyo código de identificación y título es "80020240100290UN: Laboratorios virtuales destinados a educación". Además, los investigadores del presente proyecto y colaboran estrechamente con un proyecto relacionado cuyo código de investigación y título es “80020240100288UN: La tecnología de Blockchain como un sistema descentralizado y distribuido en aplicaciones de IoT”. El plazo de ejecución es entre 2025 y 2027.

El proyecto surge como una evolución de dos proyectos de investigación anteriores y la integración de un trabajo de tesis doctoral. Los laboratorios remotos y nómadas han sido utilizados desde el año 2016 por estudiantes de las carreras Licenciatura en Ciencias de la Computación, Instituto Tecnológico Universitario (ITU) y Facultad de Ciencias Exactas y Naturales de Universidad Nacional de Cuyo y la Universidad de Mendoza. Pueden consultarse más detalles de dichos laboratorios remotos y experiencias de uso por parte de estudiantes de las carreras mencionadas en [1,2,3,4,5,6].

1. INTRODUCCIÓN

1.1 Objetivos

En los últimos años, y especialmente durante el dictado de clases en modalidad remota como consecuencia de la pandemia de COVID-19, se ha incrementado de manera significativa el uso de herramientas digitales tales como simuladores y laboratorios remotos. En este contexto, el grupo de investigación ha implementado diversas herramientas de este tipo, cuya utilización en el dictado de clases permitió evidenciar que contribuyen de forma muy significativa a facilitar el aprendizaje tanto en el aula y como

en el caso de estudio autónomo. Además, han sido valoradas de manera positiva por parte de los estudiantes.

El objetivo principal del presente proyecto es la incorporación de nuevos experimentos virtuales (o simulados) a las plataformas que se encuentran actualmente disponibles.

Como objetivo secundario se plantea la recopilación y el análisis de estadísticas de uso, así como de opiniones, críticas y sugerencias por parte de los estudiantes, con el propósito de mejorar tanto las herramientas desarrolladas como las estrategias de implementación de las mismas en el dictado de clases.

Finalmente, se espera que una de las integrantes del grupo de trabajo pueda realizar la defensa de su tesis doctoral durante el año 2026. Dicha tesis, titulada *"Desarrollo de competencias científico-tecnológicas en alumnos universitarios de carreras del área de las ciencias exactas mediante el método STEAM y el pensamiento computacional, necesarias para afrontar los desafíos del siglo XXI"*, se enmarca directamente en los objetivos del presente proyecto.

1.2 Trabajos Relacionados

Existen un número importante de simulaciones y laboratorios virtuales accesibles a través de Internet. Se mencionan a continuación 4 ejemplos. Existen varios laboratorios virtuales de física, uno de los más importantes son las Simulaciones Interactivas de la Universidad de Colorado [7], en la cual se proveen laboratorios virtuales de varios tipos, por ejemplo, laboratorios de ondas, laboratorios de termodinámica y laboratorios de óptica física. El Instituto Tecnológico de Massachusetts provee un simulador de lenguaje ensamblador de 8 bits [9]. Por último, el profesor profesor Vladimír Vaščák ha creado una gran cantidad de simuladores de mecánica, campos gravitatorios, ondas y vibraciones, termodinámica, electromagnetismo, óptica, relatividad, física atómica y nuclear, etc. [9].

A pesar de la existencia de gran cantidad de herramientas disponibles online, durante el dictado de clases se ha detectado que aún existen experiencias para las cuales no han sido creados laboratorios virtuales, o los existentes no satisfacen las necesidades particulares de las asignaturas dictadas por los profesores que forman parte del proyecto. Este proyecto pretende crear estas herramientas.

1.3 Estado actual de avance del proyecto

El proyecto se encuentra en sus etapas iniciales de ejecución. Sin embargo, es continuación de proyectos anteriores, por lo cual se han desarrollado varios laboratorios virtuales, entre ellos: laboratorio virtual de Efecto Doppler, laboratorio virtual de ondas mecánicas y laboratorio virtual de ventanas deslizantes de redes de computadoras.

Actualmente se está desarrollado un laboratorio virtual que muestra el diagrama fasorial de circuitos de corriente alterna. Este trabajo hace referencia a dicho laboratorio virtual.

El primer paso y el más importante es el diseño de su interfaz web, ya que la misma debe ser fácil de comprender, intuitiva y auto-descriptiva, de modo que el usuario puede comprender su funcionamiento a partir de la simple observación de sus elementos. Actualmente dicha interfaz web está en su etapa final de diseño.

La interfaz web del laboratorio virtual muestra:

a - El diagrama fasorial giratorio.

b - Las funciones diferencia de potencial en cada elemento del circuito y la corriente en función del tiempo.

c - Controles para que los estudiantes puedan configurar los valores de los elementos del circuito y sus unidades.

d - Controles para que los estudiantes puedan elegir que curvas y que fasores del diagrama fasorial mostrar. Esto permite a los estudiantes armar y analizar el diagrama paso a paso.

e - Panel de cálculos de varios parámetros del circuito.

La Figura 1 muestra la interfaz web completa. La misma puede sufrir modificaciones menores en el futuro.
La Figura 2 muestra un detalle del diagrama fasorial generado y las funciones que representan las diferencias de potencial en los elementos del circuito, la corriente y el ángulo de fase, para un circuito RLC inductivo. La figura 3 muestra la misma información pero para un circuito en resonancia.
Los diagramas fasoriales son giratorios y los gráficos son animados. Los estudiantes pueden avanzar y pausar la animación, quitar o agregar elementos.
La Figura 4 muestra el panel de resultados. Dicho panel permite a los estudiantes verificar los resultados de ejercicios que resuelvan.

Figura 1: Interfaz web completa

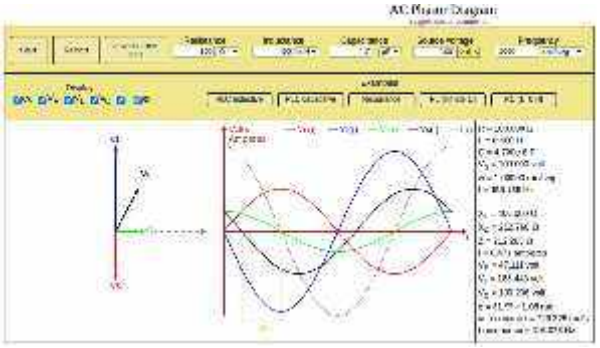


Figura 2: Interfaz web mostrando el diagrama fasorial de un circuito RLC serie inductivo

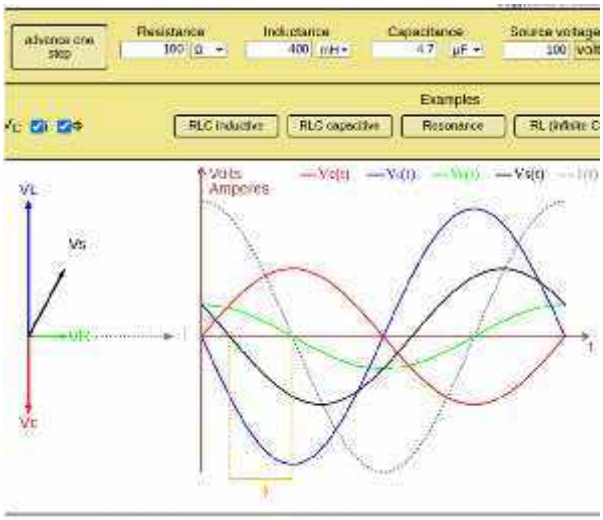
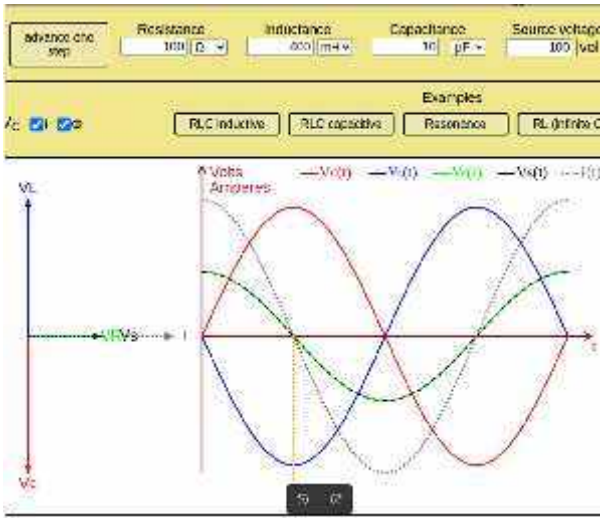


Figura 3: Interfaz web mostrando el diagrama fasorial de un circuito RLC en resonancia



2. LINEAS DE INVESTIGACION y DESARROLLO

2.1 Laboratorios virtuales

Se basan en modelos físicos implementados en software. Su ventaja es que escalan fácilmente, no tienen limitaciones de equipamiento y pueden estar disponibles online 24/7. Como desventaja, los modelos pueden no ser exactos y no se trabaja con

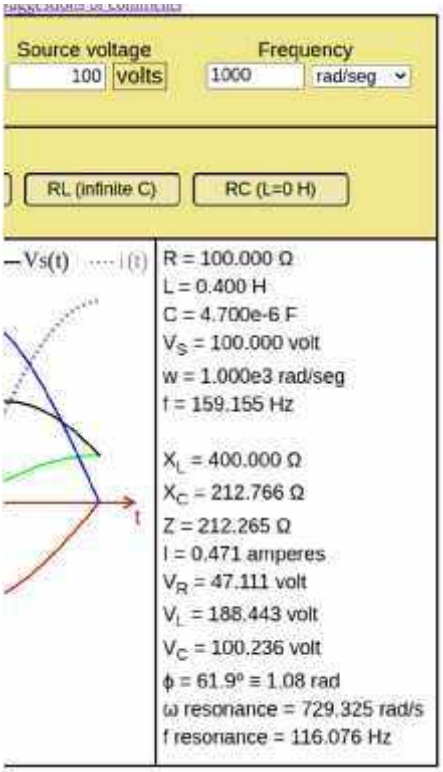
equipos reales. Aun así, han demostrado que pueden facilitar la comprensión de conceptos complejos [8], [9].

Los laboratorios virtuales implementados en este proyecto están basados en tecnología web, de manera de poder ser accedidos a través de navegadores web típicos de cualquier computadora o teléfono celular.

2.2 Prácticas educativas STEAM

Una tesis doctoral en curso, cuyo título es "Promoción de Competencias Científicas en Estudiantes de Educación Superior o Universitaria de Carreras del Área de las Ciencias Exactas, Mediante El Método STEAM Y El Pensamiento Computacional" analiza cómo la integración de herramientas digitales en prácticas STEAM contribuye al desarrollo de competencias científicas, tecnológicas y digitales en estudiantes de ciencias e ingeniería, utilizando instrumentos específicos de evaluación. Entre dichas herramientas se incluyen simuladores [10].

Figura 4: Detalle del panel de resultados



3. RESULTADOS OBTENIDOS/ESPERADOS

3.1 Contar con laboratorios virtuales disponibles a través de Internet

Se cuenta actualmente con varios laboratorios virtuales destinados a educación. Los mismos se complementan con laboratorios trasladables al curso. Se espera contar con un mayor número de estas herramientas.

3.2 Obtener estadísticas de uso y retroalimentación de los estudiantes.

En particular, el laboratorio virtual de corriente alterna presentado en este trabajo será empleado por estudiantes de la asignatura Física 2A de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de Cuyo, que se cursa en el segundo semestre del año, por lo que se espera contar con estadísticas de uso al final de 2026.

3.3 Realizar aportes en las áreas de los laboratorios virtuales y enseñanza

El proyecto tiene como objetivo el desarrollo e implementación de laboratorios virtuales en el ámbito educativo, cuyos resultados y experiencias serán difundidos en congresos y revistas científicas. Asimismo, los laboratorios estarán disponibles de forma abierta y gratuita para su uso en cualquier institución educativa.

Actualmente puede accederse a través del enlace de abajo. Sin embargo, la versión disponible es una versión en etapa de desarrollo. La versión final estará disponible a través de un servicio de web hosting. https://pablodanielgodoy.github.io/ca/

4. FORMACIÓN DE RECURSOS HUMANOS

Una de las integrantes del equipo de trabajo se encuentra realizando su doctorado. Se espera que dicha tesis sea defendida durante el año 2026.

Además, una de las integrantes de este proyecto es estudiante de grado de la carrera Licenciatura en Ciencias de la Computación de la Universidad Nacional de Cuyo. Su proyecto final de carrera, cuyo título tentativo es "Alfabetizar jugando - juego educativo web para Contribuir en la alfabetización de niños" será desarrollada en el marco del proyecto.

AGRADECIMIENTOS

Los autores agradecen el financiamiento recibido de la Secretaría de Investigación, Internacionales y Posgrado (SIIP), Universidad Nacional de Cuyo, a través de los proyectos "80020240100290UN - Laboratorios virtuales destinados a educación " y "80020240100288UN - La tecnología de Blockchain como un sistema descentralizado y distribuido en aplicaciones de IoT".

5. BIBLIOGRAFIA

[1] Godoy, Pablo; Marianetti, Osvaldo; García Garino, Carlos; "Experiences With Computer Architecture Remote Laboratories"; en Handbook of Research on Software Quality Innovation in Interactive Systems", Chapter 5; IGI Global, Hershey, Pennsylvania, USA, 2020.

[2] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "A WSN testbed for teaching purposes"; IEEE Latin America Transactions. 14(7),3351–3357 (2016).

[3] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "Laboratorio remoto para la formación de usuarios basado en el cloud"; Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación 18, 7–18 (2016).

[4] Godoy, Pablo; García Garino, Carlos; Cayssials, Ricardo; "A nomadic testbed for teaching computer architecture"; In: XVII Workshop Tecnología Informática Aplicada en Educación (WTIAE), XXIV, CACIC, 2018.

[5] Godoy Pablo Daniel, Marianetti Osvaldo Lucio, Fontana Daniel, Manganelli Silvina; "Aplicación de laboratorios virtuales en educación de ciencias exactas", en XXVII Workshop de Investigadores en Ciencias de la Computación (WICC) 2025, Mendoza.

[6] Darrah, M., Humbert, R., Finstein, J. et al. Are Virtual Labs as Effective as Hands-on Labs for Undergraduate Physics? A Comparative Study at Two Major Universities. J Sci Educ Technol 23, 803–814 (2014).

[7] Universidad de Colorado, PhET Interactive Simulations website. Accedido en febrero de 2022. https://phet.colorado.edu/es/

[8] Marco Schweighauser. Simulador de ensamblador simple de 8 bits website. Accedido en 2026. https://schweigi.github.io/assembler-simulator/

[9] Website del profesor profesor Vladimír Vaščák. Accedido en febrero de 2026. https://www.vascak.cz/physicsanimations.php

[10] Manganelli, Silvina. Aportes de las herramientas digitales a STEM, durante la investigación científica, para fomentar el desarrollo de competencias científicas, tecnológicas y digitales. Congreso Argentino de Ciencias de la Computación (CACIC). Tecnología Informática Aplicada en Educación (WTIAE). 2021.

Microservicios en Aplicaciones Móviles: Arquitectura, Desafíos y Buenas Prácticas

Adriana E. Martín[1] , Susana Chavez[1], Sergio Flores[1], Joaquin Albornoz[1,2] , Nelson Rodriguez,

[1]Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
[2] Alumno Avanzado Licenciatura en Ciencias de la Computación - F.C.E.F. y N. - U.N.S.J.
Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

arianamartinsj@gmail.com, schavez@iinfo.unsj.edu.ar, sergiorflores@gmail.com, joaquinalbornoz1207@gmail.com, arianamartinsj@gmail.com

Resumen.

Las aplicaciones móviles actuales exigen backends escalables, resilientes y de despliegue frecuente para millones de usuarios con conectividad variable. La arquitectura de microservicios descompone el backend en servicios autónomos con bases de datos independientes. Este trabajo analiza su integración con clientes móviles mediante el patrón Backend for Frontend (BFF) y API Gateway. Se comparan con arquitecturas monolíticas y se detallan ventajas, desafíos y buenas prácticas. Basado en revisión sistemática y casos reales.

Palabras clave: *Distributed Applications. Microservices Architecture. Mobile Computing. Best Practices*.

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es la línea marco del proyecto de investigación Aplicaciones Distribuidas en Arquitectura de Microservicios. Desafíos técnicos y mejores prácticas, cuya propuesta se presentó en convocatoria para el período 2026-2027 y se encuentra en evaluación. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil, distribuida y de alta performance, desde hace más de 25 años, al que se han incorporado estudiantes de informática. Similarmente a proyectos anteriores, se continúa el trabajo con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

El mercado móvil supera los 7 mil millones de usuarios en 2026. Las apps requieren alta disponibilidad 24/7, escalado elástico ante picos de tráfico y optimización para redes 4G/5G con latencia variable y batería limitada. Las arquitecturas monolíticas tradicionales generan cuellos de botella en mantenimiento, escalabilidad y tiempo de despliegue cuando múltiples equipos trabajan en paralelo. (Ver figura 1)

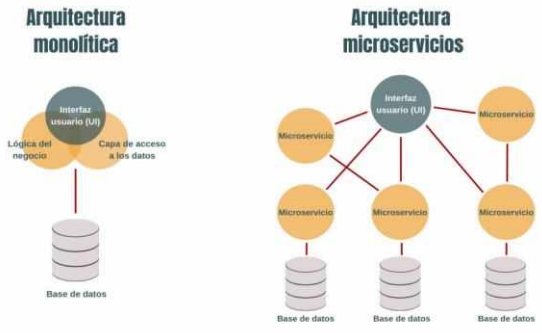


Figura 1

Microservicios es una arquitectura donde una aplicación se divide en pequeños servicios independientes, cada uno responsable de una funcionalidad específica (ej.: autenticación,

pagos, catálogo de productos, notificaciones). [1].Cada microservicio tiene su propia base de datos. Se desarrolla, despliega y escala de forma autónoma, Se comunica con otros microservicios mediante APIs ligeras.

A diferencia de la arquitectura monolítica (todo el código en un solo bloque), los microservicios permiten mayor flexibilidad y agilidad. (ver figura 2)

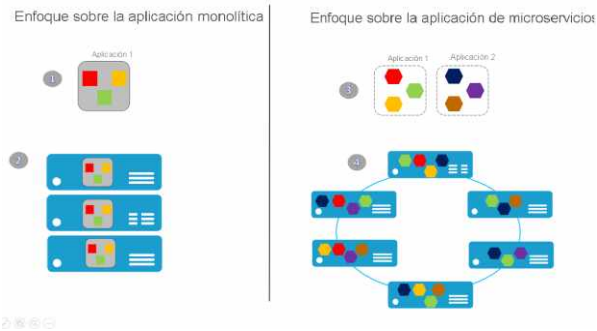


Figura 2

La arquitectura de microservicios (popular desde 2014) divide la aplicación en servicios independientes que se comunican mediante APIs ligeras (REST, GraphQL, gRPC). Cada servicio tiene su propia base de datos siguiendo una estrategia de persistencia políglota y ciclo de vida autónomo. (figura 3)

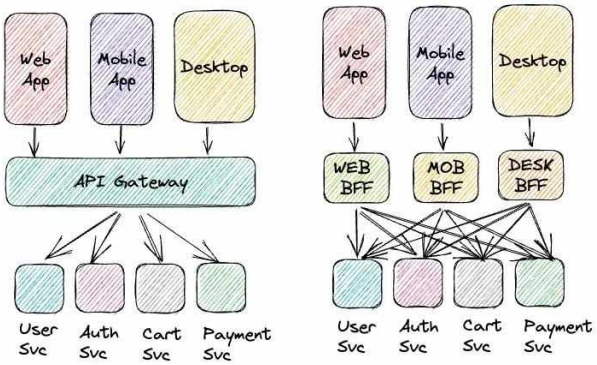


Figura 3

En el contexto móvil, la app actúa como cliente ligero (Retrofit en Android, URLSession en iOS). Sin embargo, llamar directamente a decenas de microservicios genera latencia inaceptable

La app móvil no usa microservicios internamente. En su lugar:

- Utiliza librerías como Retrofit (Android), Alamofire o URLSession (iOS) para llamar a las APIs del backend.
- El backend se construye con microservicios (Spring Boot, Node.js, .NET, Go, etc.) y se despliega en contenedores (Docker + Kubernetes).

En lugar de que la app móvil llame directamente a decenas de microservicios (lo que genera latencia y complejidad), se crea un BFF dedicado para móviles. El cual

- Agrega datos de varios microservicios.
- Optimiza respuestas (menos datos, formato ligero para móviles).
- Maneja lógica específica (paginación, caching, offline-first).

Por lo tanto el patrón Backend for Frontend (BFF), propuesto por Sam Newman, resuelve esto creando un backend específico por cliente (móvil/web) que agrega, filtra y optimiza datos.[2][3] (Figura 4 y 5) Ejemplos reales: Netflix (miles de microservicios), Uber (Kubernetes + Kafka) y la referencia eShopOnContainers de Microsoft. [4][5]

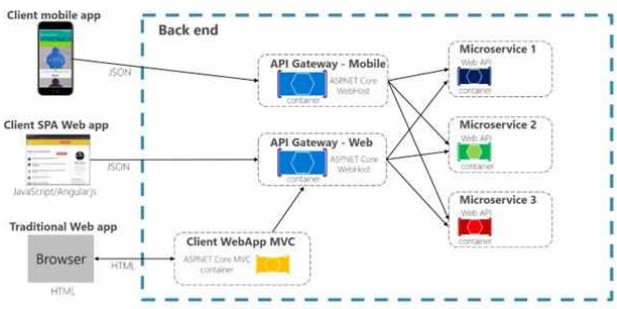


Figura 4

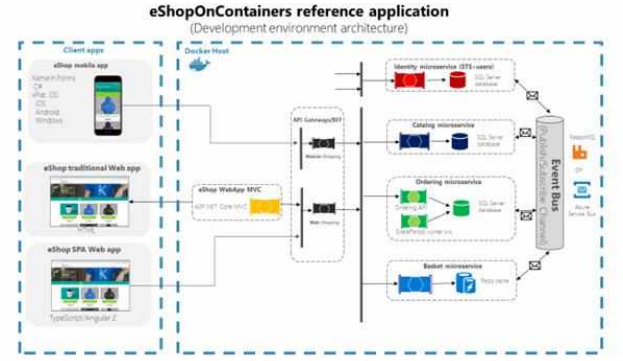


Figura 5

.Objetivos

Los objetivos del grupo de investigación en esta línea de conocimiento son los siguientes:

Objetivo general:

Analizar y proponer un conjunto de buenas prácticas para el diseño, implementación y mantenimiento de aplicaciones distribuidas basadas en arquitectura de microservicios, considerando los principales desafíos técnicos actuales.

Objetivos específicos:

Objetivos teóricos

- Diferenciar arquitecturas monolíticas vs. Microservicios.
- Analizar el rol del BFF/API Gateway en reducción de latencia y optimización de payloads.

Objetivos experimentales

- Identificar ventajas / desventajas específicas para entornos móviles.
- Proponer buenas prácticas y guía práctica para diferentes proyectos.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los investigadores que figuran en este trabajo de la Universidad Nacional de San Juan, Universidad de San Luis, Universidad Nacional de Salta y Universidad de Champagnat de Mendoza y ocho alumnos de grado, de la Universidad Nacional de San Juan..

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis de maestría en áreas afines y tres tesinas de grado en el área de Serverless computing, Virtualización y Computación distribuida.

Por otro lado se espera aumentar el número de publicaciones, y además también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.

Referencias

[1] Iranzo Jiménez, V. A. (2018). Desarrollo de software basado en microservicios. Universitat Politècnica de València.

[2] Newman, S. (2015). Building Microservices (2ª ed.). O'Reilly.

[3] Newman, S. (2015). Backends For Frontends. https://samnewman.io/patterns/architectural/bff/

[4] Microsoft. (2025). eShopOnContainers. https://github.com/dotnet-architecture/eShopOnContainers

[5] Richardson, C. (2018). Microservices Patterns. Manning.

Tecnologías de IoT y Visión Artificial aplicadas a sistemas inteligentes

Jorge Osio[1,2], Juan Salvatore[1], Mauro Salina[1] , Marcelo Cappelletti[1,2] , Diego Montezanti[1,3], Juan Piñeiro[1], Juan Carlos Czerwien[1], Braian Pezet[1], Lucia Osés[1], Martín Paez[1],Emmanuel Julio[1], Matías Busum Fradera[1], Facundo Chazarreta[1], Lucas Olivera[1,4],Patricio Gross[1], Diego Encinas[1,3], Martín Morales[1,5]

[1] Programa TICAPPS, Univ. Nac. Arturo Jauretche, Florencio Varela (1888), Argentina.
[2] Grupo de Control Aplicado, Instituto LEICI (UNLP-CONICET), La Plata (1900), Argentina.
[3] Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC.
[4] Becario de la Comisión de Investigaciones Científicas, Pcia. Bs. As. (CICPBA), La Plata (1900), Argentina.
[5] Centro UTN CODAPLI-FRLP, La Plata (1900), Argentina.
{josio, mcappelletti, dmontezanti, jsalvatore, mdsalina, dencinas, martin.morales}@unaj.edu.ar

Resumen

El presente proyecto se enfoca en líneas de investigación que involucran nuevas tecnologías como Internet de las Cosas (IoT), visión artificial y robótica. La investigación se centra especialmente en las técnicas de Smart IoT, las técnicas de aprendizaje automático aplicadas sobre dispositivos de bajas prestaciones de procesamiento y el aprendizaje profundo para el reconocimiento de objetos en aplicaciones de robótica. Se espera que los resultados obtenidos permitan proporcionar soluciones a problemáticas socio-productivas presentes en la región, como son las medioambientales, la innovación en el sector productivo mediante tecnologías de la industria 4.0.

Entre los temas de investigación que se desarrollan en esta línea, se incluyen el diseño e implementación de técnicas de visión por computadora, con los objetivos de agregar funcionalidades a dispositivos robóticos (como la clasificación de residuos) y el de análisis para descontaminación de suelos en aplicaciones agrícolas, entre otros. Además, mediante el agregado de control y supervisión remota (utilizando herramientas de IoT), se busca proporcionar a estos dispositivos la autonomía suficiente para la realización de tareas en entornos abiertos.

En el área de Smart IoT, se ha desarrollado una plataforma IoT que incluye los servicios esenciales, como un *broker* MQTT para la gestión de las comunicaciones, una base de datos, un servicio de visualización de datos y una plataforma de programación orientada a flujos (NodeRed) [1]. Esta plataforma brinda las herramientas necesarias para la obtención y almacenamiento de información, y se encuentra en desarrollo el agregado de un servicio de análisis de datos que utiliza aprendizaje automático para predicción y toma de decisiones.

Palabras clave: *Aprendizaje automático, procesamiento de imágenes, Smart IoT, Visión artificial y robótica.*

Contexto

Las líneas de I/D presentadas en este trabajo son desarrolladas por el grupo IoT&IA y están incluidas dentro del Programa TICAPPS (TIC con aplicaciones de interés social) de la Universidad Nacional Arturo Jauretche (UNAJ), Resolución N° 064/17, bajo la dirección del Dr. Ing. Martín Morales.

1. Introducción

Internet de las Cosas (IoT) ha transformado radicalmente la interacción entre el mundo físico y el digital, mediante redes de objetos interconectados capaces de recolectar y transmitir datos en tiempo real. Esta tecnología se basa en la ubicuidad de sensores y actuadores que, mediante diversos protocolos de comunicación, permiten el monitoreo de entornos complejos [2]. La esencia de IoT reside en su capacidad para generar visibilidad sobre procesos que antes sobre se solo podían monitorear de forma local, sentando las bases para la toma de decisiones basada en el análisis de datos en tiempo real.

Por otro lado, la **visión artificial** se ha consolidado como una de las disciplinas más disruptivas de la informática, permitiendo que las máquinas interpreten y comprendan el entorno en base a las características físicas. Mediante el procesamiento de imágenes estas tecnologías pueden identificar patrones, detectar objetos y segmentar escenas con una precisión que supera la capacidad humana. Al dotar a los sistemas de un "sentido de la vista", la visión artificial expande las fronteras de la automatización, permitiendo una interacción mucho más rica y contextual con el entorno físico.

La fusión entre la **Inteligencia Artificial (IA) e IoT (AIoT)** representa un salto evolutivo hacia sistemas verdaderamente autónomos. Mientras que IoT proporciona los datos y la conectividad, la IA aporta la capacidad analítica para transformar esa información en conocimiento accionable de forma inmediata. Esta integración permite optimizar el consumo energético, predecir fallos antes de que ocurran (mantenimiento predictivo) y personalizar servicios a una escala masiva, eliminando la necesidad de intervención humana constante y aumentando drásticamente la eficiencia operativa de los ecosistemas inteligentes [4-5].

Este avance es posible gracias a la sofisticación de los sistemas embebidos modernos, cuya creciente potencia de cómputo y eficiencia energética permiten ejecutar algoritmos complejos directamente en el hardware local. Esta evolución ha facilitado el despliegue de IA en las capas de Edge y Fog Computing, procesando la información en la periferia de la red en lugar de depender exclusivamente de la nube [3]. Al desplazar la inteligencia hacia los nodos locales, se reduce la latencia, se optimiza el ancho de banda y se mejora la privacidad de los datos, permitiendo respuestas en milisegundos que son críticas para aplicaciones en tiempo real.

En el ámbito de la percepción avanzada, las técnicas de visión artificial basadas en Deep Learning [6], particularmente las Redes Neuronales Convolucionales (CNN) y los Transformadores de Visión, han revolucionado la robótica y la visión 3D [8-9]. Estas herramientas permiten a los robots navegar de forma autónoma, realizar tareas de picking de alta precisión y reconstruir entornos tridimensionales con fidelidad milimétrica. La integración de mapas de profundidad y nubes de puntos con modelos de aprendizaje profundo facilita que los sistemas no solo vean, sino que comprendan la volumetría y la disposición espacial de los objetos, algo esencial para la colaboración humano-robot.

Finalmente, el surgimiento de Internet de las Cosas Inteligentes (AIoT) demanda una infraestructura de soporte robusta, donde la disposición de un servidor configurado a medida se vuelve fundamental [1]. Dado que estas aplicaciones manejan grandes volúmenes de datos heterogéneos y requieren una orquestación precisa entre el Edge y la Nube, un servidor optimizado garantiza la escalabilidad, la seguridad y la disponibilidad del sistema. Una configuración personalizada permite ajustar los recursos de procesamiento, almacenamiento y gestión de bases de datos a las necesidades específicas del despliegue, asegurando que la inteligencia distribuida funcione de manera eficiente.

Aplicaciones y desarrollos en curso

Entre las aplicaciones que se encuentran en desarrollo, combinando técnicas de IoT, procesamiento de datos e imágenes y *Machine Learning* se destacan:

1) Detección automática de residuos reciclables.
2) Detección de gestos para el control de dispositivos por parte de personas con problemas motrices.
3) Selección de tipo de cultivo en base a la época del año y las características del suelo.
4) Medición de la calidad del aire, monitoreo de la calidad del agua y obtención de magnitudes climatológicas
5) Análisis automático de parámetros del suelo mediante un robot móvil.
6) Visión artificial aplicada a robótica [10].

El reciclado de desechos es de suma importancia para el cuidado del medioambiente, debido a que posibilita la reutilización de elementos u objetos que de otro modo serían desechados y contribuirían al incremento de la basura y al daño ambiental permanente. Sin embargo, el problema radica especialmente en el esfuerzo requerido para clasificar y separar los residuos inorgánicos. Debido a esto, en la línea de investigación se está desarrollando un sistema basado en redes neuronales artificiales que permite detectar y clasificar los objetos reciclables más comunes para su posterior separación (típicamente se separa plástico, vidrio, metal y papel [8-9]). Además, se está implementando un brazo robótico que permite separar los objetos previamente identificados mediante visión artificial [9].
Respecto a la visión por computadora, se está llevando adelante la implementación de un sistema robótico que sea capaz de detectar obstáculos e identificar señales, mediante procesamiento de imágenes, para dotarlo con capacidades autónomas y utilizarlo en aplicaciones de calidad de aire y desinfección de suelo [10].
Por último, se está desarrollando una aplicación de SIoT, que obtiene imágenes provenientes de un sistema embebido, las procesa mediante un algoritmo de detección de gestos faciales y permite accionar diferentes dispositivos por cada gesto detectado, generando así comandos y acciones sin la necesidad de emisión de ningún sonido ni acción física. Esto podría resultar de gran utilidad para personas con problemas motrices y dificultades para el habla [11].

2. Líneas de Investigación y Desarrollo

El grupo de investigación que se ha constituido en la UNAJ es multidisciplinario, y sus miembros cuentan con experiencia en sistemas de cómputo de altas prestaciones, arquitecturas distribuidas, resiliencia, sistemas embebidos, IoT, aprendizaje automático y robótica.
Una de las líneas de I/D que se presentan en este trabajo está basada en el estudio y desarrollo de herramientas alternativas para el procesamiento de imágenes a partir de algoritmos de aprendizaje automático y visión por computadora. La línea referente a sistemas AIoT pretende, en primera instancia, evaluar las herramientas tecnológicas disponibles para determinar las posibles aplicaciones, pero también sus limitaciones.
En cuanto a la línea vinculada con la robótica, se encuentra en desarrollo la búsqueda de aplicaciones innovadoras, como la separación automática de objetos reciclables y la capacidad de detección de gestos y obstáculos, mediante visión artificial, para lograr autonomía en los robots móviles.
Actualmente se encuentra en proceso de evolución el nuevo proyecto denominado: ”Tecnologías de IoT y Visión Artificial aplicadas a sistemas inteligentes“.

Temas de Estudio e Investigación

Los datos recopilados a través de robots y nodos sensores en diferentes aplicaciones pueden resultar en grandes cantidades de información. Si bien esto requiere de involucrar técnicas de *Big Data* para su tratamiento, también permite obtener una

mayor precisión en las predicciones de los sistemas de *Machine Learning*, ya que los modelos se alimentan con mayor volumen de datos útiles para su aprendizaje.
Los aportes originales no sólo tienen que ver con el manejo de la información, sino también con las formas innovadoras en las que se pueden combinar y aplicar el aprendizaje automático, la visión por computadora y la robótica. El ejemplo más claro consiste en la posibilidad de disponer de un sistema de detección de objetos reciclables para la separación automática de residuos.

Resultados alcanzados

Se optimizaron los algoritmos de detección de objetos reciclables, logrando una transición de detecciones únicas a un sistema de detección múltiple simultánea. Asimismo, se implementó el procesamiento en video en tiempo real, lo que derivó en un incremento sustancial en la precisión y tasa de aciertos del sistema de clasificación automatizada.
Mediante la integración de la librería MediaPipe, se desarrolló un módulo de detección de gestos faciales para sistemas de monitoreo remoto. Las pruebas demostraron una alta eficiencia y robustez del algoritmo, logrando resultados satisfactorios incluso bajo las limitaciones de resolución de hardware de la cámara ESP32-CAM. Las pruebas realizadas permitieron verificar la precisión en la detección de gestos como guiños, apertura de boca y muecas, con una latencia aceptable en para una aplicación de tiempo real. Sin embargo, se identificaron limitaciones en escenarios con iluminación variable, uso de accesorios (como por ejemplo lentes) o ángulos pronunciados (>30°), donde la precisión disminuye, debido a errores en la generación de la malla facial.
En el ámbito del monitoreo de parámetros hídricos en arroyos, se diseñó e implementó una red de sondas de largo alcance. La arquitectura utiliza tecnología LoRa bajo un protocolo propietario configurado en topología **Daisy Chain**. Esta configuración permite extender significativamente el área de cobertura, superando las limitaciones de distancia convencionales en entornos abiertos.

Se alcanzaron hitos clave en la integración de hardware y software en robótica:

- **Integración de Visión:** Se vinculó con éxito el sistema de detección de objetos (procesado en una Raspberry Pi) con el control cinemático de un **brazo robótico**.
- **Movilidad Autónoma:** Se desarrolló un prototipo de **robot móvil con conectividad IoT**, capaz de desplazarse de forma autónoma para realizar el análisis de características físico-químicas del suelo.

Objetivos esperados

- Promover la generación de conocimiento y el desarrollo de instrumentos que puedan aplicarse en la formación de recursos humanos para investigación.
- Fortalecer la actividad de investigación y vinculación que contribuya a remediar las problemáticas del territorio de influencia de la UNAJ.
- Diseñar y optimizar diferentes arquitecturas y redes para IoT, en función de la aplicación.
- Evaluar modelos en entornos restrictivos: Analizar el desempeño de los modelos de visión computacional diseñados específicamente para sistemas embebidos, priorizando arquitecturas de bajo consumo energético y capacidad de procesamiento limitada.
- Implementación de estrategias de mapeo adaptativo: Desarrollar diferentes técnicas de generación de mapas (ocupación por rejilla, mapas topológicos y nubes de puntos) para un sistema robótico móvil. Esto permitirá evaluar la eficacia del robot al navegar de forma autónoma tanto en interiores estructurados como en terrenos irregulares, garantizando una localización robusta mediante la fusión de datos de sensores inerciales y odometría.

Formación de Recursos Humanos

Uno de los principales objetivos del Programa TICAPPS, dentro de la temática de las líneas de I/D presentadas en este trabajo, es la formación de recursos humanos, tanto de docentes investigadores como de estudiantes.
Dentro de la temática de la línea de I+D, todos los miembros del proyecto participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.
En este proyecto existe cooperación a nivel Nacional. Hay tres Doctores, dos Magisters, un Especialista, cuatro Doctorandos y un integrante realizando su Maestría en temas relacionados.
Además, hay 4 becarios y estudiantes avanzados realizando las PPS de final de carrera en la temática. Actualmente, se encuentran en curso tres Tesis de Doctorado y tres becas de Estímulo a las Vocaciones Científicas del Consejo Interuniversitario Nacional (Becas EVC – CIN) y 2 becas UNAJ de investigación relacionadas directamente con las líneas de I/D presentadas.

3. Bibliografía

[1] F. Chazarreta, M. Busum Fradera, J. R. Osio, J. E. Salvatore, Morales D. M. “Diseño de una plataforma de servicios IoT para aplicaciones de interés social”, IX CONAIISI, 2022.

[2]L. Fatmasari Rahman,” Choosing your IoT Programming Framework: Architectural Aspects”, 2016 IEEE 4th International Conference on Future Internet of Things and Cloud

[3] Medina, S., Montezanti, D., Gómez D’Orazio, L., Garay, F., De Giusti, A., & Naiouf, M. (2023, June). Distributed Architectures Based on Edge Computing, Fog Computing and End Devices: A Conceptual Review Incorporating Resilience Aspects. In Conference on Cloud Computing, Big Data & Emerging Topics (pp. 31-44). Cham: Springer Nature Switzerland

[4] J. Osio, J. Salvatore, M. Salina, M. Cappelletti, D. Montezanti, L. Olivera, M. Paez, F. Palacio, M. Busum Fradera, F. Chazarreta, “Tecnologías de Smart IoT para la solución de problemas en el medio socio productivo”, UNNOBA, Ciudad de Junín, WICC 2023.

[5] J. Osio, J. Salvatore, M. Salina, M. Cappelletti, D. Montezanti, L. Olivera, M. Paez, F. Palacio, M. Busum Fradera, F. Chazarreta, “Nuevas tecnologías de IoT y visión artificial para el diseño de sistemas con aplicaciones de interés social”, UNP, Puerto Madryn, WICC 2024.

[6] N. Buduma, “Fundamentals of Deep Learning”, Editorial O’reilly. 2017

[7] A. Gerón, “Hands-On Machine Learning with scikit Learn & TensorFlow”, Editorial O’reilly, 2017

[8] L. Osés, B. Pezet, M Salina, JR Osio, MA Cappelletti, M Morales. Técnicas de Deep Learning aplicadas a un sistema de clasificación de objetos para un recolector de residuos inteligente. 52 Jornadas Argentinas de Informática Congreso Argentino de Ciencias de la Computación (JAIIO). 2023

[9] L. Osés, B. Pezet, M. Salina, J. Osio, M. Cappelletti, M. Morales. Sistema de clasificación de objetos para un recolector de residuos inteligente mediante visión artificial y aprendizaje profundo. 11° Congreso Nacional de Ingenierías en Informática y sistemas de Información (CoNaIISI 2023)

[10] M. Jamshidi y P. J. Eicker, Robotics and remote system for hazardous environments, Prentice Hall, New Jersey, Estados Unidos (1993)

[11] M. Busum, F. Chazarreta, J. Salvatore, J. Osio, M. Paez, M. Morales. Desarrollo de técnicas de Smart IoT aplicadas a la detección

de gestos para el accionamiento de dispositivos. 11° Congreso Nacional de Ingenierías en Informática y sistemas de Información (CoNaIISI 2023)

UM-CLOUD: INFRAESTRUCTURA DE IA ON-PREMISE PARA EDUCACIÓN SUPERIOR

Diego Navarro[1], Juan José Ciarlante[1], Ignacio Bosh[1], Mariela Asensio[1]
[1]UM-Cloud, Facultad de Ingeniería, Universidad de Mendoza
diego.navarro@um.edu.ar

RESUMEN

Las universidades enfrentan una brecha significativa en recursos de infraestructura de IA comparadas con la industria. Este trabajo presenta UM-Cloud, infraestructura on-premise de IA de la Universidad de Mendoza que implementa arquitectura híbrida sobre Kubernetes utilizando KubeAI con motores duales Ollama y vLLM. La solución sirve a 6 modelos LLM en producción sobre 5 nodos GPU heterogéneos (RTX 3090, RTX 5060 Ti, RTX 2080 Super), con contexto extendido de 32K tokens que habilita casos de uso educativos avanzados. El enfoque prioriza soberanía tecnológica e independencia operativa. Los resultados validan la arquitectura como alternativa viable para instituciones con recursos limitados, posicionando a UM-Cloud como primera universidad privada latinoamericana con infraestructura Kubernetes de IA documentada.

Palabras clave: Infraestructura IA, Kubernetes, LLM, Soberanía Tecnológica, Cloud Computing.

CONTEXTO

UM-Cloud es una línea de investigación consolidada de la Facultad de Ingeniería de la Universidad de Mendoza iniciada en 2015, enfocada en arquitecturas de servicios cloud. El proyecto responde al principio institucional de entender la tecnología construyéndola para luego enseñarla, priorizando independencia tecnológica y control sobre herramientas educativas

La evolución histórica demuestra adaptación continua: 2015 (OpenStack), 2019 (refactorización a Kubernetes + Rook + OpenStack), 2020-2021 (Laboratorio Remoto COVID-19), 2022 (ampliación x4), 2023 (kros'k + cátedra Cloud Computing), 2024 (extensión a servicios IA con KubeAI). El equipo incluye al Mg. Ing. Diego Navarro (Director de la Carrera de Ingeniería en Informática) y Mg. Ing. Juan José Ciarlante (arquitecto KROS, ex-Google SRE, CNCF CKA, docente de la Facultad de Ingeniería). Infraestructura: https://cloud.um.edu.ar, repositorio: https://github.com/umcloud.
La extensión 2024 a servicios de IA mantiene principios de soberanía e independencia que han caracterizado el proyecto desde sus inicios, objeto central de este trabajo.

1. INTRODUCCIÓN

1.1 Panorama Global y Regional

La infraestructura de IA universitaria evidencia una brecha con la industria. Universidades tier-1 como Stanford (Marlowe: 248 H100, US$30M) [1], Harvard (Kempner Institute: 144 A100 + 384 H100, US$500M total instituto) [2] y Princeton (336 H100, US$9.6M) [3] representan el extremo superior. Según Kantrowitz (2024), las universidades están "woefully under-resourced" comparadas con empresas que disponen de 10K-100K GPUs [4].
En Latinoamérica, la infraestructura GPU se concentra en centros HPC nacionales con stacks tradicionales. Argentina opera

Clementina XXI (296 Intel Ponte Vecchio, 15.3 PFLOPS) [5] y la UNC-CCAD (44 A30, cluster Mendieta) [6]. Brasil lidera con Santos Dumont (248 H100 + 144 GH200, 18.85 PFLOPS) [7]. Chile tiene NLHPC-CMM, México LANCAD, Colombia Guane-1. Todos usan SLURM sobre GPUs datacenter. Alternativas incluyen National Research Platform (420+ nodos, 75+ ubicaciones, 5M GPU-hours/año, solo EE.UU.) [8][9] y proveedores cloud (costos recurrentes, latencia variable, datos externos) [10].

1.2 Propuesta UM-Cloud

Este trabajo presenta un enfoque pragmático: maximizar hardware disponible mediante arquitectura cloud-native. La propuesta implementa stack de producción validado (Kubernetes + KubeAI + Ollama/vLLM) sirviendo modelos LLM sobre GPUs heterogéneas, manteniendo soberanía de datos y operaciones. UM-Cloud se posiciona como primera universidad privada latinoamericana con infraestructura Kubernetes de IA, diferenciándose en cuatro dimensiones: (1) institución privada (vs. centros públicos/nacionales regionales); (2) orquestación Kubernetes (vs. SLURM regional); (3) GPUs consumer-grade (vs. datacenter); (4) orientación educativa directa.

Objetivos: (1) Proveer acceso LLM de producción a comunidad académica; (2) Mantener control total sobre infraestructura y datos; (3) Validar arquitectura replicable para instituciones similares; (4) Documentar aprendizajes como referencia regional.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1 Arquitectura Técnica

Stack completo sobre Kubernetes: KubeAI Operator (orquestación) [11], motores duales Ollama (GGUF optimizado) [12] y vLLM (AWQ 4-bit) [13], Open WebUI (OAuth @um.edu.ar, @alumnos.um.edu.ar), CephFS (200Gi persistente). La arquitectura dual-engine permite flexibilidad operativa: Ollama para modelos en producción con contexto extendido (32K tokens en gpt-oss-20b, habilitando casos educativos avanzados), vLLM para modelos AWQ de alto rendimiento.

Infraestructura: 5 nodos GPU (AMD Ryzen 9 5950X, 128GB RAM): 2x RTX 3090 24GB (Ampere) modelos grandes, 2x RTX 5060 Ti 16GB (Blackwell) medianos + HA, 1x RTX 2080S 8GB (Turing) ligeros. Esta heterogeneidad optimiza distribución por VRAM.

Integración GPU-Kubernetes mediante NVIDIA GPU Operator v25.3.4 [15], configurado con time-slicing para compartición de recursos: RTX 3090 (3 réplicas), RTX 5060 Ti (2 réplicas), RTX 2080 Super (2 réplicas). El device plugin habilita scheduling granular mediante etiquetado por modelo de GPU, permitiendo asignación específica según requerimientos de VRAM y compute.

2.2 Modelos Desplegados

Criterios: VRAM footprint, performance, casos de uso complementarios, licencias open-source. Producción actual (Ollama): 6 modelos con contexto 32K en gpt-oss-safe (clave uso agentes desarrollo caso Opencode). Capacidad adicional (vLLM): 4 modelos AWQ preparados (glm-4.6V-Flash, glm-4.7-Flash, qwen3-30b, qwen3-coder-30b). gpt-oss-20b: 2 réplicas para alta disponibilidad, balanceo automático. Contexto 32K tokens permite análisis de documentos extensos, código completo, casos educativos avanzados.

2.2.1 Desglose Técnico del Contexto Extendido

El despliegue de gpt-oss-20b con contexto de 32K tokens sobre RTX 5060 Ti (16GB VRAM) es posible mediante cuantización dual: modelo en GGUF Q4_K_M y KV cache en Q4_0/Q8_0. El cálculo de memoria es:

VRAM Total ≈ Pesos modelo (11.5 GB) + KV cache cuantizado (3-4 GB) + Overhead sistema (0.5 GB) = 15.5 GB / 16 GB

Modelo	GPU	VRAM	Caso de uso
gemma3-4b	RTX 2080S 8GB	~6GB	Chat general ligero
gemma3n-e2b	RTX 2080S 8GB	~2GB	Embeddings/RAG
nemotron-3-nano-30b	RTX 3090 24GB	~20GB	Razonamiento MoE
glm47-flash	RTX 3090 24GB	~8GB	Multimodal
nomic-embed-text	CPU	~0.3GB	Embeddings
gpt-oss-20b	2x RTX 5060Ti 16GB	~14GB	Seguridad 32K ctx (HA)

Tabla 1. 6 modelos con contexto 32K.

La cuantización Q4_K_M reduce los pesos del modelo 20B de ~40GB (FP16) a ~11.5GB (4 bits por peso con compensación mixta). La cuantización del KV cache permite almacenar las representaciones de atención de 32K tokens en ~3-4GB versus ~12-14GB sin cuantización. El margen residual (0.5GB) garantiza estabilidad operativa. Esta configuración maximiza el contexto disponible para aplicaciones educativas que requieren análisis de documentos extensos o bases de código completas.

2.3 Análisis Comparativos

Comparación multi-dimensional con alternativas académicas:

Fortalezas diferenciadas: (1) Soberanía operativa - datos académicos on-campus; (2) Independencia absoluta; (3) Stack moderno producción (KubeAI adoptado en CNCF Sandbox) [11]; (4) Arquitectura documentable y transferible. En la tabla 2 se muestra el panel de resultados.

Aspecto	UM-Cloud	Tier-1 US	Cloud
Soberanía datos	Total	Total	Terceros
Latencia	LAN ~1ms	LAN ~1ms	50-200ms
Disponibilidad	24/7 garantizada	24/7	24/7
Control stack	Total	Total	Vendor lock-in
Escalabilidad	Moderada (5)	Alta (248-384)	Muy alta
Dependencias	Ninguna	Ninguna	Proveedor

Tabla 2. Fortalezas diferenciadas.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1 Infraestructura Validada

En producción estable: 6 modelos LLM (uptime 8+ días), OAuth institucional (@um.edu.ar, @alumnos.um.edu.ar), API OpenAI-compatible, storage CephFS persistente. Open WebUI provee acceso intuitivo sin requerimientos técnicos.

3.2 Capacidad Operativa (Estimaciones)

Basado en VRAM y características de modelos:

Escenario	Usuarios estimados
Uso ligero (1-2 req/min)	40-60
Uso medio (5-10 req/min)	25-35
Uso intenso (15+ req/min)	12-20

Tabla 3. Estimaciones teóricas pendientes de validación con uso real.

3.3 Validación Tecnológica

Stack alineado con industria 2025: KubeAI es framework recomendado para LLM serving [11]; Ollama validado para GGUF [12]; vLLM

es referencia para inferencia eficiente [13]; arquitectura híbrida maximiza hardware heterogéneo, validado por investigación sobre GPUs consumer en inferencia [14].

3.4 Trabajos futuros:

Obtener métricas de uso real, evaluación de expansión GPU, documentación de casos educativos, publicación como referencia regional. A mediano plazo: explorar iniciativas regionales compartidas tipo NRP para LATAM.

4. FORMACIÓN DE RECURSOS HUMANOS

Equipo: Mg. Ing. Diego Navarro (Director de carrera de Ingeniería en Informática, UM-Cloud desde 2015), Mg. Ing. Juan José Ciarlante (arquitecto UM-Cloud, KROS, ex-Google SRE, CNCF CKA). Capacidades desarrolladas: diseño infraestructura Kubernetes AI/ML, operación LLM producción, gestión GPU compartida, integración OAuth institucional. Bioingeniero Ignacio Bosh, Lic. Mariela Acencio (Profesores de la Carrera de Ingeniería en Informática, Sedes Mendoza y San Rafael)

Potencial formativo: Trabajos Integradores Finales sobre optimización serving LLM, integración sistemas académicos, benchmarking hardware heterogéneo. Infraestructura como laboratorio vivo para Ingeniería en Informática (cloud computing, arquitecturas distribuidas, IA).

AGRADECIMIENTOS

Los autores agradecen el financiamiento recibido de la Dirección de Investigación de la Universidad de Mendoza, correspondiente a la convocatoria PIUA 2025 de la Facultad de Ingeniería.

5. BIBLIOGRAFIA

[1] Stanford University IT. "Unlock Marlowe's Potential for Breakthrough Research". uit.stanford.edu, 2024.

[2] Kempner Institute. "Kempner Institute Adds ~400 H100 GPUs to Its Computing Cluster". kempnerinstitute.harvard.edu, 2023.

[3] Princeton AI. "Princeton invests in new 300 GPU cluster for academic AI research". ai.princeton.edu, 2024.

[4] Kantrowitz, A. "Universities Are Woefully Under-Resourced For AI Research". Big Technology/TheWrap, 2024.

[5] Argentina.gob.ar. "Argentina pone en funcionamiento a la supercomputadora Clementina XXI". 2023.

[6] UNC-CCAD. "Cluster Mendieta Fase 2". supercomputo.unc.edu.ar, 2024.

[7] Eviden. "Santos Dumont, LNCC supercomputer,

[8] National Research Platform. nrp.ai, 2025.

[9] Weitzel, D. et al. "The National Research Platform: Stretched, Multi-Tenant, Scientific Kubernetes Cluster". PEARC '25, ACM, DOI: 10.1145/3708035.3736060.

[10] Northflank. "7 Cheapest Cloud GPU Providers". northflank.com, 2025.

[11] KubeAI Official Documentation. kubeai.org, 2025.

[12] Lambda Labs. "Using KubeAI to deploy LLMs". docs.lambda.ai, 2025.

[13] Kwon, W. et al. "Efficient Memory Management for LLM Serving with PagedAttention". SOSP '23, DOI: 10.1145/3600006.3613165.

[14] Song, Y. et al. "PowerInfer: Fast Large Language Model Serving with a Consumer-grade GPU". arXiv:2312.12456, 2024.

[15] NVIDIA. "GPU Operator Documentation". docs.nvidia.com/datacenter/cloud-native/gpu-operator, 2025.

Visual CRANE. Implementación visual para el despliegue simplificado de aplicaciones contenerizadas en entornos locales.

Thomas Aguado [1][0009-0005-2867-9718], Patricia Bazán [3][0000-0001-6720-345X], Alejandra B. Lliteras [2,4][0000-0002-4148-1299]

UNLP, Facultad de Informática [1], LINTI [2] UNLP, Facultad de Informática, LIFIA [3] UNLP, Facultad de Informática, LINTI, [4] CICPBA

taguado@info.unlp.edu.ar, pbaz@info.unlp.edu.ar, alejandra.lliteras@lifia.info.unlp.edu.ar

RESUMEN

CRANE es una herramienta para el despliegue local de aplicaciones en contenedores, liviana, de propósito general y con capacidades de escalado automático.

El diseño de CRANE presentado en [Arcidiacono et al., 2022] contempla la implementación de la herramienta con un componente de back-end para la creación y despliegue de contenedores y un componente de front-end orientado a un usuario final que debe desempeñar el rol de DevOp dentro de un proyecto [Httermann M, 2012].

Este trabajo presenta la implementación del componente de front-end con el objetivo de mejorar la experiencia del usuario tanto en contextos educativos como profesionales vinculados al área de DevOps, permitiendo la creación y administración de infraestructuras contenerizadas de forma visual, ágil e intuitiva. El sistema ofrecerá funcionalidades como la manipulación gráfica de instancias y servicios, la visualización de métricas en tiempo real, la configuración de reglas de escalado y enrutamiento, la gestión de usuarios y roles, y la generación automatizada de archivos de configuración. En conjunto, buscará emular un entorno tipo PaaS de bajo costo, accesible y alineado con estándares actuales de despliegue de software.

CONTEXTO

CRANE es una herramienta diseñada para el despliegue local de aplicaciones en contenedores, enfocada en simplificar las pruebas de entornos distribuidos de forma local. CRANE ofrece una solución liviana y de propósito general con capacidades de ruteo, escalado y monitoreo automático. Orientada a estudiantes, docentes y desarrolladores que necesiten crear y desplegar servicios en contenedores simulando las características básicas que ofrece un entorno cloud de plataforma como servicio (PaaS) [Pavon, 2023].

CRANE ya dispone de una API funcional que permite gestionar servicios contenerizados de forma eficiente, su uso actual implica una barrera técnica considerable.

Para poder interactuar con la herramienta, es necesario no solo instalar y configurar el entorno, sino también contar con conocimientos específicos sobre APIs y herramientas como Postman.[1] Esta falta de una interfaz visual limita su accesibilidad, especialmente para estudiantes, docentes y profesionales que podrían beneficiarse de sus funcionalidades, pero que no poseen formación técnica avanzada.

Frente a esta situación, se propone el desarrollo de un front-end accesible, visualmente intuitivo y centrado en la experiencia del

[1] Postman Learning Center https://learning.postman.com/docs/introduction/overview/

usuario. Esta interfaz permitirá ampliar el alcance de CRANE y facilitar su adopción en entornos educativos y profesionales, simulando un entorno PaaS local con una experiencia de uso cercana a la de plataformas en la nube, pero sin los costos ni la complejidad que estas suelen implicar. Está pensada para lograr una mejor inclusión de profesionales interesados en el mundo devops sin sumergirse en un conjunto de herramientas y conceptos complejos.

1- INTRODUCCIÓN

La cultura DevOps comenzó a consolidarse en 2009 como una respuesta a los desafíos del modelo tradicional de desarrollo y operaciones en TI. Antes de su adopción, los equipos trabajaban de manera aislada y con responsabilidades claramente divididas: el equipo de desarrollo se encargaba de escribir el código y luego lo "arrojaba por encima del muro" al equipo de operaciones, que tenía la responsabilidad de desplegarlo en producción y garantizar la estabilidad del sistema.

Este enfoque generaba numerosos problemas. El equipo de operaciones no tenía injerencia sobre la calidad del código que recibía, lo que producía conflictos frecuentes, ciclos de entrega largos, fallos reiterados en producción y, en consecuencia, una baja confiabilidad en los sistemas.

En la actual era digital, donde la capacidad de adaptarse rápidamente al cambio es crítica, las organizaciones que no logran entregar software de forma ágil y confiable corren el riesgo de perder competitividad o incluso desaparecer.

En este contexto, DevOps surge como una práctica esencial, ya que permite entregar valor al cliente de manera más rápida, segura y eficiente.

Según estudios citados en [Kim, 2021], las organizaciones que adoptaron exitosamente prácticas DevOps obtuvieron mejoras significativas en su desempeño, entre ellas:

- 46 veces más frecuencia de despliegue.
- 440 veces mayor velocidad en la entrega de cambios.
- 5 veces menos probabilidad de fallos en producción.
- 96 veces menor tiempo medio de recuperación ante incidentes.

Estos resultados reflejan el impacto transformador de DevOps en la forma en que las organizaciones diseñan, construyen y mantienen sus sistemas de software

Por otra parte según [Fernandes, 2022], en los últimos años la industria del software ha adoptado diversas tecnologías DevOps relacionadas con prácticas como la integración continua y la entrega continua. Esta creciente demanda de profesionales con conocimientos en DevOps ha generado la necesidad de modificar profundamente los cursos y metodologías tradicionales de enseñanza en ingeniería de software.

Este estudio constató que la enseñanza de DevOps enfrenta múltiples obstáculos, incluyendo:

- Complejidad en la preparación de entornos y herramientas.
- Falta de estandarización en los entornos de aprendizaje, lo que dificulta la replicabilidad.
- Escasa documentación y recursos accesibles para aprender herramientas emergentes.
- Necesidad de abordajes prácticos, hands-on, más allá de la teoría tradicional

En este contexto y considerando estas particularidades de la práctica de DevOps, la implementación visual de CRANE incluye el diseño de una interfaz para el consumo de la API de CRANE identificando las funcionalidades principales y armando los mockups necesarios utilizando Figma[2] como herramienta de diseño de interfaces.

La implementación se lleva a cabo utilizando una interfaz web con tecnologías como React[3] con Next.js[4].

Entre las prestaciones previstas se encuentra armar un tablero para la visualización gráfica de métricas y alertas integradas con Prometheus[5] y AlertManager[6] e incorporar un panel de administración para la configuración de reglas de escalado y ruteo automático, basado en políticas definidas con Open Policy Agent (OPA)[7].

Además, se prevé integrar asistentes visuales para la generación y validación de archivos docker-compose con base en las especificaciones de Docker[8].

Como aporte a la propuesta original, se agregan a la funcionalidad la generación de prototipos (estructuras pre-configuradas para el deployment rápido), la construcción de un historial de las operaciones que hace el usuario para crear acciones rápidas asi agilizar sus operaciones al mediano/largo plazo y la definición de tutoriales, ayudas, tips para asistir a los nuevos usuarios a familiarizarse con el entorno.

2- LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación presentada apunta a la creación de aplicaciones seguras, con reglas claramente definidas y en donde toda la complejidad del despliegue sea transparente al desarrollador, proponiendo una serie de configuraciones básicas, las cuales permitirán ganar tiempo y seguridad a la hora de crear entornos de desarrollo.

Como base de este proyecto, se ha tomado lo propuesto en [Arcidiacono et al., 2022] que diseña CRANE y lo desarrollado en [Pavon et al., 2023] que implementa el componente back-end del diseño de CRANE.

La solución propuesta se organiza en capas que integran el front-end desarrollado en React/Next.js, un cliente API que abstrae las operaciones expuestas por CRANE y los servicios de infraestructura encargados del despliegue y monitoreo de contenedores. La Figura 1 ilustra la arquitectura general del sistema y el flujo de interacción entre sus componentes.

3- RESULTADOS ESPERADOS/OBTENIDOS

Se espera como resultado principal el desarrollo de un sistema web funcional, accesible y visualmente intuitivo que complemente a la herramienta CRANE, permitiendo su utilización de forma más amigable por parte de usuarios no especializados.

La interfaz estará orientada al enfoque low-code y al aprendizaje práctico, con el objetivo de reducir la curva de entrada al mundo DevOps y facilitar la comprensión de

[2] Figma, Figma Inc. https://www.figma.com/
[3] React Documentation. https://es.react.dev/reference/react
[4] Next.js Documentation,.https://nextjs.org/
[5] Prometheus Documentation. https://prometheus.io/docs/introduction/overview/
[6] Alertmanager Documentation. https://prometheus.io/docs/alerting/latest/alertmanager/
[7] Open Policy Agent Documentation. https://www.openpolicyagent.org/docs/latest/
[8] Docker Documentation. https://docs.docker.com/

conceptos asociados al despliegue de servicios en contenedores.

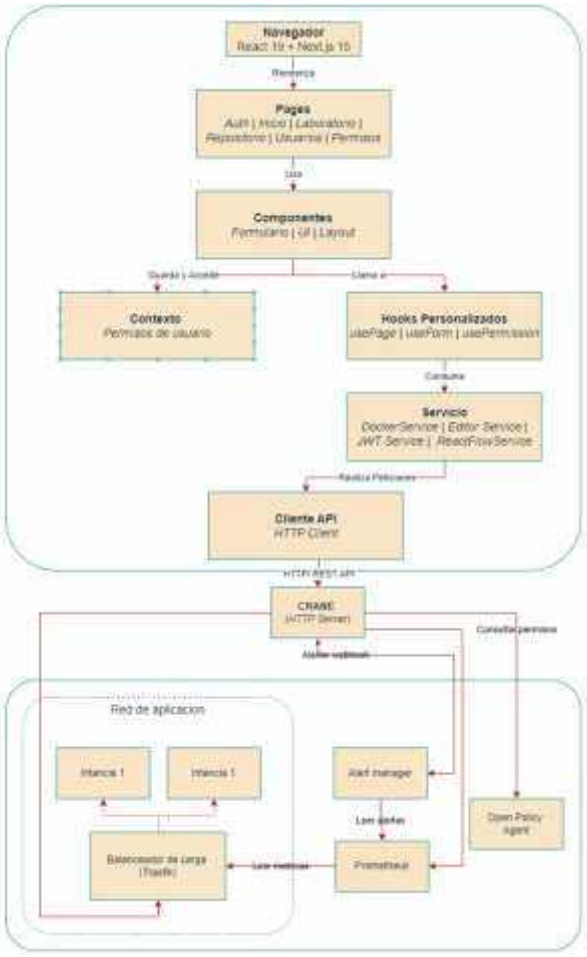


Figura 1. Arquitectura general del sistema CRANE.

Este front-end contribuirá a una mejora sustancial en la experiencia de usuario durante las tareas de despliegue, monitoreo y gestión de aplicaciones contenerizadas en entornos locales, disminuyendo la dependencia de herramientas externas como Postman o del uso directo de la línea de comandos.

Asimismo, se prevé un incremento en el uso de la herramienta en contextos académicos y de formación, favoreciendo su integración en prácticas docentes, actividades de laboratorio y proyectos educativos vinculados al desarrollo y operación de servicios en contenedores. En este sentido, se busca acortar la brecha existente entre la complejidad del ecosistema DevOps y la accesibilidad para quienes recién se inician en el mismo.

Por último, se contempla la validación de la solución mediante pruebas con usuarios finales, cuyas observaciones permitirán realizar ajustes en aspectos de diseño, usabilidad y funcionalidad, asegurando que la herramienta cumpla con los objetivos propuestos en términos de accesibilidad, eficiencia y facilidad de uso.

4- FORMACIÓN DE RECURSOS HUMANOS

La diversidad de líneas de investigación enunciadas, así como la gran variedad de aplicación de los resultados esperados, propone un foco de investigación claro y concreto para la formación de recursos humanos, de una manera directa en el desarrollo de conocimientos en Docker y el manejo de contenedores, así como, conceptos de seguridad, comunicación entre contenedores o simplemente ser capaz de analizar las métricas obtenidas por un determinado contenedor y tomar decisiones fundamentadas.

Debido a que este trabajo está cimentado sobre un proxy reverso, esto brinda muchos beneficios en la gestión y optimización de la infraestructura de la aplicación, tales como: balanceos de carga para optimizar al máximo los recursos disponibles, gestión de DNS, escalabilidad y también monitoreo.

También, al tener integrado un sistema de alertas, el recurso podrá recopilar métricas de rendimiento de los sistemas y configurar manualmente alertas personalizadas que, en caso de ser activadas, responderán dependiendo de unas políticas definidas.

Todo esto aporta conocimientos amplios sobre cómo se crean contenedores, cómo se comportan en cada momento y qué se puede

hacer para mejorar su rendimiento o qué hacer en caso de fallo.

De una manera indirecta, es posible apoyarse en CRANE para el desarrollo de aplicaciones, que facilita un despliegue rápido del entorno mediante una simple petición http, lo cual permitirá a los recursos humanos involucrados, formarse de una manera más rápida y efectiva, sin necesidad de pasar tiempo configurando un entorno que, en general, nada tiene que ver con el fin de su trabajo específico de desarrollador.

5- BIBLIOGRAFÍA

[Arcidiacono et al., 2022] Arcidiacono, J., Bazán, P., del Río, N., & Lliteras, A. B. (2022). Crane: A Local Deployment Tool for Containerized Applications. In Conference on Cloud Computing, Big Data & Emerging Topics (pp. 58-71). Springer, Cham.

[Arcidiacono et al., 2020] Arcidiacono, J., Lliteras, A. B., & Bazán, P. A. (2020). DEHIA, una plataforma para la generación y ejecución de actividades de recolección de datos con intervención humana aplicada en el Programa E-Basura. In VI Simposio Argentino de Ciencia de Datos y GRANdes DAtos (AGRANDA 2020)-JAIIO 49 (Modalidad virtual).

[Fernandes, 2022] Fernandes, M., Ferino, S., Fernandes, A., Kulesza, U., Aranha, E., & Treude, C. (2022, May). Devops education: An interview study of challenges and recommendations. In *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: Software Engineering Education and Training* (pp. 90-101).

[Httermann, 2012] Httermann, M. (2012). DevOps for developers. Apress: delivers a practical, thorough introduction to approaches, processes and tools to foster collaboration between software development and operations .

[Kim, 2021] Kim, G., Humble, J., Debois, P., Willis, J., & Forsgren, N. (2021). The DevOps handbook: How to create world-class agility, reliability, & security in technology organizations.

[Pavon, 2023] J. M. Silva Pavón, F. Bellino, P. Bazán, A. B. Lliteras y N. del Río, CRANE: Simplificando el despliegue de aplicaciones contenerizadas en entornos locales, Tesis de grado, Facultad de Informática, Universidad Nacional de La Plata, 2023. [En línea]. Disponible en: https://sedici.unlp.edu.ar/handle/10915/168369

Bases de Datos y Minería de Datos

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

Almacenamiento y recuperación para la gestión de datos masivos

V. Aragón, C. Ariza, L. Britos, N. Centorbi, M. G. Dorzán, A. Grosso, S. M. Jofré, F. Kasián, G. Leiva, S. Lososwski, V. Ludueña, N. Reyes, P. Roggero, P. Samat
LIDIC-LaBDa, Dpto. de Informática, Universidad Nacional de San Luis
{*vsaragon, crariza, mgdorzan, agrosso, fkasian, geleiva, vlud, nreyes, proggero*}@unsl.edu.ar
{*ldoorz, hnicolascentorbi, monicajofre, losowskisofia, samatpablo*}@gmail.com

K. Figueroa
Universidad Michoacana, México
karina@fismat.umich.mx

E. Chávez
CICESE, México
elchavez@cicese.mx

R. Paredes
Universidad Finis Terrae, Chile
raparede@uft.cl

C. Deco
Universidad Nacional de Rosario
deco@fceia.unr.edu.ar

Resumen

El crecimiento exponencial de datos multimedia en ecosistemas digitales ha generado la acumulación de grandes volúmenes de datos no estructurados en repositorios, cuya naturaleza no permite su gestión mediante el modelo relacional tradicional. Esta limitación ha derivado en el desarrollo de nuevos paradigmas de representación, indexación y recuperación de información. La presente investigación analiza la transición hacia herramientas de administración que no sólo aborden la complejidad de los datos no estructurados, sino que priorice la eficiencia computacional. Se enfatiza la necesidad de algoritmos optimizados que integren la jerarquía de memorias y arquitecturas de computación de alto desempeño (HPC), para lograr escalabilidad y rendimiento en el procesamiento de grandes volúmenes de datos.

Además, con el fin de capitalizar estos saberes/conocimientos en la oferta académica del departamento, se contempla el estudio, análisis y la selección de estrategias de transferencia de este conocimiento en la formación de grado y posgrado.

Palabras Claves: *bases de datos masivas, computación de alto desempeño, recuperación de información.*

1. Contexto

Las líneas "Almacenamiento y recuperación de datos" y "Diseño, desarrollo y transferencia de material didáctico para la enseñanza de tecnologías y metodologías de gestión de datos"pertenecen al Proyecto Consolidado de la UNSL: "Soluciones computacionales para la gestión de datos", que se desarrolla en el Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) y en el Laboratorio de Investigación y Desarrollo en Bases de Datos (LaBDa) de la UNSL. Se presentó en 2025, inició en 2026 y estará vigente hasta 2029.

Las líneas cuentan con 9 docentes-investigadores, tres estudiantes de maestría y tres de licenciatura y tienen como objetivo el desarrollo de herramientas eficientes para recuperar datos e información, para la administración y gestión de conjuntos de datos masivos y volcar el conocimiento adquirido en material didáctico para estudiantes de carreras de grado y posgrado. Además, se mantiene la cooperación con docentes-investigadores de otros centros: la Universidad Michoacana de San Nicolás de Hidalgo y el Centro de Investigación Científica y de Educación Superior de Ensenada - CICESE (México), la Universidad Finis Terrae (Chile) y la Universidad Nacional de Rosario.

Se están desarrollando índices diseñados para grandes volúmenes de datos no estructurados, considerando su aplicación en soluciones computacionales integrales en gestión de datos. Estos índices deben resolver eficientemente las consultas de interés sobre distintos tipos de datos no estructurados; tales como: texto, secuencias de ADN, huellas digitales, audio, vídeo, etc.; deben ser escalables y estar especialmente diseñados para memorias jerárquicas. Además, en algunos casos, se utilizan técnicas de computación de alto desempeño (HPC) para mejorar la eficiencia. Como este conocimiento puede volcarse a la enseñanza de la disciplina, se analizan estrategias y técnicas de la educación y se planea desarrollar material didáctico para la formación de estudiantes de grado y posgrado en la temática.

2. Introducción

La generación incesante de datos digitales, impulsada por la ubicuidad de dispositivos electrónicos y

la evolución de las Tecnologías de la Información y Comunicación (TIC), ha configurado un escenario caracterizado por dos pilares del Big Data: un volumen masivo y una heterogeneidad intrínseca. En este contexto, los modelos relacionales tradicionales resultan insuficientes, ya que las consultas no pueden limitarse a búsquedas por igualdad sobre datos estructurados sin incurrir en una pérdida de información relevante. Ante la necesidad de procesar activos multimedia y patrones complejos, surge el paradigma de la búsqueda basada en contenido o consultas por ejemplo (query by example), donde el dato mismo actúa como su propio descriptor.

Para abordar este desafío de manera formal e independiente del dominio, se adopta el modelo de espacios métricos. Un espacio métrico se define como un par (U, d), donde U representa el universo de objetos y $d : U \times U \rightarrow \mathbb{R}^+$ es una función de distancia que cuantifica la disimilitud, cumpliendo con las propiedades de reflexividad, positividad estricta, simetría y desigualdad triangular. Bajo este esquema, las operaciones críticas son la búsqueda por rango y la de los k vecinos más cercanos (k-NN).

Como el cálculo de d suele ser computacionalmente costoso (especialmente en biometría o análisis multimedia), la eficiencia del sistema depende de minimizar estas evaluaciones. Por ello, uno de los focos está en el diseño y optimización de *Métodos de Acceso Métricos* (MAMs), que permitan realizar búsquedas sin recurrir a revisiones secuenciales exhaustivas, garantizando la escalabilidad frente al crecimiento constante de los repositorios.

Asimismo, el desarrollo de estas herramientas considera la jerarquía de memorias como un factor determinante. Debido a que el volumen de los índices suele exceder la capacidad de la memoria RAM, es imperativo minimizar las operaciones de entrada/salida (E/S) en memoria secundaria. Por ello, se proponen soluciones que integran: construcción incremental, eficiencia en recursos y computación de alto desempeño (HPC).

En resumen, el objetivo primordial es desarrollar herramientas robustas que permitan la interacción eficiente con el usuario y la manipulación de grandes volúmenes de datos no estructurados, contribuyendo así a la resolución de problemas complejos de Big Data en entornos científicos e industriales. Sin embargo, dado que este conocimiento puede brindarse para la formación en la disciplina, se contempla su transferencia para grado y posgrado.

3. Líneas de Investigación

Como se ha mencionado, en el marco de estas líneas de investigación se pretende realizar aportes a la gestión de grandes volúmenes de datos no estructurados, brindar soluciones computacionales sobre aplicaciones reales y transferir adecuadamente el conocimiento generado a estudiantes de las carreras de grado y posgrado de la UNSL.

Métodos de acceso métricos

El modelo de espacios métricos constituye el marco formal más adecuado para gestionar repositorios masivos de datos no estructurados, donde el objetivo principal no es la coincidencia exacta, sino la recuperación de información basada en consultas por similitud. Bajo este marco, la eficiencia se logra con la implementación de MAMs.

Desde una perspectiva operativa, los MAMs aprovechan las propiedades intrínsecas de la función d (particularmente la simetría y la desigualdad triangular) para optimizar el recorrido del índice. Estas propiedades permiten estimar distancias entre un objeto de consulta q y los elementos almacenados a partir de distancias previamente computadas. Así, es posible determinar que ciertos objetos no satisfacen el criterio de búsqueda y pueden ser descartados sin evaluar directamente la función d, proceso conocido como poda del espacio (pruning). Como el cálculo de la métrica suele ser el mayor cuello de botella en términos de tiempo de CPU, reducir estas evaluaciones impacta en el rendimiento global del sistema.

En la literatura actual, los MAMs se clasifican de acuerdo a la estrategia usada para organizar y referenciar los datos en el espacio métrico: pivotes, particiones compactas o permutaciones.

El diseño de un MAM eficiente debe considerar factores críticos que determinan su viabilidad en escenarios de Big Data. Entre ellos destaca el dinamismo, es decir, la capacidad del índice para soportar operaciones frecuentes de inserción y eliminación sin requerir una reestructuración total. Asimismo, la precisión de la búsqueda define si el índice ofrece una respuesta exacta o una aproximada, siendo esta última una alternativa eficaz para reducir tiempos en volúmenes de datos masivos.Un aspecto determinante es la jerarquía de memorias. Debido a que el volumen de datos suele exceder la capacidad de la memoria RAM, los índices deben almacenarse en memoria secundaria (disco). En esta situación, la eficiencia no se limita a minimizar los cálculos de distancia, sino que requiere reducir drásticamente las

operaciones de E/S. El diseño debe priorizar que la memoria RAM contenga únicamente las estructuras de metadatos más críticas para decidir qué secciones del índice deben recuperarse físicamente del disco durante una búsqueda.

Se han desarrollado varios MAMs, algunos dinámicos y especialmente diseñados para grandes volúmenes de datos, basados en la *Lista de Clusters* (*LC*) [1] [1]. Uno de estos índices, que tiene una buena ocupación de página, operaciones eficientes tanto en cálculos de distancia como en operaciones de E/S y logra un buen desempeño en espacios de alta dimensión, es el *Conjunto Dinámico de Clusters* (*DSC*) [6]. El *DSC* organiza los centros de clusters en un *DSAT* en memoria RAM y sus clusters en memoria secundaria [6]. La información en el *DSAT* es clave para disminuir la cantidad de cálculos de distancia y mantener bajos los costos de E/S. También se ha diseñado el (*BOLDSC*), una variante del *DSC*, que sustituye el DSAT por un espacio en memoria donde los objetos son insertados temporalmente, además de un índice auxiliar sobre estos elementos [8].

Por otra parte, el buen desempeño del *DSC* se debe a la calidad de la partición generada por los centros de sus clusters. Cada cluster está determinado por su centro y su radio, pero puede tener "huecos" dentro. Por ello, el *Conjunto Dinámico de Clusters con Cortes de Regiones* (*DSCCR*) [10] mejora la poda en estos agrupamientos, seleccionando un conjunto de "pivotes" globales que permitan caracterizar las zonas de los clusters dentro de las cuales existan efectivamente elementos [5]. Así, se mantiene para cada cluster información sobre la mínima y la máxima distancia a la que cada pivote encuentra elementos de dicho cluster. Esta información identifica los "huecos" del cluster. Durante las consultas, se puede descartar un cluster si la bola de la consulta no posee intersección con la zona del cluster que realmente contiene elementos de interés y se recuperan desde memoria secundaria todos los clusters cuya intersección con la bola de consulta con centro q y radio r es *no vacía*. Al identificar las zonas del cluster que no contienen elementos, si la consulta intersecta al cluster en su zona "hueca" se lo puede descartar y ahorrar cálculos de d y lecturas a disco.

El *Árbol de Aproximación Espacial Distal Dinámico* (*DDiSAT*) [3, 2], emplea la técnica de inserción perezosa para amortiguar los costos de reconstrucción y mejora el rendimiento en las búsquedas sin requerir espacio adicional. En lugar de insertar elementos de manera inmediata, se retrasan y se colocan en una bolsa asociada a su nodo más cercano en el árbol. Cuando la cantidad de elementos pendientes de inserción en las bolsas alcanza el número total de nodos del árbol, se realiza una reconstrucción completa del árbol.

Como el *DDiSAT* está diseñado para memoria RAM, actualmente se está trabajando en una versión dinámica del *DDiSAT* optimizada para memoria secundaria. Esta versión busca equilibrar la eficiencia de las búsquedas y los costos de reconstrucción con un uso eficiente del espacio en disco y la minimización de operaciones de entrada/salida (E/S). El objetivo es alcanzar un rendimiento en los cálculos de distancia comparable al de la versión en RAM, asegurando una alta ocupación de las páginas de disco, reduciendo las operaciones de E/S y manteniendo la localidad de acceso. Esta versión está en etapa de evaluación de resultados para su publicación.

Como se ha mencionado,otra manera de acelerar las búsquedas es usar técnicas de HPC paralelizar el índice. Se ha tomado como base al *DSC* y se ha obtenido una versión paralela, que disminuye los cálculos de distancia y operaciones de E/S necesarias para responder a una consulta y resuelve más consultas simultáneas, aprovechando al máximo las operaciones de E/S realizadas. En esta implementación, cada procesador/nodo mantiene su propio índice local sobre sus datos locales y existe un broker, una cola de mensajería y los nodos conteniendo los índices locales. Este índice se encuentra en etapa final de evaluación, para su próxima publicación.

DBMS para bases de datos multimedia

PostgreSQL es un DBMS que permite realizar consultas por similitud sobre algunos atributos. Por ejemplo, si se han creado índices *KNN-GiST* sobre texto, se pueden resolver búsquedas de k-vecinos más cercanos, pero no son aplicables sobre todo tipo de datos métricos. Además, estos índices brindan plantillas sólo para crear *árboles balanceados* que, tal como se ha mostrado en [1], no son adecuados para espacios métricos generales. Así, se debe incorporar a un DBMS la capacidad de manejar distintos tipos de datos no estructurados y sus operaciones.

Las operaciones habituales sobre bases de datos multimedia son las búsquedas por rango o de k-vecinos más cercanos. Sin embargo, en un DBMS para bases de datos multimedia [9], hay otras operaciones a considerar. Por ejemplo, se debe incorporar

[1] Se utiliza aquí la palabra *cluster* en inglés, en lugar de agrupamiento en espanõl, ya que es de uso habitual en computación.

la operación de *join* por similitud [4]. Dadas dos bases de datos A y B, con $A, B \subseteq \mathcal{U}$, existen distintas variantes del join por similitud entre A y B. Estas variantes dependen del criterio de similitud Φ utilizado. Todas ellas devuelven un conjunto de pares (x, y), $x \in A$ e $y \in B$, tales que se satisface Φ entre x e y. Las variantes más conocidas son: el join por rango, el join de k-vecinos más cercanos y el join de k pares de vecinos más cercanos [7]. Formalmente, dadas $A, B \subseteq \mathcal{U}$, el *join por similitud* entre A y B es $(A \underset{\Phi}{\bowtie} B) = \{(x, y)/x \in A \wedge y \in B \wedge \Phi(x, y)\}$. Si $A = B$, la operación se denomina *auto-join* [4].

Una manera de acelerar los tiempos de respuesta de los joins consiste en disminuir los tamaños de los conjuntos obtenidos, a través de diversificar las respuestas [11]. De esta manera, el resultado es un conjunto más pequeño de pares, con respuestas útiles y diversificadas, y en menor tiempo. Permitir respuestas aproximadas a las consultas es otra manera de acelerar la respuesta. Se ha propuesto un algoritmo simple y eficiente para responder consultas de auto-join por similitud de k vecinos más cercanos aproximados, con una razonable precisión en la respuesta [4]. Con estos índices y algoritmos disponibles, se espera que *PostgreSQL* se transforme en un DBMS eficiente para datos multimedia masivos.

Gestión de datos agropecuarios

La incorporación tecnológica en el ámbito agropecuario ha dado lugar a un paradigma que optimiza la producción mediante herramientas tecnológicas y Ciencia de Datos.

Se está trabajando en diseñar e implementar una aplicación para un dispositivo móvil, que puede ser un teléfono celular inteligente, que permita comandar un instrumento de medición de nivel freático de uso agropecuario, de forma clara y sencilla. Esta aplicación debe servir de interfaz con el operario y además tener la capacidad de registrar el dato obtenido en la adquisición del nivel freático, como así también la ubicación georeferenciada, la fecha y la hora de la medición.

Sinconización de bases de datos

Otra línea considerada es la sincronización robusta y consistente de bases de datos distribuidas para un problema de venta de productos online (local, nube y web) mediante un esquema maestro-maestro para entornos de alta disponibilidad y tolerancia a fallos, considerando modelos de datos estructurados y no estructurados. Esta propuesta de investigación busca abordar el desafío crítico de mantener la coherencia y la disponibilidad de datos en un entorno distribuido para un sistema de venta de productos online, donde la información reside en bases de datos dispares: una local, una en la nube y una accesible vía web. La implementación de un esquema maestro-maestro permitirá que todas las instancias de la base de datos se puedan leer y escribir, optimizando así la resiliencia del sistema y la continuidad operativa. Se considerará, en caso de ser necesario, la integración de bases de datos relacionales y NoSQL para manejar distintos tipos de datos.

Diseño de material didáctico

En el marco de la línea "Diseño, desarrollo y transferencia de material didáctico para la enseñanza de tecnologías y metodologías de gestión de datos", la producción de recursos pertinentes para las carreras del departamento, plantea desafíos significativos. Para esta etapa inicial, resulta imperativo investigar, evaluar materiales y estrategias didácticas ya existentes en el contexto de la Educación Superior. En este sentido, se pretende realizar un relevamiento de herramientas y metodologías didácticas vigentes aplicables a la gestión de datos.

Por ejemplo, en la asignatura Arquitectura de Procesadores (UNSL), se utiliza la arquitectura M0-Extendida para ilustrar cómo la evolución de los modos de direccionamiento y los mecanismos de gestión de memoria (como la segmentación y el paginado) permiten la reubicación de código, la protección de memoria y la multiprogramación segura. Estos avances, en sinergia con el desarrollo de compiladores, el procesamiento multihilo y las arquitecturas multinúcleo, han allanado el camino para la administración masiva de datos (Big Data) y dispositivos de IoT.

Actualmente, el ecosistema de desarrollo incluye las siguientes herramientas y avances:

- Entorno de software: se dispone de un cross-assembler y un depurador bajo Linux que facilitan la experimentación estudiantil con la arquitectura definida.
- Implementación en hardware (FPGA): se desarrolla un kit de evaluación sobre FPGA que ya cuenta con interfaces funcionales para RS232 y teclado de membrana, mientras que la interfaz I2C se encuentra en fase de simulación.
- Desarrollo futuro: se proyecta la incorporación de un depurador integral en el kit de evalua-

ción que permitirá el monitoreo del estado de la CPU, la modificación de registros y celdas de memoria, y la carga de código objeto mediante la interfaz RS232.

Si bien en esta primera etapa el objetivo es utilizar el kit de evaluación para la enseñanza de la evolución de las arquitecturas de computadoras, la experiencia lograda en el desarrollo de hardware está encaminada a pensar en arquitecturas de computadoras orientadas al procesamiento de consultas a bases de datos no convencionales.

4. Resultados

Se han publicado los resultados de *DSCCR* [10], se ha finalizado el diseño e implementación de la versión para memoria secundaria del *DDiSAT* [3, 2], la cual está en etapa de evaluación de los resultados, para su publicación.

Además, se está finalizando la evaluación de la nueva versión paralela del índice *DSC* y se espera también publicar estos resultados a la brevedad.

La extensión de *PostgreSQL*, para administrar y resolver consultas sobre bases de datos multimedia, pronto estará disponible, publicando los nuevos aportes al área.

5. Formación de recursos

Se están realizando tesis de Maestría: (1) "Estructuras eficientes sobre datos masivos para búsquedas en espacios métricos" (UNSL), (2) "Sistema administrador para bases de datos métricas" (UNSL), (3) "Índices Métricos – Optimización del *DSC* usando cortes de regiones" (UNSJ)y (4) "Diseño e implementación de una versión para memoria secundaria del *DDiSAT*" (UNSJ).

Además, se está definiendo una tesis de Maestría en Sistemas Embebidos (UNSL), dos trabajos finales de la Especialización en Educación Superior (UNSL), dos trabajos finales de grado y una pasantía de investigación (UNSL).

Referencias

[1] E. Chávez and G. Navarro. A compact space decomposition for effective metric indexing. *Pattern Recognition Letters*, 26(9):1363–1376, 2005.

[2] E. Chávez, M. E. Di Genaro, and N. Reyes. An efficient dynamic version of the distal spatial approximation trees. In *Actas del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022)*, pages 468–477, Oct. 2022.

[3] E. Chávez, M. E. Di Genaro, and N. Reyes. Dynamic distal spatial approximation trees. In Patricia Pesado, editor, *Computer Science – CACIC 2022*, pages 175–189, Cham, 2023. Springer Nature Switzerland.

[4] S. Ferrada, B. Bustos, and N. Reyes. An efficient algorithm for approximated self-similarity joins in metric spaces. *Information Systems*, 91:101510, 2020.

[5] J. Lokoč, J. Moško, P. Čech, and T. Skopal. On indexing metric spaces using cut-regions. *Information Systems*, 43:1–19, 2014.

[6] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.

[7] R. Paredes and N. Reyes. Solving similarity joins and range queries in metric spaces with the list of twin clusters. *JDA*, 7:18–35, March 2009. doi:10.1016/j.jda.2008.09.012.

[8] R. Paredes, N. Reyes, K. Figueroa, and M. Hoffhein. A new dynamic, secondary-memory metric index. In P. Pesado, P. Thomas, and F. Chichizola, editors, *Actas del XXX Congreso Argentino de Ciencias de la Computación (CACIC 2024)*, pages 520–529, Oct. 2024.

[9] C. Rong, C. Lin, Y. N. Silva, J. Wang, W. Lu, and X. Du. Fast and scalable distributed set similarity joins for big data analytics. In *2017 IEEE 33rd Int. Conference on Data Engineering (ICDE)*, pages 1059–1070, April 2017.

[10] P. Samat, V. Ludueña, N. Reyes, R. Paredes, K. Figueroa, and M. Lagos. An optimization of the dynamic set of clusters for similarity searches. In *Procs. del XXXI Congreso Argentino de Ciencias de la Computación CACIC 2025)*, pages 477–486. Universidad Nacional de Río Negro, Dic. 2025.

[11] L. Santos, L. Carvalho, W. Oliveira, A. Traina, and C. Jr. Traina. Diversity in similarity joins. In G. Amato, R. Connor, F. Falchi, and C. Gennaro, editors, *Similarity Search and Applications*, volume 9371 of *Lecture Notes in Computer Science*, pages 42–53. Springer International Publishing, 2015.

Bases de Datos Espaciales y Espacio-Temporales y su aplicación en Gestión Académica Universitaria

Javier Apolloni, Berta García, Pablo Palmero y María Teresa Taranilla

Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
{javierma, bgarcia, prpalmero, tarani}@unsl.edu.ar

RESUMEN

En el Proyecto Tecnologías Avanzadas de Bases de Datos de la Universidad Nacional de San Luis se investiga sobre Bases de Datos Espaciales y Espacio-Temporales. Esta línea de investigación se centra en el estudio de modelos, técnicas y herramientas para el almacenamiento, consulta y análisis de información espacial y temporal abordando problemáticas reales. En particular, se propone la incorporación sistemática de dimensiones espaciales y temporales en la gestión académica universitaria para aportar en el fortalecimiento de la toma de decisiones estratégicas.

Palabras clave: Bases de Datos Espaciales y Espacio-Temporales, Trazabilidad, Geometría Computacional, Gestión Académica**.**

CONTEXTO

El trabajo de investigación se enmarca en el Proyecto Tecnologías Avanzadas de Bases de Datos que integra el Laboratorio de Investigación y Desarrollo en Base de Datos (LaBDa) de la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis en colaboración con investigadores de otros proyectos de esta Universidad

El proyecto se enfoca en el estudio de modelos de bases de datos no convencionales y en el desarrollo de estrategias para gestionar y recuperar información almacenada en repositorios de datos heterogéneos y no estructurados. La disponibilidad creciente de datos provenientes de dispositivos móviles, sensores, sistemas institucionales y plataformas digitales, entre otros, plantea nuevos desafíos, y requiere incorporar modelos capaces de manejar información que evoluciona en el tiempo y en el espacio. En este contexto, la generación de información confiable resulta esencial para respaldar la toma de decisiones cuando se busca comprender fenómenos dinámicos y multidimensionales.

En este marco, se propone el estudio de técnicas y herramientas que faciliten la gestión, integración y análisis de información espacial y temporal. Esto abarca el diseño de aplicaciones para modelar y administrar datos complejos, así como la obtención de información que permita realizar análisis y generar conocimiento para respaldar la toma de decisiones en diferentes ámbitos de aplicación**.**

1. INTRODUCCIÓN

El avance de las Tecnologías de la Información y la Comunicación (TIC) ha impulsado la generación de grandes volúmenes de datos vinculados con la localización de objetos, sus desplazamientos y su evolución temporal. Diversas tecnologías como Sistemas de Posicionamiento Global (GPS), dispositivos móviles, sensores remotos y sistemas institucionales producen información que permite describir procesos dinámicos en diversos ámbitos de aplicación. Este escenario plantea desafíos en el almacenamiento,

integración, consulta, análisis y visualización de datos espaciales y temporales. Los modelos tradicionales de bases de datos presentan limitaciones para representar fenómenos que integran ambas dimensiones, lo que requiere enfoques específicos para su gestión eficiente.

En este contexto, las bases de datos espaciales permiten representar y almacenar la ubicación espacial de objetos, sus atributos y las relaciones espaciales entre ellos. Las bases de datos temporales incorporan el componente temporal y permiten conservar estados actuales y pasados. Las bases de datos espacio-temporales integran ambas dimensiones, permitiendo el análisis de la evolución de objetos y eventos en el espacio a lo largo del tiempo, un aspecto clave para tareas como monitoreo, trazabilidad, detección de patrones y seguimiento de tendencias [1] [2] [3].

La integración de las TIC con las bases de datos espaciales y espacio-temporales es una herramienta fundamental para gestionar información compleja en dominios como planificación urbana, gestión de recursos naturales, trazabilidad, agricultura y ganadería, seguridad, gestión de instituciones, entre otros.

En el marco del proyecto de investigación se han abordado problemáticas en distintos ámbitos donde la dimensión espacial y temporal de los datos resulta de interés.

En producción agropecuaria, se requieren soluciones para seguimiento de indicadores productivos y ambientales, trazabilidad y monitoreo de rodeos.

En la gestión de incendios forestales, es necesario integrar información geoespacial para documentar eventos, analizar su dinámica espacial y temporal, y evaluar su impacto.

En el ámbito de la seguridad y prevención del delito, disponer de información precisa sobre localización y desplazamientos de vehículos permite realizar análisis de trazabilidad y detección de situaciones potencialmente sospechosas.

En el ámbito universitario, los sistemas de gestión académica almacenan información relevante para la planificación institucional. La incorporación de dimensiones espaciales y temporales en su análisis permite identificar patrones, evaluar cobertura e impacto institucional, fortalecer los procesos de planificación y realizar aportes para la toma de decisiones en el sistema universitario.

En todos estos escenarios se presenta la necesidad de representar y gestionar tanto datos tradicionales como datos complejos que integran dimensiones espaciales y temporales, incluyendo información actual e histórica. En este contexto, las bases de datos espaciales y espacio-temporales realizan aportes significativos para su almacenamiento, organización y consulta.

En este marco, la línea de investigación se centra en el modelado y la gestión eficiente de datos espaciales y espacio-temporales aplicados a distintos dominios de aplicación.

2. LÍNEA DE INVESTIGACIÓN

La línea de investigación aborda el estudio, diseño e implementación de modelos y estructuras para la representación, almacenamiento, indexación y consulta eficiente de datos espaciales y espacio-temporales, enfocado en aplicaciones reales.

Entre los temas de estudio se destaca el diseño e implementación de índices espaciales y espacio-temporales adaptados a diferentes contextos, el desarrollo de herramientas y aplicaciones que optimicen el rendimiento y la utilidad de las bases de datos espaciales y espacio-temporales. Se incorporan conceptos y técnicas de Geometría Computacional, con el objetivo de abordar aspectos específicos de los problemas analizados y aportar soluciones eficientes [4].

Los objetivos de investigación propuestos son los siguientes:

- Analizar y evaluar técnicas de indexación espacial y espacio-temporal aplicadas a objetos en movimiento en distintos escenarios, con el diseño e implementación de estructuras de almacenamiento y algoritmos de consulta con evaluación experimental.

- Aplicar bases de datos espaciales y espacio-temporales en la resolución de problemáticas reales
- Investigar técnicas de modelado e integración que permitan articular datos tradicionales con datos espaciales y temporales, para el análisis de fenómenos dinámicos en diversos contextos.
- Diseñar herramientas informáticas que implementen soluciones prácticas en diversos dispositivos, incorporando funcionalidades de consulta y visualización que apoyen los procesos de toma de decisiones.

En esta línea se busca consolidar un marco de estudio que articule modelado, almacenamiento e indexación eficiente de datos espaciales y temporales, con aplicación concreta en escenarios específicos.

3. RESULTADOS OBTENIDOS/ ESPERADOS

En el marco de esta línea de investigación se ha trabajado en distintos dominios de aplicación, abordando problemáticas específicas que requieren la integración de dimensiones espaciales y temporales.

En el dominio de la producción agropecuaria se propuso una plataforma integral orientada a la gestión eficiente del sector. En el ámbito ganadero se identificaron requerimientos vinculados con el seguimiento de indicadores productivos y ambientales, la trazabilidad animal y el control de la alimentación [5] [6]. En este contexto, se desarrolló una herramienta orientada al seguimiento espacial y temporal de rodeos, integrando datos tradicionales y georreferenciados. La solución permite ejecutar consultas espaciales y espacio-temporales sobre los objetos de estudio e incorpora una interfaz de comunicación con la base de datos y funcionalidades de visualización en dispositivos móviles, facilitando la toma de decisiones en tiempo real.

Posteriormente, la herramienta fue adaptada para procesar lotes de datos provenientes del monitoreo de fauna silvestre, incorporando funcionalidades para visualizar y analizar información de distintas especies, sus desplazamientos y comportamientos mediante consultas espacio-temporales. Este desarrollo corresponde a una tesis de Maestría en Ciencias de la Computación finalizada en 2025 [7].

El proyecto multidisciplinario *Gestión Integral de Incendios Forestales* se desarrolla en el marco de una colaboración entre distintas unidades académicas de la Universidad Nacional de San Luis, integrando diversas áreas del conocimiento. En este proyecto se realizaron aportes en el diseño de bases de datos espaciales y en la optimización de consultas asociadas al análisis de eventos. Se trabajó en la recopilación y procesamiento de datos espaciales y temporales provenientes de imágenes multiespectrales (Landsat 7-8 y 9, Sentinel-2 MSI, MODIS) y productos derivados (Landsat BA, MODIS BURN, FireCCI), generando una base de datos con información de incendios ocurridos en la provincia de San Luis entre 2000 y 2023 [8] [9].

En el ámbito de la seguridad se propuso el diseño de un sistema para analizar la circulación y trazabilidad de vehículos, para detectar aquellos que puedan estar vinculados con actividades ilícitas. La propuesta se basa en el seguimiento de movimientos mediante la lectura de patentes, utilizando recursos disponibles en la provincia de San Luis [10].

En relación con el ámbito universitario se realizó un trabajo centrado en la migración del sistema SIU-Guaraní versión 2 a versión 3 en la UNSL, consolidando una base de datos centralizada para las unidades académicas, preservando integridad y consistencia de información, y adaptando reportes a estándares institucionales [11].

Los sistemas de gestión académica almacenan información estructurada y valiosa para la planificación institucional. Sin embargo, su análisis suele realizarse desde enfoques tradicionales, sin incorporar la dimensión espacial ni su evolución temporal. En este sentido, se identifica una oportunidad de investigación relevante, dado que la educación

superior constituye un sistema complejo donde los componentes espaciales (lugar de origen, distribución geográfica, accesibilidad y movilidad) y temporales (trayectorias, cohortes, evolución de matrícula, permanencia y egreso) resultan esenciales para comprender y gestionar los procesos del estudiantado. Como eje de trabajo, se propone avanzar en la articulación de bases de datos espaciales y espacio-temporales con problemáticas de gestión universitaria en escenarios en los que estas dimensiones adquieren relevancia. Se espera que esta línea de trabajo permita transformar datos administrativos tradicionales en información estratégica para la planificación institucional.

Además, se plantea continuar promoviendo actividades que fomenten la relevancia y la transferencia de proyectos de investigación, desarrollo e innovación (I+D+i) en las áreas específicas de intervención. Se espera potenciar la colaboración entre distintos equipos de trabajo, consolidar la naturaleza interdisciplinaria de la investigación y mejorar las competencias de los investigadores participantes en este proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

La formación de recursos humanos se refleja en la dirección de tesis de maestría y trabajos finales de licenciatura en Ciencias de la Computación, así como en la dirección de becas y dictado de cursos de posgrado.

Los investigadores fortalecen su formación con capacitación académica continua, cursos de especialización, participación en eventos de divulgación científica, conferencias y otras actividades vinculadas con la investigación. Además, se desarrollan actividades de colaboración e intercambio científico y tecnológico con investigadores de la Universidad Nacional de San Luis y de otras instituciones académicas.

5. BIBLIOGRAFÍA

[1] Elmasri, R. Navathe, S.B.*: Fundamentals of Database Systems.* 7ma Edition. Pearson (2015)

[2] Manolopoulos, Y., Papadopoulos, A., Vassilakopoulos, M.: *Spatial Databases: Technologies, Techniques and Trends*. Idea Group (2005)

[3] Shekhar, S.; Chawla, S.: *Spatial databases: a tour.* Prentice Hall (2003)

[4] de Berg, M., Cheong O., van Kreveld, M., Overmars, M.: *Computational Geometry: Algorithms and Applications.*, Springer-Verlag, Heidelberg (2008)

[5] Gagliardi, E.; Dorzán, M.G.; Taranilla, M.T.; Palmero, P.; Casanova, C.: *GeoSeguimiento de Rodeos, hacia una plataforma integral para el Agro*. Congreso de AgroInformática CAI 2018, 47 JAIIO (2018).

[6] Palmero, P.; Gagliardi, E; Dorzán, M. G. *Seguimiento de rodeos en establecimientos agropecuarios.* XXVI Congreso Argentino de Ciencias de la Computación CACIC (2020)

[7] Palmero, P.; *Desarrollo de una Plataforma Integral de Información relacionada al Agro,* Tesis de Maestría en Ciencias de la Computación -UNSL (2025)

[8] *Gestión Integral de Incendios Forestales* Proyecto de investigación de la Agencia I+D+i PICT Aplicados Salto Institucional, PICT-2021-CAT-II-00105

[9] *Incendios Forestales 2000-2023 en San Luis* API en Google Earth Engine (2024) https://lucasneoclasico.users.earthengine.app/view/areas-quemadas-firecci-sl

[10] Leiva Olguin G., Martinez R., Paez B., Gagliardi E. O. *Diseño de un sistema de trazabilidad de vehículos en estado de sospecha.* Congreso Nacional de Ingeniería Informática / Sistemas de Información-Universidad Tecnológica Nacional (2022)

[11] Lupi Casale, Gianni; *Migración del Sistema Informático Universitario Guaraní Versión 2 a la Versión 3 en la UNSL* – Trabajo final de Licenciatura en Ciencias de la Computación UNSL (2024)

DISEÑO DE UN PROTOTIPO PARA LA PREDICCIÓN DE RIESGO DE DESERCIÓN UNIVERSITARIA (PRDU) A PARTIR DE REGISTROS SIU-GUARANÍ

Claudia N. González [1], Andrea Villagra [2], Daniel Raul Pandolfi [2], Carlos A. Talay [1], Diego R. Rodríguez Herlein [1]

Instituto de Tecnología Aplicada (ITA) – UNPA- UACO

[1] Unidad Académica Río Gallegos, Universidad Nacional de la Patagonia Austral - UARG - UNPA

[2] Unidad Académica Caleta Olivia, Universidad Nacional de la Patagonia Austral - UACO – UNPA

{gonzalez, ctalay, *dherlein}@uarg.unpa.edu.ar; {*avillagra, dpandolfi*}*@uaco.unpa.edu.ar

RESUMEN

Esta línea de investigación se orienta al diseño y validación de un prototipo para la predicción del riesgo de deserción universitaria (PRDU) a partir del análisis de registros académicos institucionales provenientes del sistema SIU-Guaraní de la Universidad Nacional de la Patagonia Austral (UNPA). El estudio se contextualiza en la realidad territorial de la provincia de Santa Cruz, caracterizada por amplias distancias geográficas, condiciones socioeconómicas heterogéneas y trayectorias educativas no lineales, factores que inciden en la permanencia estudiantil.

La propuesta metodológica integra procesos de extracción, limpieza y preparación de datos, análisis exploratorio para la caracterización de su estructura y calidad, y la aplicación comparativa de técnicas de aprendizaje automático supervisado y no supervisado, incluyendo modelos basados en redes neuronales. El objetivo es identificar patrones, variables relevantes y perfiles de riesgo que permitan estimar la probabilidad de deserción y generar evidencia empírica contextualizada.

Como resultado, se espera desarrollar un prototipo funcional que contribuya a la toma de decisiones institucionales, facilitando la implementación de estrategias de acompañamiento y retención estudiantil basadas en analítica predictiva y conocimiento derivado de datos locales.

Palabras clave: Deserción universitaria; SIU-Guaraní; aprendizaje automático; predicción de riesgo; Redes neuronales

CONTEXTO

Esta línea de investigación se desarrolla en el Laboratorio de Tecnologías Emergentes (LabTEm) del Instituto de Tecnología Aplicada (ITA), UNPA-UACO, en articulación académica con la Unidad Académica Río Gallegos (UARG), en el marco del Proyecto de Investigación "Enfoques Inteligentes para Ciudades del Futuro: Empleo, Movilidad y Ambiente Sostenible". El trabajo contribuye al eje de empleo del proyecto mediante el desarrollo de herramientas predictivas basadas en datos institucionales para fortalecer la retención estudiantil y la formación de recursos humanos en la región patagónica.

1. INTRODUCCIÓN

La deserción universitaria constituye un fenómeno complejo y multicausal que impacta tanto en las trayectorias estudiantiles como en la planificación y eficiencia de las instituciones de educación superior. Diversos estudios señalan que el abandono resulta de la interacción de factores académicos, socioeconómicos e institucionales, lo que ha motivado el desarrollo de modelos teóricos que integran dimensiones de integración académica y social, compromiso institucional y condiciones del estudiantado no tradicional [1]–[4]. En este contexto, la analítica educativa y el aprendizaje automático han emergido como enfoques relevantes para el análisis de trayectorias académicas, permitiendo transformar registros institucionales en evidencia orientada a la toma de decisiones [5]. En la provincia de Santa Cruz, esta problemática presenta particularidades

asociadas a la dispersión territorial, las grandes distancias entre localidades y la necesidad de compatibilizar estudio, trabajo y responsabilidades familiares, factores que pueden afectar la continuidad académica. En la Universidad Nacional de la Patagonia Austral (UNPA), estas condiciones configuran un escenario donde la construcción de evidencia basada en datos institucionales resulta clave para comprender patrones de permanencia y abandono, y para orientar estrategias de acompañamiento estudiantil.

En este marco, los sistemas de alerta temprana basados en datos han demostrado potencial para estimar niveles de riesgo y apoyar intervenciones institucionales, especialmente cuando integran técnicas de aprendizaje automático e interpretabilidad que favorecen la transparencia y la adopción por parte de actores académicos [6]–[11]. La disponibilidad de registros académicos y administrativos del sistema SIU-Guaraní permite abordar este problema desde una perspectiva computacional, mediante el desarrollo de modelos predictivos que requieren un tratamiento riguroso de calidad de datos, consistencia temporal y control de sesgos.

En este contexto, el presente trabajo se inscribe en una línea de investigación orientada al diseño de un prototipo de analítica predictiva para la estimación del riesgo de deserción universitaria, integrando técnicas de aprendizaje supervisado, no supervisado y modelos basados en redes neuronales, junto con enfoques de inteligencia artificial explicable y principios de uso responsable de datos [12]–[14].

A diferencia de estudios previos basados en datasets abiertos o contextos educativos masivos, esta línea de investigación aborda la predicción de riesgo de deserción a partir de registros institucionales reales en una universidad regional con características territoriales específicas. Este enfoque permite desarrollar modelos predictivos contextualizados y generar evidencia empírica aplicable a la toma de decisiones institucionales. En este sentido, el trabajo contribuye al campo de la Inteligencia Artificial aplicada a la educación mediante el desarrollo de un prototipo predictivo interpretable orientado a fortalecer las estrategias de permanencia estudiantil en entornos universitarios regionales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de I/D se orienta al diseño y validación de un prototipo de analítica predictiva para la estimación del riesgo de deserción universitaria a partir de registros institucionales del sistema SIU-Guaraní, integrando enfoques de Learning Analytics y técnicas de Machine Learning. Desde una perspectiva computacional, el problema se formula como una tarea de clasificación supervisada, cuyo objetivo es estimar la probabilidad de pertenencia a una categoría de riesgo en función de variables académicas, administrativas y sociodemográficas. Complementariamente, se abordan tareas no supervisadas orientadas a la identificación de perfiles y patrones latentes en las trayectorias académicas.

El desarrollo tecnológico contempla la evaluación comparativa de modelos supervisados como Regresión Logística, Árboles de Decisión, Random Forest, Gradient Boosting y SVM, técnicas de clustering para segmentación (k-means, métodos jerárquicos y DBSCAN) y modelos basados en redes neuronales (MLP) cuando aporten valor en escenarios de mayor complejidad y dimensionalidad [15]–[19]. La evaluación se realizará mediante esquemas de validación robustos y métricas adecuadas ante desbalance de clases (recall, F1 y PR-AUC). En todos los casos se incorporan técnicas de interpretabilidad, tales como SHAP o métodos equivalentes, considerando además las advertencias recientes sobre los límites y riesgos de interpretación en contextos educativos [20], [21]. De manera transversal, el proyecto se alinea con principios de ética, privacidad y gobernanza de datos en analítica educativa [12]–[14].

Metodológicamente, el trabajo se estructura en cuatro etapas. La primera (etapa 1) corresponde al relevamiento, integración y control de calidad de los datos, incluyendo definición de cohortes, tratamiento de valores faltantes,

normalización y validación de consistencia temporal, con el fin de construir un dataset analítico reproducible [15], [18]. La segunda (etapa 2) consiste en un análisis exploratorio orientado al modelado (EDA técnico), centrado en cobertura, distribución y decisiones de preprocesamiento que sustenten la selección de variables y transformaciones. La tercera etapa (etapa 3) aborda el entrenamiento y evaluación comparativa de los modelos definidos, incorporando criterios de estabilidad y control de sobreajuste. Finalmente, la cuarta etapa (etapa 4) contempla la implementación de un prototipo que integre el pipeline analítico y permita visualizar indicadores de riesgo y variables asociadas, acompañado de lineamientos de uso responsable, minimización de datos y comunicación de incertidumbre.

Actualmente se encuentra finalizada la etapa de relevamiento e integración de datos, con un dataset consolidado y reglas de depuración definidas. Se desarrolla la etapa de análisis exploratorio orientado al modelado, focalizada en la carrera Analista de Sistemas en las unidades académicas UARG, UACO y UART (cortes 2018–2025), con el objetivo de cerrar criterios de preprocesamiento y evaluación previos al entrenamiento de modelos.

3. RESULTADOS OBTENIDOS/ ESPERADOS

En el marco de esta línea de I/D se completó la Etapa 1 (relevamiento, integración y preparación de datos) y se encuentra en desarrollo la Etapa 2 (análisis exploratorio orientado al modelado).

Los resultados obtenidos constituyen un aporte fundamental para el desarrollo de modelos predictivos de riesgo de deserción en contextos universitarios regionales, ya que permitieron construir un dataset analítico longitudinal validado, definir criterios reproducibles de preparación y calidad de datos, y establecer una infraestructura metodológica que habilita el entrenamiento, evaluación e interpretación de modelos de aprendizaje automático aplicados a trayectorias académicas reales. Esta base resulta esencial para el desarrollo de sistemas de alerta temprana interpretables orientados a la toma de decisiones institucionales.

Dataset consolidado (Etapa 1). Se integró un dataset analítico del sistema SIU-Guaraní para el período 2018–2025. Para el caso de estudio se delimitó la carrera Analista de Sistemas en tres unidades académicas, conformando un subconjunto de N = 7.240 registros, distribuido en UARG: n = 3.650 (50,41%), UACO: n = 2.358 (32,57%) y UART: n = 1.232 (17,02%) (Tabla 1). Este recorte se utiliza como base comparativa para evaluar cobertura, calidad y consistencia de los registros previos al modelado predictivo.

Análisis exploratorio (Etapa 2 – en curso). Indicador de riesgo académico. Para esta etapa exploratoria se construyó un indicador binario de riesgo académico (RIESGO_ACAD_BIN) a partir de variables de avance, rezago y regularidad disponibles en el dataset. Este enfoque basado en proxies de trayectoria permite caracterizar patrones y calidad de datos antes de la definición formal del evento de deserción, estrategia recomendada en contextos donde los criterios institucionales de abandono requieren validación con actores de gestión. La definición operativa final de deserción se establecerá en la Etapa 3 (modelado predictivo), integrando criterios institucionales validados y estándares de literatura.

Distribución del indicador preliminar. El análisis exploratorio sobre RIESGO_ACAD_BIN mostró evidencias relevantes para el diseño del prototipo:

- **Valores faltantes en variables socioeconómicas.** Se observó una alta proporción de valores faltantes en variables como TRABAJA_BIN (NaN: UACO 55,26%; UARG 79,84%; UART 63,72%) y HIJOS_BIN (NaN: UACO 78,50%; UARG 88,88%; UART 79,79%), lo cual condiciona el uso directo de estos atributos en modelos sin un tratamiento explícito de *missingness* (Figura 1).
- **Mejor cobertura en variables de trayectoria.** Variables derivadas de trayectoria como AVANCE_PCT, INTENSIDAD_ACAD y REZAGO_PLAN presentan, en general, mejor disponibilidad relativa que las socioeconómicas (Figura 1), por lo que se consideran candidatas

prioritarias para la construcción de *features* orientadas al modelado.

- **Contraste estadístico como apoyo al EDA.** Como respaldo cuantitativo del análisis descriptivo, se aplicaron pruebas χ^2 de independencia para evaluar asociación entre variables categóricas y unidad académica, considerando únicamente el subconjunto con datos válidos. Para TRABAJA_BIN se observó asociación estadísticamente significativa (χ^2 = 10,0113; gl = 2; p = 0,00670011), aunque con tamaño de efecto pequeño (V de Cramér = 0,0669; N válido = 2.238). En cambio, para HIJOS_BIN no se detectó asociación significativa (χ^2 = 2,2003; gl = 2; p = 0,332824; V = 0,0435; N válido = 1.162). Estos resultados refuerzan la conveniencia de reportar comparaciones normalizadas y de tratar explícitamente los valores faltantes antes del modelado.
- **Comparación normalizada por unidad (evita sesgo por tamaño).** Para evitar sesgos por diferencias de matrícula entre unidades, se reportan distribuciones en porcentajes dentro de cada unidad. La distribución normalizada de TRABAJA_BIN (0/1/NaN) por unidad permite comparar patrones relativos sin depender del tamaño muestral (Figura 2).

	Unidad académica	n (registros)	%
0	UARG	3650	50.41
1	UACO	2358	32.57
2	UART	1232	17.02
3	Total	7240	100.00

Tabla 1. Distribución del caso de estudio (Analista de Sistemas, 2018–2025) por unidad académica (N=7.240). **Fuente:** elaboración propia (SIU-Guaraní UNPA, 2018–2025).

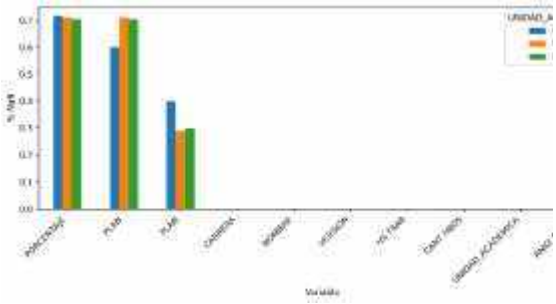


Figura 1: % de faltantes (NaN) en variables clave por unidad. *Fuente: elaboración propia (SIU-Guaraní UNPA, 2018–2025).*

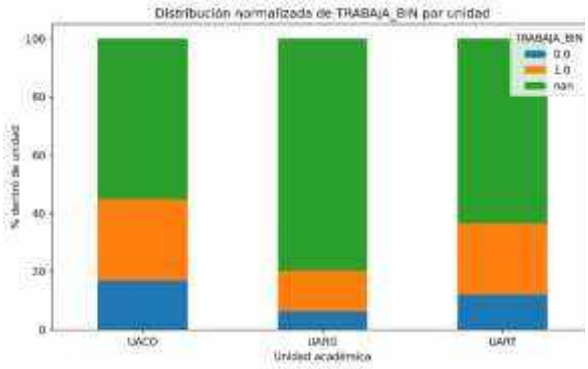


Figura 2. Distribución normalizada (%) de TRABAJA_BIN por unidad. Fuente: elaboración propia (SIU-Guaraní UNPA, 2018–2025).

4. FORMACIÓN DE RECURSOS HUMANOS

Equipo: Tesista de Maestría (Maestría en Informática y Sistemas, UNPA), dirección académica (ITA-UACO-UARG), colaboración doctoral (UNLP) para apoyo metodológico, y articulación con UARG para validación institucional.

Desarrollo de competencias en analítica educativa, ML aplicado, preprocesamiento de datos institucionales y técnicas de explicabilidad. Se proyecta ampliar la línea incorporando estudiantes de Ingeniería y Licenciatura en Sistemas de UNPA para fortalecer la formación práctica y la transferencia a gestión académica.

5. BIBLIOGRAFÍA

[1] Villegas, B. R.; Núñez L., L. A. "Deserción universitaria: revisión 2018–2023." *RIDE*, 14(28), 2024.

[2] Tinto, V. *Leaving College*. Univ. of Chicago Press, 1993.

[3] Bean, J. P.; Metzner, B. S. "Nontraditional student attrition model." *Rev. Educ. Res.*, 55(4), 1985.

[4] Spady, W. G. "Dropouts from higher education." *Interchange*, 1(1), 1970.

[5] Siemens, G.; Long, P. "Analytics in learning and education." *EDUCAUSE Rev.*, 46(5), 2011.

[6] Baker, R. S.; Inventado, P. S. "Educational data mining and learning analytics." In: Larusson, J. A.; White, B. (eds.), *Learning Analytics: From Research to Practice*. Springer, 2014.

[7] Sweeney, M.; Rangwala, H.; Lester, J. "Next-term performance prediction." *J. Educ.*

Data Mining, 8(1), 2016.
[8] Bydžovská, H. "Techniques for predicting student performance." *EDM (Proc.)*, 2016.
[9] Mustofa, S. et al. "AI-driven model for student dropout risk analysis with XAI insights." *C&Ed: AI*, 2024. doi:10.1016/j.caeai.2024.100352
[10] Rabelo, A. M.; Zárate, L. E. "Predicting dropout of higher education students." 2024. doi:10.1016/j.dsm.2024.07.001
[11] Macero Méndez, R. M. "Sistema de alerta temprana para prevención de deserción: estudio de caso." *Dilemas Contemporáneos*, 12(3), 2024. doi:10.46377/dilemas.v12i3.4644
[12] Slade, S.; Prinsloo, P. "Ethical issues in learning analytics." *Am. Behav. Sci.*, 57(10), 2013.
[13] Pardo, A.; Siemens, G. "Ethics and privacy for learning analytics." *Br. J. Educ. Technol.*, 45(3), 2014.
[14] UNESCO. *AI and Education: Guidance for Policy-makers*. UNESCO, 2021.
[15] Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*. O'Reilly, 2022.
[16] Molnar, C. *Interpretable Machine Learning*. 2022.
[17] Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*. Springer, 2009.
[18] Kuhn, M.; Johnson, K. *Applied Predictive Modeling*. Springer, 2013.
[19] Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*. MIT Press, 2016.
[20] Salih, A. M. et al. "Perspective on XAI methods: SHAP and LIME." *Advanced Intelligent Systems*, 7(1), 2025.
[21] Türkmen, G. "Review of studies on XAI in educational research." *J. Educ. Comput. Res.*, 63(2), 2025.

Gestión e Indexación de Datos No Estructurados: Técnicas y Aplicaciones

Paola Azar[1], Hugo Beilis[2], Mónica Fuentes [1], Darío Ruano [1], Carina Ruano [1], María Paula Villegas[1], Norma Herrera [1]

[1]Laboratorio de Investigación y Desarrollo en Bases de Datos
Universidad Nacional de San Luis, San Luis, Argentina
pazar18@gmail.com, fuentesmonicag@gmail.com, dmruano@unsl.edu.ar, carina.ruano@gmail.com, villegasmariapaula74@gmail.com, norma.herrera@gmail.com ,

[2] Escuela de Ingeniería y Ciencias Ambientales
Universidad Nacional de Villa Mercedes, San Luis, Argentina
hugo.beilis@gmail.com

Resumen

Los avances en las tecnologías de la información y la comunicación han permitido un crecimiento exponencial en el tamaño de las bases de datos, así como una mayor diversificación en su contenido. Un gran porcentaje de los datos disponibles son no estructurados (sonidos, imágenes, video, huellas digitales, etc), lo que dificulta su representación y gestión mediante el modelo relacional. En respuesta a esta problemática, se han propuesto nuevos modelos de datos y herramientas específicas para su almacenamiento, indexación y recuperación. En este contexto, el presente trabajo se enfoca en el estudio de modelos de bases de datos no estructurados y en el análisis de técnicas de indexación y búsqueda asociadas, con el objetivo de contribuir a una gestión más eficiente de este tipo de información.

Contexto

Este trabajo se desarrolla como una línea de investigación dentro del Proyecto Tecnologías Avanzadas de Bases de Datos, cuyo objetivo es realizar investigación sobre manejo y recuperación eficiente de información no tradicional que implique el manejo de datos no estructurados. Este proyecto forma parte de Laboratorio de Investigación y Desarrollo en Bases de Datos (LaBDa) de la Universidad Nacional de San Luis.

1 Introducción

El crecimiento exponencial de las fuentes de información ha provocado un aumento significativo tanto en el volumen como en la diversidad de los datos almacenados. En la actualidad, gran parte de la información corresponde a datos no estructurados [2], tales como texto, imágenes, audio o video, que no pueden representarse adecuadamente mediante el modelo relacional tradicional, donde los datos se organizan en nuplas (filas) en relaciones (tablas). En este contexto, los procesos de búsqueda enfrentan nuevos desafíos, tanto en términos de

rendimiento como en la naturaleza de las respuestas esperadas.

Para dar respuesta a estas necesidades, han surgido modelos de bases de datos orientados al manejo de datos no estructurados, entre los que se destacan las bases de datos de texto [6], los espacios métricos [12] y el modelo métrico-temporal [1].

En las bases de datos textuales, las consultas consisten en recuperar documentos a partir de palabras claves (si el texto está en lenguaje natural) o patrones (si es texto plano) [5, 7]. En los espacios métricos [12], las consultas se basan en la similitud entre objetos, permitiendo recuperar elementos similares a uno dado. El modelo métrico-temporal [1]extiende este enfoque incorporando la dimensión temporal, restringiendo la búsqueda a intervalos de tiempo específicos. En todos los casos, el uso de estructiras de datos o índices adecuados resulta fundamental para garantizar búsquedas eficientes.

Cuando el volumen de datos supera la capacidad de la memoria principal, los índices deben manejarse en memoria secundaria, lo que introduce nuevos desafíos. El diseño de índices en memoria principal y en memoria secundaria presenta diferencias fundamentales debido a las características de acceso y costo de cada tipo de almacenamiento. En memoria principal (RAM), el acceso es uniforme y de muy baja latencia. En cambio, en los discos rígidos tradicionales (Hard Disk Drive, HDD) el costo de acceso está dominado por las operaciones de entrada/salida (E/S), que son varios órdenes de magnitud más lentas que el acceso a RAM. Por esta razón, los índices diseñados HDD buscan reducir la cantidad de accesos a disco mediante una organización que favorezca la localidad espacial, el acceso secuencial y la transferencia eficiente de páginas completas.

La evolución de las tecnologías de almacenamiento, desde los discos rígidos tradicionales hacia las unidades de estado sólido (SSD), ha modificado las características de acceso a los datos, impactando directamente en el diseño y desempeño de los índices. Si bien ambos dispositivos, HDD y SSD, cumplen la función de almacenar información de manera persistente, difieren sustancialmente en su estructura interna, su comportamiento durante las operaciones de entrada/salida (E/S) y en los mecanismos de administración de la información.

Los HDD constituyen una tecnología basada en componentes mecánicos, lo que introduce limitaciones como tiempos de latencia más elevados, particularmente notorios en accesos aleatorios, así como susceptibilidad al desgaste y a daños por golpes o vibraciones.

Por su parte, los SSD emplean memorias flash NAND, en las cuales los datos se almacenan en celdas organizadas jerárquicamente en páginas y bloques. La ausencia de partes móviles reduce drásticamente la latencia, mejora el rendimiento en operaciones aleatorias y disminuye el consumo energético, además de aportar una mayor resistencia física. No obstante, la tecnología flash presenta limitaciones tecnológicas propias. No existe la posibilidad de cambiar el contenido de sólo una celda, dado que las lecturas y escrituras se dan a nivel de bloques de celda. Adicionalmente, los SSD presentan un número finito de ciclos de borrado y escritura, condicionando su durabilidad en entornos de uso intensivo.

Las técnicas tradicionales de indexación desarrolladas para HDD suelen ser ineficientes en SSD, debido a que generan patrones de acceso no óptimos para la memoria flash, provocando:

- Incremento de la latencia por write amplification.
- Mayor desgaste de las celdas NAND.
- Desaprovechamiento del paralelismo interno.

En el contexto de un Sistema de Gestión de Bases de Datos, las técnicas de indexación ori-

entadas a SSD deben no solo maximizar la eficiencia de los accesos al almacenamiento, sino también contribuir a prolongar la vida útil del dispositivo, evitando operaciones de escritura innecesarias y adaptándose a la naturaleza interna de la memoria flash.

En este contexto, este trabajo se centra en el estudio de modelos de bases de datos orientados al manejo de datos no estructurados, con especial énfasis en las técnicas de indexación en memoria secundaria y en el análisis de su desempeño sobre tecnologías de almacenamiento actuales, considerando su aplicación en escenarios reales de gran volumen de información.

2 Líneas de Investigación

En este trabajo se aborda el estudio, diseño y optimización de algoritmos de indexación y búsqueda orientados al manejo de datos no estructurados, así como el análisis de su desempeño en distintos escenarios. Asimismo, se exploran aplicaciones concretas que permiten validar las técnicas propuestas y favorecer la transferencia tecnológica al medio. En particular, se desarrollan soluciones aplicadas a problemáticas del ámbito de la medicina forense.

A continuación, se describen las principales líneas de investigación del proyecto.

Indexación en Memoria Secundaria

Una de las líneas de investigación del proyecto se centra en el diseño y evaluación de estructuras de indexación eficientes para bases de datos no estructurados.

En particular, se aborda el desarrollo de técnicas de indexación para HDD, lo que implica la definición de estrategias de paginación que permitan organizar los índices en bloques almacenados en disco y minimizar la cantidad de accesos necesarios durante el proceso de búsqueda. Dado que el costo de las consultas está directamente relacionado con el número de accesos a disco, el objetivo es optimizar la organización de las estructuras de datos para mejorar el desempeño global del sistema.

En este contexto, se trabaja en la implementación de dos índices HDD: el Trie de Sufijos [4], aplicado a bases de datos textuales, y el Historical-FHQT [3], orientado a bases de datos métrico- temporales. Actualmente, se han desarrollado versiones funcionales de ambos índices y se encuentran en fase de evaluación experimental, con el objetivo de analizar su eficiencia en términos de tiempo de búsqueda y costo de acceso a memoria secundaria.

Asimismo, se investiga la adaptación de técnicas de indexación clásicas para su uso en unidades de estado sólido (SSD). Específicamente, el objetivo es proponer estrategias que permitan adaptar tablas de hashing a este tipo de dispositivos de almacenamiento. El punto de partida es el índice Log-Structured Merge-trees(LSM-tree) [8, 10], un índice que se caracteriza por su capacidad para optimizar operaciones de escritura y mantener un rendimiento consistente en entornos con grandes volúmenes de datos.

La elección de LSM-trees como punto de partida se fundamente en su arquitectura jerárquica y su enfoque orientado a escrituras secuenciales, características que resultan especialmente adecuadas para las propiedades físicas y lógicas de los SSD. A partir de este modelo, se busca estudiar cómo las técnicas de hashing pueden adaptarse o rediseñarse para aprovechar la baja latencia, el acceso aleatorio rápido y la alta paralelización de operaciones que ofrecen los SSD, mitigando al mismo tiempo los efectos adversos del desgaste por escritura (write amplification). En este sentido, el trabajo pretende identificar los desafíos técnicos que surgen al combinar estructuras de datos dispersas, como las tablas hashing, con los patrones de acceso y actualización característicos de los LSM-trees, evaluando el im-

pacto de distintas estrategias de indexación compactación y administración de niveles sobre el rendimiento global del sistema.

Aplicaciones en Medicina Forense

De forma complementaria, se desarrollan trabajos orientadas al diseño de sistemas de soporte para la medicina forense. Específicamente nos encontramos trabajando en dos problemas: identificación de cadáveres y datación de la muerte [9, 11].

En el primer caso, se propone un sistema de base de datos basado en el modelo de espacios métricos, en el cual cada cadáver se representa mediante un vector característico que describe sus rasgos físicos relevantes. A partir de las características de una persona buscada, el sistema genera su representación correspondiente y realiza una búsqueda por similitud, obteniendo posibles coincidencias ordenadas según su grado de similitud. Se ha definido un conjunto inicial de características y se ha generado un espacio sintético de datos que permite evaluar el desempeño del sistema.

El trabajo ha avanzado significativamente en dos áreas claves. En primer lugar, se ha creado un espacio sintético de datos para realizar pruebas exhaustivas y evaluar el rendimiento del sistema bajo condiciones controladas. En segundo lugar, se ha definido el conjunto de experimentos a realizar, que se enfocarán en la determinación del radio de búsqueda y en el análisis de la distribución de los datos dentro del espacio métrico.

Estos experimentos son cruciales para validar la eficiencia y precisión de nuestra propuesta. Aunque la experimentación está en una fase avanzada, no se ha completado en su totalidad. Sin embargo, los resultados preliminares son altamente prometedores, lo que sugiere la viabilidad de nuestra propuesta. Planeamos presentar y publicar estos resultados en un congreso futuro, una vez que la fase de experimentación haya concluido.

En el caso de la datación de la muerte, el objetivo es desarrollar un sistema de base de datos que permita registrar casos forenses y recuperar aquellos con características similares, contribuyendo a mejorar la estimación del intervalo post mortem. Para ello, se investiga la aplicación de técnicas de clustering, que permiten agrupar datos según su similitud e identificar casos comparables a partir de un conjunto de registros con valores conocidos. En - particular, se estudian los algoritmos K-Means y DBSCAN, ampliamente utilizados en el análisis de datos multidimensionales. Actualmente, se dispone de un prototipo del sistema y los algoritmos de clustering se encuentran en etapa de implementación y evaluación experimental, con el objetivo de analizar su utilidad en la organización de la información y el soporte a la toma de decisiones en el ámbito forense.

3 Resultados Esperados

Como resultado de las líneas de investigación en curso, se espera obtener estructuras de indexación eficientes para su uso en memoria secundaria HDD, orientadas a bases de datos textuales y métrico-temporales. El desarrollo e implementación de estos índices permitirá evaluar su desempeño en términos de costo de acceso a disco y eficiencia de búsqueda, así como adquirir experiencia para el diseño de nuevas estructuras de indexación adaptadas a tecnologías de almacenamiento actuales, como los SSD.

En cuanto a la transferencia tecnológica, se prevé el desarrollo de un sistema web para la identificación de cadáveres en el contexto de búsqueda de personas desaparecidas y un sistema de soporte para la datación de la muerte. Ambos desarrollos cuentan actualmente con prototipos funcionales, que serán ampliados y evaluados en escenarios reales.

4 Formación de Recursos Humanos

Dentro de esta línea se forman docentes y alumnos de la Universidad Nacional de San Luis y de la Universidad Nacional de Villa Merecedes. Tres alumnos de la Licenciatura en Ciencias de la Computación han realizando una pasantía de investigación en este Proyecto, colaborando con los trabajos aquí expuestos. Actualmente hay en desarrollo 2 Tesis de Maestría y un Trabajo Final de la Licenciatura, todos de la Universidad Nacional de San Luis.

Bibliografia

[1] A. De Battista, A. Pascal, N. Herrera, and G. Gutierrez. Metric-temporal access methods. *Journal of Computer Science & Technology,*, 10(2):54–60, 2010.

[2] H.M. Blanken, A.P. de Vries, H.E. Blok, and L. Feng. *Multimedia Retrieval*. Data-Centric Systems and Applications. Springer Berlin Heidelberg, 2007.

[3] A. De Battista, A. Pascal, G. Gutierrez, and N. Herrera. Un nuevo índice métrico-temporal: el historical-fhqt. In *Actas del XIII Congreso Argentino de Ciencias de la Computación*, Corrientes, Agentina, 2007.

[4] G. H. Gonnet, R. Baeza-Yates, and T. Snider. *New indices for text: PAT trees and PAT arrays*. Prentice Hall, New Jersey, 1992.

[5] Xiang Li et al. Design issues of jpq: A pattern-based query language for document databases. *arXiv preprint arXiv:1504.03770*, 2015.

[6] Navarro. *Text Databases*, pages 688–694. Idea Group Inc., Pennsylvania, USA, 2005.

[7] Gonzalo Navarro. Spaces, trees and colors: The algorithmic landscape of document retrieval on sequences. *ACM Computing Surveys*, 46(4):52:1–52:47, 2014.

[8] Patrick O'Neil, Edward Cheng, Dieter Gawlick, and Elizabeth O'Neil. The log-structured merge-tree (lsm-tree). *Acta Informatica*, 33(4):351–385, 1996.

[9] H. Stigter, T. Krap, and W. L. J. M. Duijst. Estimation of the post-mortem interval; added value of mechanical excitability of human skeletal muscle. *Journal of Forensic and Legal Medicine*, 103:102664, 2024.

[10] Jian Wang, Yu Lu, Qingsong Wang, Yanyan Zhang, and Jiwu Shu. Revisiting secondary indexing in lsm-based storage systems with persistent memory. In *Proceedings of the USENIX Annual Technical Conference (USENIX ATC)*. USENIX Association, 2023.

[11] Katherine E. Weisensee et al. geofor: A collaborative forensic taphonomy database for estimating the postmortem interval. *Forensic Science International*, 355:111934, 2024.

[12] Pavel Zezula, Giuseppe Amato, Vlastislav Dohnal, and Michal Batko. *Similarity Search: The Metric Space Approach*. Springer Publishing Company, Incorporated, 1st edition, 2010.

Integración de métricas de calidad de datos en procesos de minería de datos: Validación metodológica aplicada al caso INCUCAI

Roxana Martínez, Santiago Acha, Agustín Luján, Ornella Mansilla, Nicolas Censi

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{roxana.martinez, santiago.acha}@uai.edu.ar
{agustinezequiel.lujanbidondo, ornellalujan.mansilla, nicolasmartin.censi}@alumnos.uai.edu.ar

RESUMEN

La calidad de los datos constituye un factor determinante en la confiabilidad de los procesos de minería de datos y en la robustez de los modelos predictivos resultantes. No obstante, en numerosos proyectos analíticos, las actividades de evaluación y aseguramiento de calidad se analizan de manera aislada, es decir, sin una integración sistemática dentro del pipeline de análisis. En este trabajo se presenta un modelo metodológico desarrollado en el Laboratorio de Calidad y Ciencia de Datos del CAETI (UAI), orientado a incorporar métricas formales de calidad de datos en cada etapa del proceso de minería de datos. El modelo se fundamenta en dimensiones establecidas por la norma ISO/IEC 25012, adaptadas al contexto de proyectos de Ciencia de Datos, y propone un esquema estructurado que contempla perfilado, evaluación dimensional, limpieza, validación y retroalimentación iterativa. Como instancia de validación empírica, se aplica la propuesta a un conjunto de datos del Instituto Nacional Central Único Coordinador de Ablación e Implante (INCUCAI), previamente analizado en el marco de un trabajo presentado en la Conferencia Latinoamericana de Informática (CLEI) en el año 2025 [1], lo que permite dar continuidad y profundización metodológica a los resultados obtenidos en el análisis del impacto de las dimensiones de calidad sobre el desempeño de modelos predictivos. Los resultados preliminares evidencian mejoras en consistencia, completitud y estabilidad de los modelos, lo que refuerza la necesidad de integrar formalmente la calidad de datos como componente transversal en proyectos de Ciencia de Datos.

Palabras clave: Calidad de Datos; Minería de Datos; ISO/IEC 25012; Ciencia de Datos; Modelos Predictivos; Gobernanza de Datos.

CONTEXTO

El presente trabajo se enfoca en la línea de investigación en Calidad y Ciencia de Datos desarrollada en el ámbito del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI). Esta línea se articula con el proyecto denominado "*Investigación y desarrollo de software para la validación de la calidad de datos abiertos e identificación de patrones para predicciones*", orientado al diseño de enfoques y herramientas que integren evaluación de calidad y análisis predictivo en distintos dominios de aplicación.

En este contexto, el Laboratorio de Calidad y Ciencia de Datos desarrolla actualmente una línea orientada a la formalización de un modelo metodológico que incorpora dimensiones de calidad de datos dentro de procesos de minería de datos, promoviendo prácticas reproducibles, medibles y alineadas con estándares internacionales. El trabajo que aquí se presenta consolida dicha línea, proponiendo un enfoque estructurado para la integración sistemática de métricas de calidad en el ciclo de vida de proyectos de análisis de

datos, fortaleciendo así la confiabilidad y gobernanza de los resultados obtenidos.

1. INTRODUCCIÓN

El crecimiento del volumen y diversidad de datos en entornos organizacionales, científicos y gubernamentales ha consolidado la minería de datos y el aprendizaje automático como herramientas centrales para la generación de conocimiento y soporte a la toma de decisiones. En este contexto, el enfoque empresarial denominado *data-centric AI* [2], en el que los datos se consideran el activo más valioso de la organización, ha puesto en evidencia que el rendimiento de los modelos no depende exclusivamente del algoritmo utilizado, sino fundamentalmente de la calidad del conjunto de datos utilizados para dicho análisis [3].

Diversos estudios recientes han demostrado que deficiencias en dimensiones de calidad tales como completitud, consistencia, exactitud o actualidad pueden afectar de manera significativa el desempeño de los modelos predictivos, generando degradación en métricas de evaluación, sesgos no deseados y pérdida de capacidad de generalización [4] [5]. Además, es importante mencionar que las investigaciones en el ámbito del ciclo de vida de sistemas de aprendizaje automático destacan que los problemas de calidad suelen propagarse a lo largo del pipeline analítico, generando lo que se ha denominado como "*data cascades*" [6] o efectos acumulativos derivados de decisiones tempranas sobre los datos [4] [7].

Desde una perspectiva normativa y de gobernanza, los modelos formales de calidad de datos, como el propuesto en la norma ISO/IEC 25012 [7], establecen dimensiones estructuradas para evaluar atributos de los datos en contextos organizacionales y gubernamentales. Sin embargo, la literatura reciente señala que la adopción de estos marcos no siempre se articula de manera explícita con la evaluación del impacto de la calidad sobre el rendimiento de modelos de minería de datos [9]. Por lo que, esta desconexión metodológica dificulta la trazabilidad entre métricas de calidad y resultados predictivos, afectando la reproducibilidad y la transparencia de los procesos analíticos [10].

En el ámbito sanitario, donde la calidad de los datos incide directamente en la confiabilidad de los análisis y en la toma de decisiones de alto impacto, la necesidad de enfoques estructurados resulta aún más evidente [11] [12]. En este sentido, la integración sistemática de métricas de calidad [13] dentro del ciclo de vida de proyectos de minería de datos se presenta como un desafío metodológico y práctico que requiere un especial cuidado en el análisis de los datos provenientes de sistemas nacionales de información en salud.

En función de lo expuesto, el presente trabajo propone un modelo metodológico orientado a integrar dimensiones de calidad de datos en procesos de minería de datos, estableciendo una secuencia estructurada de evaluación, validación y retroalimentación que permita vincular explícitamente métricas de calidad con el desempeño de modelos predictivos. La propuesta se valida empíricamente a partir de un conjunto de datos derivados del sistema SINTRA (Sistema Nacional de Información de Procuración y Trasplante de la República Argentina) [14], administrado por el Instituto Nacional Central Único Coordinador de Ablación e Implante (INCUCAI) [15], los cuales fueron sometidos a un proceso sistemático de curación, tratamiento, normalización y evaluación de calidad por parte del equipo de investigación, previo a su incorporación al pipeline analítico. Este enfoque permite fortalecer la confiabilidad, gobernanza y reproducibilidad de los resultados obtenidos en el ámbito sanitario.

2. LÍNEAS DE INVESTIGACIÓN DESARROLLO

El presente trabajo se enmarca en la línea de investigación en Ingeniería de Software del Centro de Altos Estudios en Tecnología Informática (CAETI), con foco en Calidad de Datos y Ciencia de Datos aplicada a contextos

gubernamentales y de gestión pública. En particular, se articula con el desarrollo de enfoques metodológicos y algoritmos orientados a la validación, análisis y explotación de datos públicos abiertos en el ámbito de la salud.

En esta instancia del proyecto, los ejes de investigación y desarrollo se estructuran de la siguiente manera:

- *Desarrollo de algoritmos de validación y evaluación dimensional de calidad*, orientados a detectar inconsistencias, valores faltantes, anomalías estructurales y problemas de integridad en datasets públicos abiertos, con especial atención a datos vinculados a procesos de donación y trasplante de órganos y tejidos.
- *Diseño e implementación de métricas específicas para el dominio sanitario*, incluyendo dimensiones como completitud, consistencia semántica, integridad geoespacial y coherencia temporal, integradas dentro de un modelo metodológico formal.
- *Integración de mecanismos de calidad en pipelines de minería de datos*, permitiendo analizar el impacto de las métricas de calidad sobre el desempeño de modelos predictivos desarrollados mediante técnicas de machine learning.
- *Análisis y validación de modelos predictivos en contextos gubernamentales*, evaluando la estabilidad, robustez y capacidad de generalización frente a distintas condiciones de calidad de datos.
- *Desarrollo de prototipos y funcionalidades para herramientas de validación de datasets públicos*, incorporando criterios de trazabilidad, reproducibilidad y gobernanza de datos.
- *Elaboración de lineamientos y buenas prácticas* para la aplicación de técnicas de aprendizaje automático en procesos de validación y análisis de datos abiertos en salud, promoviendo un enfoque data-centric orientado a la calidad.

3. RESULTADOS OBTENIDOS/ESPERADOS

Enfoque del proyecto:

En el contexto del Gobierno Abierto y la disponibilidad de datos públicos en salud, el presente trabajo aborda el desafío de integrar formalmente métricas de calidad de datos dentro de procesos de minería de datos aplicados a información sanitaria. A diferencia de enfoques centrados exclusivamente en la extracción o explotación analítica, la propuesta incorpora una capa metodológica explícita de evaluación y validación de calidad previa y transversal al modelado predictivo. El caso de estudio se basa en datos derivados del sistema SINTRA (Sistema Nacional de Información de Procuración y Trasplante de la República Argentina) [14], administrado por el Instituto Nacional Central Único Coordinador de Ablación e Implante (INCUCAI) [15], vinculados a procesos de donación y trasplante de órganos y tejidos en la República Argentina. Sobre este dominio, se implementa el modelo metodológico propuesto, incorporando dimensiones de calidad alineadas con estándares internacionales y evaluando su impacto en el desempeño de modelos predictivos.

Objetivos principales:

Integrar métricas formales de calidad de datos dentro del ciclo de vida de procesos de minería de datos, validando metodológicamente su impacto en el desempeño, estabilidad y confiabilidad de modelos predictivos aplicados al dominio sanitario.

Objetivos específicos:

- Definir y operacionalizar dimensiones de calidad de datos alineadas con estándares internacionales, adaptadas al contexto de datos abiertos en salud.
- Diseñar e implementar algoritmos de evaluación y validación dimensional sobre datasets derivados del sistema SINTRA.

- Integrar los resultados de la evaluación de calidad dentro del pipeline de minería de datos, estableciendo vínculos explícitos entre calidad y desempeño predictivo.
- Comparar el comportamiento de modelos predictivos antes y después del proceso de curación y validación de datos.
- Analizar la trazabilidad y reproducibilidad de los resultados obtenidos bajo un enfoque data-centric.

Metodología y Técnicas:

El proceso metodológico adoptado combina enfoques cuantitativos y experimentales, estructurados en fases secuenciales e iterativas: Perfilado y caracterización del dataset, identificando estructura, distribución de variables y potenciales anomalías; Evaluación dimensional de calidad, mediante métricas asociadas a completitud, consistencia, coherencia temporal, integridad y exactitud; Proceso de limpieza y transformación, orientado a corregir inconsistencias, tratar valores faltantes y normalizar estructuras; Integración en el pipeline de minería de datos, aplicando técnicas de aprendizaje automático para la construcción de modelos predictivos; Evaluación comparativa, analizando métricas de desempeño (precisión, estabilidad, validación cruzada y capacidad de generalización) antes y después de la intervención de calidad.

Este enfoque permite medir de manera objetiva el impacto de las mejoras en calidad sobre el comportamiento de los modelos.

Resultados obtenidos:

En esta etapa del proyecto se ha logrado la implementación operativa del modelo metodológico propuesto, permitiendo su aplicación sobre el caso de estudio mencionado. A partir de su análisis, se realizó una identificación sistemática de deficiencias estructurales en los datos analizados, tales como valores faltantes, inconsistencias semánticas y desalineaciones temporales. Asimismo, se desarrollaron métricas cuantificables para evaluar dimensiones críticas de calidad en el dominio sanitario, posibilitando una medición objetiva de los atributos del dataset. Los análisis comparativos realizados evidencian mejoras preliminares en la estabilidad y coherencia de los modelos predictivos tras la aplicación de procesos de curación y validación de datos, así como una mayor trazabilidad entre las decisiones de preprocesamiento y el comportamiento final del modelo.

Resultados esperados

En las siguientes etapas del proyecto se espera profundizar el análisis cuantitativo del impacto de cada dimensión de calidad sobre métricas específicas de desempeño predictivo, estableciendo relaciones formales entre niveles de calidad y variaciones en precisión, estabilidad y capacidad de generalización de los modelos. Asimismo, se prevé consolidar una guía metodológica replicable que permita integrar métricas de calidad de datos en proyectos de minería de datos en contextos gubernamentales y sanitarios, promoviendo prácticas alineadas con principios de gobernanza y reproducibilidad. Finalmente, se espera extender la validación del modelo a otros conjuntos de datos abiertos en salud pública orientados a diabetes, con el fin de evaluar su transferibilidad y robustez en distintos escenarios de aplicación.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo se encuentra conformado por una docente investigadora, Doctora en Ciencias Informáticas, quien dirige el proyecto; un docente auxiliar de la carrera Ingeniería en Sistemas Informáticos; y estudiantes de grado y posgrado que participan activamente en las distintas etapas metodológicas y experimentales del trabajo. En relación directa con la línea de investigación presentada, actualmente se desarrollan 2 tesis doctorales y 2 tesis de maestría vinculadas a la integración de métricas de calidad de datos y su impacto en modelos predictivos aplicados a datos abiertos.

Por otro lado, participan estudiantes colaboradores de la carrera de Ingeniería en Sistemas Informáticos de la UAI de distintas sedes, quienes intervienen en tareas de implementación de métricas, validación experimental y desarrollo de prototipos, fortaleciendo su formación en metodologías de investigación, gobernanza de datos y aplicaciones de aprendizaje automático en contextos gubernamentales.

La articulación entre investigación y formación de recursos humanos contribuye a consolidar una línea académica sostenida en el tiempo, promoviendo la generación de capacidades en calidad de datos, minería de datos y Ciencia de Datos aplicada al sector público.

5. BIBLIOGRAFÍA

[1] Martínez, R., et al. (2025). *Towards a scalable open data platform in public health: Evolving the INCUCAI case for organ donation*. En Actas de la 51ª Conferencia Latinoamericana de Informática (CLEI 2025). Valparaíso, Chile.

[2] Kumar, S., Datta, S., Singh, V., Singh, S. K., & Sharma, R. (2024). Opportunities and challenges in data-centric AI. *IEEE Access*, *12*, 33173-33189.

[3] Zha, D., Bhat, Z. P., Lai, K. H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2025). Data-centric artificial intelligence: A survey. *ACM Computing Surveys*, *57*(5), 1-42.

[4] Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Aroyo, L. M. (2021, May). "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

[5] Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., ... & Harmouch, H. (2022). The effects of data quality on machine learning performance. *arXiv preprint arXiv:2207.14529*.

[6] Gentyala, R. (2025). Ethical Artifacts: Engineering Verifiable Audit Trails for Human-in-the-Loop Decisions in ML Data Pipelines. *Journal of Scientific and Engineering Research*, *12*(10), 240-251.

[7] Ashmore, R., Calinescu, R., & Paterson, C. (2021). Assuring the machine learning lifecycle: Desiderata, methods, and challenges. *ACM computing surveys (CSUR)*, *54*(5), 1-39.

[8] ISO 25012 (2008). Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model. Disponible en: https://www.iso.org/standard/35736.html

[9] Polyzotis, N., Zinkevich, M., Roy, S., Breck, E., & Whang, S. (2019). Data validation for machine learning. *Proceedings of machine learning and systems*, *1*, 334-347.

[10] Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, *4*(9).

[11] Hosseinzadeh, E., Afkanpour, M., Momeni, M., & Tabesh, H. (2025). Data quality assessment in healthcare, dimensions, methods and tools: a systematic review. *BMC Medical Informatics and Decision Making*, *25*(1), 296.

[12] Mashoufi, M., Ayatollahi, H., Khorasani-Zavareh, D., & Boni, T. T. A. (2023). Data quality in health care: main concepts and assessment methodologies. *Methods of Information in Medicine*, *62*(01/02), 005-018.

[13] Martínez, R. et al. (2021). Metrics proposal to measure the quality of governmental datasets. *IEEE Latin America Transactions*, *20*(2), 301-308.

[14] INCUCAI - SINTRA (2026). "*SINTRA - El Sistema Nacional de Información de Procuración y Trasplante de la República Argentina*". Disponible en: https://sintra.incucai.gov.ar/

[15] Argentina.gob.ar - INCUCAI (2026). "*INCUCAI*". Disponible en: https://www.argentina.gob.ar/salud/incucai

Medición de inflación local con datos en línea y herramientas computacionales

E. Gutiérrez[1,2], J.Larrosa[1,2], G. Ramírez[3,4], J.I Uriarte[3,4], V. Diez[5], M. Seffino[5], F. Castagna[5], Pedro Fernández[5], J.F. Pérez Gerez[5], B. Castaño Rossi[5], Gino Pennacchioni[5], L. Fernández[2]

[1] Instituto de Ciencias e Ingeniería de la Computación (ICIC-UNS CONICET)
[2] Departamento de Economía. Universidad Nacional del Sur.
[3] Departamento de Derecho. Universidad Nacional del Sur
[4] Hyperia Big-Data
[5] Facultad de Ciencias Económicas, Universidad Nacional del Centro de la Provincia de Buenos Aires

CONTEXTO

La línea de I/D se inserta en el desarrollo y sostenimiento de un sistema de medición de precios recopilados mediante web scraping para la construcción de indicadores locales y para aplicaciones de ciencia de datos sobre paneles de alta frecuencia. El proyecto es coordinado institucionalmente por la Universidad Nacional del Sur (UNS), a través del Departamento de Economía, bajo la figura de Proyecto General de Investigación (PGI), donde se ejecutan las tareas de captura, procesamiento, control de calidad, cálculo y difusión pública de los indicadores.

RESUMEN

Esta línea de investigación se centra en construir una base de microprecios online para medir inflación local. En un contexto de alta inflación y cuestionamientos al dato oficial tras la intervención del INDEC en 2007, se implementó web scraping para relevar precios a alta frecuencia. En Bahía Blanca, el IPC Online se publica desde septiembre de 2014 y adapta la estructura del IPC nacional al contexto local. El pipeline trata faltantes y rotación de productos con canastas comparables entre períodos y automatiza la categorización con minería de texto; la cobertura online se complementa con relevamiento manual en rubros con baja presencia digital. Desde fines de 2023 se incorporó una interfaz de monitoreo para detectar fallas del scraping y apoyar recategorización. En 2015 se añadió el ICC Online para costos de construcción. La infraestructura habilita paneles semanales y aplicaciones de aprendizaje automático en alimentos y bebidas, farmacéuticos y canastas de alta frecuencia.

En 2024 se replicó este desarrollo en Tandil a fin de estudiar la evolución general de precios en una ciudad adicional, mediante un proyecto JOVIN.

Palabras clave: Ciencia de datos, inflación, Argentina, precios, web scraping.

1. INTRODUCCION

En las últimas dos décadas, Argentina ha sido una de las economías con mayor inflación del mundo, con episodios recientes en los que registró tasas anuales entre las más altas a nivel global.

Adicionalmente, la intervención del organismo del Instituto Nacional de Estadísticas y Censos (INDEC), en 2007 fue motivo de cuestionamiento respecto de la veracidad de datos y afectando a la credibilidad del dato de inflación informado por el organismo.

Estas motivaciones, generación de una base de datos de precios, relevados mediante web scraping permitiendo entonces obtener datos de alta frecuencia y que serían costoso de obtener mediante los métodos de recolección tradicionales.

1.1. IPC Online

En septiembre de 2014 fue oficialmente lanzado IPC Online (Índice de Precios Online), publicando de forma abierta información sobre inflación local para la ciudad de Bahía Blanca. El proyecto se gestó a partir del trabajo de dos emprendedores locales (Ing. Juan Ignacio Uriarte y Gonzalo Ramírez Muñoz de Toro), quienes desde fines de 2012 venían recolectando precios de comercios de la ciudad en un contexto de cuestionamiento a la credibilidad de estadísticas oficiales. En esa etapa inicial, la base de precios crecía sin un esquema metodológico completo de indexación (canasta, ponderaciones, reglas de empalme y tratamiento sistemático de faltantes). En 2014, tras buscar asesoramiento para poner en valor esa información, se incorporó un economista (Juan Manuel C. Larrosa) y el esfuerzo se orientó a construir un instrumento de medición del costo de vida con criterio estadístico.

La construcción de un índice de precios al consumidor requirió adaptar la arquitectura metodológica del INDEC al contexto local, tomando como referencia la estructura por capítulos/categorías del IPC nacional e introduciendo ajustes de pertinencia (por ejemplo, excluir servicios inexistentes a escala ciudad). Una vez lanzado, el IPC Online en septiembre 2014 evolucionó en frecuencia de recolección: comenzó con capturas mensuales y luego pasó a cuatro corridas al mes (con capacidad técnica de mayor frecuencia).

En términos operativos, el proyecto enfrentó dos desafíos centrales. Primero, el tratamiento de datos faltantes y la rotación de productos, fenómeno típico de fuentes web donde los ítems aparecen y desaparecen y las ofertas modifican la presencia en los listados. Para lidiar con esto se adoptó una regla pragmática: comparar canastas idénticas entre el período corriente y el anterior, trabajando con el subconjunto estrictamente comparable, aunque esto limita comparaciones directas en horizontes largos cuando la canasta cambia sustantivamente. Segundo, la categorización masiva de ítems y su asignación a la estructura del IPC: dado el volumen y la heterogeneidad de la información online, se automatizó la clasificación mediante algoritmos del tipo "bag-of-words" para mapear cada producto a la taxonomía del IPC. En la práctica, no todo lo recolectado se procesa: de alrededor de 200 mil precios mensuales se utiliza una fracción para el cálculo del índice. La cobertura de categorías del IPC lograda con fuentes web suele ubicarse entre 76% y 90%, incorporando típicamente miles de ítems y decenas de miles de precios por toma. Para completar capítulos con baja disponibilidad online (en particular servicios y algunos consumos no presentes en comercios online), se complementa con recolección tradicional de un conjunto acotado de ítems. Una mayor profundidad metodológica del indicador se encuentra publicada en [1] y [2].

A partir de finales del año 2023, se desarrolló tambien una interfaz web de uso para quienes participan en la estimación del indicador que permite monitorear en vivo la inflación como proceder a la carga de precios como su descarga de series de precios. Esta mejora ha permitido detectar con mayor anticipación dinámicas en la evolución de la inflación, prever errores asociados al algoritmo de web scraping y optimizar la clasificación de precios ya que adiciona un componente de categorización manual.

Actualmente, la base propia del IPC Online, dispone de 35 millones de precios, constituyendo entonces una fuente de datos con un profundo potencial a ser explotada. Se espera desarrollar en los próximos meses modelos de predicción basados en series temporales aplicando métodos econométricos y de aprendizaje automática.

El proyecto se ha enmarcado desde sus comienzos bajo la órbita institucional de la Universidad Nacional del Sur, concretamente con la figura de Proyecto General de Investigación (PGI), siendo declarado de interés por el Departamento de Economía de dicha casa de estudios.

Por otra parte, dado que el relevamiento tradicional del INDEC, de carácter nacional, demora 10 a 15 días en difundir los resultados, mientras que el IPC Online se publica el primer

día del mes, los resultados son seguidos tanto en su página web www.ipconlinebb.wordpress.com y su cuenta en X (@ipconlinebb) por la prensa tanto local como nacional.

1.2 ICC Online

Como continuidad del enfoque de medición online, el proyecto desarrolló el ICC Online (Índice de Costo de Construcción Online) para seguir la evolución del costo de la construcción a partir de fuentes digitales. La iniciativa comenzó a inicios de 2015 con la recolección automatizada de precios desde sitios web de ferreterías y comercios de materiales de construcción con presencia online en Bahía Blanca, y hacia mediados de 2015 se inició la publicación regular del índice [1].

En lo metodológico, el ICC Online se estructura siguiendo el esquema clásico de un índice de costos de construcción: separa componentes. En particular, se organiza en tres capítulos: materiales de construcción, salarios y costos laborales, y gastos generales. El potencial en este sentido y a diferencia del IPC Online reside en que este indicador permite relevar dinámicas propias de la construcción dentro de Bahía Blanca, mientras que el ICC relevado por el INDEC se ajusta específicamente a Gran Buenos Aires.

1.3 Estudios a partir de microprecios online

La infraestructura de captura y procesamiento de microprecios online permitió generar una base longitudinal lista para análisis, permitiendo realizar paneles semanales con trazabilidad y reproducibilidad, y por lo tanto habilitando estudios de micro-determinación de precios y promociones a alta frecuencia en múltiples rubros. En este sentido, sobresale la escasa literatura aplicada al caso argentino con datos de alta frecuencia.

Esto permitió el desarrollo aplicaciones empíricas aplicando métodos econométricos y de aprendizaje automático que explotan la dimensión temporal y transversal del panel (efectos fijos y aleatorios, modelos dinámicos, eventos temporales, promociones y aplicación de teoría de grafos), manteniendo como denominador común el uso de datos web semanales y su curado computacional. Concretamente, se han publicado hallazgos de yerba mate [3], cortes vacunos [4], té (promociones y liderazgo de precios) [5,6], galletitas [7], yogurt (atributos de marketing) [8], café [9], vino (dinámica de precios y promociones) [10,11], pastas secas [12] y cerveza [13].

Asimismo, se efectuaron aplicaciones fuera de los rubros alimenticios, concretamente para el sector farmacéutico.

La primera de ellas consistió en el estudio y diseño de modelo predictivos de aprendizaje automático que evalúen la probabilidad de que un medicamento ajuste su precio entre semanas usando variables observables y/o construidas en el pipeline [14].

Segundo, un módulo de interdependencia basado en redes: con variaciones semanales de precios se construyen redes de causalidad de Granger para mapear relaciones líder–seguidor entre medicamentos para enfermedades respiratorias, concretamente EPOC y Asma [15].

La disponibilidad de datos, también permite la construcción de canastas de alta frecuencia y estudiar mecanismos de traslado a precios del tipo de cambio. En esa línea, es posible reconstruir y valorizar semanalmente de la Canasta Básica Alimentaria (CBA), difundida mensualmente por INDEC, permitiendo evaluar el impacto del dólar y otros determinantes mediante modelos econométricos encontrando evidencia de influencia tanto de variaciones del tipo de cambio como del precio doméstico de combustibles en el corto y en el largo plazo. Además, la distinción entre tipo de cambio oficial y paralelo permite discutir asimetrías relevantes del *pass-through* en contextos inflacionarios [16].

1.4 Expandiendo la propuesta: El caso de Tandil.

Durante el año 2024, a partir del financiamiento de un proyecto de Jóvenes Investigadores (JOVIN), por parte de la Universidad Nacional del Centro de la

Provincia de Buenos Aires, surgió la posibilidad de replicar local con datos online, construyendo un IPC específico para la ciudad de Tandil.

El IPC Online Tandil, presentó sus primeros datos publicados para Diciembre 2024, constituyendo entonces una herramienta adicional para relevar variaciones de precios en un contexto de inflación local. Dispone de su sitio web (https://ipctonline.blogspot.com/), como su cuenta en la red social X @IPCT_Online, donde se difunde a comienzo de mes la evolución general del nivel de precios para la localidad. Actualmente, la cantidad de datos para esta ciudad es de aproximadamente supera los 1.2 millones de precios recopilados.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

a. Estimación del Índice de Precios al Consumidor tanto para la localidades de Bahía Blanca y Tandil

b. Estudio de dinámicas de determinantes evolución propias de precios de bienes y servicios.

c. Modelos aplicados de serie temporales para la aplicación de predicción de estimaciones de inflación

3. RESULTADOS OBTENIDOS/ESPERADOS

a. Mejorar categorización, mediante la implementación de un algoritmo de clasificación de texto con perfomance superior a bag-of-word

b. Desarrollar un módulo simple de predicción/nowcasting de inflación (general y por capítulos) mediante herramientas propias de aprendizaje automático

c. Extender canastas de alta frecuencia y reportes automáticos por rubro como productos derivados del sistema.

4. FORMACIÓN DE RECURSOS HUMANOS

En la actualidad, el grupo está integrado por 12 personas. De los mismos, 2 cuentan con doctorado completo y otros 2 se encuentran realizando el mismo. 4 estudiantes de grado, avanzados se encuentran participando activamente

Se destaca una participación multidisciplinaria donde confluyen integrantes con formación en ciencias de computación, economía, derecho y turismo.

A largo del proyecto, se han realizados dos tesis de grado y una tesis doctoral cuyos datos se han obtenido exclusivamente a partir de las bases generadas por el grupo de investigación. Se espera en los próximos meses concluir con una tesis de grado adicional.

La producción del grupo con más de 20 publicaciones indexadas, 2 libros y 5 documentos de trabajo, permite un importante acercamiento a los estudiantes de grado a la academia y una experiencia aplicado en el ámbito de la consultoría vinculado a la ciencia de datos.

5. BIBLIOGRAFÍA

[1] Uriarte JI, Ramírez Muñoz de Toro GR, Larrosa JM. Web scraping based online consumer price index: The "IPC Online" case. Journal of Economic and Social Measurement. 2019 May;44(2-3):141-59.

[2] Ramírez Muñoz de Toro GR, Uriarte JI, Larrosa J, Gutiérrez E. Measuring inflation in Argentina: IPC Online Project. InV Simposio Argentino de Ciencia de Datos y GRANdes DAtos (AGRANDA 2019)-JAIIO 48 (Salta) 2019.

[3] Gutiérrez E, Tedesco L, Larrosa JMC, Ramírez Muñoz de Toro G, Uriarte JI. Determinants of Weekly Price Variations of Yerba Mate in Argentina. RIVAR (Santiago). 2023;10(28):160-180.

[4]Finucci M, Giordano V, Glave U, Delbianco F. Determinantes de la variación de precio semanal de cortes de carne vacunos en una región de Argentina. Revista Nicolaita

de Estudios Económicos. 2022;16(1):51-67. doi:10.33110/rnee.v16i1.321.

[5]Larrosa JMC, Gutiérrez EM, Uriarte JI, Ramírez Muñoz de Toro GR. Granger causality networks of price leadership in the retail tea market of Argentina. J Revenue Pricing Manag. 2025;24(1):50-59.

[6]Meller LM, Larrosa JM, Delbianco F, Ramírez Muñoz de Toro G, Uriarte JI. Inflación semanal en galletitas: un enfoque de datos de panel. Rev Métodos Cuantitativos Econ Empresa. 2021;31:417-440.

[7]Giordano V, Larrosa JMC, Uriarte JI, Ramírez Muñoz de Toro GR. Pricing and Promotions in Retail Tea Market in Argentina. J Int Food Agribus Mark. 2024;36(2):298-315.

[8] Larrosa JMC, Giordano V, Ramírez Muñoz de Toro GR, Uriarte JI. Marketing attributes in yogurt weekly pricing in Argentina. J Revenue Pricing Manag. 2022;21(3):332-343.

[9] Larrosa JMC, Meller L, Uriarte JI, Ramírez Muñoz de Toro GR. Retail coffee pricing dynamics in Argentina. SN Bus Econ. 2023;3:155.

[10] Larrosa JMC, Gutiérrez EM, Ramírez Muñoz de Toro GR, Uriarte JI. A dynamic analysis of wine pricing in Argentina. Int J Wine Bus Res. 2023;35(2):189-211. doi:10.1108/IJWBR-09-2021-0052.

[11] Larrosa JMC, Gutierrez EM, Uriarte JI, Ramírez Muñoz de Toro GR. An In-Depth Analysis of Wine Price Promotions in Argentina. J Glob Bus Trade. 2025;21(1):17-43.

[12]Larrosa JMC, Gutiérrez EM, Uriarte JI, Ramírez Muñoz de Toro GR. Dry pasta in Argentina: drivers of pricing and promotions. Stud Econ Econom. 2025;49:1-17. doi:10.1080/03796205.2025.2582093.

[13] Larrosa JMC, Fernández LJ, Allerborn J, Gutierrez EM, Uriarte JI, Ramírez Muñoz de Toro GR. Drivers of beer pricing in Argentina. Int J Food Syst Dyn. 2024;15(5):540-556.

[14] Gutiérrez E, Virdis JM, Meller L, Domínguez DL. Estrategias de precios e incertidumbre económica: Un estudio de caso para la industria farmacéutica argentina usando aprendizaje automatizado. Revista de Métodos Cuantitativos para la Economía y la Empresa. 2024;38:1-16..

[15] Ramos G, Gutiérrez EM, Durán F. Interdependent pricing in the Argentine pharmaceutical market: COPD and asthma. In: Larrosa JMC, comp. Pricing Micro-Determinants in Retail Market of Bahía Blanca. Bahía Blanca: EDIUNS; 2025. p. 135-152. ISBN 978-987-655-387-2.

[16] Gutiérrez E, Delbianco F, Larrosa JMC. Determinantes de la canasta básica alimentaria en Argentina: un estudio econométrico. El Trimestre Económico. 2025

Metodología de Análisis y Mejora de la Calidad de Datos en Sistemas de Información

Mg. Rodolfo Bertone [1] C.C Marcelo Taruschio [2] Juan Ignacio Torres[3]
Ezequiel Iván Reale[4] Gonzalo Francisco Lara[5] Mariana Eduardo[6] Ezequiel Melía[7] Valentino Castellano[8]

Facultad de Ciencias Exactas e Ingeniería. Universidad Católica de La Plata.
[1] rodolfo.bertone@ucalp.edu.ar [2] marcelo.taruschio@ucalp.edu.ar
[3] juanignacio.torres@ucalpvirtual.edu.ar [4] ezequielreale33@gmail.com
[5] gflara6@gmail.com [6]marianaeegea@gmail.com [7]ezequielmelia@hotmail.com
[8]valentinocastellano@hotmail.com

Resumen

El presente proyecto propone el diseño, implementación y validación de una metodología integral orientada al análisis, diagnóstico y mejora sistemática de la calidad de datos en sistemas de información. La propuesta se fundamenta en el modelo de calidad de datos establecido por la norma ISO/IEC 25012 [12], integrando dimensiones críticas como exactitud, completitud, consistencia y trazabilidad. Asimismo, se incorporan enfoques recientes sobre gestión de calidad en entornos Big Data y arquitecturas heterogéneas, considerando los desafíos asociados a volumen, variedad y complejidad de los datos [2], [3], [10].

La investigación contempla la definición de métricas cuantitativas y cualitativas, la aplicación de técnicas automatizadas de data profiling [7], el uso de herramientas de monitoreo y medición continua [4], y la implementación de procedimientos sistemáticos de verificación y validación. El modelo metodológico será evaluado mediante estudios de caso en bases de datos relacionales y no relacionales, permitiendo medir su impacto en la confiabilidad de la información, la reducción de inconsistencias y el fortalecimiento de la toma de decisiones basada en evidencia.

Finalmente, la propuesta pretende ofrecer un marco metodológico adaptable a distintos contextos organizacionales y niveles de madurez en gobernanza de datos, alineado con modelos de gestión y certificación reconocidos [8], [11]. Su aplicación resulta especialmente pertinente en entornos latinoamericanos que presentan limitaciones estructurales y bajo grado de formalización en procesos de aseguramiento de calidad.

Palabras clave: Calidad de datos; Bases de datos; Data profiling; Gobernanza de datos; Ingeniería de software.

Contexto e Introducción

La calidad de los datos constituye un factor crítico para garantizar la confiabilidad de los sistemas de información y sustentar procesos analíticos robustos. La presencia de registros incompletos, inconsistentes, redundantes o desactualizados impacta directamente en la validez de los resultados, afectando modelos predictivos, reportes gerenciales y

procesos de inteligencia de negocios. Diversos estudios han demostrado que una calidad deficiente incrementa costos operativos, deteriora la interoperabilidad entre sistemas y reduce la eficiencia organizacional [5]. Desde una perspectiva técnica, la ausencia de controles sistemáticos de integridad, validación semántica y reglas de negocio formalizadas profundiza estos riesgos.

En investigaciones internacionales, se han identificado limitaciones estructurales en prácticas de gobernanza y aseguramiento de calidad de datos, vinculadas tanto a restricciones presupuestarias como a baja formalización de procesos. Investigaciones en dominios como salud pública evidencian deficiencias significativas en dimensiones de completitud, consistencia y actualización, detectadas mediante métricas estandarizadas y herramientas de análisis exploratorio [6], [9]. Estos antecedentes resaltan la necesidad de metodologías sistemáticas que contemplen no solo la evaluación técnica, sino también la madurez organizacional y la trazabilidad de los datos.

A nivel internacional, el estándar ISO/IEC 25012 [12] define un modelo conceptual que distingue características intrínsecas y dependientes del sistema, proporcionando una base formal para la medición de calidad. En paralelo, enfoques recientes promueven la incorporación de requisitos de calidad desde las etapas iniciales del desarrollo de software, bajo el paradigma de "data quality by design" [1]. Sin embargo, persiste una brecha entre estos marcos teóricos y su aplicación efectiva en arquitecturas distribuidas, bases NoSQL y entornos Big Data, donde la heterogeneidad y el volumen de información complejizan la implementación de controles consistentes [3], [10].

En este contexto, el presente proyecto se propone diseñar y validar una metodología integral que articule métricas formales, herramientas automatizadas de data profiling y monitoreo, y lineamientos de integración en el ciclo de vida del software. El objetivo es consolidar un enfoque sistemático, reproducible y adaptable que permita institucionalizar el aseguramiento de calidad de datos como componente estructural de los sistemas de información.

Objetivos de la Investigación

Objetivo General

Desarrollar, implementar y validar un modelo integral de análisis y aseguramiento de la calidad de datos en sistemas de información que permita evaluar de manera sistemática dimensiones como exactitud, completitud, consistencia, unicidad y trazabilidad. El modelo integrará métricas formales, indicadores cuantificables y procedimientos de control automatizados, aplicables tanto a bases de datos relacionales como a entornos NoSQL y arquitecturas distribuidas. Asimismo, contemplará mecanismos de monitoreo continuo y retroalimentación que permitan incorporar la mejora de la calidad de datos dentro del ciclo de vida del software y de los procesos organizacionales.

Objetivos Específicos

Definir y estandarizar métricas técnicas alineadas con el modelo ISO/IEC 25012 [12] y con estudios comparativos de marcos de calidad de datos [5],

estableciendo fórmulas de medición, umbrales de aceptación y criterios de evaluación reproducibles.

Diseñar procedimientos automatizados de diagnóstico basados en técnicas de data profiling, análisis estadístico descriptivo, validación de reglas de integridad y monitoreo continuo de indicadores [4], [7], incorporando scripts y consultas parametrizables que permitan su aplicación en distintos motores de bases de datos.

Proponer estrategias de mejora, corrección y certificación de datos compatibles con procesos de ingeniería de software y enfoques de calidad por diseño [1], así como con esquemas de certificación y evaluación organizacional [11], integrando controles en fases de análisis, desarrollo y mantenimiento.

Evaluar empíricamente la efectividad del modelo mediante estudios de caso comparativos, utilizando métricas de desempeño antes y después de la aplicación de la metodología, con análisis cuantitativo de reducción de inconsistencias y mejora en niveles de completitud y coherencia.

Elaborar una guía de buenas prácticas estructurada según distintos niveles de madurez en gestión de datos [8], incluyendo recomendaciones técnicas, organizacionales y de gobernanza que faciliten su adopción progresiva en instituciones con diferentes capacidades tecnológicas.

Desarrollo e Innovación

El proyecto se fundamenta en principios formales de aseguramiento de calidad de datos y análisis exploratorio, incorporando técnicas automatizadas para la detección de anomalías estadísticas, valores atípicos, registros incompletos y duplicidades mediante procesos de data profiling y validación de reglas de integridad [7]. Se integrarán herramientas de monitoreo continuo y evaluación comparativa de calidad [4], complementadas con lineamientos de gestión específicos para entornos de datos masivos y arquitecturas distribuidas [3].

Asimismo, se analizarán enfoques recientes aplicados a Big Data e IoT, donde la heterogeneidad estructural, la variabilidad semántica y el alto volumen de información requieren métricas escalables y adaptables a distintos modelos de almacenamiento [2], [10]. La metodología incorporará criterios vinculados a niveles de madurez organizacional y gobernanza de datos [8], promoviendo un esquema progresivo de adopción y mejora continua.

La principal innovación reside en la integración sistemática de métricas formales, automatización de procesos de evaluación y lineamientos de implementación dentro del ciclo de vida del software, consolidando el enfoque de calidad de datos incorporada desde etapas tempranas del diseño [1]. Este planteo reduce la brecha entre marcos conceptuales y su aplicación operativa en sistemas reales.

Plan de Trabajo

Fase I – Revisión y Diagnóstico (Meses 1–2):
Se realizará una revisión sistemática de literatura científica especializada [2], [3], [5] bajo criterios PRISMA y un relevamiento comparativo de herramientas de data profiling y

monitoreo [4], [7], identificando KPIs asociados a dimensiones como completitud, consistencia, exactitud, unicidad y actualidad, junto con sus métodos de medición cuantitativa.

Fase II – Diseño Metodológico (Meses 3–4):
Se formalizarán métricas, umbrales de aceptación y funciones de evaluación alineadas con ISO/IEC 25012 [12], definiendo modelos de puntuación compuesta, matrices de trazabilidad y correspondencia con esquemas de certificación y aseguramiento de calidad [11].

Fase III – Desarrollo de Herramientas (Meses 5–8):
Se implementarán módulos en Python y SQL orientados a análisis automatizado, validación de reglas de integridad, cálculo de métricas, generación de reportes estructurados y monitoreo continuo mediante procesos programados y control de versiones.

Fase IV – Evaluación Empírica (Meses 9–11):
El modelo será aplicado sobre bases de datos relacionales y NoSQL piloto, ejecutando análisis comparativos pre y post intervención, con validación estadística de resultados y medición objetiva del impacto sobre indicadores de calidad.

Fase V – Documentación y Difusión (Mes 12):
Se consolidará una guía metodológica técnica con especificaciones replicables, acompañada de publicaciones académicas y reportes técnicos orientados a transferencia y adopción institucional.

Resultados Esperados

Se espera obtener un modelo metodológico formalmente validado para la evaluación y mejora sistemática de la calidad de datos, junto con un conjunto de métricas estandarizadas y operacionalizadas conforme a ISO/IEC 25012 [12]. Asimismo, se desarrollarán herramientas automatizadas de análisis y monitoreo continuo [4], [7], respaldadas por evidencia empírica que demuestre reducción cuantificable de inconsistencias y aumento de confiabilidad. El proyecto aportará lineamientos técnicos de implementación progresiva según niveles de madurez organizacional [8], además de producción científica y transferencia tecnológica.

Formación de Recursos Humanos

El proyecto será ejecutado por un equipo interdisciplinario integrado por dirección académica, docentes investigadores y estudiantes avanzados, bajo un esquema de trabajo colaborativo con asignación formal de roles técnicos. La participación estudiantil en tareas de experimentación, instrumentación de métricas y validación empírica fortalecerá competencias en ingeniería de datos, aseguramiento de calidad y análisis estadístico, promoviendo formación en investigación aplicada y desarrollo tecnológico.

Conclusiones

La calidad de datos constituye un componente estructural en la confiabilidad, interoperabilidad y sostenibilidad de los sistemas de información contemporáneos. La propuesta integra marcos normativos, métricas formales y automatización operativa en un modelo metodológico coherente y empíricamente validado,

reduciendo la brecha entre teoría y práctica.

Se prevé que la metodología fortalezca la gobernanza de datos, mejore la trazabilidad y auditabilidad de la información y optimice procesos de ingeniería de software tanto en entornos relacionales como NoSQL. Su arquitectura conceptual y técnica permitirá escalar el modelo hacia ecosistemas Big Data y sistemas basados en inteligencia artificial, contribuyendo a la calidad de datos como prerrequisito para analítica avanzada y toma de decisiones basada en evidencia [3], [10].

Referencias

[1] C. Guerra-García et al., "ISO/IEC 25012-based methodology for managing data quality requirements in the development of information systems: Towards data quality by design", Data & Knowledge Engineering, vol. 145, p. 102152, 2023. DOI: 10.1016/j.datak.2023.102152

[2] A. Ridzuan et al., "A Review on Data Quality Dimensions for Big Data", Procedia Computer Science, vol. 234, pp. 341–348, 2024. DOI: 10.1016/j.procs.2024.03.008

[3] T. Nguyen et al., "Data quality management in big data: Strategies, tools, and educational implications", Journal of Parallel and Distributed Computing, 200, Article 105067, 2025. DOI: 10.1016/j.jpdc.2025.105067

[4] L. Ehrlinger et al., "A Survey of Data Quality Measurement and Monitoring Tools", Frontiers in Big Data, vol. 5, p. 850611, 2022. DOI: 10.3389/fdata.2022.850611

[5] R. Miller et al., "A Comparison of Data Quality Frameworks: A Review", Big Data and Cognitive Computing, vol. 9, no. 4, p. 93, 2025. DOI: 10.3390/bdcc9040093

[6] S. Cho et al., "A New Strategy for Evaluating the Quality of Laboratory Results for Big Data Research: Using External Quality Assessment Survey Data (2010-2020)", Annals of Laboratory Medicine, vol. 43, no. 5, pp. 425–433, 2023. DOI: 10.3343/alm.2023.43.5.425

[7] N. Shivaprasad et al., "Enhancing Data Quality through Automated Data Profiling", International Journal for Research Publication and Seminar, vol. 15, no. 4, pp. 108–117, 2024. DOI: 10.36676/jrps.v15.i4.17

[8] L. Pörtner et al., "Data Management Maturity Model—Process Dimensions and Capabilities to Leverage Data-Driven Organizations Towards Industry 5.0", Applied System Innovation, vol. 8, no. 2, p. 41, 2025. DOI: 10.3390/asi8020041

[9] B. Gordon et al., "Evaluation of freely available data profiling tools for health data research application: a functional evaluation review", BMJ Open, 12(5):e054186, 2022. DOI: 10.1136/bmjopen-2021-054186

[10] I. Taleb, N. Dahmani, K. Hayawi, and S. S. Mathew, "Big Data Quality Assessment in the IoT Era", en M. A. Serhani, Y. Xu, y Z. Maamar (Eds.), Empowering IoT with Big Data Analytics, Academic Press/Elsevier, 2025. DOI: 10.1016/B978-0-443-21640-4.00003-X

[11] F. Gualo et al., "Data quality certification using ISO/IEC 25012: Industrial experiences", Journal of Systems and Software, vol. 176, p. 110938, 2021. DOI: 10.1016/j.jss.2021.110938

[12] ISO/IEC 25012:2008, "Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model", International Organization for Standardization, Geneva, Switzerland, 2008.

MODELOS GENERATIVOS INTERPRETABLES PARA DISEÑO DE NOVO DE MOLÉCULAS

Trinidad Crozes[1,2], Vir Sabando[1], Axel J. Soto[1,2], and Ignacio Ponzoni[1,2]

[1]Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, Argentina
[2]Instituto de Ciencias e Ingeniería de la Computación, CONICET–UNS, Argentina
{trinidad.crozes,vir.sabando,axel.soto,ip}@cs.uns.edu.ar

RESUMEN

La línea de investigación propone desarrollar y evaluar estrategias avanzadas de modelado molecular basadas en algoritmos generativos para el diseño de fármacos. En primer lugar, se realizará una revisión sistemática de modelos generativos aplicados al diseño molecular, analizando su representación química, capacidad generativa y posibilidades de integración de múltiples fuentes de datos. Luego, se compararán experimentalmente distintos enfoques y se implementarán mejoras, incluyendo la incorporación de nuevas métricas de evaluación. En paralelo, se estudiarán estrategias de modelado de relaciones cuantitativas estructura-actividad (*Quantitative Structure-Activity Relationship, QSAR*) para validar las moléculas generadas, comparando modelos de consenso basados en redes neuronales profundas individuales frente a modelos multitarea, considerando propiedades ADMET, actividad biológica y otras características relevantes. Los modelos generativos y predictivos más prometedores se integrarán mediante aprendizaje por refuerzo para optimizar simultáneamente múltiples propiedades. Asimismo, se incorporarán estrategias de explicabilidad (XAI) y analítica visual para facilitar la interpretación de resultados y la integración del conocimiento experto. Finalmente, se desarrollará una solución quimioinformática integral orientada al diseño y validación de quimiotecas virtuales, evaluando su desempeño mediante cribado virtual en proyectos de descubrimiento de fármacos.

Palabras clave: quimioinformática, modelos generativos, analítica visual.

CONTEXTO

La presente línea de investigación se enmarca en el proyecto de investigación “Aprendizaje automático y analítica visual para el uso de modelos generativos y predictivos aplicados a quimioinformática y el procesamiento de lenguaje natural”, dirigido por los Drs. Axel J. Soto e Ignacio Ponzoni. El proyecto es coordinado y financiado por la Universidad del Sur, con la contribución del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) quien financia la beca doctoral de la Lic. Crozes.

1 INTRODUCCIÓN

En el proceso de I+D de un fármaco hay que diferenciar cuatro fases: descubrimiento, preclínica, clínica y, finalmente, aprobación y registro. La fase de descubrimiento de nuevos fármacos constituye un proceso largo y complejo, con una tasa de éxito muy baja. Aproximadamente, se estima que solo 250 de cada 10.000 moléculas de la etapa de investigación básica pasan a la fase de investigación preclínica. Las técnicas computacionales de modelado molecular más representativas para conducir la fase de descubrimiento incluyen al

modelado *QSAR*, el cribado virtual de compuestos y el diseño *de novo*. La técnica *QSAR* define modelos matemáticos para predecir la relación entre la estructura química de un compuesto y su actividad biológica[12]; el cribado virtual identifica los candidatos más prometedores de una base de datos de compuestos o una biblioteca virtual[3]; y el diseño *de novo* genera desde cero nuevas moléculas activas en las líneas de desarrollo de fármacos[6].

Una forma de abordar la implementación de estas técnicas de modelado molecular es a través de algoritmos de aprendizaje automático[7]. Gracias al creciente avance de los recursos computacionales disponibles, los algoritmos basados en aprendizaje profundo (*deep learning, DL*) han logrado avances notables en quimioinformática[8]. En particular, estos algoritmos de *DL* han sido capaces de aprender modelos *QSAR* de alto desempeño[12], contribuir al desarrollo de estrategias de cribado virtual[4], y definir modelos generativos para el diseño *de novo* y optimización de moléculas[10]. En relación a esto último, las moléculas se suelen representar mediante cadenas simplificadas de caracteres, las cuales codifican la estructura química bidimensional de un compuesto mediante la notación SMILES[11]. Sobre la base de dicha representación, se han estudiado modelos de generación de moléculas que buscan aprender la distribución de probabilidad de los caracteres mediante una gran variedad de arquitecturas neuronales entrenadas con grandes quimiotecas de estructuras químicas[9]. Estos modelos buscan generar pequeñas moléculas similares a las presentes en el conjunto de entrenamiento, las cuales exhiban características estructurales consistentes con el perfil deseado para el nuevo fármaco. Por ende, se espera que los compuestos generados mediante estos métodos posean una estructura química que les permitan ser buenos candidatos.

De este modo, los algoritmos de DL pueden ofrecernos soluciones de calidad para todos los problemas de modelado molecular antes mencionados[1]. No obstante, existen limitaciones en los modelos de DL a la hora de su adopción por expertos en química medicinal, siendo la falta de explicabilidad de las decisiones tomadas por estos modelos una de sus principales críticas, lo cual vuelve necesario incorporar estrategias de XAI para afrontar este desafío[2]. Esta novedosa y creciente área de la inteligencia artificial tiene por objetivo presentar el aprendizaje maquinal como un proceso comprensible para el ser humano, tratando de construir explicaciones sobre cómo estos modelos alcanzan un determinado resultado o predicción [5], y están empezando a ser utilizados en el ámbito del descubrimiento de fármacos[2]. En el marco de este plan de beca nos centraremos en el desarrollo de metodologías para diseño *de novo* de moléculas mediante modelos generativos utilizando redes neuronales profundas y estrategias de XAI que permitan obtener modelos robustos, confiables e interpretables para los expertos en el dominio farmacológico.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Nuestra propuesta de investigación busca determinar qué tipo de algoritmo generativo es más adecuado para el diseño de nuestros enfoques de modelado molecular y, además, si resultaría deseable extender algunas de las metodologías revisadas o diseñar un enfoque generativo completamente nuevo. Uno de los aspectos que se desea evaluar es el impacto de la representación molecular elegida por cada técnica. Además, se analizará la viabilidad de incorporar otras fuentes de datos, además de la estructura química de los compuestos, al proceso de entrenamiento de los modelos generativos. Las actividades relacionadas con este objetivo son:

Actividad 1.a: Efectuar una revisión de la literatura referida al uso de modelos generativos en diseño de fármacos, con el objetivo de organizar taxonómicamente las distintas familias de métodos, junto con sus ventajas, limitaciones y nichos de uso reportadas por los autores. Como parte de esta revisión, se buscará identificar las diferentes estrategias empleadas para guiar a los modelos generativos en términos de los objetivos de diseño.

Actividad 1.b: Como resultado de la actividad anterior, tomar modelos disponibles, o reproducir el entrenamiento, de los principales algoritmos generativos de moléculas y compararlos para diversos escenarios a fin de establecer experimentalmente sus ventajas y limitaciones, tratando de

identificar sus nichos de uso. Analizar y comparar los algoritmos teniendo en cuenta su capacidad generativa y la posibilidad de integrar diversas fuentes de información química y física al proceso de entrenamiento.

Actividad 2: Evaluar e implementar para los algoritmos generativos más prometedores posibles mejoras identificadas en el proceso de revisión. Incorporar nuevas fuentes de datos. Analizar el compendio de métricas existentes aplicables a los compuestos generados por modelos generativos y evaluar la necesidad de proponer métricas adaptadas o novedosas. Comparar el desempeño de estas nuevas versiones respecto de las originales.

Otra de nuestras preguntas de investigación busca determinar cuál estrategia de modelado *QSAR* permite una mayor efectividad al integrar la evaluación de desempeño de varias propiedades fisicoquímicas y de actividad biológica como forma de validación del modelo generativo. En particular, se desea comparar el desempeño de modelos *QSAR* de consenso basados en redes neuronales profundas entrenadas de forma individual para cada propiedad respecto de modelos *QSAR* basados en redes neuronales profundas multitarea.

Actividad 3: Identificar propiedades *ADMET*, de actividad biológica, y otras propiedades de interés (como por ejemplo la clasificación Anatómica Terapéutica Química (ATC) de una droga) para construir casos de estudio a fin de evaluar las estrategías de modelado generativo molecular propuestas. Las propiedades *ADMET* serán las mismas para todos los casos de estudio, mientras que las actividades biológicas y afinidad con proteínas dependerán de cada tipo de tratamiento farmacológico estudiado. Para las propiedades elegidas, se recopìlará y procesará los datos requeridos para entrenar sus modelos predictivos.

Actividad 4: Diseñar, implementar y comparar el desempeño de los modelos *QSAR* de consenso basados en redes neuronales profundas entrenados de forma individual para cada propiedad respecto de modelos *QSAR* basados en redes neuronales profundas multitarea, para cada caso de estudio definido en la Actividad 3. De manera similar, se entrenarán modelos para predecir las demás propiedades moleculares.

Actividad 5: Integrar en una metodología basada en aprendizaje por refuerzo a los modelos generativos más prometedores (ver Actividades 1 y 2) con los modelos *QSAR*, y otros modelos predictivos de caracterización de moléculas, más prometedores (ver Actividades 3 y 4) y evaluar el desempeño de la/s estrategia/s por refuerzo resultante/s frente a los modelos generativos del estado del arte. Seleccionar la metodología generativa que finalmente quedará incorporada a nuestra propuesta de modelado por refuerzo.

Asimismo, otra de nuestras preguntas de investigación se orienta a evaluar cómo la incorporación de una estrategia de *XAI* puede mejorar la interpretabilidad de los resultados de nuestra propuesta de modelado por refuerzo.

Actividad 6: Adaptar y/o diseñar una estrategia de *XAI* que pueda facilitar la interpretación de los resultados obtenidos por el proceso de generación y optimización de moléculas por parte de expertos en el dominio químico.

Nuestra siguiente pregunta de investigación se orienta a evaluar cómo la incorporación de una estrategia de analítica visual a nuestra propuesta puede mejorar la incorporación del conocimiento experto y brindar apoyo a la estrategia para interpretar los resultados obtenidos por el proceso de modelado molecular.

Actividad 7: Diseñar una estrategia de analítica visual que pueda facilitar la interpretación de los resultados obtenidos por el proceso de generación y optimización de moléculas por parte de expertos en dominio químico.

Finalmente, se busca investigar cómo validar el desempeño de nuestra propuesta en el diseño de quimiotecas virtuales para proyectos de descubrimiento de fármacos.

Actividad 8: Integrar las técnicas de modelado generativo, interpretación y visualización resultantes de los objetivos específicos anteriores en una solución quimioinformática final.

Actividad 9: Diseñar una estrategia de analítica visual que facilite el análisis e interpretación de resultados en una etapa de cribado virtual de compuestos. Este cribado se realizará sobre las quimiotecas virtuales creadas mediante nuestra propuesta, aprovechando los modelos *QSAR* generados (ver Actividades 3 y 4), y permitirá evaluar la calidad del modelo generativo en función de las

tasas de éxito alcanzadas durante el cribado.

3 RESULTADOS OBTENIDOS/ESPERADOS

El objetivo general de esta línea de investigación es el desarrollo de nuevas tecnologías basadas en aprendizaje profundo que permitan generar y optimizar compuestos químicos con las propiedades físicas, químicas y/o de bioactividad deseadas, usando modelos generativos guiados por objetivos de diseño. Una de las estrategias que evaluaremos consiste en combinar aprendizaje por refuerzo con el uso de modelos generativos con modelos discriminantes, estos últimos basados en *QSAR*, que orienten la generación *de novo* hacia regiones del espacio químico donde las moléculas posean los valores de propiedad y actividad biológica buscados para un proyecto específico de desarrollo de fármacos. Otro aspecto central para el desarrollo de nuestras técnicas es el uso de XAI [2], la cual permita que el experto en farmacología pueda interpretar y comprender las decisiones tomadas por el modelo computacional. Esto es clave para la adopción de estos modelos generativos en el ámbito de la química médica.

Explicitada la meta general de la presente línea de investigación, se enuncian los **resultados logrados y aquellos esperados**:

R1. Se realizó una revisión exhaustiva de las técnicas propuestas para generación de moléculas basadas en inteligencia artificial, buscando contrastarlas en términos de sus fortalezas y debilidades en sus desempeños, clasificándolas además según los tipos de datos y representaciones moleculares que éstas emplean. El objetivo fue determinar cuáles incorporar a nuestros enfoques de modelado molecular o, inclusive, si es necesario diseñar un nuevo método generativo. Sobre esta revisión buscaremos publicar un artículo de tipo *survey* en los próximos meses.

R2. Se planifica diseñar modelos *QSAR*, junto con otros tipos de modelos predictivos, para evaluar moléculas en el contexto de propiedades fisicoquímicas y de actividad biológica de interés para un proyecto de diseño de fármacos específico. Integrar en una metodología basada en aprendizaje por refuerzo a los modelos generativos y *QSAR* más prometedores, y evaluar sus desempeños contra los modelos generativos del estado del arte.

R3. Se planifica diseñar una estrategia de inteligencia artificial explicable que permita evaluar e interpretar los resultados obtenidos por los modelos desarrollados para cumplir los objetivos específicos anteriores.

R4. Se planifica diseñar una estrategia de analítica visual interactiva que brinde soporte a las estrategias de inteligencia artificial interpretable propuestas para cumplir el objetivo específico anterior.

R5. Se planifica integrar las técnicas de modelado generativo, interpretación y visualización resultantes de los objetivos específicos anteriores en una herramienta computacional que facilite a expertos en química medicinal hacer diseño *de novo* de compuestos. Validar la nueva herramienta con especialistas del dominio químico.

4 FORMACIÓN DE RECURSOS HUMANOS

El doctorado de la Lic. Trinidad Crozes se desarrolla en el marco de la presente línea de investigación, radicada en el grupo BioChemTICs del Instituto de Ciencias e Ingeniería de la Computación (CONICET-UNS). La temática central de esta línea aborda el diseño de metodologías de aprendizaje automático orientadas a la inferencia y análisis de redes jerárquicas a partir de datos de expresión génica, en el ámbito de la bioinformática. A su vez esta línea de investigación dio lugar a tres tesis doctorales previas. El trabajo se inscribe en la disciplina Informática y Comunicaciones, con aplicación en ciencias exactas y naturales.

REFERENCIAS

[1] Anighoro, A. Deep learning in structure-based drug design. *Artificial Intelligence in Drug Design* (2022), 261–271.

[2] Jiménez-Luna, J., Grisoni, F., and Schneider, G. Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence 2*, 10 (2020), 573–584.

[3] Lin, X., Li, X., and Lin, X. A review on applications of computational methods in drug screening and design. *Molecules 25*, 6 (2020), 1375.

[4] Maia, E. H. B., Assis, L. C., De Oliveira, T. A., Da Silva, A. M., and Taranto, A. G. Structure-based virtual screening: from classical to artificial intelligence. *Frontiers in chemistry 8* (2020), 343.

[5] Minh, D., Wang, H. X., Li, Y. F., and Nguyen, T. N. Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review* (2022), 1–66.

[6] Mouchlis, V. D., Afantitis, A., Serra, A., Fratello, M., Papadiamantis, A. G., Aidinis, V., Lynch, I., Greco, D., and Melagraki, G. Advances in de novo drug design: from conventional to machine learning methods. *International journal of molecular sciences 22*, 4 (2021), 1676.

[7] Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow Jr, R. A., Fisher, J., Jansen, J. M., Duca, J. S., Rush, T. S., et al. Rethinking drug design in the artificial intelligence era. *Nature Reviews Drug Discovery 19*, 5 (2020), 353–364.

[8] Tripathi, M. K., Shrivastava, S. K., Karthikeyan, S., Sinha, D., and Nath, A. Application of machine learning and molecular modeling in drug discovery and cheminformatics. In *Advanced AI Techniques and Applications in Bioinformatics*. CRC Press, 2021, pp. 201–214.

[9] Vanhaelen, Q., Lin, Y.-C., and Zhavoronkov, A. The advent of generative chemistry. *ACS Medicinal Chemistry Letters 11*, 8 (2020), 1496–1505.

[10] Wang, M., Wang, Z., Sun, H., Wang, J., Shen, C., Weng, G., Chai, X., Li, H., Cao, D., and Hou, T. Deep learning approaches for de novo drug design: An overview. *Current Opinion in Structural Biology 72* (2022), 135–144.

[11] Weininger, D. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *Journal of chemical information and computer sciences 28*, 1 (1988), 31–36.

[12] Xu, Y. Deep neural networks for QSAR. In *Artificial intelligence in drug design*. Springer, 2021, pp. 233–260.

Pipeline Integrativo de Machine Learning Predictivo y Genómica Funcional para la Caracterización de Firmas Moleculares en Enfermedades Complejas

Rodrigo Gómez López[1], Rocío L. Cecchini [1,2], Axel J. Soto [1,2], Ignacio Ponzoni[1,2]
[1] Instituto de Ciencias e Ingeniería de la Computación, CONICET--UNS
[2] Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS)
rodrigo.gomez@cs.uns.edu.ar

Resumen

La caracterización funcional e identificación de firmas diagnósticas en enfermedades complejas y crónicas enfrenta importantes desafíos, relacionados tanto con la dimensionalidad y variabilidad intrínseca de los datos genómicos como con la limitada generalización de los hallazgos entre estudios. En este contexto, esta línea de investigación propone el desarrollo de un marco metodológico para la identificación y caracterización de firmas génicas robustas asociadas a enfermedades complejas.

Como primera instancia de aplicación, se abordó la detección y caracterización de firmas moleculares en fibromialgia, una enfermedad crónica donde estas problemáticas se manifiestan notoriamente. Para este fin, se exploró la implementación preliminar de una estrategia predictiva multimodelo (SVM, Elastic-Net, XGBoost, Random Forest) entrenada bajo un esquema de optimización bayesiana y replicación estadística, junto con metodologías de prefiltrado de datos y acompañada finalmente de análisis funcionales y regulatorios para su contextualización biológica.

El principal aporte de esta propuesta consiste en el desarrollo de un marco de consenso basado en explicabilidad (SHAP), que agrega los resultados del entrenamiento para filtrar así marcadores consistentes y mitigar sesgos algorítmicos y de pre-filtrado, el cual, acoplado a análisis biológicos post-hoc, busca contribuir al consenso de hipótesis y afianzar bases exploratorias para futuras aplicaciones en diversas enfermedades complejas.

Palabras claves: Aprendizaje automático, Explicabilidad, Bioinformática.

Contexto

La presente línea de investigación se enmarca en el proyecto de investigación "Desarrollo de algoritmos, modelado predictivo y análisis bioinformático basado en técnicas de aprendizaje automático", dirigido por las Dras. Jessica Andrea Carballido y Rocío Luján Cecchini. El proyecto es coordinado y financiado por la Universidad del Sur, con la contribución del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Introducción

La investigación en enfermedades complejas y multifactoriales presenta importantes dificultades metodológicas relacionadas con la identificación de firmas moleculares diagnósticas y la caracterización

de los mecanismos subyacentes. En consideración a esto, el análisis de datos transcriptómicos mediante técnicas de aprendizaje automático se ha posicionado como una herramienta clave para dilucidar las bases moleculares de estas patologías. Como primera instancia de esta línea de investigación, este trabajo preliminar se enfoca en la fibromialgia, un síndrome crónico caracterizado por un dolor musculoesquelético generalizado, fatiga persistente, trastornos de sueño, problemas cognitivos y diversas manifestaciones somáticas (Rodríguez and Mendoza, 2020; Siracusa et al., 2021; Al Sharie et al., 2024). Esta condición se caracteriza por tener una etiología y fisiopatología compleja, donde confluyen distintos factores genéticos, neurológicos e inmunológicos. Si bien numerosas investigaciones recientes han logrado elucidar varios de los mecanismos moleculares subyacentes a su desarrollo y progresión, las causas específicas de la enfermedad todavía permanecen sin esclarecer por completo (Al Sharie et al., 2024; Jurado-Priego et al., 2024)..

Diversos trabajos abordaron la fibromialgia desde un enfoque bioinformático, combinando técnicas de análisis de expresión diferencial con modelos predictivos, y culminando con validaciones biológicas post-hoc. Estos estudios se realizan a partir de datos transcriptómicos, representados por matrices de expresión génica, donde las filas representan genes y las columnas corresponden a muestras de pacientes y controles sanos, con valores numéricos que indican los niveles de expresión de cada gen.

En este sentido, en el trabajo de Zhao y colaboradores (Zhao et al., 2025) se entrenó un modelo XGBoost sobre genes detectados por análisis de expresión diferencial, identificando así genes con potencial discriminativo. Alineado a estos hallazgos, otras investigaciones entrenaron un modelo SVM sobre un conjunto de datos transcriptómicos, identificando firmas asociadas a procesos inmuno-inflamatorios y neuromoduladores (Jones et al., 2016). Adicionalmente, diversos autores analizaron también datos transcriptómicos, ocupando un criterio basado en la magnitud de cambio de expresión entre controles y pacientes para pre-filtrado, y finalmente aplicando KNN, identificando así firmas asociadas a procesos inmunes (Lukkahatai et al., 2018).

Paralelamente a estos trabajos, otros estudios exploraron la hipótesis de subtipificación de la enfermedad, en contraste a un único espectro clínico difuso. Por ejemplo, en investigaciones previas se realizó un análisis de componentes principales y posteriormente clustering, identificando cuatro subtipos con firmas génicas diferenciales (Mohapatra et al., 2024). También, trabajos similares aplicaron clustering sobre genes diferencialmente expresados para identificar dos subtipos moleculares, y posteriormente construyeron un clasificador SVM con alta especificidad para la enfermedad (Yao et al., 2023).

En conjunto, estos antecedentes reflejan una creciente implementación del aprendizaje automático aplicado al análisis de datos transcriptómicos, dado su potencial para la identificación de firmas diagnósticas. Sin embargo, todavía persisten limitaciones relacionadas con la interpretación, replicación y generalización de resultados, generando falta de consenso en torno a marcadores diagnósticos (Mohapatra et al., 2024). Además, a pesar del creciente desarrollo de técnicas de explicabilidad, su aplicación se concentra predominantemente en análisis post-hoc con fines de validación o contextualización (Zhao et al., 2025), no así como elemento metodológico central para la identificación de mecanismos asociados.

En estas circunstancias, el objetivo de esta línea de investigación es el desarrollo de un marco metodológico integrativo para la identificación de firmas moleculares y la caracterización funcional y regulatoria en enfermedades complejas. Utilizando a la fibromialgia como caso de estudio preliminar, se propone un flujo de trabajo que busca superar las limitaciones metodológicas antes mencionadas. Desde esta perspectiva, el enfoque combina dos estrategias distintas de prefiltrado, una basada en detección génica por análisis de expresión diferencial, y otra con filtrado exclusivamente por biotipo génico. Se emplean múltiples técnicas de aprendizaje automático supervisado (XGBoost, Random Forest, SVM lineal y Elastic-Net), bajo un marco de optimización bayesiana, e implementadas bajo esquemas de réplica para asegurar estabilidad estadística y superar posibles sesgos algorítmicos. Subsecuentemente, sobre estos modelos se calculan medidas de importancia mediante técnicas de explicabilidad (SHAP), se construyen consensos y se identifican conjuntos génicos consistentes entre enfoques predictivos. Finalmente, los genes consensuados son utilizados para análisis biológicos post-hoc, permitiendo contextualizar las firmas identificadas y validar los resultados. De este modo, la línea propone una aproximación metodológica sistemática y transparente, aplicable a distintas patologías, para la identificación y validación de biomarcadores moleculares.

Línea de investigación

La presente línea de investigación se centra en el desarrollo e implementación de metodologías human-in-the-loop que integren técnicas de aprendizaje automático con estrategias de interpretabilidad y visualización, aplicadas al análisis de datos de expresión génica en el contexto de la medicina de precisión. Se exploran cuatro ejes complementarios: (i) la incorporación de técnicas de inteligencia artificial explicable (XAI) y de inferencia causal para explicar los agrupamientos generados por algoritmos no supervisados; (ii) el uso de lógica difusa para modelar módulos génicos superpuestos y capturar grados de pertenencia de genes y condiciones; (iii) el diseño de estrategias de analítica visual que faciliten la incorporación de conocimiento experto y la validación de los patrones descubiertos; y (iv) la adaptación de estos enfoques a distintos problemas biológicos y biomédicos, buscando contribuir a una caracterización más precisa y personalizada de enfermedades complejas.

Resultados obtenidos

Se logró implementar un marco metodológico preliminar que integra modelos de aprendizaje supervisado, técnicas de explicabilidad, metodologías de consenso y análisis biológicos post-hoc para identificar genes discriminantes y módulos funcionales en enfermedades complejas, empleando a la fibromialgia como instancia de validación. Para esto, se incorporaron cuatro modelos (Support Vector Machine, XGBoost, Random Forest, Logistic Regression con penalización Elastic-Net) bajo dos estrategias distintas de prefiltrado. Utilizando técnicas de explicabilidad (SHAP), a fin de poder explicar cada réplica en términos de la importancia de sus características y, sobre la base de estas, se busca generar un consenso intra-modelos e inter-modelo, para obtener así un conjunto reducido de genes consenso.

A fin de contextualizar biológicamente los resultados, se aplicarán algoritmos de

inferencia de redes regulatorias y consultas en bases de datos de interacciones proteicas, buscando identificar de esta manera moléculas reguladoras significativamente enriquecidas y propuestas como potenciales reguladores maestros de los genes detectados. Asimismo, se construirá una red de interacción proteína-proteína (PPI) que integre los genes identificados, de la cual se puedan extraer distintos módulos funcionales mediante técnicas de clustering y consultas a bases de datos curadas. Finalmente, tanto los módulos como los reguladores serán interpretados dentro del marco fisiopatológico de la enfermedad a partir de la literatura científica. Este enfoque busca obtener un conjunto de genes discriminantes, replicable entre modelos y, por lo tanto, independiente de sesgos algorítmicos específicos.

Como trabajo futuro, se plantea complementar el marco metodológico de exploración y extracción de firmas moleculares con un nuevo flujo de trabajo basado en aprendizaje no supervisado y orientado específicamente a la identificación de subgrupos clínicos en enfermedades complejas, asociados con perfiles transcriptómicos diferenciales, y consecuentemente, a firmas moleculares distintivas. Este nuevo esquema de trabajo sería acompañado además con técnicas de explicabilidad post-hoc para interpretar las firmas diagnósticas de cada subtipo, evaluar su estabilidad mediante validación cruzada, y contrastar los resultados con bases de datos clínicas. En virtud de que el pipeline incorpora estrategias para abordar la dimensionalidad y la inherente variabilidad génica, su implementación no es exclusiva de la fibromialgia, sino potencialmente aplicable a distintas enfermedades complejas con características biológicas semejantes.

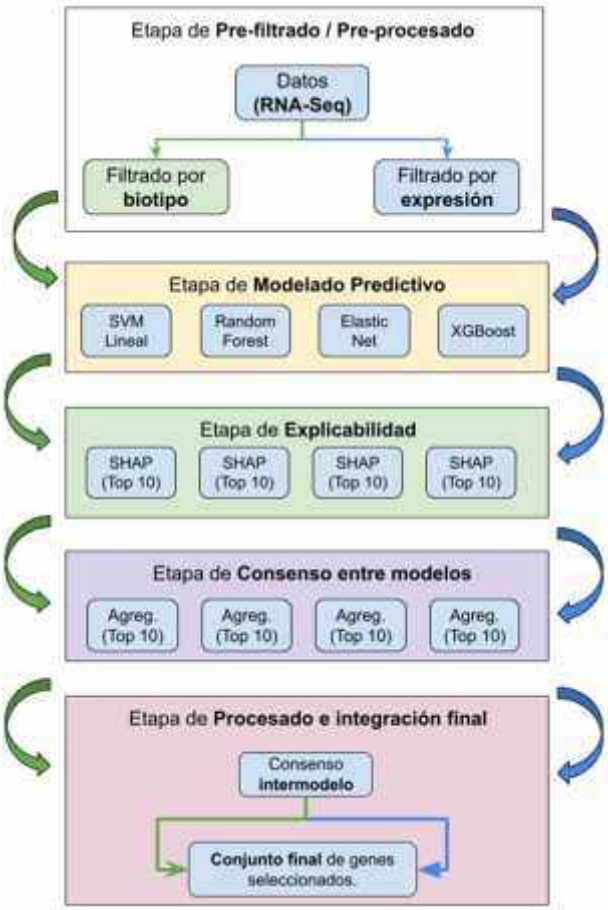


Figura 1: Representación visual del pipeline propuesto.

Formación de Recursos Humanos

El doctorado del Lic. Rodrigo Gómez López se desarrolla en el marco de la presente línea de investigación, radicada en el grupo BioChemTICs del Instituto de Ciencias e Ingeniería de la Computación (CONICET-UNS). La temática central de esta línea aborda el diseño de metodologías de aprendizaje automático orientadas a la inferencia y análisis de redes jerárquicas a partir de datos de expresión génica, en el ámbito de la bioinformática. El trabajo se inscribe en la disciplina Informática y Comunicaciones, con aplicación en ciencias exactas y naturales.

Bibliografía

Al Sharie, S., Varga, S. J., Al-Husinat, L., Sarzi-Puttini, P., Araydah, M., Bal'awi, B. R., and Varrassi, G. (2024). Unraveling the complex web of fibromyalgia: A narrative review. Medicina, 60(2):272.

Jones, K. D., Gelbart, T., Whisenant, T. C., Waalen, J., Mondala, T. S., Iklé, D. N., Salomon, D. R., Bennett, R. M., and Kurian, S. M. (2016). Genome-wide expression profiling in the peripheral blood of patients with fibromyalgia. Clinical and Experimental Rheumatology, 34(2 Suppl 96):S89–S98.

Jurado-Priego, L. N., Cueto-Ureña, C., Ramírez-Expósito, M. J., and Martínez-Martos, J. M. (2024). Fibromyalgia: A review of the pathophysiological mechanisms and multidisciplinary treatment strategies. Biomedicines, 12(7):1543.

Lukkahatai, N., Walitt, B., Deandrés-Galiana, E. J., Fernández-Martínez, J. L., and Saligan, L. N. (2018). A predictive algorithm to identify genes that discriminate individuals with fibromyalgia syndrome diagnosis from healthy controls. Journal of Pain Research, 11:2981–2990.

Mohapatra, G., Dachet, F., Coleman, L. J., Gillis, B., and Behm, F. G. (2024). Identification of unique genomic signatures in patients with fibromyalgia and chronic pain. Scientific Reports, 14(1).

Rodríguez, D. F. G. and Mendoza, C. A. (2020). Physiopathology of fibromyalgia. Reumatología Clínica (English Edition), 16(3):191–194.

Siracusa, R., Di Paola, R., Cuzzocrea, S., and Impellizzeri, D. (2021). Fibromyalgia: Pathogenesis, mechanisms, diagnosis and treatment options update. International Journal of Molecular Sciences, 22(8):3891.

Yao, M., Wang, S., Han, Y., et al. (2023). Microinflammation related gene signatures are associated with clinical features and immune status of fibromyalgia. Journal of Translational Medicine, 21:594.

Zhao, F., Zhao, J., Li, Y., Song, C., Cheng, Y., Li, Y., Wu, S., He, B., Jiao, J., and Chang, C. (2025). Identification of diagnostic biomarkers for fibromyalgia using gene expression analysis and machine learning. Frontiers in Genetics, 16.

Sistemas inteligentes. Técnicas, modelos y aplicaciones

W. Hasperué[1], L. Lanzarini[1], C. Estrebou[1], E. Rucci[1], F. Ronchetti[1], G. Reyes[2,3]
E. Gutiérrez[4,5], A. Fernandez Bariviera[6]

[1] Instituto de Investigación en Informática LIDI*, Facultad de Informática, UNLP, La Plata, Argentina
[2] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Guayaquil, Ecuador
[3] Artificial Intelligence Research Group, Universidad Bolivariana del Ecuador, Yaguachi, Durán, Ecuador
[4] Instituto de Ciencias e Ingeniería de la Computación (ICIC-UNS CONICET)
[5] Departamento de Economía. Universidad Nacional del Sur.
[6] Dpto. de Economía, Universitat Rovira i Virgili, Reus, España
*** Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)**

{whasperue, laural, cesarest, erucci, fronchetti}@lidi.info.unlp.edu.ar, gxreyesz@ube.edu.ec
emiliano.gutierrez@uns.edu.ar, aurelio.fernandez@urv.net

CONTEXTO

Esta presentación corresponde a las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto "Inteligencia de Datos. Técnicas y Modelos de Machine Learning" perteneciente al Programa de Incentivos (2023-2026).

RESUMEN

Esta línea de investigación se centra en el desarrollo de Sistemas Inteligentes para la resolución de problemas de Minería de Datos utilizando técnicas de Aprendizaje Automático. Los sistemas desarrollados se aplican particularmente al procesamiento de grandes volúmenes de información y al procesamiento de flujo de datos.

Una de las líneas de investigación llevadas a cabo está enfocada en el estudio y desarrollo de modelos de IA para el desarrollo de chatbots, en particular para responder consultas en el ámbito universitario (alumnos, docentes, personal administrativo, etc.).

Por otro lado, y desde el punto de vista de la salud, se está trabajando con investigadores del CENEXA (CONICET-UNLP-CIC) en la obtención de modelos de predicción de diabetes y prediabetes.

Así mismo, se trabaja con investigadores del GIIA- UBE de Ecuador en modelos adaptativos sobre flujos de datos multidimensionales estudiando técnicas capaces de analizar datos en tiempo real provenientes de múltiples fuentes heterogéneas (sensores, dispositivos IoT, cámaras, GPS y reportes ciudadanos).

También se investigan aspectos de explicabilidad e interpretabilidad de modelos de machine learning. En particular, el caso de estudio actual es la valoración de inmuebles con foco no sólo en la precisión de la predicción sino en la comprensión de los factores que determinan el precio de una vivienda usando técnicas de explicación independientes del modelo.

Palabras clave: LLM, Minería de Datos, Explicabilidad de modelos, Flujos de datos dinámicos, TinyML, Transformers, Diabetes.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos tipos de estrategias adaptativas. Los resultados obtenidos han sido aplicados en la solución de problemas de distintas áreas, como salud, interpretabilidad de modelos, implementación de chatbots, y sistemas embebidos. A continuación, se detallan las investigaciones realizadas durante el último año.

1.1. BIG DATA

Desarrollo de un chatbot para el ámbito universitario

En [1] se presentó el desarrollo de un chatbot conversacional basado en BERT para agilizar la atención de consultas en la Universidad Nacional Arturo Jauretche. La base de datos con la cual fue creado el modelo, se construyó mediante técnicas de web crawling y web scraping del sitio oficial, estructurando la información en formato SQuAD y aplicando técnicas de aumentación con sinónimos.

El modelo fue generado a partir de un fine-tuning de un modelo BERT-Base en español, optimizando sus hiperparámetros a través de la biblioteca Optuna. Para la validación, se integró métricas de coincidencia exacta, similitud semántica y una evaluación humana asistida. Los resultados obtenidos permiten observar un desempeño robusto ya que el chatbot resolvió correctamente el 84,61% de consultas que pueden ser encontradas en la página de preguntas frecuentes del sitio web de la universidad; asimismo resolvió el 83,02% de las 1389 interacciones reales analizadas.

El modelo construido es altamente escalable y efectivo para la transformación digital universitaria, pudiendo ser replicado en otras universidades. No obstante, se identificaron debilidades para mejorar como la incorporación de memoria de contexto y la actualización dinámica de datos ante cambios en la web institucional.

Técnicas adaptativas para el procesamiento de flujos de datos multidimensionales

El trabajo estratégico sobre "Aprendizaje Inteligente sobre Flujos de Datos Multidimensionales" [11] constituye el eje conceptual para comprender la evolución y proyección del aprendizaje en flujos dinámicos. A partir de su análisis bibliométrico, se han identificado tendencias y vacíos que orientan nuevas líneas de investigación en aprendizaje incremental y detección de *concept drift*.

En coherencia con esta estrategia, se están iniciando estudios exploratorios en tres frentes complementarios. Uno aborda la incorporación de la dimensión emocional mediante el análisis de voz en flujos de conductores. Otro se enfoca en la clasificación inteligente de incidentes de tráfico en tiempo real. Finalmente, se plantea la detección adaptativa de anomalías en movilidad urbana, configurando un programa de investigación aún en desarrollo con proyección hacia sistemas urbanos inteligentes.

Identificación de personas con riesgo de diabetes y prediabetes

La Diabetes Tipo 2 (DT2) es una enfermedad crónica que afecta a la regulación del nivel de azúcar en sangre. Debido a su creciente prevalencia en combinación con su elevado costo de atención, constituye un serio problema de salud pública, por lo que se han realizado grandes esfuerzos por desarrollar estrategias efectivas para su prevención y tratamiento así como para evitar sus complicaciones crónicas. En ese sentido, resulta importante reconocer que las consecuencias negativas de esta enfermedad comienzan en una etapa previa conocida como prediabetes, la cual implica un riesgo elevado de desarrollar DT2 en los siguientes años [4].

El desarrollo de la DT2 es un proceso lento y progresivo condicionado por factores genéticos, ambientales y de comportamiento. La detección de DT2 y prediabetes representa un verdadero desafío para la medicina debido a la ausencia de síntomas patogenómicos y/o la falta de conocimiento de los factores de riesgo asociados. Es por lo que frecuentemente una persona pueda pasar meses (o incluso años) sin saber que se encuentra en riesgo.

Brevemente, esta línea propone desarrollar y validar modelos predictivos de diabetes y prediabetes específicos para la población argentina utilizando técnicas de Aprendizaje Automático. En particular, ya se ha iniciado el trabajo experimental orientado al desarrollo de modelos que permitan identificar personas con alta probabilidad de tener estas enfermedades, contando con resultados preliminares [2][3].

También interesa explorar modelos que sean capaces de predecir la aparición de complicaciones microvasculares en personas que ya tienen la enfermedad.

1.2. VALORACIÓN INMOBILIARIA CON APRENDIZAJE AUTOMATICO

Esta línea de investigación se orienta al estudio del mercado inmobiliario argentino mediante el desarrollo y aplicación de modelos cuantitativos y técnicas avanzadas de aprendizaje automático para la valoración de propiedades. El objetivo central es generar herramientas robustas y socialmente relevantes que contribuyan a mejorar la precisión en la estimación de precios, promoviendo mayor transparencia, eficiencia en la asignación de recursos y competencia en el sector. Asimismo, se busca aportar evidencia empírica que sirva de insumo para el diseño y evaluación de políticas públicas de vivienda más efectivas y basadas en datos.

El enfoque metodológico integra modelos tradicionales de regresión hedónica con algoritmos de aprendizaje automático, aplicados sobre grandes volúmenes de datos provenientes de publicaciones inmobiliarias en línea, enriquecidos con información geográfica y contextual. Esta combinación permite capturar tanto relaciones estructurales clásicas entre atributos y precios como patrones no lineales y efectos complejos que emergen de los datos. La comparación sistemática de distintos métodos predictivos posibilita identificar aquellos con mejor desempeño, así como analizar su estabilidad y capacidad de generalización en distintos contextos urbanos.

Un componente central de la línea es la interpretabilidad de los modelos desarrollados. Más allá de maximizar la precisión predictiva, se prioriza la comprensión de los factores que inciden en la formación de precios. Para ello, se emplean técnicas de explicación independientes del modelo, tales como gráficos de dependencia parcial, curvas de expectativa condicional individual y métodos basados en valores de Shapley. Estas herramientas permiten transparentar la contribución de cada atributo, identificar efectos marginales y explorar heterogeneidades, complementando los enfoques econométricos tradicionales. En conjunto, la investigación articula ciencia de datos, economía urbana y análisis territorial para fortalecer la toma de decisiones en el ámbito inmobiliario. Los resultados obtenidos hasta el momento han sido publicados en [10].

1.3. TINYML/EDGEML

El desafío fundamental de TinyML/EdgeML radica en adaptar modelos de aprendizaje automático para su ejecución en microcontroladores (MCUs), dispositivos caracterizados por memorias de pocos kilobytes, capacidad de cómputo reducida y presupuestos energéticos estrictos.

Para abordar este desafío, el grupo desarrolló EmbedIA [5, 6], un framework de código abierto que automatiza la conversión de modelos de aprendizaje automático y redes neuronales en código C/C++/Arduino optimizado para MCUs [6]. Una característica diferencial respecto a otras soluciones —como TensorFlow Lite Micro o ARM CMSIS-NN— es su compatibilidad con arquitecturas sin soporte nativo para FPU, instrucciones DSP o extensiones SIMD, lo que lo hace apto para una amplia variedad de microcontroladores. Evaluaciones comparativas han mostrado ventajas en uso de memoria y velocidad de inferencia frente a frameworks alternativos [6, 7, 8].

EmbedIA soporta modelos de TensorFlow/Keras, TFLite y redes binarias Larq, así como modelos de Scikit-Learn incluyendo SVM, árboles de decisión, KNN, Random Forest y regresión. Incorpora además optimizaciones en punto fijo de 8, 16 y 32 bits, e integración transparente de etapas de preprocesamiento —como normalización y transformadas de señal— dentro del pipeline de inferencia. Recientemente se incorporó soporte para instrucciones DSP/SIMD en ESP32 y sus variantes, con trabajo en curso para extender esta capacidad a microcontroladores ARM Cortex-M.

Entre las aplicaciones desarrolladas, EmbedIA fue empleado en el despliegue de un clasificador CNN basado en espectrogramas para clasificar el tipo de actividad física en personas con Diabetes Tipo 1 [9]. El sistema fue desplegado en cuatro plataformas (ESP32, ESP32-C3, ESP8266 y Raspberry Pi Pico), manteniendo la precisión original del modelo (87.64%) con las variantes float y fixed32, y con una leve reducción aceptable en fixed16 (85.16%). En el marco de proyectos de innovación con alumnos, se está desarrollando además un prototipo de vehículo miniatura con conducción autónoma sobre ESP32-CAM, integrando visión por computadora y un modelo entrenado mediante aprendizaje por refuerzo en Unity, exportado en formato ONNX.

En cuanto a trabajo en curso, se avanza en la integración nativa de PyTorch y ONNX como fuentes de modelos. Paralelamente, se realizaron exploraciones sobre MCUs con unidades de procesamiento neuronal (NPUs) utilizando la placa Luckfox Pico (SoC RV1103), donde la comparación entre CPU ARM Cortex-A7 y NPU integrada arrojó tiempos de inferencia de 1.48 s frente a 14.9 ms respectivamente. Estas experiencias sientan las bases para una futura integración de soporte NPU en EmbedIA.

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Construcción de grandes modelos del lenguaje para la personalización de chatbots.
- Identificación de personas con riesgo de diabetes y prediabetes.
- Técnicas adaptativas para el procesamiento de flujos de datos multidimensionales.
- Estudio de modelos de Deep Learning para series temporales.
- Estudio de aplicaciones y bibliotecas diseñadas para realizar benchmarks de modelos de deep learning aplicados a series temporales.
- Estudio modelos y técnicas de compresión de modelos para microcontroladores.
- Evaluación y análisis de bibliotecas y frameworks de aprendizaje automático para microcontroladores.
- Desarrollo de un Framework de código abierto para inferencia de modelos de aprendizaje automático y redes neuronales en microcontroladores.
- Evaluación de microcontroladores con unidades de procesamiento neuronal (NPUs).
- Aprendizaje automático interpretable aplicado al análisis de mercados inmobiliarios

3. RESULTADOS OBTENIDOS

- Chatbots específicos para el uso en la vida universitaria.
- Modelos de predicción de diabetes y prediabetes
- Análisis comparativo de modelos de Deep Learning para series temporales.
- Desarrollo y validación de un modelo predictivo para la estimación de precios inmobiliarios.
- Implementación de métodos de interpretabilidad para identificar y cuantificar los principales factores que inciden en la valoración inmobiliaria

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: cuatro profesores doctores con dedicación exclusiva, un profesor con dedicación exclusiva, un alumno diplomado, tres tesistas de Doctorado en Cs. Informáticas (uno con beca de postgrado de la UNLP) y dos profesores extranjeros.

En el último año se han finalizado 2 tesis de maestría, 1 tesis de especialista, 3 tesinas de grado de Licenciatura y 3 prácticas profesionales supervisadas.

Actualmente se están desarrollando 3 tesis de doctorado, 7 tesis de maestría, 1 trabajo final de especialista y 2 tesinas de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

[1] Leguizamón, M; Hasperué, W. Uso de BERT para la creación de un chatbot de atención de consultas en la UNAJ. XXXI Congreso Argentino de Ciencias de la Computación. Viedma. Argentina. 2025.

[2] Rucci, E.; Tittarelli, G.; Ronchetti, F.; Elgart, J.F.; Lanzarini, L.; Gagliardino, J.J. "Primeras Experiencias en la Identificación de Personas con Riesgo de Diabetes en la Población Argentina utilizando Técnicas de Aprendizaje Auto-mático". XXIX Congreso Argentino de Cien-cias de la Computación. Octubre de 2023.

[3] Rucci, E., Tittarelli, G., Ronchetti, F., Elgart, J. F., Lanzarini, L. y Gagliardino, J. J. First Experiences with the Identification of People at Risk for Diabetes in Argentina Using Machine Learning Techniques, en Computer Science – CACIC 2023, P. Pesado, W. Panessi, y J. M. Fernández, Eds., Cham: Springer Nature Switzerland, 2024, pp. 227-239. doi: 10.1007/978-3-031-62245-8_16

[4] Federación Internacional de Diabetes, "Atlas de la diabetes de la fid," tech. rep., Federación Internacional de Diabetes, 2017. https://diabetesatlas.org/resources/2017-atlas.html.

[5] Estrebou C., Feming M., Saavedra M. D., Adra F. MbedML: A Machine Learning Project for Embedded Systems. IX Jornadas de Cloud Computing, Big Data & Emerging Topics. ISBN: 978-950-34-2016-4 Pp. 25-28. Facultad de Informática. UNLP. Junio 2021.

[6] Estrebou, C., Saavedra, M. D., Adra, F., Fleming, M. TinyML for Small Microcontrollers. X Jornadas de Cloud Computing, Big Data & Emerging Topics. ISBN 978-950-34-2126-0. Pp 42-46. Facultad de Informática. UNLP. Mayo 2022.

[7] Estrebou C., Feming M., Saavedra M. D., Adra F. A Neural Network Framework for Small Microcontrollers. XXVII Congreso Argentino de Ciencias de la Computación. ISBN: 978 -987-633-574-4. Pp. 51-60. Univ. Nacional de Salta. Octubre 2021.

[8] Estrebou C., Feming M., Saavedra M. D., Adra F., De Giusti, A. E. Lightweight Convolutional Neural Networks Framework for Really Small TinyML Devices. Second International Conference on Smart Technologies, Systems and Applications. SmartTech-IC 2021. Quito, Ecuador. Diciembre 2021.

[9] Saavedra, M.D., Estrebou, C.A., Inthamoussou, F.A., Garelli, F.: Deployment of a Spectrogram-Based Classifier for Physical Activity in Type 1 Diabetes on Embedded Systems. CACIC 2025.

[10] Gutiérrez E. Lanzarini L., Delbianco F., Cecchini R. (2025). Real Estate Market in Buenos Aires (Argentina): A Hedonic and Interpretable Machine Learning Approach. Cloud Computing, Big Data and Emerging Topics. 13th Conference, JCC-BD&ET 2025, La Plata, Argentina, June 24–26, 2025, Revised Selected Papers. ISBN: 978-3-032-06336-6. Págs. 61-76. Springer Cham. Octubre 2025.

[11] Reyes, G.; Tolozano-Benites, R.; Lanzarini, L.; Hasperué, W.; Barzola-Monteses, J. Intelligent Learning on Multidimensional Data Streams: A Bibliometric Analysis of Research Evolution and Future Directions. Information. 2025, 16, 1067. https://doi.org/10.3390/info16121067

Uso de Técnicas de Explotación de Información para Análisis de Similitudes y Diferencias entre las Aguas del Río de la Reconquista y Matanza-Riachuelo

Carlos Gustavo Lopez, Enrique Fernandez, Marcelo Fabián Esperanza, Martin Pastorini, Elda Melania Valenzuela, María Alejandra Ochoa

Universidad Nacional del Oeste

Instituto de Ingenierías y Nuevas Tecnologías

Especialidad en Ciencia de Datos

gustavolopez/ enfernandez/ mpastorini/ mochoa@uno.edu.ar, maresperanza9@hotmail.com, valenzuelamelania2023@gmail.com

Resumen

El objeto del presente trabajo es analizar mediante técnicas de explotación de información diferencias y similitudes en los componentes del agua del Río de la Reconquista y MatanzaRiachuelo.
Para ello se utilizarán datos del estado del agua del Río de la Reconquista generados por la Universidad Nacional del Oeste, obtenidos mediante muestreos del agua del Río y posterior análisis en laboratorio propio, y datos públicos generados por ACUMAR sobre la cuenca Matanza Riachuelo.

Palabras clave: Explotación de Información, Minería de datos, Medio ambiente, recurso hídrico, Calidad del agua, Rio de la Reconquista.

Contexto

El proyecto "Análisis de calidad del agua del Rio de la Reconquista en la zona adyacente del partido de Merlo utilizando tecnologías basadas en Inteligencia Artificial." que lleva adelante el Instituto de Ingenierías y Nuevas Tecnologías de la Universidad Nacional del Oeste, es continuación de dos proyectos anteriores llevados adelante por esta Institución y buscan desarrollar un modelo predictivo del comportamiento del agua del Río de la Reconquista [1]. Para ello se realizaron tareas de muestreo y análisis químico de las propiedades del agua en cinco puntos del Río, todos vinculados a su paso por el Partido de Merlo [1]. En la actualidad el Instituto cuenta con un set de 10.000 muestras generadas mediante un proceso de generación de datos sintéticos [2, 3] que partió de un proceso de muestreo real.

El objetivo central del presente trabajo es aplicar técnicas de minería de datos sobre los datos mencionados y los publicados por ACUMAR para determinar que similitudes y diferencias existen entre ambos ríos.

Introducción

La contaminación de algunos recursos hídricos es mayor de lo esperado para un país que se encuentra en etapa de desarrollo como es Argentina [4]. Este hecho es uno de los problemas percibidos en el Plan Nacional Federal de Recursos Hídricos (PNFRH) [5], que tiene como uno de sus objetivos "reducir y prevenir la contaminación de los recursos hídricos, y proteger y preservar los ecosistemas vinculados a los mismos" [6]. Si bien los recursos hídricos se ven degradados a lo largo de todo su curso, puede distinguirse el tipo de

contaminación de acuerdo a la altura de su cuenca. Así, por ejemplo, encontramos ríos cuya contaminación varía de acuerdo al siguiente esquema [7]:

- En la cuenca alta, las aguas se encuentran afectadas por la contaminación de origen agrícola.
- En la cuenca media, se evidencia una degradación progresiva de la calidad del agua con aumento del grado de afectación de las aguas, por el mayor aporte de vertidos de sustancias contaminantes de origen domiciliario e industrial.
- En la cuenca baja, las aguas del río se encuentran severamente degradadas debido al aporte acumulado de contaminantes vertidos en toda la cuenca y por los vertidos de contaminantes de origen domiciliario e industrial.

La evaluación de la salud ecosistémica de la Cuenca Matanza Riachuelo se fundamenta, entre otros procedimientos, en un esquema de monitoreo estacional ejecutado por ACUMAR [7]. Este diseño experimental busca capturar la variabilidad fisicoquímica y microbiológica relacionada a los ciclos hidrológicos anuales, con el objetivo de determinar la evolución cronológica del cuerpo de agua y la eficacia de las medidas de remediación. Los datos obtenidos en estas campañas de campo son procesados, informados y publicados de forma abierta, además sirven como insumo para el cálculo del Índice de Calidad de Agua (ICA). El ICA funciona como un indicador que agrega múltiples parámetros complejos (tales como Oxígeno Disuelto, DBO5, metales pesados y nutrientes) en un valor adimensional, permitiendo categorizar el estado ecológico del recurso en una escala preestablecida.

En este marco de control regional, la Universidad Nacional del Oeste (UNO) también desempeña un rol técnico al realizar un monitoreo específico en el tramo del río que atraviesa el Partido de Merlo. Esta labor académica aporta una mayor densidad de datos en el tramo que comprende a la cuenta media del Rio Reconquista, permitiendo identificar focos de contaminación puntuales y generar datos de rigor científico que complementan el diagnóstico oficial, fortaleciendo el conocimiento territorializado del recurso hídrico.

Líneas de Investigación y Desarrollo

Desde una perspectiva metodológica, la investigación se apoya en la integración y armonización de datos heterogéneos, combinando datos primarios del Río Reconquista generados por la Universidad Nacional del Oeste mediante muestreos [3] y análisis en laboratorio, y datos secundarios de acceso público elaborados por ACUMAR para la cuenca Matanza–Riachuelo, disponibles en su portal de datos abiertos y en su Base de Datos Hidrológicos [8]

En el plano del desarrollo, se contempla la construcción de un pipeline reproducible para: depuración (control de calidad), tratamiento de valores faltantes, normalización de unidades, estandarización de variables y generación de indicadores derivados, con el objetivo de asegurar la comparabilidad entre campañas, estaciones y puntos de muestreo. Este enfoque resulta consistente con prácticas actuales de análisis multivariado en monitoreo de calidad de agua, donde la multidimensionalidad de los parámetros exige procedimientos sistemáticos de preparación y validación del dato. [mdpi.com]

En cuanto a las técnicas analíticas, el estudio incorporará:

1. **Análisis descriptivo y exploratorio** para caracterizar distribuciones, variabilidad y tendencias preliminares de los parámetros fisicoquímicos y/o microbiológicos considerados [9].
2. **Análisis de correlación** (p. ej., Pearson o Spearman, según supuestos) para identificar asociaciones entre variables, detectar colinealidad y explorar relaciones potencialmente indicativas de procesos comunes (por ejemplo, co-variación entre nutrientes, demanda de oxígeno y sólidos). Este tipo de análisis es habitual como paso

previo y complementario en estudios multivariados de calidad de agua [10].

3. Análisis multivariado para reducción de dimensionalidad y extracción de patrones:
    - Análisis de Componentes Principales (PCA) para sintetizar la variabilidad explicada por múltiples parámetros y facilitar la comparación global entre sitios/cuencas, técnica ampliamente usada en evaluación de calidad de agua. [9] [11] [12]
    - Análisis de conglomerados (Cluster Analysis / HCA o k-means) para agrupar estaciones o períodos en función de similitud multivariada y detectar segmentos con perfiles de contaminación comparables. [11] [13]
    - (Opcional, según disponibilidad de variables) Análisis discriminante / análisis factorial para identificar subconjuntos de parámetros con mayor capacidad de separar grupos espaciales/temporales, tal como se reporta en estudios recientes sobre ríos con monitoreos extensos. [12]

Finalmente, los resultados se orientarán a producir evidencia empírica y trazable para el diagnóstico comparativo de ambas cuencas, aportando insumos analíticos para el diseño de estrategias de monitoreo y gestión ambiental basadas en datos, en línea con recomendaciones contemporáneas sobre el uso de métodos multivariados en el análisis de calidad de agua [12] [14].

Formación de Recursos Humanos

El presente trabajo da origen al trabajo final de especialidad en Ciencias de Datos que se dicta en la Universidad Nacional del Oeste. El mismo será realizado por un único estudiante con el apoyo del equipo de investigación del proyecto.

Bibliografía

[1] Carlos Lopez, Enrique Fernandez, Natanael Ferrán, Martín Pastorini, Nicolás Coronel, Marcelo Esperanza; “Análisis del Río de la Reconquista”; Ingeniería Sanitaria y Ambiental, Nro. 149; ISSN 0328 2937; octubre 2023

[2] Carlos Gustavo Lopez, Enrique Fernandez, Marcelo Fabián Esperanza, Martin Pastorini, Natanael Ferrán. 2024. Simulador de Ensayos de Laboratorio para Análisis de Calidad del Agua. Proceedings XXVI Workshop de Investigadores en Ciencias de la Computación. ISBN 978-987-3984-85-3

[3] Carlos Gustavo Lopez, Enrique Fernandez, Marcelo Fabián Esperanza, Martin Pastorini, Natanael Ferrán. María Alejandra Ochoa. 2024. Contrastación de Métodos de Simulación para Ampliar Evidencias con Vistas a Desarrollar un Proceso de Minería de Datos (Un Testigo Basado en los parámetros del Índice de Calidad del Agua). XXX Congreso Argentino de Ciencias de la Computación (CACIC). ISBN: 978-950-34-2428-5

[4] Castilla V., Canevaro, S., López, B.. Migración, degradación ambiental y percepciones del riesgo en la cuenca del río (Buenos Aires, Argentina), Revista de Estudios Sociales [En línea], 76 | 01 abril de 2021, Publicado el 04 abril 2021, consultado el 15 julio 2022. URL: http://journals.openedition.org/revestudsoc/49626

[5] Rodríguez, A. (Coord.). (2008). Plan Nacional Federal de los Recursos Hídricos de la República Argentina. Ediciones Gráficas Especiales

[6] Consejo Hídrico Federal. Plan Nacional Federal de los Recursos Hídricos. 2007

[7] ACUMAR, Autoridad de Cuenca Matanza-Riachuelo, https://www.acumar.gob.ar/wp-content/uploads/2016/12/CDCA_ACUMAR_INFO_AGUA_abril-2022_completo-web-con-anexos_compressed.pdf Accesed 2022.

[8] ACUMAR. Autoridad de Cuenca Matanza Riachuelo. *Redes de Monitoreo de Agua*

[Dataset]. Portal de Datos Abiertos ACUMAR. 2025.

[9] Benkov, I., Varbanov, M., Venelinov, T., & Tsakovski, S. Principal Component Analysis and the Water Quality Index—A Powerful Tool for Surface Water Quality Assessment: A Case Study on Struma River Catchment, Bulgaria. *Water, 15*(10), 1961. 2023.

[10] Ding, H., Niu, X., Zhang, D., Lv, M., Zhang, Y., Lin, Z., & Fu, M. Spatiotemporal analysis and prediction of water quality in Pearl River, China, using multivariate statistical techniques and data-driven model. *Environmental Science and Pollution Research, 30*, 63036–63051. 2023.

[11] Maiolo, M., & Pantusa, D. Multivariate Analysis of Water Quality Data for Drinking Water Supply Systems. *Water, 13*(13), 1766. 2021.

[12] Souza, A. T., Carneiro, L. A. T. X., Silva Junior, O. P., Carvalho, S. L., & Américo-Pinheiro, J. H. Assessment of water quality using principal component analysis: a case study of the Marrecas stream basin in Brazil. *Environmental Technology*. 2020.

[13] Liu, J., Zhang, D., Tang, Q., Xu, H., Huang, S., Shang, D., & Liu, R. Water quality assessment and source identification of the Shuangji River (China) using multivariate statistical methods. *PLOS ONE, 16*(1), e0245525. 2021.

[14] Muniz, D. H. F., & Oliveira-Filho, E. C. (2023). Multivariate Statistical Analysis for Water Quality Assessment: A Review of Research Published between 2001 and 2020. *Hydrology, 10*(10), 196. 2023.

Computación Gráfica, Imágenes y Visualización

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

ANÁLISIS VISUAL JERÁRQUICO PARA EL MANEJO Y GESTIÓN DE INCERTIDUMBRE APLICADO A DATOS MULTIDIMENSIONALES

Luque, Leandro E.[1,3], **Antonini, Antonella S.**[1,2,3], **Raidou, Renata G.**[4], **Gröller, Eduard**[4], **Castro, Silvia M.**[1,2,3], **Ganuza, M. Luján**[1,2,3]

[1] Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (CIC-UNS)
[2] Dpto. de Cs. e Ing. de la Computación, Universidad Nacional del Sur (DCIC-UNS)
[3] Instituto de Ciencias e Ingeniería de la Computación (CONICET-UNS)
[4]Institute of Visual Computing & Human-Centered Technology (TU Wien)
{ leandro.luque, antonella.antonini, mlg, smc }@ cs.uns.edu.ar
{ rraidou, groeller }@cg.tuwien.ac.at

RESUMEN

El análisis visual bajo incertidumbre requiere un diseño iterativo entre especialistas en visualización y expertos del dominio para comprender las tareas y requisitos que guiarán el análisis. Integrar la incertidumbre en el proceso visual agrega complejidad, ya que exige su modelado, gestión y comunicación de manera perceptible y cognitivamente efectiva.

Para ello, se necesitan nuevos paradigmas visuales e interacciones capaces de representar distintos niveles de detalle, soportar datos multidimensionales o espacio-temporales y preservar la trazabilidad de la incertidumbre durante todo el flujo analítico. Asimismo, es fundamental desarrollar mecanismos teóricos y prácticos que permitan transformar los datos y evaluar la calidad y confiabilidad de las soluciones visuales bajo incertidumbre. Este proyecto tendrá aplicaciones en distintos campos, entre ellos las ciencias geológicas, el eye tracking y los datos de acceso abierto.

Palabras claves: Incertidumbre, Análisis Visual, Datos Multidimensionales

CONTEXTO

Esta línea de trabajo se realiza en el ámbito del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC), el Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS) y el Institute of Visual Computing & Human-Centered Technology (TU WIEN), instituciones de investigación de reconocido prestigio nacional e internacional.

1. INTRODUCCIÓN

El análisis visual ha cobrado relevancia en los últimos años debido al crecimiento exponencial de los datos generados por la sociedad, lo cual hace necesario que se deba contar con herramientas que permitan explorar y analizar dicho volumen de información. Sin embargo, un aspecto que resulta importante y ha cobrado especial interés es cómo detectar, modelar, trazar y cuantificar los distintos tipos de incertidumbre asociada. Esta tarea no resulta trivial, ya que para ciertos dominios de aplicación no existe una certeza completa sobre los datos utilizados. Esto se origina como consecuencia de datos incompletos, parámetros desconocidos, problemas de reconstrucción, entre otros. Por otro lado, no solo los datos introducen incertidumbre al problema a resolver, sino también, los expertos y/o usuarios finales a través de sus habilidades cognitivas y sesgos personales.

Esto da lugar a una necesidad de herramientas de análisis visual adecuadas que tengan en cuenta el aspecto de la incertidumbre como eje central, con el fin de poder obtener respuestas más completas a un determinado problema. Muchas de las herramientas actuales solo cubren un aspecto reducido del *pipeline* de visualización, haciendo que el análisis final no sea suficientemente completo.

Como explica Hägele [14], *"incertidumbre se refiere a la información de la cual el desacuerdo y la credibilidad pueden ser inferidos"*. Si se tiene en cuenta el ciclo de analítica visual [18] tradicional, es posible notar que cada uno de sus componentes se ve afectado por la incertidumbre, lo cual origina que la misma deba ser cuantificada, propagada y comunicada a lo largo de todo el ciclo. De forma resumida, la incertidumbre impacta en el fenómeno del mundo real bajo estudio, las observaciones y medidas realizadas, el modelado y procesamiento de los datos, las transformaciones y filtrados, el mapeo visual, la percepción y cognición de los usuarios, y finalmente las interacciones definidas.

Los desafíos necesarios para poder abarcar cada uno de los aspectos previamente mencionados requieren un esfuerzo coordinado entre los expertos de visualización y los expertos de otros campos. Si bien existen modelos de cómo la incertidumbre puede ser incorporada al ciclo de analítica visual [19], cada dominio de aplicación representa un desafío en sí mismo que debe ser abordado de forma particular. Estos modelos resultan útiles como referencias para identificar características estructurales de la incertidumbre, mientras que las cuestiones más específicas deben ser abordadas según el campo de aplicación.

Esto conlleva a que no existe una única solución para un problema que resuelva todos los requerimientos de los expertos. Por tal motivo, resulta necesario ofrecer múltiples representaciones visuales bajo incertidumbre que le permitan a los expertos contar con diferentes vistas complementarias de la información. Sumado a esto, es necesario definir técnicas de interacción que permitan manipular estas representaciones con el fin de adecuar la herramienta a las necesidades de los expertos.

No cabe duda que la incertidumbre hoy en día juega un papel fundamental en cómo los expertos de distintos campos de aplicación perciben el mundo. El desafío clave en este sentido consiste en cómo modelar y mapear visualmente un concepto complejo como la incertidumbre (en todas sus etapas), para facilitar el entendimiento de distintos problemas, desde una perspectiva más global y precisa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El diseño y desarrollo de herramientas de analítica visual es un proceso iterativo entre los expertos en visualización y los profesionales de un dominio en particular, que requiere un conocimiento profundo de los requerimientos y/o tareas que se busca resolver. La incorporación del concepto de incertidumbre como elemento principal en el proceso de visualización, agrega una dificultad adicional respecto a cómo aprovechar el manejo y gestión de la misma para guiar la toma de decisiones.
En ciertos escenarios, los resultados alcanzados pueden ser incompletos e insuficientes cuando los expertos necesitan comprender y explorar cómo la incertidumbre afecta a los datos. En este punto, es importante destacar que las nuevas herramientas de analítica visual que se propongan deberán proveer mecanismos para simplificar el análisis e impacto de la incertidumbre en cada una de las etapas del proceso de resolución del problema.

Dada la amplia variedad de dominios de aplicación donde se puede aplicar soluciones de análisis visual jerárquico considerando el aspecto de incertidumbre, es posible detallar

una serie de tareas necesarias para resolver los desafíos planteados:

1. Crear nuevas soluciones visuales que permitan modelar, gestionar, propagar, componer y trazar la incertidumbre a lo largo del todo proceso de visualización.

2. Desarrollar un conjunto de paradigmas visuales que soporten el aspecto de incertidumbre:

- Facilitar el aspecto del modelado de la incertidumbre, referido a su comportamiento jerárquico o no.
- Facilitar el aspecto vinculado a los mecanismos de manejo, gestión y trazabilidad de la incertidumbre.
- Facilitar el aspecto vinculado al manejo perceptual y cognitivo de la incertidumbre por parte de los usuarios finales.
- Proveer representaciones que permitan tener en cuenta el contexto y las tareas, y como la incertidumbre impacta en las mismas.
- Soportar múltiples niveles de datos y de abstracción de la información.
- Contextualizar los resultados y/o *insights*, encontrados mediante el uso de la solución propuesta, con respecto a la incertidumbre y cómo impacta en la toma de decisiones.

3. Desarrollar las interacciones que soportan cada uno de los aspectos que describen y caracterizan a la componente vinculada a la incertidumbre. Dado que la misma puede tener un aspecto jerárquico, es importante que dichas interacciones puedan manipular diferentes niveles de detalle de análisis para los expertos.

El análisis visual busca generar soluciones para explorar interactivamente los datos y evaluar la calidad de las mismas. La calidad de la solución se verá reflejada por la calidad de las representaciones visuales empleadas y las interacciones desarrolladas para manipular las mismas. Para su representación, los datos sufren inicialmente una serie de transformaciones que resultan convenientes tanto para los expertos de visualización como para los usuarios finales. En la mayoría de los escenarios que se buscan analizar en este proyecto, la incertidumbre es una componente que se encuentra vinculada con los datos iniciales. Esto hace que la misma sufra también transformaciones que puedan afectar la toma de decisiones.

También es importante tener en cuenta que la naturaleza multidimensional de los datos puede afectar el entendimiento de la incertidumbre y su impacto en los distintos componentes del proceso de visualización. Es por ello, que se debe trabajar para:

1. Desarrollar los elementos teóricos y prácticos adecuados para transformar los datos multidimensionales en nuevas representaciones teniendo en cuenta el componente de incertidumbre.
2. Desarrollar métodos y principios que permitan evaluar la calidad de las soluciones de análisis visual de datos multidimensionales bajo incertidumbre.

Por lo tanto, el ciclo de análisis visual debe proveer entornos donde los expertos de dominio dispongan de:

1. La posibilidad de emplear distintas técnicas de analítica visual que permita obtener *insights* basadas en incertidumbre de los datos.
2. Un marco de referencia que permita cubrir los distintos aspectos vinculados al proceso de vincular la visualización y la incertidumbre, teniendo en cuenta los aspectos analíticos.

3.RESULTADOS OBTENIDOS Y ESPERADOS

Esta línea de investigación ha comenzado a ser desarrollada en el último tiempo en conjunto entre el VyGLab y el Institute of Visual Computing and Human-Centered Technology, abordando las nuevas tendencias de análisis visual.

Empleando los trabajos anteriores realizados entre ambos grupos [3,7,8,9] y de forma individual por el VyGLab [1,2], se busca mejorar y extender las soluciones propuestas con el objetivo de adaptarlas al contexto del manejo y gestión de la incertidumbre. Además, el esquema de trabajo interdisciplinario con otros grupos de investigación como el INGEOSUR (Instituto Geológico del Sur) y el LDNC (Laboratorio de Desarrollo en Neurociencias Cognitivas) permitirá evaluar y aplicar en distintos dominios los distintos resultados alcanzados durante el desarrollo de la línea de investigación.

4. FORMACION DE RECURSOS HUMANOS

Proyectos de Investigación

-PGI 24/ZN42 *Análisis Visual de Datos Multidimensionales*. Directora: Dra. M. Luján Ganuza.

Becas

-Leandro Luque. *Desarrollo de una plataforma de análisis visual de datos multidimensionales y espacio-temporales asistida por modelos generativos.* Beca postdoctoral CONICET. Adjudicada a partir de agosto de 2025. Directora: Dra. Silvia Castro

-Antonella Antonini. *Análisis visual guiado para datos multidimensionales.* Beca postdoctoral CONICET. Adjudicada a partir de agosto de 2025. Directora: Dra. Silvia Castro

5. BIBLIOGRAFÍA

[1] Antonini, A. S., Luque, L., Ferracutti, G., Bjerg, E. A., Castro, S. M., & Ganuza, M. L. (2024). SpinelVA. A new perspective for the visual analysis and classification of spinel group minerals. Earth Science Informatics, 17(4), 3851–3861.

[2] Antonini, A. S., Tanzola, J., Asiain, L., Ferracutti, G., Castro, S. M., Bjerg, E. A., & Ganuza, M. L. (2024). Machine Learning model interpretability using SHAP values: Application to Igneous Rock Classification task. Applied Computing and Geosciences, 23, 100178.

[3] Antonini, A. S., Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Matković, K., Gröller, E., Bjerg, E. A., & Castro, S. M. (2021). Spinel web: An interactive web application for visualizing the chemical composition of spinel group minerals. Earth Science Informatics, 14(1), 521–528.

[4] Correa, C. D., Chan, Y.-H., & Ma, K.-L. (2009). A framework for uncertainty-aware visual analytics. 2009 IEEE Symposium on Visual Analytics Science and Technology, 51–58.

[5] Dasgupta, A., Chen, M., & Kosara, R. (2012). Conceptualizing visual uncertainty in parallel coordinates. In Computer Graphics Forum (Vol. 31, No. 3pt2, pp. 1015-1024). Oxford, UK: Blackwell Publishing Ltd.

[6] Evers, M., Huesmann, K., & Linsen, L. (2021). Uncertainty‐aware Visualization of Regional Time Series Correlation in Spatio‐temporal Ensembles. In Computer Graphics Forum (Vol. 40, No. 3, pp. 519-530).

[7] Ganuza, M. L., Castro, S. M., Ferracutti, G., Bjerg, E. A., & Martig, S. R. (2012). SpinelViz: An interactive 3D application for visualizing spinel group minerals. Computers & Geosciences, 48, 50–56.

[8] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., K. & Matkovic, K., (2017). Interactive Visual Categorization of Spinel-Group Minerals. 33rd Spring Conference on Computer Graphics, pp. 1-10. Mikulov, República Checa, 15-17.

[9] Ganuza, M.L., Ferracutti, G., Gargiulo, M.F., Castro, S., Bjerg, E. A., Matkovic, K., & Gröller, E., (2014). The spinel explorer - interactive visual analysis of spinel group minerals. IEEE Transactions on Visualization and Computer Graphics, 20(12): 1912-1922. ISSN: 1077-2626.

[10] Gillmann, C., Maack, R. G. C., Raith, F., Pérez, J. F., & Scheuermann, G. (2023). A taxonomy of uncertainty events in visual analytics. *IEEE Computer Graphics and Applications*, *43*(5), 62-71

[11] Gschwandtner, T., Bögl, M., Federico, P., & Miksch, S. (2016). Visual Encodings of Temporal Uncertainty: A Comparative User Study. IEEE Transactions on Visualization and Computer Graphics, 22(1), 539–548.

[12] Görtler, J., Schulz, C., Weiskopf, D., & Deussen, O. (2017). Bubble treemaps for uncertainty visualization. IEEE Transactions on Visualization and Computer Graphics, *24*(1), 719-728.

[13] Görtler, J., Spinner, T., Streeb, D., Weiskopf, D., & Deussen, O. (2019). Uncertainty-aware principal component analysis. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 822-831.

[14] Hägele, D., Schulz, C., Beschle, C., Booth, H., Butt, M., Barth, A., Deussen, O., & Weiskopf, D. (2022). Uncertainty visualization: Fundamentals and recent developments. It - Information Technology, 64(4–5), 121–132.

[15] Hullman, J. (2020). Why Authors Don't Visualize Uncertainty. IEEE Transactions on Visualization and Computer Graphics, 26(1), 130–139.

[16] Hullman, J., Qiao, X., Correll, M., Kale, A., & Kay, M. (2019). In Pursuit of Error: A Survey of Uncertainty Visualization Evaluation.
ansactions on Visualization and Computer Graphics, 25(1), 903–913.

[17] Karami, A. (2015, November). A framework for uncertainty-aware visual analytics in big data. In *CEUR Workshop Proceedings* (Vol. 1510, pp. 146-155). CEUR Workshop Proceedings.

[18] Keim, D.A., Mansmann,F ., Schneidewind, J., Thomas, J., & Ziegler,H.: Visual analytics: scope and challenges. In: Visual data mining (2008)

[19] Maack, R. G. C., Raith, F., Pérez, J. F., Scheuermann, G., & Gillmann, C. (2025). A workflow to systematically design uncertainty-aware visual analytics applications. The Visual Computer, 41(3), 1485–1498.

[20] Maack, R. G. C., Scheuermann, G., Hagen, H., Peñaloza, J. T. H., & Gillmann, C. (2023). Uncertainty-aware visual analytics: Scope, opportunities, and challenges. The Visual Computer, 39(12), 6345–6366.

[21] Maack, R. G., Raymer, M. L., Wischgoll, T., Hagen, H., & Gillmann, C. (2021). A framework for uncertainty-aware visual analytics of proteins. *Computers & Graphics*, *98*, 293-305.

[22] MacEachren, A. M., Roth, R. E., O'Brien, J., Li, B., Swingley, D., & Gahegan, M. (2012). Visual semiotics & uncertainty visualization: An empirical study. IEEE Transactions on Visualization and Computer Graphics, *18*(12), 2496-2505.

[23] Matzen, L. E., Howell, B. C., Tuft, M., & Divis, K. M. (2024). Transparent risks: The impact of the specificity and visual encoding of uncertainty on decision making. In Computer Graphics Forum (Vol. 43, No. 3, p. e15094).

[24] Musleh, M., Ceneda, D., Ehlers, H., & Raidou, R. G. (2025, February). ConAn: Measuring and Evaluating User Confidence in Visual Data Analysis Under Uncertainty. In *Computer Graphics Forum* (Vol. 44, No. 1, p. e15272).

[25] Musleh, M., Muren, L. P., Toussaint, L., Vestergaard, A., Gröller, E., Raidou, R. G. (2023). Uncertainty guidance in proton therapy planning visualization. Computers & Graphics, 111, 166-179.

[26] Padilla, L., Kay, M., & Hullman, J. (2022). Uncertainty visualization. Computational Statistics in Data Science. Wiley.

[27] Peischl, J. and Raidou, R. (2025). Investigating the Propagation of CT Acquisition Artifacts along the Medical Imaging Pipeline. In Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - IVAPP; ISBN 978-989-758-728-3; ISSN 2184-4321, SciTePress, pages 752-764

[28] Potter, K., Kniss, J., Riesenfeld, R., & Johnson, C. R. (2010). Visualizing summary statistics and uncertainty. In Computer Graphics Forum (Vol. 29, No. 3, pp. 823-832). Oxford, UK: Blackwell Publishing Ltd.

[29] Pöthkow, K., & Hege, H. C. (2013). Nonparametric models for uncertainty visualization. In Computer Graphics Forum (Vol. 32, No. 3pt2, pp. 131-140). Oxford, UK: Blackwell Publishing Ltd.

[31] Sacha, D., Senaratne, H., Kwon, B. C., Ellis, G., & Keim, D. A. (2015). The role of uncertainty, awareness, and trust in visual analytics. IEEE Transactions on Visualization and Computer Graphics, *22*(1), 240-249.

[32] Senaratne, H. (2017). Uncertainty-aware visual analytics for spatio-temporal data exploration. PhD Thesis. Universität Konstanz. Germany

[33] Skeels, M., Lee, B., Smith, G., & Robertson, G. G. (2010). Revealing uncertainty for information visualization. Information Visualization, 9(1), 70–81.

[34] Tastekin, O. (2025). *Uncertainty-aware PCA for nonnormally distributed data*. PhD Thesis. University of Stuttgart, Germany.

[35] Tominski, C., Behrisch, M., Bleisch, S., Fabrikant, S. I., Mayr, E., Miksch, S., & Purchase, H. (2023). Visualizing uncertainty in sets. *IEEE Computer Graphics and Applications*, *43*(5), 49-61.

[36] Weiskopf, D. (2022). Uncertainty visualization: Concepts, methods, and applications in biological data visualization. Frontiers in Bioinformatics, 2, 793819.

APLICACIÓN DE REDES CNN PARA LA DETECCIÓN DE ÓRGANOS REPRODUCTIVOS DE KIWI EN IMÁGENES Y VIDEOS

Silvia N. Pérez[1], Dante Mendoza[1], Alejandra Yommi[2], M. Angela David[2],

Jorge Lozano[2], Camila Pagano[1]

[1]Instituto de Ingenierías y Nuevas Tecnologías, Universidad Nacional del Oeste
[2]Instituto Nacional de Tecnología Agropecuaria. Estación Experimental Balcarce

RESUMEN

La aplicación de técnicas de aprendizaje profundo en el ámbito agrícola ha experimentado un notable crecimiento, impulsado por la necesidad de optimizar la producción y mejorar la eficiencia en la gestión de cultivos. En particular, los algoritmos de detección de objetos basados en redes neuronales convolucionales (CNN) se han consolidado como una herramienta de valor para el monitoreo en sistemas frutihortícolas.

Este trabajo continúa una línea de investigación desarrollada en proyectos previos, orientada a la detección automática de órganos vegetativos y reproductivos en plantas de kiwi a partir de imágenes. Se tiene como objetivo evaluar el desempeño de modelos de detección previamente entrenados sobre imágenes de alta calidad, al ser aplicados a imágenes capturadas con teléfonos móviles. Esto permitirá analizar su robustez frente a variaciones en la calidad de adquisición y reducir los costos asociados al equipamiento de captura.

Se propone además extender el enfoque a secuencias de video, incorporando arquitecturas de detección en tiempo real con el fin de avanzar hacia sistemas de monitoreo continuo de cultivos. Esta estrategia permitiría la detección y el conteo dinámico de órganos vegetativos durante recorridos en la plantación, sentando las bases para aplicaciones prácticas de agricultura de precisión en escenarios reales de campo.

Palabras clave: *Visión artificial. Redes neuronales convolucionales. Producción frutihortícola*

CONTEXTO

Esta presentación es parte de la investigación llevada a cabo en el proyecto "Aplicación de modelos de aprendizaje profundo en imágenes para la estimación de producción frutihortícola", proyecto acreditado por la Universidad Nacional del Oeste (UNO) y desarrollado en colaboración con un equipo del Instituto Nacional de Tecnología Agropecuaria (INTA-Balcarce) especializado en producción de kiwi.

1. INTRODUCCIÓN

En la actividad frutícola, una estimación precisa de la cosecha es crucial para la planificación y gestión agrícola. Agricultores y cooperativas dependen de esta información para optimizar la recolección, el almacenamiento y la distribución de los productos. Tradicionalmente, estas estimaciones se realizaban de forma manual, lo que introducía un margen de error importante (Apolo et al, 2020). En la actualidad las redes neuronales convolucionales, basadas en el aprendizaje profundo, ofrecen una alternativa más precisa y objetiva al permitir el reconocimiento y conteo automático de frutos (Williams et al, 2019).

Particularmente en la producción de kiwi, que es intensiva en mano de obra y en prácticas realizadas mayormente de modo manual, se requiere estimar el número de botones florales, así como también el volumen de la fruta cosechada al contar los frutos por unidad de

superficie en una etapa avanzada de desarrollo de los mismos (David et al, 2020). Estos procedimientos apuntan a mejorar el manejo de la plantación por lo que resulta de interés introducir herramientas que permitan automatizar la operatoria.

La línea de investigación apunta a desarrollar un sistema basado en el procesamiento de imágenes que, a través de un dispositivo acoplable a un vehículo, capture imágenes al recorrer una plantación y entregue al productor resultados de la detección de tales órganos frutales, permitiendo de este modo ajustar la carga frutal y estimar su producción tempranamente.

En un proyecto previo se desarrollaron modelos CNN para contar órganos vegetativos en imágenes obtenidas en sitios prefijados de una plantación de kiwi. Dichos conteos se compararon con conteos in situ evaluando la factibilidad de reemplazar el conteo manual por los conteos automáticos provistos por los modelos. Los resultados fueron auspiciosos y dieron lugar a nuevos objetivos, comprendidos en el proyecto actual y presentados en las próximas secciones.

1.1. Detección de botones florales en imágenes

Una de las etapas importantes en que se realizan los conteos de campo con el fin de estimar la cosecha, es la etapa de desarrollo de botones florales. En investigaciones previas se probaron y evaluaron modelos preentrenados para detección de objetos de la API Tensorflow 2.8.1, resultando elegida para la detección de botones florales la familia Faster R-CNN Inception ResNet V2. Tanto el entrenamiento como la validación se realizó con imágenes de la plantación tomadas por una fotógrafa profesional, con condiciones controladas de luz y distancia a la canopia. Para la comparación con el conteo de campo, se fijaron marcos en sitios especificados donde se realizó el conteo manual y se tomaron imágenes de calidad profesional. Al mismo tiempo, se tomaron también imágenes con un teléfono móvil, de menor resolución, con tamaño promedio de 3MB versus 10 MB promedio de las imágenes profesionales.

En esta etapa se propone evaluar la detección en imágenes de teléfonos móviles, utilizando como base los modelos ajustados previamente. Esto permitiría reducir costos en el equipamiento del dispositivo de captura de imágenes. Dado que las redes neuronales convolucionales del tipo Faster R-CNN pueden presentar un rendimiento de detección insatisfactorio cuando el objeto presenta problemas de oclusión, deformación o tamaño pequeño (Xiao et al, 2020), un objetivo a seguir es comparar la performance de los modelos desarrollados en nuevas imágenes de menor calidad.

Como ensayo preliminar, se consideraron imágenes obtenidas utilizando teléfonos móviles en la campaña 2023-2024 sobre las que se aplicaron los modelos desarrollados en las campañas anteriores. En esa campaña se obtuvieron también imágenes de la fotógrafa profesional, por lo que puede compararse la detección de un mismo modelo sobre imágenes de distinta prestación de un sitio específico. A modo de ejemplo, las Fig. a) y b) muestran la detección de botones florales en imágenes sobre un mismo sitio, la primera tomada con teléfono móvil y la segunda la imagen de mayor calidad tomada por fotógrafa. Se observa que la detección da resultados similares, lo que va en dirección de los objetivos propuestos. Se plantea entonces mejorar las CNN ajustadas, agregando imágenes de este tipo al conjunto de entrenamiento y empleando estrategias como aumento artificial de datos y combinación de metodologías de entrenamiento.



Fig.1 a): Detección en imagen de teléfono móvil.



Fig.1 b): Detección en imagen de mayor calidad.

1.2. Detección de botones florales en videos

El desarrollo de sistemas de visión artificial para la detección de objetos en tiempo real ha cobrado gran relevancia en los últimos años.

La detección en video con este tipo de sistemas da un gran avance para la monitorización continua de cultivos, al permitir un seguimiento en tiempo real de su evolución. Una de las herramientas más avanzados y ampliamente adoptados para este tipo de detección es la familia de algoritmos YOLO (por sus siglas en inglés: You Only Look Once), que ha demostrado su efectividad en la detección de objetos en tiempo real. A diferencia de métodos tradicionales basados en ventanas deslizantes o propuestas de regiones, YOLO ofrece la capacidad de procesar imágenes completas en una única pasada, mejorando considerablemente la velocidad de detección sin sacrificar la precisión (Redmon, 2016). Las versiones más estables, como YOLOv8 y YOLOv10, han introducido mejoras en términos de eficiencia computacional y capacidad de generalización, haciéndolas especialmente adecuadas para aplicaciones en el campo agrícola.

Investigaciones previas han explorado el uso de YOLO en diversas áreas de la agricultura, como la detección de plagas, la clasificación de cultivos y la estimación de rendimientos. Por ejemplo, estudios recientes han demostrado la capacidad de modelos como YOLOv4 y YOLOv5 para identificar diferentes tipos de frutas en imágenes capturadas desde drones o cámaras fijas, alcanzando precisiones superiores al 90% en la identificación de cultivos como manzanas, uvas y tomates (Ziyue, Bo y Boyun, 2024).

Adaptaciones de YOLOv3, como OrangeYolo, creada para la detección de cítricos y YOLO-Tomato, para la detección en plantaciones de tomates, han demostrado ser altamente eficaces en la detección y conteo de frutas en condiciones complejas de campo, incluso en escenarios de oclusión de frutas. La incorporación de redes densas y técnicas de pooling espacial ha permitido que estos modelos sean más robustos ante variaciones en el entorno (Zhang et al, 2022). Estas investigaciones destacan que los modelos avanzados de IA no solo mejoran la precisión de la estimación del rendimiento, sino que también optimizan los recursos agrícolas, haciendo que la producción sea más eficiente y sostenible.

En este contexto, la combinación de YOLO con técnicas de seguimiento de objetos ha mostrado resultados prometedores, permitiendo la cuantificación dinámica y el análisis de crecimiento a lo largo del ciclo productivo. Este enfoque minimiza los errores de duplicación en la detección y ofrece una visión más precisa del estado de los cultivos en entornos naturales.

En esta investigación se propone desarrollar modelos de detección de objetos en secuencias de video, orientados a aplicaciones de monitoreo en tiempo real, empleando arquitecturas de la familia YOLO. Para el desarrollo e implementación de los modelos se utilizarán herramientas de código abierto en el lenguaje de programación Python, apoyadas en bibliotecas de visión por computador y aprendizaje profundo que facilitan el entrenamiento, la evaluación y la inferencia.

Además, se emplearán conjuntos de datos de video obtenidos en una plantación de kiwi para evaluar la detección de órganos vegetativos en secuencias temporales, al contrastar los resultados con las estimaciones de conteo obtenidas a partir de imágenes estáticas. Esta comparación permite analizar las ventajas y limitaciones de cada enfoque, así como su potencial aplicación en escenarios reales de monitoreo agrícola.

2. LINEAS DE INVESTIGACION Y DESARROLLO

- Aplicaciones de redes neuronales convolucionales.
- Detección de objetos en imágenes y videos.
- Estimación de producción frutihortícola.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los modelos CNN entrenados sobre imágenes de alta calidad y aplicados sobre imágenes de este tipo sobre sitios especificados, lograron una buena aproximación a los obtenidos en campo en las investigaciones previas (Pérez et al, 2024). En la Tabla 1 se muestra promedio y desvío de los errores relativos de los conteos automáticos contrastados con el conteo de campo.

Tabla 1: Errores relativos promedio y desvío estándar

	Campaña 2022	Campaña 2023
Error relativo	-0.053	0.004
Desvío estándar	0.25	0.31

Tales resultados fueron auspiciosos y motivaron ensayos preliminares aplicando las redes CNN desarrolladas a imágenes capturadas con teléfonos móviles durante la campaña 2023–2024. Los resultados observados sugieren que los modelos de detección mantienen un desempeño consistente al ser aplicados sobre imágenes de menor calidad, al lograr identificar los botones florales de forma visualmente comparable a las detecciones obtenidas en imágenes profesionales. Esto permite esperar una adecuada capacidad de generalización del modelo, lo que resulta relevante para su aplicación en escenarios reales de campo y con equipamiento de menor costo.

En relación con la detección en secuencias de video, se realizaron pruebas iniciales sobre registros audiovisuales de recorridos en la plantación, observándose la factibilidad de aplicar modelos de detección en tiempo real para la identificación dinámica de órganos vegetativos. Estos ensayos permiten validar la viabilidad del enfoque propuesto y constituyen un punto de partida para el desarrollo de sistemas de monitoreo continuo basados en video.

Como resultados esperados, se prevé mejorar el desempeño de los modelos mediante la incorporación de imágenes de menor calidad al conjunto de entrenamiento, el uso de técnicas de aumento artificial de datos y la evaluación comparativa de distintas arquitecturas de detección. Asimismo, se espera avanzar en la cuantificación objetiva de los resultados mediante métricas de precisión y comparación con conteos manuales de campo. A largo plazo, se proyecta el desarrollo de un sistema de visión artificial que asista al productor en la estimación temprana de la producción de kiwi, al contribuir a la optimización del manejo del cultivo.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto se compone de dos investigadores del Instituto de Ingenierías y Nuevas Tecnologías de la UNO, a cargo de las aplicaciones de tecnologías de aprendizaje profundo aquí presentadas, y tres investigadores de INTA-EE Balcarce quienes

se encargan de la supervisión en la plantación y análisis de campo.

Se suma a estos Camila Pagano, alumna de Informática de la UNO, quien desarrolla su plan de trabajo de Beca de Estímulo a las Vocaciones Científicas. Se prevé incorporar alumnos de la Licenciatura en Informática de la UNO para su formación en el área de aplicaciones de *deep learning*.

5. BIBLIOGRAFÍA

Apolo, O. E. A., Guanter, J. M., Cegarra, G. E., Raja, P., & Ruiz, M. P. (2020). Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. *European journal of agronomy: the official journal of the European Society for Agronomy*, *115*(4), 183-194.

David, M.A., Yommi, A., Sanchez, E. (2020). *Elección del terreno y plantación del cultivo de kiwi.* 1a ed. Balcarce, Buenos Aires: Ediciones INTA. Libro digital, 39 p.

Pérez, S. N., Yommi, A., David, M. A.,Lozano, J., Massacesi, A.., Mendoza, D. H. (2024). *Inteligencia artificial aplicada al conteo de órganos vegetales en plantas de kiwi.* Segunda Jornada AgTech. UNMDP-INTA. Balcarce, abril 2024.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788.

Williams, H. A. M., Jones, M. H., Nejati, M., Seabright, M. J., Bell, J., Penhall, N. D., Barnett, J. J., Duke, M. D., Scarfe, A. J., Ahn, H. S., Lim, J., & MacDonald, B. A. (2019). Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. Biosystems Engineering, 181, 140-156. https://doi.org/10.1016/j.biosystemseng.2019.03.007

Xiao, Y., Wang, X., Zhang, P., Meng, F., & Shao, F. (2020). Object Detection Based on Faster R-CNN Algorithm with Skip Pooling and Fusion of Contextual Information. *Sensors*, *20*(19), 5490. https://doi.org/10.3390/s20195490

Zhang, W.; Wang, J.; Liu, Y.; Chen, K.; Li, H.; Duan,Y.; Wu, W.; Shi, Y.; Guo, W. (2022) Deep-learning-based in-field citrus fruit detection and tracking, Horticulture Research, Volume 9, 2022, uhac003, https://doi.org/10.1093/hr/uhac003

Ziyue, J., Bo, Y., & Boyun, L. (2024). *Precise Apple Detection and Localization in Orchards using YOLOv5 for Robotic Harvesting Systems*. arXiv preprint arXiv:2405.06260.

Avances en Experiencias de Realidad Extendida Multisensorial Accesible

R. Guerrero[(1)], M. J. Abásolo[(2)(3)], J. Benitez[(2)], F. Puricelli[(2)]
(1) Laboratorio de Computación Gráfica / Dpto. Informática / FCFMyN
Universidad Nacional de San Luis
rag@unsl.edu.ar
(2) Facultad de Informática-Universidad Nacional de La Plata
(3) Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
mjabasolo@lidi.info.unlp.edu.ar

RESUMEN

El desarrollo reciente de las tecnologías de Realidad Extendida (XR) ha permitido la creación de entornos tridimensionales interactivos con alta fidelidad visual y auditiva. No obstante, la experiencia inmersiva continúa centrada en estos canales, dejando subrepresentadas otras dimensiones perceptuales clave para la construcción de la presencia. En este marco, el olfato adquiere relevancia por su vínculo con la memoria, la emoción y la orientación espacial, lo que evidencia su potencial para enriquecer el realismo y la conexión cognitiva en entornos virtuales.

A pesar de los avances en dispositivos de emisión aromática, persisten limitaciones vinculadas a su integración en motores de interacción en tiempo real, la sincronización de estímulos y la ausencia de estándares de representación digital. Frente a este escenario, la investigación propone explorar la convergencia entre plataformas XR abiertas e interfaces multisensoriales programables.

Se proponen dos líneas complementarias: el desarrollo de plataformas XR web accesibles escalables, y el diseño de interfaces olfativas portátiles sincronizadas con eventos virtuales. Su articulación busca avanzar hacia sistemas inmersivos multisensoriales que mejoren la experiencia, amplíen su aplicabilidad en educación y simulación, y acerquen la interacción virtual a formas perceptuales más próximas a la experiencia humana.

Palabras clave: Realidad Extendida, Multimodalidad, Interacción Olfativa, Framework XR Web, Computación Gráfica.

CONTEXTO

La presente investigación se desarrolla en el marco de una colaboración estratégica entre la Universidad Nacional de San Luis (UNSL) y la Universidad Nacional de La Plata (UNLP), orientada a la formación de recursos humanos en la carrera de Especialización en Computación Gráfica e Imágenes Visuales (CGIV). Esta articulación interinstitucional permite abordar la multimodalidad en Realidad Extendida (XR) desde dos frentes complementarios: la creación de frameworks de visualización avanzados y el desarrollo de hardware sensorial periférico, garantizando una formación integral y multidisciplinaria de los especialistas.

Por parte de la UNSL, las actividades se inscriben en el proyecto “La Realidad Extendida en la Comunicación de Información para el Metaverso”, con base en el Laboratorio de Computación Gráfica, donde se exploran nuevas interfaces de interacción. En paralelo, desde la Facultad de Informática de la UNLP, el trabajo se integra a una trayectoria consolidada en Realidad

Virtual, Aumentada y Sistemas Híbridos, actualmente bajo el ala de los proyectos: “F031 - Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual”, dirigido por C. Sanz; “F032 - Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, Fog-Edge-Cloud, Big Data, Robótica, y Tiempo Real”, dirigido por M.Naiouf.

1. INTRODUCCIÓN

Las tecnologías de Realidad Extendida (XR), que comprenden la Realidad Virtual, la Realidad Aumentada y la Realidad Mixta, han experimentado un crecimiento sostenido impulsado por avances en dispositivos de visualización, procesamiento gráfico y plataformas de desarrollo interactivo [1]. Estos entornos permiten actualmente la interacción tridimensional, la navegación espacial y la simulación de escenarios complejos. Sin embargo, su evolución ha estado centrada predominantemente en la estimulación visual y auditiva, mientras que otros canales perceptuales fundamentales para la experiencia humana han permanecido subrepresentados. Desde el campo de la interacción humano-computadora se sostiene que la percepción del entorno emerge de la integración simultánea de múltiples estímulos sensoriales, por lo que la inmersión puede entenderse como el grado en que un sistema logra generar en el usuario la sensación de presencia en un entorno distinto del físico inmediato. Tradicionalmente, esta sensación ha sido promovida mediante la fidelidad visual, el audio espacial y, en menor medida, la retroalimentación háptica. No obstante, investigaciones recientes indican que los sentidos operan de manera interdependiente, conformando experiencias perceptuales unificadas.

Dentro de esta perspectiva, el olfato adquiere particular relevancia debido a su estrecha vinculación con la memoria episódica, la emoción y la orientación espacial, ya que los estímulos olfativos poseen la capacidad de activar asociaciones cognitivas profundas, contribuyendo a incrementar el realismo de las experiencias, mejorar la memoria contextual y fortalecer la conexión emocional con el entorno virtual. En consecuencia, la incorporación de estímulos sensoriales adicionales se presenta como una vía para aproximar las experiencias inmersivas a formas de percepción más cercanas a la experiencia natural [2].

El desarrollo de entornos inmersivos ha sido posible gracias al avance de motores de creación como *Unity* y *Unreal Engine*, así como a la expansión del ecosistema XR impulsada por dispositivos promovidos por organizaciones como *Meta Platforms*, *HTC* y *Sony*, lo que ha permitido el pasaje desde interfaces bidimensionales hacia experiencias tridimensionales interactivas y ha consolidado nuevos modos de interacción digital [3-8]. Paralelamente, el diseño de experiencias inmersivas ha evolucionado desde enfoques centrados en la simulación hacia paradigmas orientados a la vivencia, mientras producciones culturales como *Ready Player One* han contribuido a instalar en el imaginario social la posibilidad de entornos virtuales plenamente sensoriales [9].

En la actualidad, la investigación en entornos XR ha comenzado a explorar la incorporación de modalidades sensoriales complementarias, tales como interfaces hápticas, simulaciones térmicas y estímulos olfativos concebidos como componentes contextuales de la experiencia. Se han desarrollado prototipos de dispositivos de emisión olfativa orientados a simuladores de entrenamiento, experiencias museográficas, aplicaciones terapéuticas y entornos de entretenimiento, los cuales han demostrado que los estímulos olfativos mantienen una fuerte asociación con la memoria episódica, la evocación emocional y la percepción de realismo ambiental. No obstante, el campo presenta limitaciones estructurales significativas: la mayoría de los dispositivos olfativos funcionan como periféricos aislados, sin una integración profunda con los motores de interacción, mientras que la emisión,

persistencia y disipación de aromas introduce problemas de latencia y sincronización que dificultan su alineación temporal con eventos virtuales. A ello se suma la escasa estandarización existente para la representación digital de estímulos olfativos y las limitaciones de escalabilidad que mantienen a muchos sistemas en fases experimentales.

En este escenario, persiste una falta de articulación entre interfaces olfativas programables, motores de interacción en tiempo real, arquitecturas abiertas accesibles y soluciones técnicamente escalables, lo que evidencia que el desafío actual no se limita a la generación de estímulos sensoriales, sino a su integración coherente dentro de ecosistemas interactivos capaces de sincronizar eventos, modular intensidades y sostener experiencias reproducibles. Paralelamente, la democratización de las tecnologías XR ha impulsado el desarrollo de plataformas basadas en la web que reducen barreras de acceso y permiten despliegues multiplataforma, aunque la incorporación de nuevas interfaces sensoriales dentro de estos entornos accesibles continúa siendo limitada. En este contexto, se vuelve relevante avanzar hacia modelos que integren plataformas abiertas con dispositivos sensoriales programables, con el objetivo de mejorar la calidad de la experiencia inmersiva, ampliar su aplicabilidad y favorecer su adopción en contextos educativos, formativos y de simulación, proyectando así el desarrollo de sistemas inmersivos multisensoriales accesibles como una línea estratégica orientada a entornos virtuales más naturales, significativos y transferibles al mundo real.

En este contexto, la investigación se sitúa en la convergencia entre tecnologías XR abiertas, interacción multisensorial e interfaces olfativas. Se parte de la hipótesis de que la integración de plataformas accesibles con dispositivos sensoriales programables puede mejorar la calidad de la experiencia inmersiva, ampliar su aplicabilidad y favorecer su adopción en contextos educativos, formativos y de simulación.

De este modo, el desarrollo de sistemas inmersivos multisensoriales accesibles se proyecta como una línea estratégica orientada a avanzar hacia entornos virtuales más naturales, significativos y transferibles al mundo real.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El desarrollo contemporáneo de entornos inmersivos plantea nuevos desafíos vinculados no solo a la calidad de la simulación visual, sino también a la posibilidad de ampliar las formas de interacción hacia experiencias más accesibles, integradas y perceptualmente enriquecidas. En este contexto, se presentan dos líneas de investigación complementarias que abordan, desde perspectivas tecnológicas y experienciales, la evolución de las plataformas XR hacia modelos abiertos y multisensoriales.

Plataformas XR Web Accesibles y Escalables

Esta línea se orienta al estudio y desarrollo de plataformas XR web accesibles y escalables, con el objetivo de posibilitar la creación y despliegue de experiencias inmersivas de bajo costo que prioricen la interoperabilidad, la facilidad de uso y la visualización en múltiples modos (como entornos tridimensionales, realidad virtual o realidad mixta) sin requerir software propietario. En este marco, se investigará la evolución de frameworks XR hacia arquitecturas capaces de incorporar interacción multisensorial sin afectar su accesibilidad ni su capacidad de expansión [1,10,11].

Interfaces Multisensoriales y Dispositivos Olfativos

Esta línea se centra en el diseño e integración de interfaces multisensoriales, particularmente mediante el desarrollo de dispositivos portátiles de emisión aromática que puedan sincronizarse con eventos

virtuales, modular su intensidad en función del contexto espacial y actuar como guías perceptuales o indicadores de estado dentro del entorno. Esta línea también abordará el estudio del olfato como canal de interacción en entornos virtuales, evaluando su impacto en la experiencia del usuario y en procesos vinculados al aprendizaje o la toma de decisiones [12,13,14].

A partir de la articulación entre ambas líneas, se espera avanzar en la construcción de una plataforma XR multisensorial basada en tecnologías web abiertas, junto con un prototipo optimizado de interfaz olfativa portable y un modelo de integración entre eventos virtuales y estímulos sensoriales. Asimismo, se prevé generar evidencia empírica sobre el impacto de la multisensorialidad y explorar casos de uso en ámbitos educativos y de simulación.

3. RESULTADOS

La investigación planteada presenta dos avances complementarios en el ámbito de la computación inmersiva:

Hardware Sensorial: Se realizó el diseño, la implementación y la validación de un dispositivo de hardware emisor de aromas, diseñado específicamente para su integración en aplicaciones móviles. Este periférico permite expandir la respuesta sensorial de los entornos virtuales, aportando una capa de interacción olfativa controlada.

Arquitectura de Software: Se realizó el desarrollo de un framework de software orientado a la creación y gestión de presentaciones visuales en entornos de Realidad Extendida (XR). Este framework abstrae la complejidad técnica de la visualización inmersiva, facilitando la generación de contenidos dinámicos y adaptativos.

A futuro se espera poder alcanzar alguno de los siguientes objetivos:

- Prototipo optimizado de interfaz olfativa portable.
- Plataforma XR multisensorial basada en tecnologías web abiertas.
- Modelo de integración entre eventos virtuales y estímulos sensoriales.
- Evidencia empírica sobre el impacto de la multisensorialidad.
- Casos de uso en educación y simulación

4. FORMACIÓN DE RECURSOS HUMANOS

Participan dos investigadores formados, uno de la UNLP y otro de la UNSL, que codirigen a dos alumnos de la Especialización en CGIV que han completado durante el 2025 los siguientes trabajos:

- Julián Benitez "*Diseño de dispositivo emisor de aromas y su integración en un juego de realidad virtual inmersiva*" [15].
- Fernando Puricelli "*Diseño de un Framework de Realidad Mixta de bajo costo para Entornos de Instrucción*" [16].

5. BIBLIOGRAFÍA

1. Speicher, M., Hall, B. D., & Nebeling, M. (2019). What is Mixed Reality? Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1-15. https://doi.org/10.1145/3290605.3300767
2. Y. A. Yarzagaray De León. (2023). Metaverso y sus implicaciones en la comunicación.
3. Unity Industry. (s. f.). https://unity.com/industry
4. Epic. (s. f.). La herramienta de creación 3D en tiempo real más potente. Unreal Engine. Recuperado 23 de febrero de 2025, de https://www.unrealengine.com/es-ES/home
5. Meta Quest Pro. 2025, https://www.sciencefocus.com/reviews/meta-quest-pro-review, last accessed 2025/06/07.
6. Borges, Miguel & Symington, Andrew & Coltin, Brian & Smith, Trey & Ventura,

Rodrigo. 2018. HTC Vive: Analysis and Accuracy Improvement. 2610-2615. 10.1109/IROS.2018.8593707.
7. Sony. 2025, https://itc.ua/es/noticias/sony-ha-mostrado-un-concepto-de-tecnologia-que-anade-olores-reales-a-los-juegos-en-el-ejemplo-de-the-last-of-us/, last accessed 2025/06/07.
8. Khronos, G. (2019). OpenXR - High-performance access to AR and VR —collectively known as XR— platforms and devices. The Khronos Group. https://www.khronos.org/OpenXR/
9. Spielberg, S. (Director). (2018). Ready Player One [Película]. Warner Bros. Pictures; Amblin Entertainment; Village Roadshow Pictures.
10. Wani, S. S., Kamble, S. B., & Raman, R. (2020). A Study and comparative analysis of development tools for Virtual Reality applications. 2020 Seventh International Conference on Information Technology Trends (ITT), 212-217. https://doi.org/10.1109/ITT51279.2020.9320893.
11. Bauer, J., & Ebert, A. (2015). Mobile Devices for Virtual Reality Interaction. A Survey of Techniques and Metaphors. En Virtual Realities: International Dagstuhl Seminar, Dagstuhl Castle, Germany, June 9-14, 2013, Revised Selected Papers (pp. 91-107). Springer International Publishing.
12. Archer NS, Bluff A, Eddy A, Nikhil CK, Hazell N, Frank D, et al. 2022. Odour enhances the sense of presence in a virtual reality environment. PLoS ONE 17(3): e0265039,
13. Garcia-Ruiz, M. A., Kapralos, B., & Rebolledo-Mendez, G. 2021. An Overview of Olfactory Displays in Education and Training. Multimodal Technologies and Interaction, 5(10), 64.
14. Valerie Ann Worwood.The Complete Book of Essential Oils and Aromatherapy, Revised and Expanded (25th Anniversary Edition) (2016). ISBN: 978-1-57731-139-3
15. Benítez, JS, Abásolo, MJ y Guerrero, RA (2025) Design of an Scent-Emitting Device and Its Integration into an Immersive Virtual Reality Game. En Conference on Cloud Computing, Big Data & Emerging Topics (pp. 171-186). Cham: Springer Nature, Suiza.
16. Puricelli, F, Abásolo, MJ y Guerrero, RA (2026) Framework liviano y abierto para el desarrollo de Presentaciones interactivas de Realidad Extendida. Descripción, desarrollo y despliegue basado en tecnologías web de código abierto.. En XIV Iberoamerican Conference on Applications and Usability of Interactive Media jAUTI2025, Universidad de Aveiro, (en prensa)

INTEGRACIÓN DE TECNOLOGÍAS INMERSIVAS CON EMERGENTES

Y. Alvarado[(1)], N. Jofré[(1)], M. Lopresti[(1)], N. Miranda[(1)], G. Rodríguez[(1)], J. Fernandez[(1)], M. F. Piccoli[(1,2)], R. Guerrero[(1)]

(1) Laboratorio de Computación Gráfica / Dpto. Informática / FCFMyN
Universidad Nacional de San Luis
Ejército de los Andes 950, Tel: 0266 4420823, San Luis, Argentina

(2)Facultad de Ciencia y Tecnología-Universidad Autónoma de Entre Ríos
25 de Mayo 385 - Concepción del Uruguay, Entre Ríos, Argentina

{ymalvarado, npasinetti, omlopres, ncmiran, gbrodriguez, jmfer, mpiccoli, rag}@unsl.edu.ar

RESUMEN

Los entornos virtuales inmersivos constituyen un nuevo paradigma de interacción digital al integrar tecnologías de Realidad Extendida, redes avanzadas y sistemas inteligentes para habilitar experiencias persistentes, orientadas a la comunicación, la educación y la colaboración remota. Su adopción masiva plantea desafíos vinculados a la accesibilidad abarcando desde los dispositivos heterogéneos, la eficiencia computacional y la calidad de experiencia. En este contexto, las arquitecturas distribuidas permiten el procesamiento remoto de escenas complejas, el renderizado en tiempo real y la transmisión adaptativa de datos tridimensionales. Asimismo, la incorporación de modelos de interpretación emocional sustentados en inteligencia artificial, visión por computadora y sensado multimodal posibilita representar la dimensión expresiva del usuario, favoreciendo interacciones más empáticas y naturales. Estas estrategias se articulan con el diseño de plataformas inmersivas escalables, interoperables e inclusivas.

En este trabajo se abordan los desafíos de la integración de estas tecnologías para el desarrollo de entornos virtuales accesibles, eficientes y centrados en el usuario.

Palabras clave: Entornos Virtuales, Realidad Virtual, Interacción Natural, Optimización de procesamientos, Tecnologias Emergentes.

CONTEXTO

La propuesta de trabajo se lleva a cabo dentro del proyecto "*Tecnologías de Computación para Optimizar la Comunicación en Entornos Metaverso*". Este proyecto es desarrollado en el ámbito del Laboratorio de Computación Gráfica de la Universidad Nacional de San Luis.

1. INTRODUCCIÓN

En los últimos años, la evolución de las tecnologías inmersivas ha dado lugar a ecosistemas digitales complejos que transforman las formas tradicionales de interacción humano-sistema. Estos entornos virtuales tridimensionales se sustentan en la convergencia de infraestructuras de comunicación de alta velocidad, como las redes 5G/6G, arquitecturas de procesamiento distribuido basadas en computación en la nube, computación en el borde, y sistemas inteligentes apoyados en inteligencia artificial y aprendizaje automático, permitiendo la construcción de ambientes virtuales persistentes para actividades sociales, educativas, productivas y sanitarias [1].

En este contexto, la comunicación mediada por avatares, la telepresencia y la colaboración remota adquieren un papel central en la transformación de los modelos tradicionales

de interacción humana. Esta evolución no sólo redefine la manera en que los individuos se vinculan entre sí y con los sistemas digitales, sino también plantea nuevos desafíos en términos de accesibilidad, inclusión y calidad de experiencia para una adopción masiva y sostenible de estas plataformas.

La expansión de estos entornos digitales ha puesto en evidencia la necesidad de garantizar experiencias inmersivas accesibles desde dispositivos heterogéneos y con distintos niveles de capacidad de procesamiento. Por consiguiente, generalizar el acceso a dichos entornos requiere el desarrollo de soluciones, las cuales reducen la dependencia al hardware de alto rendimiento, optimicen el consumo energético y permitan la utilización de tecnologías inmersivas en diferentes contextos educativos, laborales y sociales. En este sentido, la eficiencia computacional, la sostenibilidad tecnológica y la inclusión digital se convierten en factores determinantes para la consolidación de ecosistemas virtuales a gran escala [2].

En respuesta a dichas demandas, la investigación actual se orienta hacia el uso de tecnologías emergentes sustentadas en arquitecturas distribuidas que permitan el procesamiento remoto de escenas complejas, el renderizado eficiente en tiempo real y la transmisión adaptativa de datos tridimensionales [3,4]. La compresión inteligente de datos, el balanceo dinámico de carga y los mecanismos de streaming inmersivo surgen como estrategias clave para posibilitar plataformas escalables y ubicuas [5].

Sin embargo, el avance hacia entornos virtuales masivos y técnicamente eficientes debe contemplar las dimensiones comunicacionales y emocionales de la interacción humana. La sensación de presencia social y la naturalidad en la comunicación dependen en gran medida de la capacidad de los sistemas para interpretar y representar expresiones emocionales así como comportamientos verbales y no verbales [6]. Desde esta perspectiva, resulta conveniente incorporar modelos computacionales en entornos inmersivos que enriquezcan la experiencia del usuario y favorezcan interacciones más empáticas, realistas y centradas en el individuo.

El desarrollo de estas capacidades demanda la convergencia entre tecnologías inmersivas y emergentes orientadas al modelado integral del usuario dentro del entorno virtual. En este marco, la captura y procesamiento de información multimodal en tiempo real (proveniente de dispositivos de interacción natural, sistemas de seguimiento corporal, interfaces hápticas y mecanismos de sensado fisiológico) permite construir representaciones digitales dinámicas capaces de reflejar no sólo la presencia física sino también aspectos cognitivos y conductuales del individuo. Esta integración posibilita la adaptación contextual de los entornos virtuales, la personalización de la experiencia y la mejora de la comunicación mediada por avatares en escenarios tridimensionales, constituyendo un campo de investigación estratégico para las Ciencias de la Computación.

La presente investigación tiene por objetivo estudiar la sinergia de las tecnologías mencionadas mediante el desarrollo de plataformas inmersivas escalables, la optimización del procesamiento en sistemas distribuidos y la incorporación de modelos de interpretación emocional, a fin de contribuir al diseño de entornos virtuales más accesibles, eficientes y centrados en el usuario.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Con el propósito de abordar el diseño de modelos computacionales que integren la dimensión humana de la interacción junto con los requerimientos de eficiencia, escalabilidad y accesibilidad propios de estos ecosistemas digitales, se proponen las siguientes líneas de investigación y desarrollo.

- **Modelado e Interpretación de la Dimensión Expresiva en la Comunicación Multimodal**

  Esta línea se orienta al desarrollo de modelos capaces de detectar, analizar e interpretar tanto estados emocionales como comportamientos verbales y no verbales en tiempo real dentro de entornos inmersivos. La integración de técnicas de Inteligencia Artificial, Visión por Computadora y Sensores Biométricos permite la captura de información multimodal asociada a expresiones faciales, postura corporal, características de la voz y variables fisiológicas. En este marco, los gemelos digitales humanos emergen como una estrategia para representar de manera dinámica y persistente tanto los aspectos físicos como emocionales del usuario, posibilitando la reproducción de interacciones más empáticas, realistas y centradas en el individuo [6-9].

- **Optimización del Procesamiento en Sistemas Inmersivos**

  El funcionamiento de entornos virtuales complejos en dispositivos heterogéneos requiere el desarrollo de estrategias de eficiencia computacional que permitan reducir la latencia y optimizar el uso de los recursos disponibles. En este sentido, la adopción de arquitecturas distribuidas basadas en computación en la nube (*Cloud Computing*) y computación en el borde (*Edge Computing*) posibilita el procesamiento remoto de escenas tridimensionales, el renderizado en tiempo real y la transmisión adaptativa de contenidos inmersivos [3,4,10]. Asimismo, el uso de técnicas de procesamiento distribuido y paralelo, balanceo dinámico de carga y compresión inteligente de datos constituye un conjunto de mecanismos fundamentales para mejorar la calidad de la experiencia del usuario en plataformas inmersivas [5, 11].

- **Plataformas Inmersivas Escalables y Accesibles**

  Esta línea aborda el diseño de entornos virtuales capaces de crecer en número de usuarios, incorporar nuevas funcionalidades, integrar múltiples fuentes de contenido y operar en dispositivos con diferentes capacidades de procesamiento sin degradar el rendimiento ni la calidad de experiencia. La escalabilidad se vincula con la implementación de infraestructuras flexibles y adaptativas, mientras que la accesibilidad implica el desarrollo de soluciones que permitan la participación de usuarios con diversas capacidades físicas, cognitivas y tecnológicas, así como su utilización en contextos con limitaciones de conectividad o hardware. En este marco, la interoperabilidad entre plataformas, la adaptación dinámica de contenidos y el soporte para interfaces multimodales constituyen aspectos clave para la construcción de ecosistemas inmersivos inclusivos y ubicuos [12,13].

3. RESULTADOS OBTENIDOS/ESPERADOS

Las actividades realizadas hasta el momento por este grupo se han enmarcado dentro de 2 proyectos de Investigación de la UNSL, 1 proyecto de la Comunidad Europea, 4 proyectos de Desarrollo Tecnológico de la Secretaría de Políticas Universitarias y 5 proyectos de Extensión de Interés Social.

En el marco de estas investigaciones, se han desarrollado sistemas de Realidad Virtual (RV) con interacción multimodal, incorporando comandos verbales y gestuales, así como aplicaciones innovadoras en RV y Realidad Aumentada (RA).

Actualmente, las líneas de trabajo se orientan a la integración de nuevas estrategias destinadas a optimizar la percepción y la interacción del usuario, promoviendo su estimulación tanto cognitiva como física, y fortaleciendo la eficiencia computacional mediante la incorporación de tecnologías avanzadas de cómputo, tanto a nivel de

infraestructura, incluyendo arquitecturas paralelas y distribuidas, cloud y edge computing, como en el desarrollo de software para entornos de alto desempeño.

4. FORMACIÓN DE RECURSOS HUMANOS

Los trabajos realizados han permitido la definición de trabajos finales de carrera de la Licenciatura en Cs. de la Computación (ocho finalizados), tesis de Especialización en Educación Superior (uno finalizada), tesis de Especialización en Computación Gráfica (dos finalizados), tesis de Maestría en Cs. de la Computación (dos finalizada y dos en ejecución) y tesis de Doctorado en Ciencias de la Computación (cuatro en ejecución).

Adicionalmente se ha obtenido dos becas de iniciación a la investigación y una beca de perfeccionamiento en investigaciones otorgadas por la Secretaría de Ciencia y Técnica de la UNSL; tres becas doctorales: una de CONICET y dos de UTN.

5. BIBLIOGRAFÍA

1. Martínez-Cárdenas, B., Osorio Umaña, F., Bedecarratz Scholz, F., Salazar Camargo, J. C., Arancibia Obrador, M. J., Azuaje-Pirela, M., Viollier Bonvin, P., Amaya Villegas, P., Navarro, R., Bozzo Hauri, S., & Sánchez Polanco, S. (2022). Tecnologías emergentes. Qué son y cómo aprovecharlas en las industrias creativas y culturales. CERLALC y Universidad Autónoma de Chile.
2. Rivera García, D., & Rivera Berrío, J. G. (2023). Metaversos: ¡Acércate a los mundos de la realidad virtual! Córdoba, España: Red Educativa Digital Descartes. ISBN 978-84-18834-79-0.
3. Buyya, R., & Srirama, S. N. (Eds.). (2019). Fog and edge computing: principles and paradigms. John Wiley & Sons.
4. Erl, T., Mahmood, Z., Puttini, R. (2013). Cloud computing concepts, technology and architecture. Prentice Hall Press. ISBN:978-0-13-338752-0.
5. Hwang, K., Dongarra, J., Fox, G. C. (2018). Cloud Computing and Distributed Systems: From Parallel Processing to the Internet of Things. Estados Unidos: Elsevier Science.
6. Rodríguez, G., Jofré, N., Alvarado, Y., Fernández, J., & Guerrero, R. (2025). Estrategias avanzadas para lograr experiencias virtuales más inmersivas y personalizadas. En Actas del XXVII Workshop de Investigadores en Ciencias de la Computación (pp. 254–258). Universidad Nacional de Cuyo. http://sedici.unlp.edu.ar/handle/10915/183848
7. Novelli, E., Alvarado, Y., & Guerrero, R. (2025). Beyond text: Enhancing human-computer interaction through multimodal embodied conversational agents. En Actas del XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025) (pp. 405–414). Universidad Nacional de Río Negro. http://sedici.unlp.edu.ar/handle/10915/189846
8. Rodriguez, G., Castro, S., & Guerrero, R. (2023). Hand and arm gesture interaction natural user interface: A unified architecture. En Actas del 11º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI).
9. Alvarado, Y., Guerrero, R., & Serón, F. (2023). Inclusive learning through immersive virtual reality and semantic embodied conversational agent: A case study in children with autism. Journal of Computer Science and Technology, 23, 107–116. https://doi.org/10.24215/16666038.23.e09
10. Andriulo, F., Fiore, M., Mongiello, M., Traversa, E., Zizzo, V.. (2024). Edge Computing and Cloud Computing for Internet of Things: A Review. Informatics. 11. 71. 10.3390/informatics11040071.
11. Bonnin, R., Delrieux, C., Piccoli, M. F. (2025). Deep learning-based visual aid

for low vision. IEEE Embedded Systems Letters. Advance online publication. https://doi.org/10.1109/LES.2025.3548959.
12. Rodríguez, G., Jofré, N., Alvarado, Y., Fernández, J., & Guerrero, R. A. (2024). A Flexible Framework for Natural User Interface in the Metaverse. In XXX Congreso Argentino de Ciencias de la Computación (CACIC)(La Plata, 7 al 11 de octubre de 2024).
13. Rodríguez Copa, Graciela & Castro, Silvia & Guerrero, Roberto. (2024). Un Módulo Integrado de Interfaz Humano-Computadora Basada en Gestos. RevITA. 2. 10.22305/revita-unpa.v2.n1.1081.

LA VISUALIZACIÓN COMPARATIVA EN EL ANÁLISIS DE CONJUNTOS DE DATOS COMPLEJOS

Dana K. Urribarri[1,2,3], Martín Larrea[1,2,3], Leandro Luque[1,3], Matías Selzer[1,2,3], Antonella Antonini[1,2,3], María Silvia Vera Laceiras[4]

[1]Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (UNS-CIC PBA)
[2]Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
[3]Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
[4]Facultad de Exactas Químicas y Naturales. UNaM. Posadas, Misiones UNNE UTN.
[4]Facultad de Cs. Forestales. UNaM. Eldorado, Misiones.
[4]Instituto de Investigación, diseño e innovación. Facultad de Cs.Exactas QyN. UNaM.

{dku, mll, leandro.luque, matias.selzer, antonella.antonini}@cs.uns.edu.ar,
silvia.veralaceiras@fcf.unam.edu.ar

RESUMEN

La comparación es una tarea importante en el análisis de datos, en la cual dos o más conjuntos de datos deben analizarse buscando similitudes o diferencias. La visualización cumple un rol importante en el proceso de entender, analizar y comunicar datos. La comparación implica un desafío para la visualización, dado que se debe combinar tanto factores de los objetos individuales como de la relación entre ellos. Mientras que la visualización de información tradicionalmente se ha centrado en examinar objetos o conjuntos de datos individuales, muchos trabajos han resaltado la importancia de desarrollar herramientas de visualización que soporten explícitamente la tarea de comparación. El objetivo de esta línea de investigación es el estudio y desarrollo de técnicas de visualización, representaciones visuales y herramientas que asistan en la comparación de conjuntos de datos, teniendo en cuenta la heterogeneidad de datos y conjuntos de datos, y la escalabilidad visual de las técnicas.

Palabras clave: Visualización comparativa, Comparación, Visualización de información

CONTEXTO

Esta línea de investigación continúa la presentada en el proyecto PGI 24/ZN44 financiado por la Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur, Argentina, y finalizado en diciembre del 2025; actualmente, el proyecto en el cual se inserta esta investigación se encuentra en evaluación por la misma secretaría. El trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) perteneciente al Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur y centro asociado CIC-PBA. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La era del *Big Data* se caracteriza por el crecimiento exponencial de la información digital. Sin embargo, si bien la abundancia de recursos favorece el descubrimiento de conocimiento y la innovación, trae aparejada la complejidad derivada de la sobrecarga de

información. Debido a las consecuencias negativas de tal sobrecarga [1] –la principal consecuencia es la disminución de la calidad y la eficiencia en la toma de decisiones; quienes toman las decisiones necesitan consumir más tiempo y recursos cognitivos para abordar la tarea–, los sistemas inteligentes, el filtrado de datos y la visualización de datos cobran relevancia al momento de facilitar la interpretación de estos grandes volúmenes de información [2, 3]. En este escenario, las tecnologías de ciencias de datos y, en particular, las herramientas de análisis visual juegan un papel importante en la generación de conocimiento valioso para apoyar la toma de decisiones críticas [4].

La comparación es una tarea importante en el análisis de datos, en la cual dos o más conjuntos de datos deben analizarse buscando similitudes o diferencias; aparece a lo largo de áreas de aplicación tan diversas como epidemiología [5], oceanografía [6], biología [7], meteorología [8, 9], ciencias forestales [10], etc. y, a su vez, involucra diversos tipos de datos, tales como imágenes [11], grafos [12, 13], superficies [14], series de tiempo [15], etc. El mayor desafío de la comparación es que no sólo debe combinar factores de los objetos individuales sino también de la relación entre ellos.

La visualización cumple un rol importante en el proceso de entender, analizar y comunicar datos. Como disciplina, busca desarrollar métodos, herramientas y representaciones visuales que faciliten el análisis de datos y permitan extraer conclusiones de manera rápida y eficiente sobre las posibles relaciones existentes en los datos. Esto la hace eficaz cuando la gran cantidad de ítems y/o su dimensionalidad dificulta el análisis y entendimiento de los datos crudos [16]. Mientras que la visualización de información tradicionalmente se ha centrado en examinar objetos o conjuntos de datos individuales, muchos trabajos han resaltado la importancia de desarrollar herramientas de visualización que soporten explícitamente la tarea de comparación [17, 18, 19, 20, 21, 22, 23, 24, 25].

El diseño de visualizaciones en general y de visualizaciones comparativas en particular no está exento de desafíos. La naturaleza de los datos (datos multidimensionales, datos temporales, datos espaciales, categóricos, etc.) y la tarea de comparación que se desea realizar son determinantes a la hora de diseñar una visualización de datos. Para lograr un diseño de calidad hay que tener en cuenta varios aspectos, como por ejemplo, la escalabilidad visual de la técnica [26], la incertidumbre de los datos [27, 28], la percepción humana [19, 22], entre otros.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo de esta investigación consiste en estudiar y desarrollar técnicas de visualización, representaciones visuales y herramientas que asistan en la comparación de conjuntos de datos y se centra en dos ejes principales:

(1) Adaptación a la heterogeneidad de datos: Diseñar y desarrollar técnicas de visualización comparativa adaptadas a diversos tipos datos complejos, como series de tiempo multivariadas (incluyendo atributos categóricos y numéricos), datos espaciales y geo-espaciales (ej., imágenes satelitales y superficies), y datos multidimensionales.

(2) Escalabilidad Visual: Investigar y proponer estrategias visuales que permitan:
   - La comparación simultánea de más de dos series de tiempo.
   - La comparación de grandes series de tiempo.

3. RESULTADOS OBTENIDOS y ESPERADOS

En esta línea de investigación, previamente, se ha desarrollado una técnica de visualización comparativa de series de tiempo. Esta técnica se ha empleado tanto en la comparación de datos provenientes de movimientos de karate [29, 30] como en la comparación de datos meteorológicos [9, 31]. Sobre la base de esta técnica se han obtenido resultados parciales:

- Se ha logrado la implementación web de una versión genérica de esta técnica que permite la comparación de series de tiempo con datos numéricos [32].
- Se ha realizado una revisión de estado del arte del uso de las variantes de DTW en visualización [33].

Como resultados futuros de esta línea de investigación se espera:

- Ampliar la visualización web para la comparación de series de tiempo con datos categóricos.
- Incluir variantes multirresolución de DTW para la comparación de grandes series de tiempo.
- Incluir variantes progresivas de DTW para la comparación de streaming de datos.

Estos dos últimos puntos presentan desafíos a nivel visualización, dado que ambos requieren representaciones incrementales: el DTW progresivo, debido a la naturaleza del conjunto de datos, y el DTW multirresolución, debido a la transición entre los niveles de abstracción. Esto requiere nuevas representaciones visuales que se adapten a la naturaleza dinámica y evolutiva de los datos a visualizar.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo que concierne a la formación de recursos humanos, se incentiva la incorporación de alumnos de grado o de posgrado que deseen trabajar en temas relacionados. En particular, se ha concluido un proyecto final de carrera y se encuentra otro en desarrollo.

- Facundo Alvarado. Proyecto final de carrera. *Implementación web de una técnica de visualización para la comparación de datos temporales.* Finalizada.
- Agustín Cuello. Proyecto final de carera. *Visualización comparativa en Grafana.* En desarrollo.
- María Silvia Vera Laceiras. Doctorado en Ciencias de la Computación. *Desarrollo de un Modelo de procesamiento y análisis digital de imágenes satelitales para detección de cambios. Estudio de caso: Detección temprana de deforestación y degradación en la Reserva biosfera Yabotí.* En desarrollo.

REFERENCIAS

[1] X. Ji, “The negative psychological effects of information overload,” BCP Education & Psychology, vol. 9, pp. 250–255, Mar. 2023.

[2] S. Jauhari and S. Vobugari, “Delineating the big data era and the information overload problem,” [Preprint], 2025.

[3] M. M. Francelin and F. O. Meschini, “Knowledge organization and the technological challenges of the big data era,” Knowledge Organization, vol. 52, Apr. 2025.

[4] M. Arnold, M. Goldschmitt, and T. Rigotti, “Dealing with information overload: a comprehensive review,” Frontiers in Psychology, vol. 14, June 2023.

[5] T. Kreuz, “Comparative visualization of epidemiological data during various stages of a pandemic,” Feb. 2021.

[6] M. Yano, T. Itoh, Y. Tanaka, D. Matsuoka, and F. Araki, "A comparative visualization tool for ocean data analysis based on mode water regions," Journal of Visualization, vol. 23, no. 2, pp. 313–329, 2020.

[7] S. Nagpal, K. D. Baksi, B. K. Kuntal, and S. S. Mande, "NetConfer: A web application for comparative analysis of multiple biological networks," BMC Biology, vol. 18, no. 1, 2020.

[8] A. Dasgupta, J. Poco, Y. Wei, R. Cook, E. Bertini, and C. T. Silva, "Bridging theory with practice: An exploratory study of visualization use and design for climate model comparison," IEEE Transactions on Visualization and Computer Graphics, vol. 21, pp. 996–1014, sep 2015.

[9] M. L. Larrea and D. K. Urribarri, "Visualization Technique for Comparison of Time-based Large Data Sets," in Conference on Cloud Computing Conference, Big Data & Emerging Topics (JCC-BD&ET), vol. 1444 of Communications in Computer and Information Science, pp. 179–187, Springer International Publishing, 8 2021.

[10] J. W. Nam, K. McCullough, J. Tveite, M. M. Espinosa, C. H. Perry, B. T. Wilson, and D. F. Keefe, "Worlds-in-Wedges: Combining Worlds-in-Miniature and Portals to Support Comparative Immersive Visualization of Forestry Data," in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 747–755, mar 2019.

[11] M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend, "High-Resolution Global Maps of 21st-Century Forest Cover Change," Science, vol. 342, pp. 850–853, nov 2013.

[12] K. Andrews, M. Wohlfahrt, and G. Wurzinger, "Visual Graph Comparison," in 2009 13th International Conference Information Visualisation, vol. 00, pp. 62–67, jul 2009.

[13] Y. Naragino and K. Misue, "Visualization Techniques for the Comparative Analysis of Weighted Free Trees," in 2017 21st International Conference Information Visualisation (IV), pp. 45–51, jul 2017.

[14] K. Kim, J. V. Carlis, and D. F. Keefe, "Comparison techniques utilized in spatial 3d and 4d data visualizations: A survey and future directions," Computers & Graphics, vol. 67, pp. 138–147, oct 2017.

[15] J. Fuchs, F. Fischer, F. Mansmann, E. Bertini, and P. Isenberg, "Evaluation of alternative glyph designs for time series data in a small multiple setting," ACM, apr 2013.

[16] W. S. Cleveland, Visualizing Data. Hobart Press, 1993.

[17] M. Gleicher, "Considerations for Visualizing Comparison," IEEE Transactions on Visualization and Computer Graphics, vol. 24, no. 1, pp. 413–423, 2018.

[18] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, "Visual comparison for information visualization," Information Visualization, vol. 10, no. 4, pp. 289–309, 2011.

[19] N. Jardine, B. D. Ondov, N. Elmqvist, and S. Franconeri, "The Perceptual Proxies of Visual Comparison," IEEE Transactions on Visualization and

Computer Graphics, vol. 26, pp. 1012–1021, jan 2020.

[20] S. LYi, J. Jo, and J. Seo, "Comparative layouts revisited: Design space, guidelines, and future directions," IEEE Transactions on Visualization and Computer Graphics, vol. 27, pp. 1525–1535, feb 2021.

[21] B. Ondov, N. Jardine, N. Elmqvist, and S. Franconeri, "Face to Face: Evaluating Visual Comparison," IEEE Transactions on Visualization and Computer Graphics, vol. 25, pp. 861–871, jan 2019.

[22] B. D. Ondov, Revealing Perceptual Proxies in Comparative Data Visualization. PhD thesis, 2021.

[23] H.-G. Pagendarm and F. H. Post, "Comparative Visualization: Approaches and Examples," Visualization in Scientific Computing (H. Göbel, H. Müller and B. Urban), no. February, pp. 95–108, 1995.

[24] J. Schmidt, Scalable Comparative Visualization. Visual Analysis of Local Features in Different Dataset Ensembles. PhD thesis, Faculty of Informatics (TU Wien), 2016.

[25] A. Srinivasan, M. Brehmer, B. Lee, and S. M. Drucker, "What's the Difference?: Evaluating Variants of Multi-Series Bar Charts for Visual Comparison Tasks," in ACM Conference on Human Factors in Computing Systems (CHI), pp. 1–10, ACM, Jan. 2018.

[26] S. G. Eick and A. F. Karr, "Visual scalability," Journal of Computational and Graphical Statistics, vol. 11, pp. 22–43, mar 2002.

[27] G. P. Bonneau, H. C. Hege, C. R. Johnson, M. M. Oliveira, K. Potter, P. Rheingans, and T. Schultz, "Overview and state-of-the-art of uncertainty visualization," 2014.

[28] A. Kamal, P. Dhakal, A. Y. Javaid, V. K. Devabhaktuni, D. Kaur, J. Zaientz, and R. Marinier, "Recent advances and challenges in uncertainty visualization: a survey," Journal of Visualization, may 2021.

[29] D. K. Urribarri, M. L. Larrea, S. M. Castro, and E. Puppo, "Visualization to compare karate motion captures," in XXV Congreso Argentino de Ciencias de la Computación. CACIC 2019. Libro de Actas. (P. Pesado and M. Arroyo, eds.), pp. 446–455, Oct. 2019.

[30] D. K. Urribarri, M. L. Larrea, S. M. Castro, and E. Puppo, "Overview+detail visual comparison of karate motion captures," in Computer Science – CACIC 2019 (P. Pesado and M. Arroyo, eds.), vol. 1184 of Communications in Computer and Information Science, pp. 139–154, Cham: Springer International Publishing, 5 2020.

[31] D. K. Urribarri and M. L. Larrea, "A visualization technique to assist in the comparison of large meteorological datasets," Computers & Graphics, mar 2022.

[32] F. A. Alvarado, "EudypTS: A Web-based tool for the visual comparison of time-series." Proyecto final de carrera, 2025.

[33] D. K. Urribarri, "A Review on Dynamic Time Warping in Data Visualization," in Libro de actas - XXXI Congreso Argentino de Ciencias de la Computación - CACIC 2025, pp. 385–394, 2025.

REALIDAD EXTENDIDA APLICADA A LA EXPLORACIÓN GEOLÓGICA: UN ENFOQUE INMERSIVO

Matías N. Selzer[1,2], Rodrigo H. Herlein[1,2], M. Luján Ganuza[1,2], Antonini Antonella S.[1,2], Luque, Leandro E.[1,2], Urribarri, Dana K.[1,2], Larrea, Martín L.[1,2], Ferracutti, Gabriela R.[3], Asiain, Lucía[3], Bjerg, Ernesto A.[3], y Silvia M. Castro[1,2]

[1]Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (UNS-CIC)
Dpto. de Cs. e Ing. de la Computación, Universidad Nacional del Sur (DCIC-UNS)
{matias.selzer, rodrigo.herlein, mlg, antonella.antonini, leandro.luque, dku, mll, smc}@cs.uns.edu.ar
[2]Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
[3]INGEOSUR, Dpto. de Geología, Universidad Nacional del Sur
gferrac@uns.edu.ar; {lasiain, ebjerg}@ingeosur-conicet.gob.ar

RESUMEN

Las tecnologías inmersivas experimentan un avance sumamente rápido, propiciando constantemente el surgimiento de nuevas aplicaciones e innovaciones. De forma paralela, la visualización de datos se consolida como una herramienta esencial para interpretar grandes volúmenes de información mediante representaciones gráficas. La integración de estos entornos inmersivos con técnicas de visualización potencia sustancialmente las capacidades analíticas de los usuarios. En este contexto, consideramos que la adopción de la Realidad Extendida (XR) puede aportar beneficios sin precedentes al ámbito de las Geociencias. Por este motivo, nuestro propósito se centra en el diseño y desarrollo de nuevas metodologías de inmersión y visualización situada, con el objetivo de optimizar las tareas de exploración geológica en el campo. Se pretende crear un marco de trabajo interactivo en tres dimensiones que envuelva al especialista, facilitando el análisis y la toma de decisiones directamente en el lugar de estudio.

Palabras claves: Visualización Situada, Tecnologías Inmersivas, Tecnologías XR

CONTEXTO

Este trabajo se realiza en estrecha colaboración con investigadores de centros de investigación de reconocido prestigio como el INGEOSUR (Instituto Geológico del Sur CONICET-UNS) y el Departamento de Geología de la Universidad Nacional del Sur, y el VyGLab (Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica) del Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS).

1. INTRODUCCIÓN

Las tecnologías inmersivas abarcan aquellas herramientas capaces de integrar elementos virtuales dentro del mundo físico, como la Realidad Aumentada (RA) y la Realidad Mixta (RM), o de generar simulaciones completamente artificiales, como la Realidad Virtual (RV). El conjunto de estas tecnologías suele agruparse bajo el término de Realidad Extendida (XR). En la actualidad, su rápida evolución está generando un impacto profundo en la sociedad, expandiendo su uso mucho más allá del entretenimiento hacia áreas como la educación médica [11], la psicología y el turismo.

Por otra parte, la visualización de datos resulta indispensable para comprender interacciones complejas en grandes conjuntos de información [1, 3, 16]. Al integrar la XR en este proceso, surgen formas novedosas de percibir e interactuar con los datos [13, 17]. Esta sinergia hace posible superponer modelos digitales sobre el entorno real o construir mundos virtuales enteros dedicados a la exploración analítica multiescala [12, 15].

La fusión de técnicas inmersivas y visualización de datos representa un avance de gran utilidad para el dominio de las Geociencias [4, 5, 6, 7, 8]. El análisis geológico tradicional requiere interpretar numerosas variables en el terreno, una tarea que puede optimizarse significativamente mediante la visualización de datos in-situ. Vincular la información directamente con el entorno físico en tiempo real incrementa la precisión y eficacia del profesional, fomentando además el trabajo colaborativo [10, 14].

En esta misma línea, la Visualización Situada aporta un valor fundamental al desplegar representaciones gráficas directamente en su contexto físico de origen [20]. Esto facilita tareas como la consulta instantánea de bibliografía geológica, el registro de muestras y la aplicación de métodos de mapeo innovadores. La combinación de datos topográficos, estructurales y químicos analizados directamente en el campo permite una comprensión mucho más profunda de las relaciones espaciales subyacentes.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las propuestas de investigación y desarrollo que llevamos adelante se estructuran en los siguientes ejes:

Desarrollo de software inmersivo aplicado a las geociencias: Nuestro propósito es unificar las múltiples fuentes de datos geológicos que se manejan durante el trabajo de campo, aprovechando los beneficios de la RA, RM y RV. Si bien el uso de estas tecnologías en espacios exteriores al aire libre plantea grandes retos técnicos para la correcta visualización de los datos, es factible construir sistemas interactivos de XR utilizando dispositivos de bajo costo y posibilitando la manipulación colaborativa de objetos [9]. Contamos con los recursos y capacidades necesarias para impulsar estos desarrollos a nivel nacional.

El eje central de nuestro trabajo es la creación de una aplicación móvil basada en Realidad Aumentada que centralice mapas topográficos, anotaciones de campo y registros de análisis químicos. Esta herramienta está pensada para disminuir la dependencia del formato papel y hacer mucho más dinámica la interpretación de los datos espaciales. De este modo, el geólogo dispondrá de una plataforma integrada para gestionar su información de manera más eficiente, lo cual perfeccionará la operatividad general durante las expediciones geocientíficas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Hasta el momento hemos consolidado avances significativos en varias áreas clave:

En cuanto al software XR de bajo costo, hemos logrado progresos sustanciales en el rendimiento gráfico. Desarrollamos técnicas que permiten a los dispositivos móviles renderizar entornos virtuales con una calidad visual sobresaliente en formato estereoscópico [18]. Asimismo, logramos integrar físicas de colisión precisas y el soporte para superficies

irregulares, elevando el nivel de realismo y la inmersión [19]. Procesar mundos virtuales de gran extensión representa un desafío técnico importante cuando se utiliza hardware económico, pero resulta un requisito indispensable para poder simular adecuadamente las amplias regiones topográficas que son objeto de estudio en la geología.

Respecto a la integración visual de datos complejos, mantenemos un avance constante con aplicaciones directas en el área de Geociencias. Hemos diseñado e implementado herramientas web interactivas orientadas al análisis de composición química y gestión de recursos naturales [2]. A su vez, optimizamos diversas técnicas para la exploración de datos multidimensionales, las cuales se pueden adaptar con facilidad a los requerimientos específicos del dominio geológico [1, 16]. Este esfuerzo sostenido busca enriquecer y agilizar el manejo integral de la información dentro de la disciplina.

4. FORMACION DE RECURSOS HUMANOS

A continuación, se detallan los proyectos de investigación, las tesis finalizadas y en desarrollo y las becas obtenidas vinculadas con las líneas de investigación presentadas.

Proyectos de Investigación:

Aplicación de Tecnologías Inmersivas y Visualización Situada en Geociencias, PGI SeCyT-UNS, Código: 24/ZN43, Director: Dr. Matias Nicolás Selzer.

Análisis Visual de Datos Multidimensionales, PGI SeCyT-UNS, Código: 24/ZN42, Director: Dra. María Luján Ganuza.

Tesis en Desarrollo:

- “Realidad Aumentada Móvil en Exteriores para Visualización de Datos Geológicos”, tesis de Doctorado en Ciencias de la Computación. Alumno: Juan Manuel Trippel Nagel. Directores: Dra. Silvia Castro - Dr. Ernesto Bjerg.

-“Análisis visual de datos espacio-temporales”; Doctorado en Ciencias de la Computación. Doctorando: Rodrigo Nicolás Herlein. Directora: Dra. (Ing.) María Luján GANUZA; Codirector: Dr. (Ing.) Matias Nicolás SELZER;

5. BIBLIOGRAFÍA

[1] Antonini, A. S., Ganuza, M. L. & Castro, S. M. (2022). VISUEL - A Web Dynamic Dashboard for DataVisualization. Journal of Computer Science and Technology, 22(1), e03.

[2] Antonini, A. S., Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Matković, K., Gröller, E., Bjerg, E. A & Castro, S. M. (2021). Spinel web: an interactive web application for visualizing the chemical composition of spinel group minerals. Earth Science Informatics, 1-8.

[3] Antonini, A. S., Luque, L., Ganuza, M. L. & Castro, S. M. (2022). Towards a Taxonomy for Non-Paired General Line Coordinates: A Comprehensive Survey. International Journal of Data Science and Analytics.

[4] Ferracutti, G. R., Gargiulo, M. F., Ganuza, M. L., Bjerg, E. A., and Castro, S. M. Determination of the spinel group end members based on electron microprobe analyses. Mineral. Petrol. 109, 2 (2015), 153-160.

[5] Ganuza, M. L., Castro, S. M., Ferracutti, G., Bjerg, E. A., and Martig, S. Spinelviz: An interactive 3d application for visualizing spinel group minerals. Comput. Geosci. 48 (2012), 50–56.

[6] Ganuza, M. L., Ferracutti, G., Gargiulo, M.F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. The spinel explorer - interactive visual analysis of spinel group

minerals. IEEE TVCG. 20, 12 (2014), 1913–1922.
[7] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. Interactive visual categorization of spinel-group minerals. Proc. of the Spring Conf. on Comput. Graph. (2017).
[8] Ganuza, M. L., Gargiulo, M. F., Ferracutti, G., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. Interactive semi automatic categorization for spinel group minerals. In 2015 IEEE, VAST 2015, October 25-30, 2015 (2015), 197–198.
[9] Grandi, J.G., Debarba, H.G., Nedel, L., and Maciel, A. (2017, May). Design and evaluation of a handheld-based 3d user interface for collaborative object manipulation. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 5881-5891).
[10] Gutwin, C., Greenberg, S., 2002. A descriptive framework of workspace awareness for real-time groupware. Computer Supported Cooperative Work (CSCW) Vol. 11. n. 3-4 , pp. 411-446.
[11] Huang, H.-M., Liaw, S.-S., Lai, C.-M., 2016. Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. Interactive Learning Environments, 24(1), 3-19.
[12] Ibayashi, H., et al., 2015. Dollhouse VR: a multi-view, multi-user collaborative design workspace with VR technology. SIGGRAPH Asia 2015 Emerging Technologies. 1-2, 2015.
[13] Kutak, D., Selzer, M. N., Byska, J., Ganuza, M. L., Barisic, I., Kozlikova, B. & Miao, H. (2021). Vivern A Virtual Environment for Multiscale Visualization and Modeling of DNA Nanostructures. IEEE Transactions on Visualization and Computer Graphics, 28(12), 4825-4838.
[14] Lacoche, J., Pallamin, N., Boggini, T., and Royan, J. (2017, November). Collaborators awareness for user cohabitation in co-located collaborative virtual environments. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (pp. 1-9).
[15] Le Chénéchal, M., Lacoche, J., Royan, J., Duval, T., Gouranton, V., and Arnaldi, B. (2016, March). When the giant meets the ant an asymmetric approach for collaborative and concurrent object manipulation in a multi-scale environment. In 2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE) (pp. 18-22).
[16] Luque, L. E., Ganuza, M. L., Antonini, A. S. & Castro, S. M. (2021a). npGLC-Vis Library for Multidimensional Data Visualization. En Cloud Computing, Big Data & Emerging Topics: 9th Conference, JCC-BD&ET, La Plata, Argentina, June 22-25, 2021, Proceedings (Vol. 1444, p. 188).
[17] Sabando, M. V., Ulbrich, P., Selzer, M., Byška, J., Mičan, J., Ponzoni, I., . . . Kozlíková, B. (2021). ChemVA: Interactive Visual Analysis of Chemical Compound Similarity in Virtual Screening. IEEE Transactions on Visualization and Computer Graphics, 27(2), 891-901.
[18] Selzer, M. N., Ganuza, M. L., Urribarri, D. K., Larrea, M. L. & Castro, S. M. (2021). Stereoscopic Image-Based Rendering Technique for Low-Cost Virtual Reality. En P. Pesado & J. Eterovic (Eds.), Computer Science – CACIC 2020 (pp. 91-101). Cham: Springer International Publishing.
[19] Selzer, M. N., Ganuza, M. L. & Castro, S. M. (2021). High Visual-Quality Scenes in Low-Cost Virtual Reality With Collisions and Irregular Surfaces. IEEE Computer Graphics and Applications, 41(4), 40-51.
[20] White, S.; Feiner, S., Kopylec J., 2006. Virtual vouchers: Prototyping a mobile augmented reality user interface for botanical species identification. Proceedings of the 3D User Interfaces, p. 119–126.

Ingeniería de Software

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

ADMINISTRACIÓN DEL CONTEXTO DE DOMINIO ORGANIZACIONAL PARA LA COLABORACIÓN ENTRE MÚLTIPLES AGENTES DE INTELIGENCIA ARTIFICIAL Y PERSONAS EN PROYECTOS DE INGENIERÍA SOFTWARE

Sebastián U. Flores, Mario Berón
Departamento de Informática - Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis - Ejército de Los Andes 950 - San Luis - Argentina
s.flores@outlook.com.ar, mberon@unsl.edu.ar

RESUMEN

En los últimos años, los avances en el campo de la Inteligencia Artificial han impulsado la incorporación de agentes inteligentes en múltiples sectores. En particular, en la Ingeniería del Software, se han desarrollado soluciones orientadas al análisis, la implementación y la revisión de código fuente, la ejecución de pruebas automatizadas y la interacción con fuentes de información diversas, entre otras aplicaciones. No obstante, ciertos desafíos permanecen sin resolver. Uno de ellos es la persistencia del conocimiento organizacional de forma accesible tanto para personas como para agentes autónomos, contemplando su consulta, actualización y verificación de consistencia a lo largo del tiempo. Esta problemática se ve agravada por la dispersión del conocimiento entre múltiples actores y fuentes digitales heterogéneas —documentos textuales, contenido audiovisual, herramientas de terceros y repositorios de código, entre otros—. Dicha fragmentación obstaculiza la incorporación de nuevos desarrolladores a los proyectos y restringe a los agentes inteligentes a operar sobre un contexto limitado por el conocimiento, rol o el nivel de acceso de cada interlocutor. El presente artículo describe los avances obtenidos en la línea de investigación, y establece las bases para continuar investigando diferentes formas de abordar el desafío descrito.

Palabras clave: *Inteligencia Artificial, Agentes Autónomos, Gestión del Conocimiento, Ingeniería del Software, Fragmentación.*

CONTEXTO

La presente línea de investigación se enmarca en el proyecto de investigación denominado *"Ingeniería del Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube"*. El proyecto es reconocido por el *Programa de Incentivos PI 2026*, y es la continuación de diferentes proyectos de investigación, a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional.

1. INTRODUCCIÓN

En los últimos años, se han producido avances significativos en diversas tecnologías vinculadas a la Inteligencia Artificial, los cuales han posibilitado la incorporación de agentes inteligentes en múltiples sectores, tanto en entornos virtuales como físicos.

La evolución disruptiva de los Grandes Modelos del Lenguaje (*Large Language Models*, LLMs) ha dotado a los agentes inteligentes de capacidades para la comprensión de toda clase de conocimiento que pueda ser representada de forma estructurada [1]. Desarrollos recientes han habilitado el razonamiento secuencial sobre el contexto adquirido, favoreciendo la obtención de respuestas más precisas [2, 3]. En particular, desde el año 2023, el tamaño de la ventana de contexto se ha expandido a una tasa aproximada de 30 veces por año [4].

Tecnologías como la Generación Aumentada por Recuperación (*Retrieval-Augmented Generation*, RAG) han facilitado la integración

de agentes con fuentes de datos externas, complementando así el conocimiento incorporado durante su entrenamiento [5]. Sus iteraciones más recientes, entre las que se destaca GraphRAG, superan ciertas limitaciones del enfoque original al permitir la construcción de grafos de conocimiento, lo que habilita el análisis de relaciones complejas subyacentes entre diferentes conjuntos de información [6].

Por su parte, la consolidación del *Model Context Protocol* (MCP) como estándar universal para la conectividad entre agentes, herramientas y fuentes de datos externas, ha contribuido a reducir la fragmentación derivada de integraciones personalizadas para cada agente o proveedor, potenciando a su vez el desarrollo de nuevas herramientas ejecutables tanto en entornos locales como en la nube [7].

Sin embargo, persisten diversos desafíos sin resolver. Uno de ellos, inherente a la Ingeniería del Software desde sus orígenes, es la adquisición, integración, validación, almacenamiento y disponibilización del conocimiento organizacional hacia todos los actores involucrados a lo largo del ciclo de vida de los proyectos [8].

Habitualmente, la información se encuentra dispersa entre múltiples individuos y fuentes heterogéneas —documentos textuales, planos, herramientas de terceros, contenido audiovisual, repositorios de código fuente, actas de reuniones, entre otras—. Esta dispersión origina el fenómeno denominado *islas de información*, en el que grupos específicos desarrollan sus propios lenguajes, procesos y repositorios de conocimiento, sin comunicarlos al resto de la organización [9].

Asimismo, dentro del conjunto de participantes en los proyectos de una organización, solo una minoría concentra la mayor proporción del conocimiento crítico. Esta situación, conocida como *conocimiento tribal* [10], constituye un subproducto sistémico de presiones de entrega, deuda técnica y las dinámicas de poder organizacional, y genera consecuencias tales como la dependencia excesiva hacia dichos expertos o el riesgo de pérdida de conocimiento ante su eventual desvinculación [11].

La fragmentación descrita en los párrafos anteriores origina las siguientes consecuencias:

- **Dificultades en el proceso de onboarding:** La incorporación de nuevos desarrolladores a un proyecto se ve obstaculizada por la ausencia de un repositorio de conocimiento centralizado y accesible.
- **Ineficiencia en la recuperación de información:** El personal destina una cantidad significativa de tiempo a la búsqueda de respuestas a consultas complejas, dispersas entre múltiples personas y fuentes de información con estructuras y tipos de datos heterogéneos.
- **Sobrecarga del personal experto:** Las dinámicas de conocimiento tribal concentran la demanda de consultas en un grupo reducido de individuos, reduciendo su productividad, incrementando su estrés y favoreciendo su desvinculación voluntaria.
- **Deterioro de la retención de talento:** La incapacidad organizacional para proveer un ecosistema de conocimiento estructurado frecuentemente deriva en desconexión cultural y dificultades para retener a los profesionales calificados.
- **Restricción de los agentes inteligentes:** Los agentes autónomos quedan limitados a operar sobre un contexto acotado a la información que cada interlocutor puede proveer, condicionado por sus conocimientos y sus niveles de acceso a datos y herramientas de la organización.
- **Falta de persistencia del contexto agéntico:** La imposibilidad de conectar a los agentes inteligentes con una base de conocimiento completa y actualizada en el tiempo origina la necesidad de reintroducir un conjunto de instrucciones e información de contexto en cada nueva sesión con un agente. Esto incrementa el riesgo de suministrar datos incompletos o inconsistentes respecto del estado de conocimiento vigente en la organización.
- **Inconsistencia de la base de conocimientos:** La ausencia de mecanismos

de sincronización dificulta mantener la coherencia del conocimiento entre los diferentes equipos de la organización.

- **Obstáculos normativos:** La dispersión y falta de trazabilidad de la información genera dificultades para el cumplimiento de estándares internacionales y marcos regulatorios relativos al tratamiento de datos corporativos y personales.

El conjunto de estas dificultades se traduce en un incremento de los costos y los plazos requeridos para la ejecución de proyectos.

Los avances recientes en el ámbito de la Inteligencia Artificial proveen herramientas capaces de adquirir información desde fuentes heterogéneas para su posterior estructuración y procesamiento, tanto por parte de usuarios humanos como de agentes autónomos. En este escenario, resulta imperativo desarrollar soluciones que optimicen la interacción de las tecnologías vigentes con las fuentes de información organizacional, a fin de potenciar el crecimiento institucional y mitigar los riesgos inherentes a la gestión del conocimiento.

Investigaciones contemporáneas sugieren que no es suficiente dotar a los agentes de vastas bases de conocimiento o incorporar extensos conjuntos de instrucciones y contexto inicial en las interacciones. El incremento excesivo del contexto subyacente genera otra serie de problemáticas:

- **Ventana de contexto limitada:** A pesar de los progresos sustanciales en la capacidad de procesamiento de las ventanas de contexto —tendencia que se mantiene en crecimiento exponencial [4] —, persiste un límite físico y computacional. Al superarlo, el agente se ve obligado a descartar información previa o calificada como "menos relevante".
- **Incremento de costos operativos:** A medida que el volumen de la conversación aumenta, el modelo debe procesar una mayor carga de datos en cada iteración para garantizar respuestas precisas y evitar la omisión de detalles previos. Esta carga computacional se traduce directamente en un incremento de los costos de tokenización [12].
- **Fenómeno lost-in-the-middle:** En prompts extensos y complejos, los modelos tienden a priorizar la información ubicada en los extremos (inicio y final), lo que aumenta significativamente el riesgo de omitir detalles críticos situados en el cuerpo central del texto [12].
- **Degradación y contaminación del contexto (*Context Rot/Pollution*):** La provisión de un contexto excesivamente amplio —que integre el *prompt*, el historial de conversación, bases de conocimiento y conectores externos— puede inducir estados de confusión en el agente. Esto deriva en la generación de alucinaciones o en la propuesta de soluciones que contravienen principios arquitectónicos o de seguridad establecidos explícitamente [12].

Ante los desafíos listados, las tendencias actuales proponen una transición desde la *ingeniería de prompting* hacia la *ingeniería de contexto* [12, 13, 14, 15]. Este nuevo enfoque abarca el estudio y desarrollo de las siguientes dimensiones:

- **Almacenamiento del contexto:** Comprende el conjunto de tecnologías y metodologías destinadas a la persistencia del contexto fuera de la ventana activa del agente. El objetivo es permitir la consulta o actualización asincrónica de la información sin saturar la memoria operativa de la sesión, facilitando incluso el intercambio de contexto entre múltiples agentes.
- **Selección del contexto:** Implementación de técnicas diseñadas para identificar y extraer exclusivamente la información de mayor relevancia para la ejecución de las tareas solicitadas.
- **Aislamiento del contexto:** Consiste en la segmentación de la información para optimizar el enfoque del agente y reducir el costo computacional de la búsqueda de datos. Se identifican dos vertientes principales:
  - **Descomposición de tareas:** División de la tarea principal en subtareas ejecutadas por subagentes especializados, quienes entregan los resultados y un resumen ejecutivo al agente principal.

- **Entornos de trabajo dinámicos:** Fragmentación del contexto en entornos operativos que pueden ser cargados o liberados de la memoria a demanda del sistema.

- **Compresión del contexto:** Técnica orientada a retener únicamente los tokens esenciales para la resolución del problema. Se emplean métodos como el *resumen estratégico* (preservación de puntos de interés) y el *recorte selectivo* (eliminación de secciones redundantes o de baja prioridad).

Dado que la ingeniería de contexto es un área de estudio emergente, persisten numerosos interrogantes y desafíos técnicos que representan oportunidades significativas para la investigación científica y el desarrollo de soluciones con impacto nacional o global.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los ejes temáticos que orientan la presente investigación y desarrollo son los siguientes:

- **Indexación Semántica Multimodal:** Relevamiento del estado del arte de herramientas y técnicas para la indexación de información proveniente de fuentes de datos digitales heterogéneas, tales como documentos textuales, grabaciones de audio y video, presentaciones, sitios web, código fuente, herramientas de gestión de proyectos y actas de reuniones, entre otras. A partir de dicho relevamiento, se contempla el desarrollo de tecnologías que faciliten la integración de las soluciones analizadas.

- **Persistencia y Estructuración del Conocimiento:** Análisis de las tecnologías y estructuras empleadas para el almacenamiento de la información extraída y de las relaciones identificadas entre sus elementos. La selección de la estructura y el formato más adecuados estará orientada por el tipo de dato o relación detectados durante la indexación del contexto organizacional, priorizando criterios de accesibilidad, escalabilidad y robustez. Complementariamente, se considera la identificación de técnicas que garanticen la consistencia de la información adquirida desde fuentes diversas y en distintos momentos temporales.

- **Interoperabilidad Agéntica:** Evaluación del estado del arte de las tecnologías para la interconexión de agentes, fuentes de datos y servicios, tomando como referencia el estándar *Model Context Protocol* (MCP). Se prevé el desarrollo de herramientas que posibiliten la integración del contexto organizacional indexado y almacenado con distintos agentes inteligentes, de manera centralizada, segura y alineada con los estándares vigentes.

- **Interacción Humano-Agente:** Desarrollo de interfaces visuales que permitan a los distintos actores de la organización acceder al contexto almacenado con el propósito de consultarlo, actualizarlo o integrarlo con diversas herramientas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se propone llevar adelante las siguientes actividades:

- Realizar un análisis exhaustivo del estado del arte de la Inteligencia Artificial en el ámbito de la Ingeniería del Software, con el fin de identificar las tecnologías y herramientas de mayor impacto en la actualidad.

- Elaborar una proyección fundamentada de las tendencias emergentes en los próximos años.

- Redactar informes derivados del análisis realizado.

- Diseñar e implementar herramientas que aporten valor concreto al sector, tanto a escala nacional como internacional.

4. FORMACIÓN DE RECURSOS HUMANOS

Los progresos obtenidos en esta línea de investigación sirven como base para el desarrollo de tesis de posgrado, ya sea de doctorado o maestrías en Ingeniería de Software y desarrollo

de trabajos finales de las carreras Licenciatura en Ciencias de la Computación, Ingeniería en Informática e Ingeniería en Computación de la Universidad Nacional de San Luis, en el marco de los Proyectos de Investigación mencionados en la Sección *Contexto*.

5. BIBLIOGRAFÍA

[1] Stanford University, «AI Demystified: Introduction to large language models,» 13 12 2024. [En línea]. Available: https://uit.stanford.edu/service/techtraining/ai-demystified/llm.

[2] D. G. e. al., «DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning,» 17 09 2025. [En línea]. Available: http://dx.doi.org/10.1038/s41586-025-09422-z.

[3] OpenAI, «Aprendizaje de razonamiento con LLM,» 12 09 2024. [En línea]. Available: https://openai.com/es-419/index/learning-to-reason-with-llms/.

[4] Epoch AI, «LLMs now accept longer inputs, and the best models can use them more effectively,» 25 06 2025. [En línea]. Available: https://epoch.ai/data-insights/context-windows.

[5] P. L. e. al., «Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,» 12 04 2021. [En línea]. Available: https://arxiv.org/abs/2005.11401.

[6] H. H. e. al., «RAG vs. GraphRAG: A Systematic Evaluation and Key Insights,» 17 02 2025. [En línea]. Available: https://doi.org/10.48550/arXiv.2502.11371.

[7] The Linux Foundation Projects, «What is the Model Context Protocol (MCP)?,» 25 11 2024. [En línea]. Available: https://modelcontextprotocol.io/docs/getting-started/intro.

[8] D. Georgiev, «Exploring knowledge management from a software engineering perspective,» 05 09 2023. [En línea]. Available: https://doi.org/10.34190/eckm.24.2.1497.

[9] T. Winters, M. Tom y W. Hyrum, Software Engineering at Google, O'Reilly, 2020.

[10] S. Blea, «Tribal Knowledge – The Anti-DevOps Culture,» 27 06 2019. [En línea]. Available: https://devblogs.microsoft.com/premier-developer/tribal-knowledge-the-anti-devops-culture/.

[11] E. Jabrayilzade, M. Evtikhiev, E. Tüzün y V. Kovalenko, «Bus factor in practice,» 17 10 2022. [En línea]. Available: https://dl.acm.org/doi/10.1145/3510457.3513082.

[12] H. Lin, «Architecting efficient context-aware multi-agent framework for production,» 04 12 2025. [En línea]. Available: https://developers.googleblog.com/architecting-efficient-context-aware-multi-agent-framework-for-production/.

[13] N. Verma, «Active Context Compression: Autonomous Memory Management in LLM Agents,» 12 01 2026. [En línea]. Available: https://doi.org/10.48550/arXiv.2601.07190.

[14] Anthropic, «Effective context engineering for AI agents,» 29 09 2025. [En línea]. Available: https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents.

[15] LangChain, «Context Engineering,» 2 07 2025. [En línea]. Available: https://blog.langchain.com/context-engineering-for-agents/.

ANALÍTICAS PERMEABLES A LA INCERTIDUMBRE PARA LA TOMA DE DECISIONES EN LA INGENIERÍA DE SOFTWARE

Carlos Casanova[1], Esteban Schab[1], Lucas Prado[1], Gabriela Elizalde Caviglia[1], Lucas Forni[1], Tobías Grandi[1], Giovanni Daián Rottoli[1], Agustina Anselmino[2,1], Mariana Cóccola[2,1], Fabiana Piccoli[3]

[1]*GIICOS, Univ. Tecnológica Nacional, Facultad Regional Concepción del Uruguay (UTN-FRCU)*
[2]*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*
[3]*LIDIC- Univ. Nacional de San Luis (UNSL)*
{casanovac, schabe, pradol, elizaldeg, fornil, grandit, rottolig, anselminoa}@frcu.utn.edu.ar
marcoccola@santafe-conicet.gov.ar
mpiccoli@unsl.edu.ar

RESUMEN

Las Analíticas se definen como el proceso científico de transformar datos en conocimiento para tomar mejores decisiones. El diseño de tales analíticas no es una tarea sencilla, ya que deben considerar aspectos como la toma de decisiones multicriterio y multipersona, la fusión de información de diferentes fuentes y naturalezas, el análisis de datos e información, el manejo de software matemático sofisticado, la resolución de problemas, la interpretación del significado de la información para otros, etc. En este contexto es necesario considerar la posibilidad de que los datos disponibles puedan ser incompletos, imprecisos o incluso inconsistentes, posibilidad que debe ser tenida en cuenta en la generación de analíticas. Cualquiera de las formas de incertidumbre mencionadas debe ser reducida, o bien incorporada, cuantificada y analizada a la hora de producir información que facilite la toma de decisiones, ya que de otro modo la utilidad de las Analíticas sería efímera y difícilmente justifique el esfuerzo de confeccionarlas. Esto es particularmente cierto en contextos dinámicos como los procesos de desarrollo de software. Por lo tanto, se propone construir una colección de Analíticas que brinden información para facilitar la toma de decisiones científica en problemas de ingeniería de software caracterizados por información incompleta, imprecisa o inconsistente.

Palabras clave: Incertidumbre, Ingeniería de Software, Optimización, Inteligencia Computacional, Ciencia de Datos.

CONTEXTO

El presente trabajo se desarrolla en el ámbito del Grupo de Investigación sobre Inteligencia Computacional y Optimización de Sistemas (GIICOS), perteneciente al Departamento Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Concepción del Uruguay.

1. INTRODUCCIÓN

La ingeniería de software es la aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento del software [1], esto es, la aplicación de la ingeniería al software. Como toda disciplina ingenieril, en ella se combinan la técnica y el arte en la resolución de problemas relacionados con su objeto de estudio, esto es, el software. Sin embargo, existen aspectos relativos a este que introducen desafíos únicos para esta ingeniería. En este sentido la comunidad científica viene realizando esfuerzos en la definición realista de problemas propios de la ingeniería de software. En particular, la **Ingeniería de Software Basada en Búsqueda (ISBB)** [2] es una disciplina que pretende, mediante la aplicación de métodos de búsqueda, ayudar al ingeniero de software a lograr productos de calidad. El primer paso para aplicar cualquier técnica de búsqueda exitosamente es formular el problema como uno de búsqueda/optimización. Por lo tanto en esta línea pueden encontrarse muchos problemas formalizados. No obstante, enfoques como el

de la ISBB no han encontrado asidero en la industria.

1.1. Ingeniería de Software e Incertidumbre

Existen aspectos relativos a los problemas que surgen en el proceso de desarrollo de software que introducen desafíos únicos para esta ingeniería. Junto a otras características comunes se destacan la necesidad de tratar con información incompleta, imprecisa o inconsistente desde la formulación de los problemas. Este escenario es propio de entornos donde se desenvuelven los seres humanos.

La presencia de incertidumbre en estos contextos ha llevado a la industria del software a agilizar sus metodologías, descartando la prescripción de planes rígidos, priorizando la inclusión del cliente en el proceso de software y flexibilizando los contratos [3]. Una de las razones que puede explicar la escasa utilización de metodologías científicas para administrar proyectos de software es que los modelos y algoritmos que tradicionalmente se han utilizado no tienen en cuenta desde su formulación, en su gran mayoría, la incertidumbre inherente al proceso de desarrollo de software [4]. Además, la incertidumbre presente en las etapas tempranas de los proyectos de desarrollo de software difícilmente sea una incertidumbre modelable adecuadamente con variables aleatorias, ya que este contexto se caracteriza por la escasez de información, y la dificultad para realizar estimaciones precisas.

1.2. Analíticas para la toma de decisiones

Las Analíticas se definen como el proceso científico de transformar los datos en conocimiento para tomar mejores decisiones [5]. Estas son un proceso de dos pasos: el conocimiento se genera desde los datos y, luego, las decisiones se toman con base en dicho conocimiento.

El diseño de tales Analíticas no es una tarea sencilla, ya que deben considerar aspectos como la toma de decisiones multicriterio y multipersona, la fusión de información de diferentes fuentes y naturalezas, el análisis de datos e información, el manejo de software matemático sofisticado, la resolución de problemas, la interpretación del significado de la información para otros, etc. Estos aspectos introducen distintas formas de incertidumbre, las cuales deben ser reducidas, o bien incorporadas, cuantificadas y analizadas a la hora de producir información que facilite la toma de decisiones, ya que de otro modo la utilidad de las Analíticas sería efímera y difícilmente justifique el esfuerzo de confeccionarlas. Esto es particularmente cierto en contextos de gran volatilidad e incertidumbre, como el de la Ingeniería de Software, donde los requerimientos pueden ser imprecisos; las negociaciones, de mucho conflicto; las estimaciones, de baja especificidad, y la información, escasa, en términos generales. En estos contextos se suele renunciar a la generación de Analíticas y se adoptan enfoques más reactivos, heurísticos y cortoplacistas, cuyas políticas pueden resultar de muy inferior calidad a las de un enfoque científico adecuado.

1.3. Investigación Operativa

La Investigación de Operaciones (IO) es un enfoque científico en la toma de decisiones que busca el mejor diseño y operación de un sistema, por lo regular en condiciones que requieren la asignación de recursos escasos [6]. Según la *Operational Research Society*, la IO es la aplicación del método científico a los problemas complejos producidos en la dirección y gestión de grandes sistemas de personas, máquinas, materiales y dinero, en la industria, comercio, administración y defensa. El enfoque característico consiste en construir un modelo científico del sistema, incorporando medidas de factores tales como el azar, el riesgo u otros tipos de incertidumbre, mediante el cual se pueda predecir y comparar los resultados de las diferentes decisiones, estrategias o controles posibles. El objetivo es ayudar a los responsables a determinar su política y acciones en forma científica.

La IO se ocupa también de la administración práctica de la organización. Por lo tanto, para tener éxito, también debe proporcionar conclusiones claras que el tomador de

decisiones pueda usar cuando sea necesario [7].

1.4. Inteligencia Computacional

La Inteligencia Computacional (IC) toma su inspiración de la naturaleza para desarrollar sistemas inteligentes basados en computadoras. Los pilares principales que componen a la IC son las Redes Neuronales, que modelan aspectos del funcionamiento del cerebro, los sistemas difusos, que modelan la forma en la que las personas describen el mundo a su alrededor, y computación evolutiva, que modela aspectos sobre la variación y selección natural en la bioesfera [8]. Estos múltiples métodos no son competitivos entre sí, sino que son complementarios y pueden ser utilizados juntos para resolver un problema dado [9].
La complejidad de ciertos problemas produce que no sea posible utilizar técnicas tradicionales para abordarlos si lo que se pretende son buenas soluciones en lapsos razonables. Más aún, para problemáticas cercanas a la realidad, las hipótesis y supuestos de las mismas pueden hacer imposible la aplicación de técnicas de computación convencional, generalmente debido a la presencia de formas específicas de incertidumbre en los parámetros. La IC, en tal caso, apunta a resolver los problemas aprovechando la imprecisión y la incertidumbre presente en el proceso de toma de decisiones, brindando las "soluciones más satisfactorias" en tal contexto de toma de decisiones.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación se describen las líneas de investigación dentro del equipo.

2.1. Formulación realista de problemas de la IS

A nivel de industria de software mundial, ninguna técnica formal ha tenido un impacto significativo [3]. Una de las razones que puede explicar este fenómeno es que los modelos y algoritmos que se utilizan no tienen en cuenta desde su formulación, en su gran mayoría, la incertidumbre inherente al proceso de desarrollo de software (y por lo tanto, las soluciones obtenidas resultan de una vida útil casi efímera), o bien, si la contemplan, los modelos y resultados son difíciles de entender para una persona no experta en optimización o inteligencia artificial [4]. En consecuencia, se busca formular modelos matemáticos utilizando mecanismos formales de captación de incertidumbre, esto es, permitir que la incertidumbre *permee* en el proceso de generación de la analítica utilizando un modelo adecuado como los conjuntos difusos, o las distribuciones de posibilidades, modelos que pueden resultar adecuados para la realidad de un proceso de desarrollo de software. En este sentido, se trabaja junto a expertos de la industria del software en la especificación de problemas de la IS, con el objetivo de darle utilidad a las soluciones que se pueden obtener con los distintos algoritmos disponibles.

2.2. Construcción de agentes basados en aprendizaje por refuerzo

Las analíticas pueden clasificarse en descriptivas, diagnósticas, predictivas y prescriptivas. Los datos generados por los procesos de desarrollo de software pueden ser utilizados en algoritmos de aprendizaje supervisado o no supervisado para la generación de las tres primeras clases de analíticas. Sin embargo, una dificultad en lo relativo a las analíticas prescriptivas es que no cuentan con un "profesor" para decirles qué acción tomar en cada circunstancia.
Un tipo de aprendizaje, el cual no necesita de un profesor es el llamado aprendizaje por refuerzo [10]. En este esquema es el propio agente quien es capaz de juzgar y criticar sus acciones con base en sus percepciones y de alguna medida de aptitud, recompensa o refuerzo. La tarea del aprendizaje por refuerzo es usar tales recompensas para aprender una política óptima (o aproximadamente óptima) del entorno, sin asumir ningún conocimiento *a priori*. Esta política le dice al agente qué hacer en cada estado posible a alcanzar. Por lo tanto, se juzga al aprendizaje por refuerzo como una herramienta muy útil para la elaboración de

analíticas prescriptivas, por compartir ambos el mismo objetivo.

2.3. Integración con herramientas existentes de gestión de proyecto / proceso / producto

Como ya se ha dicho, los modelos y resultados de esta línea de investigación pueden resultar difíciles de entender para una persona no experta en optimización o inteligencia artificial. Se pretende, por lo tanto, construir herramientas software que soporten la toma de decisiones, que sean usables por administradores no expertos en IC. Es razonable suponer que si esta herramienta se integra a herramientas que la industria ya utiliza tendrá mayores posibilidades de ser incorporada.

2.4. Paradigmas de Cómputo Alternativos

Cada vez se requiere mayor poder de procesamiento para abordar los distintos modelos de decisión descriptos y su consecuente análisis. Es por eso que se propone la utilización de paradigmas de computación paralela y cuántica (simulada) para suplir esta necesidad creciente de poder de cómputo. La computación paralela se ha aplicado ampliamente en el campo de la IC [11] y la cuántica posee como una de sus fortalezas la resolución de problemas de optimización mediante el algoritmo de temple cuántico y otros métodos relacionados [12].

3. RESULTADOS OBTENIDOS/ESPERADOS

Hasta el momento se han resuelto satisfactoriamente varios problemas, aunque ninguno de ellos está agotado. Entre ellos se encuentran el problema del próximo lanzamiento (*Next Release Problem*, NRP), resuelto mediante múltiples algoritmos evolutivos como NSGA-II, IBEA y una versión novedosa de PSO denominada FMOPSO [13], desarrollada por los autores, escrita en C++. Otro problema es el de la planificación de lanzamientos (*Release Planning Problem*, RPP), una extensión del NRP, en este caso utilizando programación matemática difusa, incorporando restricciones blandas [14]. También se ha resuelto el problema de priorización de requerimientos de software utilizando relaciones de preferencia difusa mediante etiquetas difusas [15].

También se ha logrado una primera versión de un explorador de frentes de Pareto que pretende ser usable por usuarios no expertos en técnicas de optimización o de inteligencia computacional, haciendo uso de la librería *Dash*. Este software incorpora la posibilidad de carga de las soluciones de los algoritmos de búsqueda, su procesamiento mediante algoritmos de aprendizaje no supervisado y su presentación en un lenguaje comprensible para el tomador de decisiones haciendo uso de descriptores estadísticos y herramientas de visualización de información [16].

Además se proyecta poder construir *plugins* o mecanismos de interoperabilidad con *suites* existentes usadas en la industria.

Finalmente, se ha desarrollado un marco para la construcción de analíticas prescriptivas que están compuestas de recomendaciones de acciones a llevar a cabo, combinando prescripciones iniciales, análisis de sensibilidad, y la posibilidad de actualizar los parámetros ante la llegada de nuevos datos. En particular se ha definido un nuevo problema llamado *NRP Consonante* con todos estos conceptos aplicados al problema NRP [17].

4. FORMACIÓN DE RECURSOS HUMANOS

La investigación presentada constituye las líneas fundacionales del GIICOS. Tres investigadores se encuentran realizando su tesis de doctorado. Además participan en el proyecto cuatro becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en el contexto de esta investigación.

5. BIBLIOGRAFÍA

[1] «ISO/IEC/IEEE International Standard - Systems and software engineering--Vocabulary», *ISOIECIEEE 247652017E*, pp. 1-541, 2017, doi: 10.1109/IEEESTD.2017.8016712.

[2] T. Colanzi, W. Assunção, … S. V.-I. and, y U. 2020, «The Symposium on Search-Based Software Eengineering: Past, Present and Future», *Inf. Softw. Technol.*, vol. 127, 2020.

[3] D. C. C. Peixoto, G. R. Mateus, y R. F. Resende, «The issues of solving staffing and scheduling problems in software development projects», en *Proceedings - International Computer Software and Applications Conference*, 2014. doi: 10.1109/COMPSAC.2014.96.

[4] J. M. Rojas y G. Fraser, «Is search-based unit test generation research stuck in a local optimum?», en *Proceedings - 2017 IEEE/ACM 10th International Workshop on Search-Based Software Testing, SBST 2017*, 2017, pp. 51-52. doi: 10.1109/SBST.2017.4.

[5] C. Holsapple, A. Lee-Post, y R. Pakath, «A unified foundation for business analytics», *Decis. Support Syst.*, vol. 64, n.º C, pp. 130-141, ago. 2014, doi: 10.1016/j.dss.2014.05.013.

[6] A. Ebrahimnejad y J. L. Verdegay, *Studies in Fuzziness and Soft Computing: Fuzzy Sets-Based Methods and Techniques for Modern Analytics*. Springer International Publishing AG, 2018. doi: 10.1007/978-3-319-73903-8.

[7] F. S. Hillier y G. J. Lieberman, *Introducción a la Investigación de operaciones*. McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V., 2010.

[8] D. B. Fogel, D. Liu, y J. M. Keller, *Fundamentals of Computational Intelligence: Neural Networks, Fuzzy Systems, and Evolutionary Computation*, 1.ª ed. Wiley, 2016. doi: 10.1002/9781119214403.

[9] N. Siddique y H. Adeli, *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*, 1.ª ed. Wiley, 2013. doi: 10.1002/9781118534823.

[10] R. S. Sutton y A. G. Barto, *Reinforcement Learning, second edition: An Introduction*. MIT Press, 2018.

[11] B. S. P. Mishra y S. Dehuri, «Parallel Computing Environments: A Review», *IETE Tech. Rev.*, vol. 28, n.º 3, pp. 240-247, may 2011, doi: 10.4103/0256-4602.81245.

[12] A. Das y B. K. Chakrabarti, *Quantum Annealing and Related Optimization Methods*. Springer Science & Business Media, 2005.

[13] C. Casanova, G. D. Rottoli, E. Schab, L. Bracco, F. Pereyra, y A. De Battista, «Fuzzy Bi-objective particle swarm optimization for next release problem», en *Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE*, Lisboa, Portugal, 2019, pp. 509-512. doi: 10.18293/SEKE2019-082.

[14] C. Casanova, F. Pereyra Rausch, y L. Prado, «Modelo de Planificación de Lanzamientos de Software utilizando Restricciones Blandas», en *VII Congreso Nacional de Ingeniería Informática / Sistemas de Información*, 2019.

[15] G. D. Rottoli y C. Casanova, «Multi-criteria and Multi-expert Requirement Prioritization using Fuzzy Linguistic Labels», *ParadigmPlus*, vol. 3, n.º 1, pp. 1-18, feb. 2022, doi: 10.55969/paradigmplus.v3n1a1.

[16] C. Casanova, E. Schab, L. Prado, y G. D. Rottoli, «Hierarchical Clustering-based Framework for a posteriori exploration of Pareto fronts: application on the bi-objective Next Release Problem», *Front. Comput. Sci.*, vol. 5, 2023, doi: 10.3389/fcomp.2023.1179059.

[17] C. A. Casanova Pietroboni, E. A. Schab and L. M. Prado, «Uncertainty-permeable Analytics for the Next Release Problem through Fuzzy Mathematical Programming and Possibility Theory», 2024 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Yokohama, Japan, 2024, pp. 1-8, doi: 10.1109/FUZZ-IEEE60900.2024.10612024.

Aplicaciones móviles 3D con Realidad Virtual y Realidad Aumentada

Pablo Thomas, Federico Cristina, Sebastián Dapoto, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{pthomas, fcristina, sdapoto, ppesado}@lidi.info.unlp.edu.ar

Resumen

Se presenta una línea de investigación y desarrollo, que tiene por objeto estudiar temas relacionados con aspectos de Ingeniería de Software, orientados al desarrollo de aplicaciones móviles tridimensionales (3D) sobre diversas plataformas. En particular, se pone el foco en el desarrollo de aplicaciones 3D inmersivas, a través del uso de Realidad Virtual (RV) y Realidad Aumentada (RA), y en su rendimiento gráfico y eficiencia energética. También se realiza investigación en el desarrollo de aplicaciones móviles basadas en la ubicación, haciendo uso de métodos de posicionamiento en interiores.

Palabras claves: Dispositivos Móviles – Aplicaciones 3D – Realidad Virtual – Realidad Aumentada – M-Learning – Rendimiento gráfico 3D – Eficiencia energética 3D – Posicionamiento en interiores

Contexto

Esta línea de Investigación forma parte del proyecto "*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital.*", en particular del subproyecto "*Ingeniería de Software para escenarios híbridos"*, del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Existe una importante cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Introducción

En la actualidad los dispositivos móviles permiten ejecutar aplicaciones complejas y con exigentes requerimientos de hardware. Debido a esto, existen cada vez más alternativas de motores de juego que permiten desarrollar aplicaciones tridimensionales para dispositivos móviles.

La RV es una simulación interactiva por computadora en la cual se sustituye el mundo real con información sensorial que recibe el usuario. La RV permite generar entornos inmersivos donde el usuario puede interactuar con representaciones virtuales de objetos, que de otro modo sería difícil o imposible de acceder [1].

La RA es el conjunto de tecnologías que permiten que el usuario visualice el mundo real con información virtual añadida, mediante un dispositivo tecnológico. Los elementos físicos reales y tangibles se combinan con elementos virtuales. La RA es interactiva y en tiempo real [2].

El ámbito educativo debe adaptarse a los cambios y nuevas formas de aprendizaje. M-learning (mobile learning) plantea métodos modernos de apoyo al proceso de aprendizaje mediante el uso de dispositivos móviles. Las aplicaciones móviles 3D son una herramienta ideal para acercar a los alumnos [3] [4] [5].

Los sistemas de RA pueden transformar la forma de aprender y trabajar [6]. Mediante la utilización de gafas de RA y sensores inteligentes, es posible, por ejemplo, asistir a un técnico que realiza tareas de mantenimiento y deba realizar procedimientos con las manos libres.

A su vez, muchos de los potenciales usuarios de las aplicaciones 3D con RV o RA pueden no disponer de dispositivos de última generación, por lo que resulta de gran interés realizar un análisis en profundidad de los parámetros que inciden en el rendimiento gráfico de este tipo de aplicaciones.

De la misma forma, otro aspecto de relevancia es el consumo de energía de este tipo de aplicaciones, por lo que también es importante realizar un estudio que permita identificar cuáles son las características que inciden en la eficiencia energética.

Por otro lado, el posicionamiento en interiores (en inglés, indoor localization) puede ser esencial para actividades como la atención médica, la seguridad, los juegos de realidad aumentada y la asistencia a personas, entre muchos otros servicios basados en la ubicación. Aunque existen diferentes técnicas para estimar el posicionamiento en espacios interiores [7]; la triangulación WiFi se presenta como una tecnología ideal para abordar este desafío [8]. Esta técnica se basa en la medición de la intensidad de señales WiFi provenientes de puntos de acceso distribuidos en el entorno. Al analizar las diferencias de intensidad de señal entre los puntos de acceso detectados por un dispositivo móvil, es posible determinar la ubicación relativa del usuario en el espacio.

Una alternativa para la estimación del posicionamiento en interiores es el uso de técnicas de reconocimiento de imágenes. Este enfoque requiere la construcción de una base de datos con imágenes de los distintos espacios a identificar. A partir de ella, mediante técnicas de clasificación visual y modelos de aprendizaje automático, se extraen características distintivas de cada espacio, las cuales se asocian posteriormente a una ubicación específica. [9]

Líneas de Investigación y Desarrollo

- Metodologías y técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones móviles 3D multiplataforma
- Mobile Learning
- Frameworks para el desarrollo de aplicaciones móviles 3D
- Realidad virtual y realidad aumentada en aplicaciones móviles 3D
- Rendimiento gráfico y eficiencia energética de aplicaciones móviles 3D
- Aplicaciones móviles basadas en el posicionamiento en interiores

Resultados esperados/obtenidos

Los resultados esperados/obtenidos se pueden resumir en:

- Avanzar en la capacitación continua de los miembros de la línea de investigación.
- Avanzar en el aprendizaje de frameworks o motores para el desarrollo de aplicaciones 3D multiplataforma, particularmente para dispositivos móviles [10] [11] [12] [13].
- Avanzar en el desarrollo de aplicaciones móviles 3D con realidad virtual y realidad aumentada, teniendo como finalidad enriquecer las experiencias interactivas.
- Avanzar en el estudio del rendimiento gráfico y la eficiencia energética en aplicaciones móviles 3D con realidad virtual y realidad aumentada.
- Avanzar en la investigación y desarrollo de aplicaciones móviles basadas en el posicionamiento en interiores.
- Se ha desarrollado una aplicación móvil 3D con RV para la evaluación del rendimiento gráfico en este tipo de entornos. La aplicación está desarrollada en Unity y permite evaluar entornos con y sin RV mediante una serie de pruebas que aíslan diferentes características gráficas. Figuras 1 y 2.

- Se ha desarrollado una aplicación con RA para visualizar el funcionamiento de las compuertas lógicas de manera didáctica, interactiva y accesible. Dado que los circuitos lógicos constituyen uno de los temas introductorios en las carreras vinculadas a la disciplina informática, el objetivo de la aplicación es facilitar el proceso de aprendizaje. Figura 3.
- Se ha desarrollada una aplicación con RV para el tratamiento de la acrofobia. La aplicación favorece la desensibilización y la habituación a las alturas, permitiendo que la persona afronte sus miedos de forma segura y guiada. Figura 4.
- Se ha desarrollado un prototipo de asistencia de ubicación para personas ciegas en la Facultad de Informática de la Universidad Nacional de La Plata. El prototipo determina la ubicación del usuario de dos formas diferentes: mediante la red WiFi o a partir del análisis de una imagen capturada por el propio usuario. Figura 5.
- Se está avanzando en el desarrollo de un prototipo que permita medir el rendimiento gráfico de las aplicaciones móviles 3D con RA.

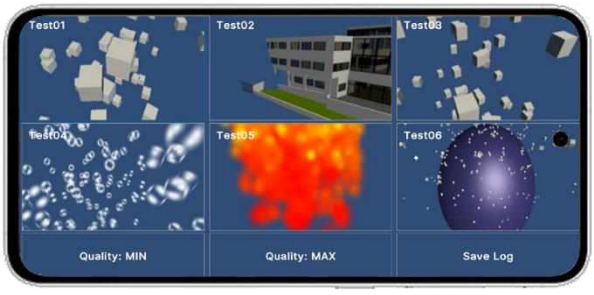


Figura 1. Aplicación móvil 3D para la evaluación del rendimiento gráfico sin RV.



Figura 2. Aplicación móvil 3D para la evaluación del rendimiento gráfico con RV.

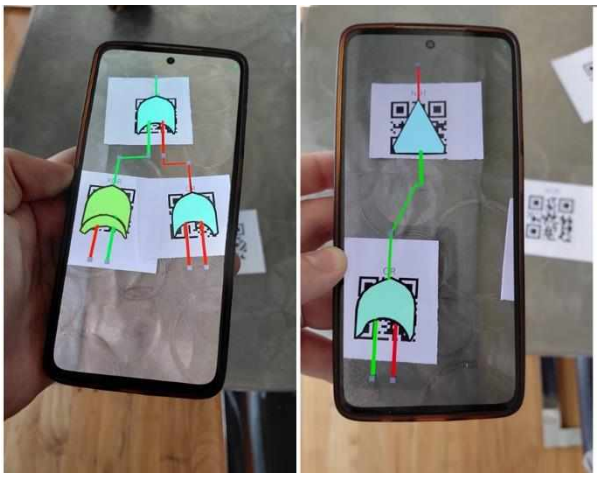


Figura 3. Aplicación móvil con RA para visualizar el funcionamiento de las compuertas lógicas.



Figura 4. Aplicación móvil con RV para el tratamiento de la acrofobia.



Figura 5. Prototipo de asistencia de ubicación para personas ciegas. Pantalla principal, captura de imagen y resultado del posicionamiento.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Prácticas Profesionales Supervisadas, Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto.

Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

▪ Referencias

1. Linowes J. “Unity Virtual Reality Projects”. 2015. ISBN-13: 978-1783988556.
2. R. Silva, J.C. de Oliveira, G. Giraldi. "Introduction to augmented reality". 2003.
3. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. Capítulo de Libro: “3D Mobile Prototype for Basic Algorithms Learning”. Libro: “Computer Science & Technology Series - XXI Argentine Congress Of Computer Science. Selected Papers” (300 páginas). EDULP. ISBN: 978-987-4127-00-6, páginas 239-247. Año 2016.
4. Kantel E., Tovar G., Serrano A.”Diseño de un Entorno Colaborativo Móvil para Apoyo al Aprendizaje a través de Dispositivos Móviles de Tercera Generación.” IEEE-RITA 5, no. 4 (2010): 146-151.
5. Rosa Paredes, J. Alfredo Sánchez, Liliana Rojas, Daniel Strazzulla, Ronel Martínez-Teutle. "Interacting with 3D Learning Objects". 2009 LA Web Congress. ISBN: 978-0-7695-3856-3/09.
6. X. Pan, X. Sun, H. Wang, S. Gao, N. Wang and Z. Lin, "Application of an assistant teaching system based on mobile augmented reality (AR) for course design of mechanical manufacturing process," 2017 IEEE 9th International Conference on Engineering Education (ICEED), 2017, pp. 192-196, doi: 10.1109/ICEED.2017.8251191.
7. Gu, Fuqiang & Hu, Xuke & Ramezani, Milad & Acharya, Debaditya & Khoshelham, Kourosh & Valaee, Shahrokh & Shang, Jianga. (2019). Indoor Localization Improved by Spatial Context - A Survey. ACM Computing Surveys. 52. 64:1-35. 10.1145/3322241.
8. S. Qiao, C. Cao, H. Zhou and W. Gong, "The trip to WiFi indoor localization across a decade — A systematic review," 2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Rio de Janeiro, Brazil, 2023, pp. 642-647, doi: 10.1109/CSCWD57460.2023.10152700.
9. Documentación TensorFlow Lite. https://www.tensorflow.org/lite/guide
10. Unity 3D Homepage: https://unity3d.com/.
11. Unreal Engine Homepage. https://www.unrealengine.com/.
12. CryEngine Homepage. https://www.cryengine.com/
13. Godot Engine Homepage. https://godotengine.org

APORTES DE LA INTELIGENCIA ARTIFICIAL A LA GESTIÓN DEL CONOCIMIENTO EN EL ÁMBITO EMPRESARIAL

Ansalas, María Eugenia[1]; Straccia, Luciano[1,2]; Vegega, Cinthia[2]

[1]Universidad Tecnológica Nacional, Facultad Regional La Plata, Programa de Maestría en Administración de Empresas, La Plata, Argentina
maria.ansalas@frlp.utn.edu.ar

[2]Universidad Tecnológica Nacional, Facultad Regional Buenos Aires, Grupo GEMIS.BA, Ciudad Autónoma de Buenos Aires, Argentina
lstraccia@frba.utn.edu.ar, cvegega@frba.utn.edu.ar

RESUMEN

La gestión del conocimiento (GC) aborda la creación y el uso del conocimiento en las organizaciones desde un enfoque multidisciplinar. El Grupo GEMIS, a través de diversos proyectos, se propone analizar la gestión del conocimiento desde diferentes enfoques y perspectivas integrales. Este trabajo presenta los proyectos de investigación desarrollados en dicho marco, los resultados obtenidos y un nuevo proyecto que busca incorporar aportes desde la Inteligencia Artificial. Además, este proyecto articula propuestas de diversas Facultades y carreras de la Universidad Tecnológica Nacional generando aportes sustantivos en la formación de recursos humanos.

Palabras clave: gestión del conocimiento, inteligencia artificial.

CONTEXTO

La Universidad Tecnológica Nacional (UTN) posee un Programa de Investigación, Desarrollo e Innovación, de Sistemas de Información e Informática. En la Facultad Regional Buenos Aires (UTN.BA), en el marco del Departamento de Ingeniería en Sistemas de Información, se llevan adelante proyectos de investigación y desarrollo. Allí desarrolla sus actividades, entre otros grupos de investigación, el grupo GEMIS.BA que trabaja en el campo de la Ingeniería en Sistemas de Información, Ingeniería de Software e Ingeniería del Conocimiento.

El Grupo GEMIS.BA realiza, en el marco de las temáticas objeto de este trabajo, diversos proyectos asociados a la línea de Gestión del Conocimiento con proyectos de investigación y desarrollo y actividades de docencia e investigación en diferentes instituciones contando con trabajos finales de especialización, tesis de maestría y un plan de tesis doctoral en proceso.

Diversos integrantes del Grupo forman parte de los equipos docentes de los programas de Maestría de la Universidad Tecnológica Nacional, Facultad Regional La Plata a través del dictado de diversos seminarios y la dirección de tesis de maestría.

1. INTRODUCCIÓN

El uso y la aplicación de la Inteligencia Artificial (IA) en el ámbito empresarial está siendo cada vez más habitual y está adquiriendo un rol diferencial por su carácter disruptivo y su potencialidad, utilizándose con el objeto de optimizar y dotar de mayor eficiencia a los

procesos del negocio y mejorar los servicios ofrecidos a sus clientes.

La Gestión del Conocimiento (GC) ha adquirido un papel cada vez más relevante como proceso clave en las empresas, para sostener la innovación, la mejora continua y la toma de decisiones estratégicas [1]. En esencia, la GC busca facilitar la creación, organización, transferencia y aplicación del conocimiento dentro de las organizaciones, permitiéndoles responder de manera ágil a los cambios del entorno, preservar su capital intelectual y potenciar su competitividad [2].

La integración de la IA en las organizaciones se refleja en los hallazgos de la Encuesta Anual Global de CEO - Capítulo Argentina [3] elaborada por la consultora PwC. El informe señala que las empresas han comenzado a evidenciar mejoras en el desempeño organizacional como resultado de su implementación y anticipan extender su aplicación en los próximos años.

Estudios recientes de McKinsey & Company [4] y Gartner [5], destacan que la automatización del conocimiento tiene un impacto significativo en la mejora de la eficiencia operativa y en la toma de decisiones estratégicas dentro de las organizaciones. Sin embargo, también advierten que su implementación comprende desafíos importantes, tanto en términos de adopción tecnológica como de gestión efectiva del cambio organizacional.

La IA, en particular a través del aprendizaje automático y el procesamiento de lenguaje natural, ofrece nuevas formas de automatizar y mejorar la manera en que las organizaciones gestionan su conocimiento. Estas tecnologías no solo permiten una mejor organización y análisis de grandes volúmenes de datos [6], sino que también pueden contribuir a la creación de conocimiento mediante la identificación de patrones ocultos, la predicción de comportamientos y la generación de perspectivas relevantes para la toma de decisiones estratégicas [7]. A su vez, Russell y Norvig [8] introducen la perspectiva desde la IA aplicada, enfocándose en su capacidad para mejorar tanto la eficiencia operativa como la creación de conocimiento organizacional.

A partir de las encuestas e informes anuales realizados por consultoras de prestigio internacional mencionadas, se puede inferir que existe un uso de las tecnologías de IA en los procesos internos de las organizaciones, por lo que este trabajo busca evaluar cómo esas herramientas tecnológicas se integran de manera efectiva y contribuyen en el proceso de Gestión del Conocimiento (GC). A pesar de las múltiples promesas que ofrece la IA, las organizaciones enfrentan desafíos importantes en su implementación, tales como la resistencia organizacional, los riesgos éticos y las limitaciones tecnológicas [9]. Vegega, Pytel y Pollo-Cattaneo [10] plantean que contar con herramientas que puedan "aprender" en forma automática a partir de los datos es muy útil, pero (al menos todavía) no reemplazan los conocimientos de los expertos del dominio. Sin embargo, se pueden aplicar modelos de IA que emulan la forma en que los expertos realizan sus actividades.

Hasta el momento y según informes recientes, los desarrollos más avanzados y significativos en la incorporación de IA en las organizaciones, se observa en empresas que han realizado inversiones estratégicas sostenidas en tecnología [3,4,5,11]. Estas organizaciones, generalmente con alta madurez digital han comenzado a integrar herramientas de IA en áreas específicas del negocio, particularmente en procesos de automatización documental, análisis predictivo, generación de reportes inteligentes y asistentes virtuales. Sin embargo, estas implementaciones suelen ser aún parciales y localizadas, lo que evidencia que el camino hacia una integración plena y transversal de la IA en la gestión del conocimiento, en particular, está en una etapa inicial.

La IA está ayudando a resolver problemas tanto a las personas como a las empresas. Alguno de los aportes más comunes es en áreas vinculadas a la sobrecarga informativa, la lentitud en los procesos de búsqueda, el acceso al conocimiento, y la falta de análisis predictivo en decisiones basadas en datos. No obstante, subsisten y persisten ciertas limitaciones como la dificultad para capturar conocimiento tácito [12,13], que sigue siendo mayormente humano y contextual, los riesgos éticos y organizacionales, como sesgos algorítmicos, pérdida de control y resistencia cultural [9,14] y la falta de integración estratégica, dado que ciertas implementaciones de IA son aisladas y no responden a una visión global de la empresa.

En un entorno de negocios en constante evolución, donde la innovación tecnológica es un factor determinante, este proyecto puede guiar a las organizaciones a comprender cómo capitalizar el uso de la IA para mejorar y potenciar su proceso de GC.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el año 2018 se llevó a cabo el primer proyecto de investigación y desarrollo (PID) del Grupo GEMIS.BA asociado a GC en la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional. Este proyecto buscaba indagar cuál era el grado de implementación de GC en las pequeñas y medianas fábricas de software del Área Metropolitana de Buenos Aires (AMBA), cuáles eran los problemas presentes en esta implementación y qué modelos se podrían proponer para que esas organizaciones pudieran implementarlos (PID UTN, 2018-2020). Este proyecto obtuvo resultados construyendo un modelo general pero que se centraba especialmente en las actividades, desconociendo (o minimizando) otros aspectos.

A partir de los resultados obtenidos en el proyecto anterior y atendiendo a la necesidad de indagar en algunos aspectos particulares de la GC, en el año 2021 se iniciaron proyectos asociados a aspectos tecnológicos y a la medición de algunos de esos aspectos. En la UTN.BA se realizó un proyecto que busca analizar modelos de representación del conocimiento y proponer tecnologías y arquitecturas tecnológicas para la implementación de GC en las organizaciones (PID UTN 2021-2024). Además, entre los años 2021 y 2023 uno de los investigadores del Grupo dirigió un proyecto en la Universidad Argentina de la Empresa (UADE) cuyo objetivo fue definir un modelo de medición de GC construyendo indicadores sobre las tecnologías para la medición de GC y proponiendo un modelo de implementación de los indicadores en las organizaciones (PID UADE, 2021-2023). Finalmente, en el año 2023 se inició el proyecto de desarrollo de un observatorio tecnológico para relevar y monitorear el estado de aplicación de la gestión del conocimiento en organizaciones de la Ciudad de Buenos Aires (PID UTN, 2023-2025).

A partir de los resultados obtenidos de los proyectos mencionados se identifica la necesidad de incorporar una perspectiva que considere las discusiones actuales en el marco de los avances tecnológicos y el impacto de las nuevas tecnologías sobre los avances de la sociedad, considerándose a la Inteligencia Artificial como un aspecto central de las nuevas discusiones tecnológicas y sociales.

Este proyecto se inscribe como una continuidad y profundización de la línea de investigación en Gestión del Conocimiento, enfocándose específicamente en el análisis crítico del impacto de la Inteligencia Artificial —especialmente el aprendizaje automático y el procesamiento de lenguaje natural— en los procesos de gestión del conocimiento dentro de organizaciones empresariales. Así, busca aportar nuevos conocimientos y lineamientos estratégicos que

orienten a las empresas en la adopción efectiva y ética de herramientas de IA.

3. RESULTADOS OBTENIDOS/ESPERADOS

En [15] se exploran diferentes vistas o dimensiones a considerar en la gestión del conocimiento a través de un modelo integral: 1) individuos/personas; 2) aspectos organizacionales; 3) actividades y procesos; 4) tecnología y representación del conocimiento; 5) medición. En [16] se hace foco en los aspectos tecnológicos analizando y categorizando los conceptos asociados a la tecnología para la gestión del conocimiento, posteriormente ampliando en [17] en un enfoque de algunas tecnologías y la innovación.

Este proyecto tiene como objetivo analizar cómo la Inteligencia Artificial, en particular el aprendizaje automático y el procesamiento de lenguaje natural, impactan en la gestión del conocimiento dentro de empresas, identificando su contribución en las distintas actividades de la gestión del conocimiento organizacional, así como los desafíos y oportunidades asociados a su implementación.

Específicamente se buscará: 1) identificar el grado de adopción y los usos actuales de la IA en la gestión del conocimiento en empresas; 2) clasificar las principales contribuciones y limitaciones de la IA en la gestión del conocimiento; 3) diseñar una propuesta de lineamientos estratégicos para la implementación efectiva de IA en la gestión del conocimiento en empresas.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo propuesto se encuentra conformado por investigadores formados, investigadores de apoyo y alumnos de la carrera de Ingeniería en Sistemas de Información. Esta línea de trabajo busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para que asciendan dentro de la carrera de investigadores, además de fomentar la aplicación de las temáticas en su propia actividad profesional.

Además el equipo cuenta con investigadores estudiantes que realizan sus primeras participaciones en el ámbito de la investigación y cuya experiencia en el proyecto propuesto contribuirá especialmente al inicio de trayectos de formación académica. Se espera la incorporación de nuevos becarios alumnos y graduados a partir de las diferentes propuestas de becas de investigación que son llevadas a cabo por la Universidad y la Facultad.

Este proyecto ya cuenta con el desarrollo de un plan de tesis para la Maestría en Administración de Negocios de la Universidad Tecnológica Nacional, Facultad Regional La Plata de la autora principal de este trabajo. De esta forma se prevé la integración en la participación de actores de diversas carreras y facultades de la Universidad.

5. BIBLIOGRAFÍA

[1] Bukowitz, W., & Williams, R. (1999). The Knowledge Management Fielbook. Finalcial Times / Prentice Hall.

[2] Wiig, K. (1997). Knowledge management: Where did it come from and where will it go? Expert Systems with Applications, 13((1)), 1-14.[7] Belly, P. (s/f). Como vincular la Inteligencia Artificial con la Gestion del Conocimiento. Recuperado el 04 de marzo de 2025, de https://bellykm.com/como-vincular-la-inteligencia-artificial-con-la-gestion-del-conocimiento/

[3] PricewaterhouseCoopers. (2024). 27ª Encuesta Anual Global de CEO – Capítulo Argentina. PwC Argentina. Obtenido de https://www.pwc.com.ar/es/encuesta-ceo/assets/27-CEO-Survey-2024.pdf

[4] McKinsey & Company. (30 de Mayo de 2024). El estado de la IA a principios de 2024: La adopción de la IA generativa aumenta y comienza a generar valor.

[5] Gartner. (17 de Octubre de 2022). Gartner Top 10 Strategic Technology Trends for 2023. Obtenido de https://www.gartner.com/en/articles/gartner-top-10-strategic-technology-trends-for-2023

[6] Holdsworth, J. (11 de Agosto de 2024). ¿Qué es el procesamiento de lenguaje natural (PLN)? Obtenido de https://www.ibm.com/mx-es/think/topics/natural-language-processing

[7] Belly, P. (s/f). Como vincular la Inteligencia Artificial con la Gestion del Conocimiento. Recuperado el 04 de marzo de 2025, de https://bellykm.com/como-vincular-la-inteligencia-artificial-con-la-gestion-del-conocimiento/

[8] Russel, S., & Norvig, P. (2021). Inteligencia artificial: un enfoque moderno (4ª edición). Francia: Pearson.

[9] Torres, M. (15 de Febrero de 2023). Conecta. Obtenido de 8 desafíos de la adopción de inteligencia artificial en las empresas.

[10] Vegega, V., Pytel, P.,; Pollo-Cattaneo, M. (2016). Sistemas Inteligentes para clasificación automática de clientes: Caso de estudio. En Desarrollo e Innovación en Ingenieria (Primera ed., págs. 459-489). Instituto Antioqueño de Investigación.

[11] Accenture. (2025). Technology Vision 2025: Una nueva era de IA que transformará la autonomía en las empresas.

[12] Polanyi, M. (1996). The tacit dimension. Doubleday & Company

[13] Nonaka, I.; Takeuchi, H. (1995). The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford: University Press.

[14] Bostrom, N. (2017). Superintelligence: Paths, Dangers, Strategies. Oxford University Press

[15] Straccia, L.; Pollo-Cattaneo, M.F.; Maulini, A. (2022). Visões da Gestão do Conhecimento para a construção de um modelo integral. En Proceeding in Congresso Brasileiro de Gestão do Conhecimento, São Paulo, Brasil.

[16] Straccia, L.; Maulini, A.; Bongiorno, M.G.; Giorda, M.; Pollo-Cattaneo, M.F. Knowledge Representation and Technologies in the Latin American Academic Literature. ICAIW 2022: Workshop at ICAI, Arequipa, Perú.

[17] Straccia, L.; Pollo-Cattaneo, M.F. (2024) Knowledge Management: Innovation, Technologiy and Ontologies. ICAIW2024: Workshop at ICAI, Viña del Mar, Chile.

ASISTENCIA DE UN AGENTE COGNITIVO EN LA ELICITACIÓN DE REQUERIMIENTOS.

Ezequiel Moyano, Daniel Aguil Mallea, Ivan D'Uva, Fabiola Horas, Nadia Ramos, Raul Espíndola, Lisandro Ureta, Camila Burgos, Ariana Mendez.
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: +54-2901-432403
{emoyano, daguil, iduva, npramos, fhoras}@untdf.edu.ar, {raule2013, lisandroureta, camilaburgos142}@gmail.com, mendezari2000@icloud.com

Resumen

La ingeniería de requerimientos es una de las fases más críticas del ciclo de vida del software. Una especificación y gestión adecuada constituye uno de los desafíos más importantes de la industria, errores, omisiones o ambigüedades son la principal causa de sobre costos, retrasos y fracasos en los proyectos.

Los avances tecnológicos, sobre todo en Inteligencia Artificial (IA), han permitido comenzar a complementar las técnicas tradicionales para la especificación de requerimientos a través de modelos generativos basados en grandes modelos de lenguaje (LLM), junto a la integración con estrategias avanzadas de recuperación aumentada por generación (RAG).

La IA Generativa utiliza técnicas sofisticadas para crear contenido lógico y coherente, mientras que el uso de RAG representa un avance fundamental abordando las limitaciones intrínsecas de los LLM como las alucinaciones, el conocimiento desactualizado y la falta de transparencia y trazabilidad.

El proyecto propone un enfoque innovador basado en el concepto del "analista aumentado", donde agentes de IA, potenciados con técnicas de RAG, colaboran con analistas humanos; buscando liberar la carga cognitiva de los especialistas, automatizando la redacción y de documentos de especificaciones y la detección de inconsistencias, y que los analistas se concentren en la negociación, la estrategia y la resolución de problemas complejos.

Palabras claves: Ingeniaría de requerimientos, Inteligencia Artificial, Agentes cognitivos, Generación Aumentada por Recuperación.

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la Universidad Nacional de Tierra del Fuego, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática vinculada a la ingeniería de software, específicamente, en ingeniería de Requerimientos y su aplicabilidad en la Industria 4.0.

Actualmente esta línea se desarrolla en el marco de un proyecto de investigación (Oct/2025-Sept/2027) El mismo ha obtenido aval UNTDF.

1 Introducción

La ingeniería de requerimientos es fundamental debido a que es la primera, y más importante, etapa de todo paradigma de ingeniería de software; según Brooks "La parte más difícil de construir un sistema de software es decidir qué construir".

A través de este proceso se obtiene una Especificación de Requerimientos (ERS) que se desarrolla sobre la base de un modelo conceptual y va a representar abstracciones, supuestos y restricciones sobre el dominio de una futura aplicación[1]. La especificación de requerimientos representa la comprensión de las

necesidades, deseos y expectativas de los diferentes interesados y/o expertos del dominio (stakeholders) que se encuentran relacionados con él.

Una especificación y gestión adecuada de requerimientos de software constituye uno de los desafíos más importantes de la industria. La dificultad de la obtención de los requerimientos radica en la incertidumbre derivada del proceso de transferencia de información desde una o varias fuentes, no siempre consistentes; errores, omisiones o ambigüedades en esta etapa temprana son la causa principal de sobre costos, retrasos y fracasos en los proyectos tecnológicos.

Esta dificultad se ve incrementada por la sub-estimación por parte de personas de formación tecnológica de las denominadas técnicas "blandas" (metodologías que tratan con las interacciones humanas y los procesos sociales) que requiere el proceso de elicitación de requerimientos. La tarea de mejorar estos procesos es difícil y, por lo general, no sale como uno quisiera[2]; mejorar la comunicación dentro del equipo, de forma que sean capaces de trabajar en forma sistemática, genera que los proyectos de software tengan más probabilidad de completarse con el tiempo, el número de errores que encuentra el cliente disminuye sustancialmente, la credibilidad de la organización de software aumenta y la administración tiene un problema menos por el cual preocuparse[3].

El proceso de ingeniería de requerimientos continúa presentando importantes desafíos, una buena especificación de requerimientos es difícil y cada vez más compleja [4]; requiere de profesionales altamente formados (o especializados) en todos los procesos que involucra.

Específicamente, nuestra provincia no escapa a esa realidad donde los recursos humanos formados en la disciplina son particularmente escasos o se ven absorbidos por tareas rutinarias; motivo por el cual es usual encontrarse con dificultades al intentar que un equipo que utilice enfoques metodológicos para obtener mejores resultados[5].

A pesar de su importancia y todo el esfuerzo, los procesos de ingeniera de requerimientos siguen siendo, en gran medida, artesanal y dependiente de la pericia individual de los analistas; presentando problemáticas claves[6]:

- Elicitación incompleta y ambigua: La captura de necesidades de los stakeholders es un proceso inherentemente propenso a malentendidos, omisiones y falta de claridad, lo que obliga a ciclos de revisión extensos.
- Especificación lenta y propensa a errores: La traducción de conversaciones y documentos a artefactos formales, como historias de usuario o casos de uso, es una tarea manual, intensiva en tiempo y que requiere un alto nivel de habilidad para evitar la introducción de inconsistencias.
- Validación tardía e ineficiente: La detección de conflictos o requerimientos faltantes suele ocurrir en etapas avanzadas, cuando el costo de la corrección es exponencialmente mayor.

Los avances tecnológicos, sobre todo en diversas herramientas basadas en Inteligencia Artificial (IA), han permitido comenzar a complementar las técnicas tradicionales para la especificación de requerimientos a través del uso de estas herramientas, sobre todo con modelos generativos basados en grandes modelos de lenguaje (LLM), junto a la integración con estrategias avanzadas de recuperación aumentada por generación (RAG).

La IA Generativa utiliza técnicas avanzadas y sofisticadas para crear contenido lógico y coherente, a través de modelos que han sido entrenados con grandes cantidades de información (como GPT-4, Gemini, Copilot, etc), conocidos por su habilidad para producir textos detallados, relevantes y adaptados a diferentes situaciones[7]. Estos modelos utilizan redes neuronales entrenadas a través de

métodos de aprendizaje que permiten representar eficientemente aspectos semánticos y sintácticos del lenguaje natural[8]. Lo más importante es que las IA pueden contribuir al desarrollo de habilidades específicas, como la elicitación de requisitos[9].

La Generación Aumentada por Recuperación (RAG) es una técnica innovadora en el campo del Procesamiento del Lenguaje Natural (PLN) y la Inteligencia Artificial (IA) generativa, diseñada para mejorar la precisión, fiabilidad y contextualización de las respuestas de los Grandes Modelos de Lenguaje (LLMs); ya que se basa en el acceso de información actualizada de fuentes externas, enriqueciendo las interacciones[10].

Esta metodología aborda limitaciones inherentes a los LLMs tradicionales, como la generación de información inexacta, denominadas alucinaciones (debilidades críticas de los LLMs es su tendencia a "inventar" respuestas o generar texto inexacto y sin sentido, especialmente cuando carecen de información relevante), la obsolescencia de los datos de entrenamiento y la falta de conocimiento especializado.

RAG supera estas limitaciones al vincular los LLM con repositorios de conocimiento externos, actualizados y específicos del dominio en tiempo real.

El objeto de estudio es el desarrollo de un framework metodológico y tecnológico que permita integrar agentes inteligentes en el flujo de trabajo de la ingeniería de requerimientos. Se propone un enfoque innovador basado en el concepto del "analista aumentado", donde agentes de inteligencia artificial, potenciados con técnicas de recuperación aumentada por generación (RAG), colaboran activamente con los analistas humanos.

Esta aproximación es estratégicamente relevante porque busca, por un lado, liberar la carga cognitiva de los especialistas, automatizando la revisión de documentos, la redacción de especificaciones estándar y la detección de inconsistencias. Esto permite que los analistas humanos se concentren en la negociación, la estrategia y la resolución de problemas complejos. Por otro lado, el enfoque busca democratizar el conocimiento experto. El agente, alimentado por la documentación de proyectos anteriores y buenas prácticas de la industria, actúa como una memoria organizacional persistente y accesible para todo el equipo.

2 Línea de Investigación y Desarrollo

Como parte de la línea de investigación y desarrollo, el equipo de investigadores lleva adelante trabajos relacionados con las temáticas vinculadas a la ingeniería de software, específicamente, la importancia de la Ingeniería de Requerimientos y su impacto en el ámbito de de la industria 4.0; se plantean los siguientes objetivos a alcanzar en el presente proyecto:

Objetivos Generales

Diseñar un framework asistido por inteligencia artificial para optimizar la fase de Elicitación del proceso de ingeniería de requerimientos de software.

Objetivos Específicos

1.- Identificar puntos críticos y métricas de rendimiento del proceso tradicional de elicitación de requerimientos

2.- Analizar un conjunto de tecnologías de inteligencia artificial (LLM, frameworks de RAG, bases de datos vectoriales) adecuadas para la construcción de un agente de asistencia.

3.- Diseñar la arquitectura de un agente de elicitación.

4.- Desarrollar un prototipo funcional del agente enfocado en la elicitación.

5.- Evaluar la eficacia del prototipo mediante un estudio experimental comparativo, a través de métricas de tiempo, completitud y calidad de los requerimientos obtenidos.

Hipótesis

Se parte de la hipótesis de que el desarrollo y la aplicación de un framework metodológico y tecnológico basado agentes inteligentes, pueda colaborar activa y eficientemente con analistas humanos; permitiendo integrar en el flujo de trabajo mejorando la eficiencia, completitud y consistencia del proceso de ingeniería de requerimientos en comparación con los métodos tradicionales.

Esta aproximación es estratégicamente relevante porque busca, por un lado, liberar la carga cognitiva de los especialistas, automatizando la revisión de documentos, la redacción de especificaciones estándar y la detección de inconsistencias; permitiendo que los analistas humanos se concentren en la negociación, la estrategia y la resolución de problemas complejos.

Por otro lado, el enfoque busca democratizar el conocimiento experto. El agente, alimentado por la documentación de proyectos anteriores y buenas prácticas de la industria, actúa como una memoria organizacional persistente y accesible para todo el equipo.

3 Resultados Esperados

El proyecto se encuentra recién iniciado por lo que a la fecha no pueden enumerarse resultados obtenidos. El grupo de investigación ha comenzado a trabajar en las primeras actividades del proyecto que consiste en obtener un resumen general de las teorías y las prácticas de las metodologías actuales, un análisis de las diversas plataformas de IA generativas, como también respecto a recuperación aumentada por generación o bases de datos vectoriales (según convenga a los propósitos del proyecto), evaluando sus capacidades, arquitecturas y casos de éxito.

Diseñar una método de evaluación técnica con diversos criterios ponderados, como la facilidad, usabilidad, la curva de aprendizaje, etc., que justifique la elección en base a una matriz de evaluación.

Para el resto del proyecto se espera obtener:

Desde el plano académico, se espera obtener un framework metodológico validado, un análisis comparativo de tecnologías de inteligencia artificial para la ingeniería de requerimientos, una tesis de grado finalizada y la base para futuras publicaciones científicas.

Desde una perspectiva práctica e industrial, el proyecto generará un prototipo funcional que servirá como prueba de concepto para empresas de desarrollo, junto con una metodología para reducir tiempos y costos en las fases iniciales de los proyectos.

Finalmente, el impacto formativo se materializará en una mejora en la capacitación de los futuros ingenieros de software, exponiéndolos a herramientas de vanguardia y a una nueva forma de concebir el análisis de requerimientos.

4 Formación de Recursos Humanos

El Equipo de Trabajo está conformado por docentes investigadores de la UNTDF, licenciados en informática y alumnos de la Carrera.

El Proyecto incluye docentes en etapa de formación de postgrado, el Director del proyecto está realizando actualmente su tesis de Magister en Ingeniería de Software en la Universidad Nacional de La Plata, denominada: “Reuso de Requerimientos Referenciales en Dominios Complejos. Caso de Estudio: Aplicaciones de Modelado de Cuencas.” Dirigida por la Mgs. Adriana Urciuolo y co-dirigida por el Dr. Leandro Antonelli.

El Licenciado Aguil Mallea, se espera que mediante la realización del trabajo pueda enfocar finalmente sus tesis de posgrado, en alguna de las dimensiones que se desarrollarán, de la Maestría en Ingeniería de Software de la UNLP.

Los docentes Ivan DÚva y Fabiola Horas se encuentran cursando Postgrado en temas afines al proyecto. A su vez, ellos como el Lic. Brian

Rigoni puedan incorporar a sus desarrollos profesionales resultado del proyecto.

Colaborará y prestará apoyo un grupo de alumnos en formación de grado, con el objetivo de formarlos en los aspectos generales de la investigación y que le sirva para el desarrollo de su futura tesis de grado.

El estudiante Raúl Espíndola ha terminado de aprobar todo el plan de estudios de la Lic. en Sistemas y realizará su Tesis de Grado respecto a generar agentes IA que ayuden a lograr una buena especificación de requerimientos.

Los estudiantes Lisandro Ureta y Ariana Mendez se encuentran en una etapa intermedia de la carrera Licenciatura en Sistemas y se espera que el proyecto sirva de incentivo para la definición de su tema de tesina de grado.

Por otra parte, a través del proyecto se espera consolidar un equipo de trabajo de la UNTDF en la temática, sumando nuevos integrantes al grupo principal (el cual ya ha ejecutado numerosos proyectos vinculados al problema).

Este proyecto permite además consolidar la cooperación con otros organismos.

5 Referencias

1. System Requirements Engineering, Loucopoulos, Pericles Karakostas, Vassilios McGraw-Hill International series in Software Engineering ISBN 0-07-707843-8. 1995.
2. Software Engineering, Sommerville, I., Pearson Education, Harlow. 2016.
3. El proceso unificado de desarrollo de software, Jacobson I, Addison-Wesley, 2013
4. EL Lenguaje de Modelado Unificado: Un manual de referencia, .Booch G., Addison-Wesley, 1994.
5. Evaluación de la viabilidad para proyectos de implementación ERP en la administración pública Argentina, Eduardo Iberti, Pablo Pytel y Florencia Pollo-Cattaneo - WICC 2015.
6. La Elicitación de Requisitos en el Contexto de un Proyecto Software Markus Manies, Uolevis Nikual Lappeenranta University of Technology, Finland Ing. USBMed, Vol. 2, No. 2, Jul-Dic 2011.
7. Language Models are Few-Shot Learners. Tom B. Brown. En: arXiv preprint arXiv:2005.14165. 2020)
8. Language Models are Unsupervised. Multitask Learners. Alec Radford, Jeffrey Wu, Dario Amodei. Inf. tec. OpenAI, 2019.
9. Virtual Stakeholder: Una herramienta basada en en IA generativa para apoyar la obtención de requisitos. Tiago Brito, Marcia Lucena, Isabel Nunes Universidad Federal de Río Grande del Norte. ISBN: 978-65-01-52831-1. WER-2025.
10. Generación aumentada por recuperación para PNL con uso intensivo de conocimiento Tareas. Lewis, P. Preimpresión de arXiv: 2005.11401. 2020.

ASPECTOS DE INGENIERÍA DE SOFTWARE Y BASES DE DATOS NOSQL PARA EL DESARROLLO DE SISTEMAS DE SOFTWARE Y APLICACIONES PARA DISPOSITIVOS MÓVILES.

Luciano Marrero, *Pablo Thomas*, *Verena Olsowy*, *Leonardo Corbalán*, *Juan Fernández Sosa*, *Fernando Tesone*, *Rodolfo Bertone*, *Eduardo Ibáñez*, *Patricia Pesado*

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC
{lmarrero, pthomas, volsowy, corbalan, jfernandez, ftesone, pbertone, eibanez,ppesado}@lidi.info.unlp.edu.ar

RESUMEN

Se presenta una línea de investigación que tiene por objeto estudiar las problemáticas que afronta la Ingeniería de Software y las Bases de Datos (Relacionales y No Relacionales) ante una era digital que crece de forma constante. Los avances tecnológicos no se detienen y cada vez son más las personas que pueden acceder a un mundo gobernado por internet y sus diferentes aplicaciones. Además, se estudian diversas tendencias en el desarrollo de aplicaciones móviles resilientes, incluyendo la utilización de tecnologías como blockchain y su posible aplicación en el contexto de las ciudades inteligentes.[1, 3, 14].

La idea de considerar que un único modelo de datos pueda adaptarse de forma eficiente a todos los requerimientos es impensada. Se plantean diversas alternativas que permiten almacenar la información de forma no estructurada (Bases de Datos NoSQL) [2, 4]

También, es necesario nuevas formas de pensar el diseño de Bases de Datos que se adapte a esta nueva dinámica del desarrollo del Software [6, 30].

Por último, el avance de la tecnología ha impulsado además aspectos que hace algunos años no eran considerados en el desarrollo del Software, tales como, movilidad, geolocalización y la diversidad de los dispositivos electrónicos involucrados [7, 11].

Palabras claves: Diseño no estructurado de datos. Procesos de diseño de Bases de Datos Relacionales. Bases de Datos NoSQL. Bases de Datos como servicio en la nube. Plataformas para Dispositivos Móviles. Aplicaciones Nativas – Aplicaciones Multiplataforma. Aplicaciones Web Progresivas. Resiliencia. Ciudades Inteligentes.

CONTEXTO

Esta línea de Investigación forma parte del Proyecto "*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital*" del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación. Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa y Latinoamérica. Se utilizan los recursos de Hardware y Software disponibles en el III-LIDI para diseñar, desarrollar y probar diferentes soluciones a problemas relacionados con escenarios a investigar. Las publicaciones científicas generadas y la transferencia continua de resultados concretos validan esta línea de investigación [28].

1. INTRODUCCIÓN

"*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital*" es un proyecto del Programa de Incentivos que propone profundizar las investigaciones que se vienen realizando en el III-LIDI y extender la mirada a nuevos desafíos y cambios que están en gestación. Se organiza en tres subproyectos que permiten atender de manera ordenada el objetivo general propuesto. "*SP1: Ingeniería de Software para escenarios híbridos*", "*SP2: Ciudades Inteligentes sostenibles (CIS). Gobernanza Digital. Buenas prácticas y calidad*" y "*SP3: Creación de tecnologías digitales para el escenario educativo".* [14, 28].

Esta línea de investigación se centra en el subproyecto "*SP1 - Ingeniería de Software para escenarios híbridos*".

Se orienta a la investigación de metodologías y técnicas de la Ingeniería de Software, Bases de Datos (Relacionales y No Relacionales), estudio de procesos de diseño de Base de Datos No Relacionales, escenarios híbridos, aplicaciones móviles, entre otras temáticas de interés [9, 10].

Actualmente, en la ciencia informática se presentan un conjunto de nuevos desafíos que se deben afrontar de forma vertiginosa. Estos desafíos van desde cualquier etapa o nivel inicial hasta que se llega a un resultado o producto final.

En lo que respecta a las Bases de Datos, principalmente, las Bases de Datos NoSQL, en ciertas ocasiones, se ha pasado de tener que pensar en un diseño o esquema detallado en la representación de sus datos a tener que plantear directamente las estructuras a nivel físico teniendo en cuenta el problema a resolver, y la flexibilidad que poseen los esquemas no estructurados de datos [3, 8].

Actualmente, no es común encontrar procesos de diseño establecidos para documentar la semántica en una Base de Datos NoSQL. No obstante, desde hace muchos años, es frecuente en la disciplina informática, agilizar los tiempos de las tareas a realizar a través de suprimir documentación que avale la toma de decisiones.

Las Bases de Datos NoSQL no cuentan un proceso masivamente aceptado para el diseño de esquemas de datos, sino que han surgido como soluciones directamente a nivel físico. No obstante, a medida que han adquirido popularidad, se han propuesto diferentes

procesos y/o metodologías de diseño NoSQL, necesarios para comprender la semántica de los datos.

Existe un amplio abanico de tecnologías, tanto en Bases de Datos (Relacionales y No Relacionales), y de lenguajes para el desarrollo de todo tipo de aplicaciones de Software [6, 29, 30].

Los avances en tecnologías de comunicación y computación han propiciado el surgimiento de la Computación móvil, la cual enfrenta importantes desafíos en el ámbito del desarrollo de software. La computación móvil ha impulsado el desarrollo de nuevas metodologías y prácticas que han contribuido al fortalecimiento de la Ingeniería de Software como disciplina, acompañando la evolución tecnológica en este ámbito [7, 11].

En un mercado cada vez más competitivo, las compañías deben reducir los tiempos de desarrollo y asegurar la compatibilidad con la mayor cantidad de dispositivos posible. No obstante, la fragmentación del ecosistema, derivada de la diversidad de hardware, sistemas operativos y plataformas de desarrollo, representa un desafío significativo.

Una forma de afrontar este desafío es optar por el desarrollo nativo para cada plataforma de destino, utilizando las herramientas, tecnologías y lenguajes específicos de cada una [14, 19, 21].

Como alternativa al desarrollo nativo, los enfoques multiplataforma permiten reutilizar el código fuente para crear versiones compatibles con distintas plataformas. Existen diversas estrategias para el desarrollo multiplataforma, entre las que se incluyen: las aplicaciones web móviles que conducen a las PWA (Progressive Web Apps); el desarrollo híbrido; el desarrollo interpretado; y, finalmente, el enfoque de compilación cruzada.

Otra temática de investigación se obtiene de analizar el proceso de despliegue del Software.

En la mayoría de las empresas el proceso de despliegue no se realiza de manera sistemática y controlada; esto impacta en la finalización del proyecto y la no aceptación del producto final, ocasionando inconvenientes que generan rehacer el trabajo y baja de productividad en su proceso. En este contexto, la comunicación y la sincronización entre los equipos de trabajo continúan siendo un pilar fundamental para el éxito de un proyecto. La utilización de repositorios de información y sincronización permiten realizar un control de versiones distribuido, trabajando en modo offline o en modo online, con la facilidad de disponer herramientas específicas para la resolución de conflictos entre versiones [11, 13, 19].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio y análisis de procesos de diseño para Bases de Datos no relacionales. Aplicación y adaptación. [6, 22, 30].
- Desarrollo de casos de estudio y métricas de performance para distintos tipos de motores de Bases de Datos No Relacionales (NoSQL) [16, 29].
- Desarrollar casos de estudio e incorporar nuevas pruebas para escalabilidad horizontal en motores de Bases de Datos No Relacionales [29].
- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles [12, 13].

- Análisis y estudio comparativo de requerimientos no funcionales, tales como rendimiento, consumo de energía, tamaño de software, entre otros, en los distintos enfoques de Aplicaciones Móviles [11, 14].
- Ciudades inteligentes: El rol de las aplicaciones móviles y de la tecnología blockchain [21].
- Internet de las Cosas (IoT), Internet de Todo (IoE), dispositivos beacons y aplicaciones de sensado móvil o sensado urbano [19].

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados esperados/obtenidos se pueden resumir en:

- Capacitación continua de los miembros de las líneas de investigación [28].
- Estudio y análisis de procesos de diseño para almacenamiento no estructurado de datos y para Bases de Datos no relacionales [6, 8].
- Análisis comparativo de aplicación de procesos de diseños de Bases de Datos NoSQL. Estudio experimental [30]
- Estudio y análisis de distintos motores de Bases de Datos para el desarrollo de aplicaciones móviles [4, 16].
- Estudio y análisis sobre escalamiento horizontal para pruebas de performance entre distintos motores de Bases de Datos no relacionales [4, 29].
- Estudio y análisis sobre el impacto del enfoque de desarrollo en el rendimiento de las aplicaciones, considerando plataformas como iOS y Android, junto con diversos frameworks multiplataforma [14, 19].
- Análisis sobre la ejecución de consultas en diversos casos de estudio para distintos tipos de motores de Bases de Datos no relacionales [5].
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles [13].
- Estudio y análisis sobre las implicaciones de la elección del enfoque de desarrollo en la eficiencia energética de las aplicaciones, abarcando aquellas con alta carga de procesamiento, reproducción de video y acceso a imágenes [13, 14].
- Estudio y análisis en la elaboración de recomendaciones para mejorar la resiliencia de las aplicaciones móviles ante un aumento repentino de la carga computacional del dispositivo y el consumo de energía [11].

Algunas de las transferencias realizadas por el III-LIDI relacionadas con este proyecto, entre otras, son:

- Mantenimiento de la aplicación móvil para la comunidad de la Facultad de Informática de la UNLP.
- Mantenimiento de un Software para la administración y/o gestión de congresos, el cual es utilizado por la RedUNCI y otras entidades.
- Mantenimiento de Sistemas de Gestión Administrativa de Instituciones Universitarias.

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

5. BIBLIOGRAFÍA

1. Introducción a la Computación Cuántica y Aspectos de Ingeniería de Software y Bases de Datos NoSQL para el desarrollo de Sistemas de Software Híbridos. XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025). Mendoza, 10 y 11 de abril de 2025. ISBN: 978-987-575-267-2. Páginas: 351-359. URL: https://sedici.unlp.edu.ar/handle/10915/182261
2. Una revisión de bases de datos que poseen multimodelos. XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025). Sede Atlántica de la Universidad Nacional de Río Negro - Laboratorio de Informática Aplicada - Viedma, 6 al 10 de octubre de 2025. URL: https://cacic2025.unrn.edu.ar/
3. Aspectos de Ingeniería de Software, Bases de Datos Relacionales, y Base de Datos NoSQL para el desarrollo de sistemas de Software Híbridos. XXVI Workshop de Investigadores en Ciencias de la Computación (WICC 2024) (Puerto Madryn, 18 y 19 de abril de 2024). ISBN: 978-987-8352-57-2. Páginas: 474-480. URL: https://sedici.unlp.edu.ar/handle/10915/1736
4. Un estudio comparativo de Bases de Datos NoSQL Orientadas a Grafos. Martinez Manuel, Mozzón Federico, Marrero Luciano, Olsowy Verena, Thomas Pablo. XXX Congreso Argentino de Ciencias de la Computación (CACIC 2024). La Plata, 7 al 11 de octubre de 2024. ISBN: 978-950-34-2428-5. Páginas: 591-602. URL: http://sedici.unlp.edu.ar/handle/10915/176466
5. Aspectos de ingeniería de software, bases de datos relacionales, y bases de datos NoSQL para el desarrollo de sistemas de software híbridos. XXV Workshop de Investigadores en Ciencias de la Computación (WICC 2023). Junín, 13 y 14 de abril de 2023. ISBN: 978-987-3724-66-4; 978-987-3724-67-1. URL: https://sedici.unlp.edu.ar/handle/10915/162004
6. Aplicación de Procesos de Diseño de Bases de Datos NoSQL. Marrero Luciano, Olsowy Verena, Tesone Fernando, Thomas Pablo, Pesado Patricia. XXIX Congreso Argentino de Ciencias de la Computación (CACIC 2023). ISBN: 978-987-9285-51-0. Páginas: 134-145. Universidad Nacional de Luján, Luján, Buenos Aires. URL: https://sedici.unlp.edu.ar/handle/10915/163107
7. Análisis comparativo de alternativas para la detección y comunicación con dispositivos BLE Beacons en Aplicaciones Móviles Multiplataforma. Fernández Sosa, J. F., Medina, S., Thomas, P. J., Corbalán, L. C. & Naiouf, M. XXIX Congreso Argentino de Ciencias de la Computación (CACIC 2023). Luján, 9 al 12 de octubre de 2023. ISBN: 978-987-9285-51-0. Páginas: 318-331. URL: https://sedici.unlp.edu.ar/handle/10915/163107
8. NoSQL database design processes: a comparative study. Marrero Luciano, Olsowy Verena, Thomas Pablo. XI Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC 2023). La Plata, 27 al 29 de junio de 2023. ISBN: 978-950-34-2271-7. Páginas: 58-64. URL: http://sedici.unlp.edu.ar/handle/10915/155446
9. Aspectos de ingeniería de software, bases de datos relacionales, no relacionales y como servicios en la nube para el desarrollo

de sistemas de software híbridos. Marrero Luciano, Thomas Pablo, Pasini Ariel, Bertone Rodolfo, Ibañez Eduardo, Aguirre Verónica, Panizzi Marisa, Olsowy Verena, Tesone Fernando, Pesado Patricia. XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022) 28 y 29 de abril, Mendoza. ISBN: 978-987-48222-3-9. Páginas: 370-376. http://sedici.unlp.edu.ar/handle/10915/144325

10. Un estudio de procesos de diseño de bases de datos NoSQL. Marrero Luciano, Osowy Verena, Tesone Fernando, Thomas Pablo Javier, Corbalán Leonardo, Fernández Sosa Juan, Pesado Patricia. XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022) (La Rioja, 3 al 6 de octubre). ISBN: 978-987-1364-31-2. Páginas: 404-414. http://sedici.unlp.edu.ar/handle/10915/149452
11. Recommendations for Resilient App Development. Corbalán, L. C., Fernández Sosa, J. F., Tesone, F., Dapoto, S. H., Cristina, F., Thomas, P. J., Pesado, P. M. X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022). ISBN: 978-950-34-2126-0. Páginas: 81-86. URL: http://sedici.unlp.edu.ar/handle/10915/139373
12. Aspectos de ingeniería de software, bases de datos relacionales, y bases de datos no relacionales y bases de datos como servicios en la nube para el desarrollo de software híbrido. Marrero Luciano, Thomas Pablo, Olsowy Verena, Pasini Ariel, Bertone Rodolfo, Ibañez Eduardo, Aguirre Verónica, Panizzi Marisa, Pesado Patricia. XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021). ISBN: 978-987-24611-3-3; 978-987-24611-4-0. Páginas: 404-409. Chilecito, La Rioja. URL: http://sedici.unlp.edu.ar/handle/10915/120139
13. Un Análisis Experimental de Sistemas de Gestión de Bases de Datos para Dispositivos Móviles. 2021. Tesone Fernando, Thomas Pablo, Marrero Luciano, Olsowy Verena, Pesado Patricia. ISBN: 978-987-633-574-4. Páginas: 343-355. XXVII Congreso Argentino de Ciencias de la Computación (CACIC 2021). Modalidad virtual, 4 al 8 de octubre de 2021. Universidad Nacional de Salta (UNSA), Argentina. URL: http://sedici.unlp.edu.ar/handle/10915/130353
14. Sensado móvil como estrategia de participación ciudadana en Ciudades Inteligentes. Fernández Sosa, J. F., Aguirre, V., Corbalán, L. C., Delía, L. N., Thomas, P. J., Pesado, P. M. XXVII Congreso Argentino de Ciencias de la Computación (CACIC 2021). Modalidad virtual, 4 al 8 de octubre de 2021. ISBN: 978-987-633-574-4. Páginas: 723-734. URL: http://sedici.unlp.edu.ar/handle/10915/129809
15. Aspectos de ingeniería de software, bases de datos relacionales y bases de datos no relacionales para el desarrollo de sistemas de software en escenarios híbridos. Luciano Marrero, Pablo Thomas, Ariel Pasini, Rodolfo Bertone, Eduardo Ibáñez, Verónica Aguirre, Verena Olsowy, Fernando Tesone, Patricia Pesado. ISBN: 978-987-3714-82-5. Páginas: 494-498. XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020). Universidad Nacional de la Patagonia Austral (UNPA). El Calafate, Santa Cruz (Mayo 2020). URL: http://sedici.unlp.edu.ar/handle/10915/104026
16. Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales. Marrero Luciano, Olsowy Verena, Tesone Fernando, Thomas Pablo, Delia Lisandro, Pesado Patricia. XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020). ISBN: 978-987-4417-90-9. Páginas: 299-309. Universidad Nacional de La Matanza, del 5 al 9 de octubre del 2020. URL: http://sedici.unlp.edu.ar/handle/10915/114202
17. Performance Analysis in NoSQL Databases, Relational Databases and NoSQL Databases as a Service in the Cloud. Marrero Luciano, Olsowy Verena,

Tesone Fernando, Thomas Pablo, Delia Lisandro, Pesado Patricia. https://link.springer.com/chapter/10.1007/978-3-030-75836-3_11
18. Aspectos de ingeniería de software y bases de datos para el desarrollo de sistemas de software en escenarios híbridos. Patricia Pesado, Rodolfo Bertone, Pablo Thomas, Ariel Pasini, Luciano Marrero, Eduardo Ibáñez, Alejandra Rípodas, Verónica Aguirre, Verena Olsowy, Fernando Tesone. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019). ISBN: 978-987-3984-85-3. Universidad Nacional de San Juan (UNSJ). Abril 2019. URL: http://sedici.unlp.edu.ar/handle/10915/77088
19. Aplicaciones para Dispositivos Móviles. Estrategias y enfoques de desarrollo. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019). Universidad Nacional de San Juan. ISBN: 978-987-3984-85-3. URL: http://sedici.unlp.edu.ar/handle/10915/76941
20. Un estudio comparativo de bases de datos relacionales y bases de datos NoSQL. Pesado Patricia, Thomas Pablo, Delía Lisandro, Marrero Luciano, Olsowy Verena, Tesone Fernando, Fernández Juan Sosa. XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019). Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019. ISBN 978-987-688-377-1. URL: http://sedici.unlp.edu.ar/handle/10915/91403
21. A Study of Non-Functional Requirements in Apps for Mobile Devices. Corbalán, L., Thomas, P., Delía, L., Cáseres, G., Sosa, J. F., Tesone, F., & Pesado, P. (2019, June). In Conference on Cloud Computing and Big Data (pp. 125-136). Springer, Cham
22. *NoSQL*: modelos de datos y sistemas de gestión de bases de datos. Migani Silvina, Vera Cristian, Lund María Inés. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018). 26 y 27 de abril, Universidad Nacional del Nordeste (UNNE), Corrientes. ISBN: 978-987-3619-27-4. Páginas: 225-228. URL: http://sedici.unlp.edu.ar/handle/10915/67258
23. Ingeniería de Software. Un Enfoque Práctico. Novena Edición. Roger S. Pressman, Bruce R. Maxim. McGraw-Hill Interamericana Editores, S.A. 2021. ISBN: 978-1-4562-8772-6.
24. Ingeniería de Software Teoría y Práctica. Pfleegger Shari Lawrence. Pearson / Prentice Hall. 2002. ISBN: 9789879460719.
25. Ingeniería de Software. Novena Edición. Ian Sommerville. Addison Wesley / Pearson. 2011. ISBN: 978-607-32-0603-7.
26. Diseño Conceptual de Bases de Datos, un enfoque de entidades-interrelaciones. Carlo Batini, Stefano Ceri, Shamkant B. Navathe. Addison-Wesley/Díaz de Santos. ISBN 0-201-60120-6 (1994).
27. Model Driven Extraction of NoSQL Databases Schema: Case of MongoDB. Brahim, A.; Ferhat, R. and Zurfluh, G. (2019). In Proceedings of the 11th International Joint Conference on Knowledge Discovery. V.1: KDIR, ISBN 978-989-758-382-7, pages 145-154.
28. Instituto de Investigación en Informática LIDI (III-LIDI). Facultad de Informática. Universidad Nacional de La Plata (UNLP). URL: https://weblidi.info.unlp.edu.ar/proyectos/investigacion/
29. Un estudio comparativo de Bases de Datos NoSQL. Trabajo de Especialista en Ingeniería de Software. Marrero Luciano. Universidad Nacional de la Plata (UNLP). 29 de mayo de 2023. URL: https://sedici.unlp.edu.ar/handle/10915/155904
30. Un estudio de procesos de diseño de bases de datos NoSQL. Trabajo de grado. Olsowy Verena. Facultad de Informática. Universidad Nacional de La Plata (UNLP). 24 de octubre de 2023. URL: https://sedici.unlp.edu.ar/handle/10915/159625

AUTOMATIZACIÓN DE ESTIMACIONES

Nicolás Hernández, Julio Castillo, Marina Cardenas,Emanuel Ravera, Sofia Britos
Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/Facultad Regional Córdoba/Universidad Tecnológica Nacional
{damiannicolas05, dr.jotacastillo, ing.marinacardenas, ravera.emanuel, soffibritos01}@gmail.com

RESUMEN

El proyecto de investigación que se expone en este trabajo trata la problemática de estimar el tamaño de los sistemas a ser desarrollados y cómo mediante el uso de métodos de estimación se puede mitigar los riesgos en el desarrollo de software en las etapas iniciales de un proyecto.

El proyecto tiene como finalidad desarrollar una herramienta que contribuya a simplificar la gestión de proyectos de software.

Existen diversos métodos de estimación, cada uno con características, fortalezas y limitaciones propias. Entre ellos, se seleccionaron y aplicaron los enfoques basados en puntos de función, debido a las ventajas que ofrecen y a su mayor viabilidad para ser automatizados.

El proceso de desarrollo se dividió en dos etapas. La primera se centró en la estimación de proyectos que emplean metodologías tradicionales, mientras que posteriormente se abordó la estimación de iniciativas desarrolladas bajo enfoques ágiles.

En ambas etapas se llevaron a cabo actividades similares: análisis e investigación de métodos de estimación, estudio de fuentes de datos pertinentes, diseño e implementación de un sistema automatizado de estimación y, finalmente, evaluación de la exactitud del sistema desarrollado.

CONTEXTO

El presente trabajo se llevó a cabo dentro del grupo de investigación denominado Grupo de Aprendizaje Automático, Lenguajes y Autómatas (GA2LA) en donde se desarrollan proyectos relacionados con autómatas y lenguajes formales, procesamiento del lenguaje natural, y aprendizaje automático, y en especial proyectos orientados a la aplicación de la inteligencia artificial. Este grupo realiza sus actividades en el Laboratorio de Ingeniería de Software (LIS) del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN - FRC).

El mismo se desarrolló dentro de los proyectos homologados por la SCyT de la UTN denominados “Modelo para el procesamiento de textos estructurados Fase 2” (cód. SIECACO0008518) y “Desarrollo de aplicaciones de análisis de texto para Ingeniería de Software” (cod. COSITCI569).

1. INTRODUCCIÓN

Todo proyecto de desarrollo de software para cumplir con sus metas debe equilibrar tres variables fundamentales las

cuales se conocen como la triple restricción denominada así por Martin Barnes [1][2].

Estas tres variables las cuales son la funcionalidad que el sistema debe ofrecer (Alcance), el equipo, hardware, herramientas necesarias y costos (Recursos) y por último el periodo necesario para el desarrollo del proyecto (Tiempo) son importantes para que el proyecto finalice de manera exitosa.

Solo el 31% de los proyectos son exitosos, el 19% falla y el 50% falla en alcanzar alguna de las variables de la triple restricción [3][4]. En la actualidad existe una gran incertidumbre al momento de estimar el tamaño del software y la cantidad de recursos. Con el uso de prácticas como el vibe-Coding [5], automatización de procesos con herramientas n8n [6] y el cambio de arquitecturas de sistemas tradicionales a sistemas agenticos [7] es cada vez más difícil estimar el tamaño del software, por esta razón, métodos de estimación como aquellos basados en puntos de función son esenciales ya que miden el tamaño basándose en la funcionalidad requerida por los usuarios del sistema software.

2. LÍNEAS DE INVESTIGACIÓN

En este trabajo están presentes tres líneas de investigación.

La primera línea de investigación se enmarca en el área de administración de proyectos, más precisamente en la estimación en etapas iniciales, y control de calidad de sistemas de información para lograr cumplir con los requerimientos. En esta línea de investigación se puso especial interés en la investigación de métodos de estimación [8][9].

Estos métodos se dividen en dos grupos básicos, el primero basado en el conocimiento de expertos o la experiencia y el segundo grupo [4] son los métodos que se basan en modelos matemáticos o métodos algorítmicos. Del segundo grupo se seleccionaron los métodos IFPUG [10] y NESMA [11] para su automatización.

La segunda línea de investigación se enmarca en el área de análisis de textos estructurados. En este proyecto se analizaron diferentes archivos con formato estructurado como archivos XML, XMI, JSON, HTML, y formatos propietarios de diferentes herramientas de modelado de desarrollo software o herramientas CASE.

La tercera y última línea de investigación se enmarca en el área de la inteligencia artificial y el uso de grandes modelos de lenguaje (LLM) [12]. En esta línea de investigación se analizaron y desarrollaron métodos para la obtención de información de archivos de texto que contenían historias de usuarios mediante el uso de grandes modelos de lenguaje como la familia de modelos mistral [13].

Por último, se planea continuar con la investigación mediante la generación de material de entrenamiento y el desarrollo de un agente y/o modelo de IA que permita estimar proyectos tradicionales, ágiles y basados en las nuevas tecnologías emergentes.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos fueron los siguientes:

3.1. Banco de conocimientos

Para realizar las pruebas de precisión del sistema de estimación se confeccionó un banco de conocimientos en estas áreas:

- Conocimientos sobre métodos de estimación, técnicas, ventajas y desventajas, y metodologías de desarrollo en los cuales se pueden aplicar.
- Conocimientos sobre análisis de texto estructurados como XML, XMI, JSON, HTML y formatos propietarios

como UML los cuales por su estructura son variaciones de los 4 formatos mencionados. Además, se estudiaron las técnicas de procesamiento de estos archivos como el análisis de las etiquetas (XML, XMI), uso del DOM (HTML) y acceso a objetos (JSON).

- Se confeccionó una lista con 35 herramientas de modelado de software o herramientas CASE con las características de cada herramienta. De este conjunto de herramientas se seleccionaron 16 que soportaban total o parcialmente a UML, además de disponer de una versión gratuita, de prueba o ser open source. Por último, las herramientas que cuentan con la opción de exportar los diagramas a uno de los formatos de archivos de texto estructurado siendo esta característica una de las más importantes.

3.2. Banco de pruebas

Para validar la precisión de las estimaciones se confeccionó un banco de pruebas dividido en metodologías tradicionales y metodologías ágiles.

Cada una de las categorías se subdivide en cuatro categorías de proyectos de desarrollo de software las cuales son Gestión de Servicios, Gestión de la Producción, Comercialización y Administración, tal como se muestra en la figura 1.

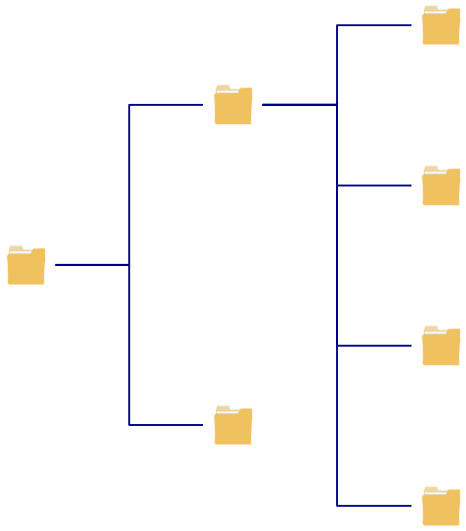


Figura 1. Estructura del banco de pruebas.

Cada proyecto de software cuenta con su correspondiente documentación. En el caso de las metodologías tradicionales, cuentan como mínimo con el plan de desarrollo con los tiempos y el documento donde se detallan los requerimientos, diagramas de casos de uso, modelo de objetos y diagramas de clases. En el caso de las metodologías tradicionales los proyectos cuentan con el backlog, las historias de usuarios y los tiempos de desarrollo.

3.3. Web Service (Metodología tradicional)

El Web Service ofrece la siguiente funcionalidad

- Realizar estimaciones.
- Consultar estimaciones realizadas.
- Modificar estimaciones
- Visualizar estimaciones.
- Eliminar estimaciones.

El sistema se estructura en tres sub-proyectos: el primero actúa como end-point, permitiendo el acceso a los servicios expuestos; el segundo consiste en una librería auxiliar que define las clases y estructuras necesarias para encapsular y transportar la información; y el tercero constituye el core del sistema, responsable del procesamiento y análisis de los archivos que contienen los

diagramas. En la Figura 2 se presenta la interfaz de acceso al servicio web, documentada y visualizada mediante Swagger

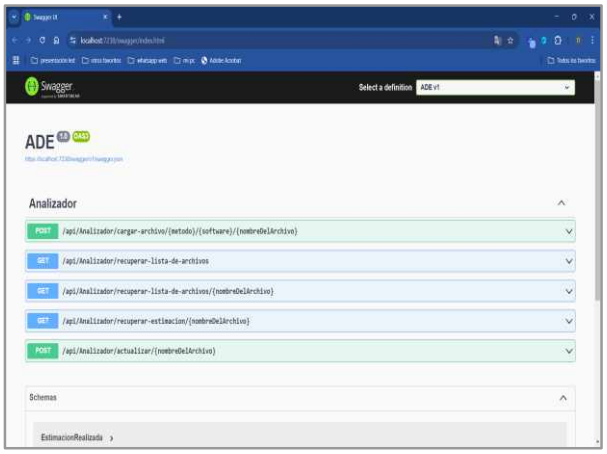


Figura 2. Interfaz que muestra los servicios de Web Service "ADE".

3.4. Sitio Web

El sitio web fue desarrollado con el objetivo de integrar los servicios web correspondientes a la metodología tradicional y a las metodologías ágiles en una única plataforma.

La aplicación presenta una estructura compuesta por una página principal desde la cual se puede acceder a cuatro funcionalidades: (1) realizar nuevas estimaciones; (2) consultar el listado de estimaciones existentes; (3) visualizar en detalle una estimación previamente generada, con posibilidad de modificar los objetos contabilizados y sus parámetros, entre ellos: tipo de objeto, complejidad, tiempo de desarrollo por punto de función y puntos de función de conversión. Así también permite agregar o eliminar objetos (estimación realizada); y (4) acceder a una sección de soporte o información general.

La organización estructural del sitio web se ilustra en la Figura 3.

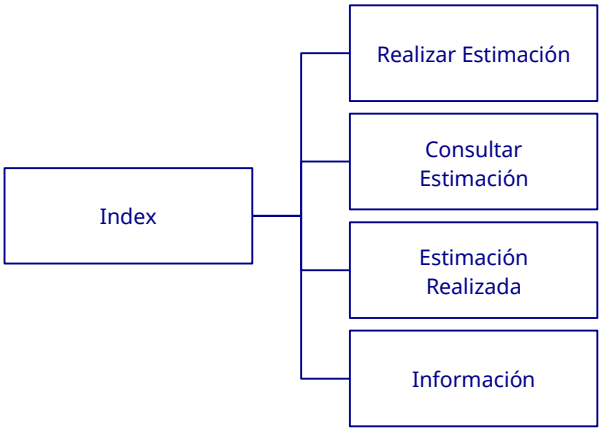


Figura 3. Estructura del Sitio Web del Sistema.

3.5. Web Service (Metodologias agiles)

El servicio web se desarrolló con Spring Boot, con una arquitectura monolítica. La API expone distintos endpoints que permiten procesar archivos PDF para extraer historias de usuario, recibir historias individuales y gestionar múltiples historias de usuario en una misma solicitud.

Una vez obtenidos los datos, estos son enviados a un servicio auxiliar externo implementado en Google Colab, cuyo propósito es aprovechar las capacidades que dicha plataforma ofrece para la implementación y ejecución de grandes modelos de lenguaje (LLM).

Este servicio externo procesa la información mediante un prompt, con el objetivo de analizar los datos, identificar y clasificar los objetos correspondientes, asignar los puntos de función a cada uno y, finalmente, realizar una depuración del resultado generado por el modelo. Esta limpieza incluye la identificación del bloque en formato JSON, la eliminación de texto adicional y la validación de su estructura, a fin de retornar una estimación correctamente formateada.

En la etapa actual del proyecto se prevé la integración de este servicio web con el sitio web mencionado previamente, mediante la incorporación de una nueva opción funcional.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto cuanta con la participación de un doctor en Ciencias de la Computación y una magíster en Ing. En Sistemas de Información, quienes actuaron como directores en el desarrollo de una tesis de maestría realizada por un especialista en Ingeniería en Sistemas de Información. Asimismo, la iniciativa sirvió como marco de trabajo para que estudiantes llevaran adelante las actividades correspondientes a sus becas de servicio e investigación.

5. BIBLIOGRAFÍA

[1] Barnes, M. (1988). Construction project management. International Journal Of Project Management, 6(2), 69-79. https://doi.org/10.1016/0263-7863(88)90028-2.

[2] Vahidi, R., & Greenwood, D. (2009, septiembre). Triangles, tradeoffs and success: A critical examination of some traditional project management paradigms. In Construction Facing Worldwide Challenges: CIB Joint International Symposium (pp. 27-30). Sept.

[3] Barnes, M. (1988). Construction project management. International Journal Of Project Management, 6(2), 69-79. https://doi.org/10.1016/0263-7863(88)90028-2.

[4] Sommerville, I. (2011). *Ingeniería de software* (9.ª ed.). Pearson. ISBN 9780137035151.

[5] Sarkar, A., & Drosos, G. (2025). *Vibe coding: Programming through conversation with artificial intelligence*. *ResearchGate*.

[6] Ahmad, R. D., Hadi, E. A., & Rusnandi, E. (2025). *Automation of technical documentation validation workflows based on GitHub using n8n during software development stages*. *Journal of Multimedia Technology and Applied Software, 2*(2), 64–69. https://doi.org/10.71266/jmtas.v2i2.62

[7] Abou Ali, M., Dornaika, F., & Charafeddine, J. (2025). *Agentic AI: a comprehensive survey of architectures, applications, and future directions*. *Artificial Intelligence Review, 59*.

[8] Hernández, N., Castillo, J., Cardenas, M., Albarracín, G., & Victoria, J. (2025). *Análisis de métodos de estimación de puntos de función y fuentes de datos compatibles para la automatización del proceso de estimar*. WICC 2025, Facultad de Ingeniería, Universidad Nacional de La Plata, Buenos Aires, Argentina.

[9] Cardenas, M., Hernández, N., Castillo, J., Cuevas, J. C., & Páez, N. del V. (2025). *Análisis de textos estructurados mediante el desarrollo de herramientas para generación de material de entrenamiento*. WICC 2025, Facultad de Ingeniería, Universidad Nacional de La Plata, Buenos Aires, Argentina.

[10] Bundschuh, Manfred & Dekkers, Carol. (2008). The IFPUG Function Point Counting Method. https://doi.org/10.1007/978-3-540-68188-5_12. https://www.researchgate.net/publication/314391334_The_IFPUG_Function_Point_Counting_Method.

[11] Morrow, P., Wilkie, F. G., & McChesney, I. R. (2014). Function point analysis using NESMA: simplifying the sizing without simplifying the size. Software Quality Journal, 22, 611-660.

[12] Hernández, N., Castillo, J., Cardenas, M., Ravera, G., Koroch, A., Liendo, S., Santino, F., Negrelli, M., & Britos, P. (2025). *Propuesta de arquitectura para el desarrollo de un sistema de agentes de IA basada en protocolos*. CONAIISI 2025. Córdoba, Argentina.

[13] Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Singh Chaplot, D., de las Casas, D., Bressand, F., Lengyel, G., Lample, G., Saulnier, L., Renard Lavaud, L., Lachaux, M.-A., Stock, P., Le Scao, T., Lavril, T.,

Wang, T., Lacroix, T., & El Sayed, W. (2023). *Mistral 7B*.
[14] Hasan, A. (2024, 18 noviembre). Using Hugging Face Models in Google Colab: A Beginner's Guide. DEV Community. https://dev.to/ajmal_hasan/using-hugging-face-models-in-google-colab-a-beginners-35ll

Automatización de HuerTIC con framework de desarrollo basado en WS-CDL

Oscar Testa[1]; Rubén Pizarro[1]; Pablo García[1]; Leandro Castro[1]; Sonia Ponteprimo[1]

[1] Facultad de Ciencias Exactas y Naturales
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-245220– Int. 7124
[otesta, ruben,pablogarcia,castro.leandro]@exactas.unlpam.edu.ar

Resumen

Este proyecto de investigación, surge como continuidad del proyecto "Ingeniería de Software: Composición de Servicios en Ambientes Ubicuos", desarrollado también en el ámbito de la Facultad de Ciencias Exactas y Naturales de la UNLPam.

Tomando como resultado el framework de coordinación de dispositivos ubicuos, basado en la utilización de los estándares de SOA para servicios web, se plantea como líneas futuras de investigación implementar el citado framework en ambientes reales, donde se pueda poner a prueba dicho entorno.

Se están realizando los trabajos iniciales sobre el prototipo del framework que se instalará en la Huerta Tecnológica Escolar "HuerTIC", como una estrategia pedagógica que contribuye en muchos aspectos a mejorar la calidad de vida de los alumnos participantes, como así también la de la comunidad en la que se encuentra inserta la Escuela Secundaria que realizará la experiencia. Además de poder implementar en un ambiente real el framework para coordinar los distintos sensores y actuadores para automatizar las tareas de riego y control de temperatura y humedad del cultivo.

El ambiente de puesta en marcha del prototipo real será en la escuela secundaria, donde se encuentra físicamente la huerta.

Palabras clave: sistemas ubicuos, SOA, servicios, composición de servicios, coreografías.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: "Coordinación de dispositivos de IoT para resolver problemas específicos de la distribución de energía eléctrica" – Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa. Las líneas aquí presentadas actualmente forman parte de las bases de un anteproyecto de tesis doctoral en ingeniería de software por la Universidad Nacional de San Luis, además de un proyecto de Tesis de Maestría en Enseñanza en Escenarios Digitales en AUSA.

Introducción

Los avances de las comunicaciones entre dispositivos han permitido que estos sean generadores y consumidores de servicios al mismo tiempo, es decir, de acuerdo a las capacidades del dispositivo puede no solo obtener, sino también ofrecer a otros equipos sus funciones y así cooperar entre ellos. La tendencia actual es hacia los ambientes ubicuos, los cuales se caracterizan por estar poblados de numerosos dispositivos que, gracias a la integración extrema de los elementos electrónicos, son invisibles al usuario y están en permanente rastreo de la actividad humana [2].

Dispositivos ubicuos son todos aquellos dispositivos que pueden existir en todas partes, es decir, son dispositivos electrónicos que tienen capacidad de procesamiento y comunicación y pueden ser encontrados en lugares diversos de la vida cotidiana.

La computación ubicua es un desarrollo tecnológico que intenta que las computadoras no se perciban en el entorno como objetos

diferenciados, y que la utilización por parte de los seres humanos sea lo más transparente y cómoda posible, facilitando de esta manera la integración en la vida cotidiana. Desde hace varios años los dispositivos ubicuos han ganado importancia y presencia en la vida cotidiana de las personas, debido principalmente a que: poseen distintos tipos de sensores (posicionamiento, proximidad, luminosidad, temperatura, etc.), facilitan la conectividad incluso en áreas con poca señal o acceso a las redes, permiten la convergencia tecnológica (computo, medios, telefonía, etc) y brindan acceso a servicios de distinta índole (mapas, ayudas, etc).

Por composición entendemos la forma en que se pueden combinar o enlazar un número indeterminado de dispositivos para llevar adelante una tarea determinada. En ambientes ubicuos, la composición de dispositivos, presenta nuevos desafíos tales como: la heterogeneidad (ya sea por la diversidad de dispositivos involucrados, como por la presencia de dispositivos de varios fabricantes), las contingencias de los dispositivos y la personalización de los mismos (por ej. provisión de servicios de acuerdo a las preferencias del usuario). Dado que los dispositivos en donde los servicios son ejecutados poseen limitaciones de recursos (ej. poca memoria y batería), se deben hacer consideraciones especiales respecto a la eficiencia y rendimiento de la composición de servicios [3].

La composición en este tipo de ambientes implica que los dispositivos deben dialogar entre ellos para poder compartir los servicios que ofrecen con la finalidad de obtener un servicio con valor agregado, o bien para abordar la solución de una problemática particular, como podría ser la seguridad de un hogar, o la seguridad vial, por mencionar algunos ejemplos.

Si bien hoy en día podemos decir que distintos sensores o dispositivos se pueden comunicar entre ellos, compartiendo de alguna manera sus servicios, generalmente lo realizan a partir de protocolos propietarios y sin seguir definiciones estándares, provocando que otros componentes de otros proveedores (o incluso de los mismos) no puedan ser utilizados. Esto obviamente representa una importante limitación en la composición de dispositivos ubicuos. Adicionalmente la composición de dispositivos ubicuos presenta un nuevo desafío. Los mecanismos de composición en ambientes masivos, necesitan hacer frente a las distintas contingencias que pueden ocurrir con estos dispositivos. Los dispositivos ubicuos tienen distintas limitantes como son la cantidad de memoria disponible, la durabilidad de la batería, la disponibilidad de acuerdo a la red del lugar donde se encuentre en un momento determinado. Todas estas variantes hacen que la composición de dispositivos[1] ubicuos se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

Líneas de Investigación y Desarrollo

La orientación a servicios proporciona estrategias bien conocidas (por ejemplo, las orquestaciones) para la composición de servicios. Tales estrategias resultan herramientas de gran utilidad para la construcción de aplicaciones complejas mediante la combinación de servicios heterogéneos, los cuales presentan la característica muy favorable de su simplicidad. Además, su carácter distribuido los convierte en una solución especialmente aplicable a ambientes ubicuos. Sin embargo, en contextos donde los servicios son dinámicos, móviles, menos fiables y dependientes del dispositivo, los mecanismos de composición establecidos para servicios web no pueden aplicarse directamente [8].

[1] Si bien los autores se refieren a la composición de servicios, se hace dentro de un contexto de dispositivos ubicuos, lo cual a los fines de este trabajo se puede interpretar como composición de dispositivos, haciendo que la terminología para este caso particular sea más adecuada.

Por otro lado, la composición de múltiples dispositivos ubicuos presenta características propias que la separan de la composición tradicional de servicios web. Es necesario lidiar con las características heterogéneas que cada uno de ellos podría presentar, como por ejemplo la cantidad de memoria y la duración de la batería.

Los dispositivos ubicuos se conectan en contextos diferentes en cada ocasión y, en consecuencia, su disponibilidad es muy variable y la confiabilidad se vuelve extremadamente dinámica. Todas estas dificultades hacen que la composición de dispositivos se transforme en un área de investigación que aún presenta muchos aspectos a dilucidar y que resulta de interés en términos de investigación [3].

En la actualidad, existen múltiples proyectos donde se integran sensores y dispositivos ubicuos a situaciones problemáticas del mundo real. Por ejemplo, la domótica busca que dispositivos y sensores deban actuar en coordinación para prevenir incidentes de seguridad diversos en contextos domésticos.

El concepto de Industria 4.0 [9], busca integrar dentro de una planta fabril la intercomunicación de todos los dispositivos que componen la cadena de producción, buscando, fundamentalmente, una coordinación de tareas que atienda las variables que afectan al proceso productivo (tiempos, stocks, demandas).

Es inmediato detectar múltiples tipos de procesos en los cuales los dispositivos ubicuos adquieren trascendencia: aeronáutica, automotriz, alimentos, etc.

En consecuencia, esta investigación en particular aborda la implementación de un framework de coreografía de servicios con dispositivos ubicuos, la cual está definida por Testa, et all. en[11]. A partir de la utilización de este framework es que se propone implementarlo sobre sistemas reales, y específicamente para automatizar las tareas de riego, temperatura y humedad ambiente y del suelo en una Huerta Tecnológica Educativa, en un ambiente educativo de la ciudad de Dorila, Provincia de La Pampa.

Resultados y Objetivos

La proliferación de dispositivos de computación ubicuos e interconectados (PDAs, tabletas, móviles, etc), así como los recientes avances en la tecnología de radio frecuencia y las redes de sensores están fomentando la creación de ambientes donde las aplicaciones de internet y los servicios se están haciendo más populares y necesarias para los usuarios de móviles. La composición de servicios a través de múltiples dispositivos móviles presenta un nuevo desafío el cual no es compatible con la composición de servicios como se plantea actualmente. En particular, los mecanismos de composición en ambientes masivos, como lo es el de dispositivos móviles, necesita hacer frente a las distintas contingencias que pueden ocurrir con estos elementos (desconexión, falta de batería, etc), así como también contemplar la heterogeneidad de los mismos. Estos dispositivos tienen distintas limitantes como son la cantidad de memoria disponible, la durabilidad de la batería, la disponibilidad de acuerdo a la red del lugar donde se encuentre en un momento determinado. Todas estas variantes hacen que la composición de servicios incluyendo dispositivos móviles se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy [3].

En este tipo de ambientes, mecanismos automáticos y dinámicos son necesarios para la composición de servicios, ya que de esta forma se puede compensar la falta de disponibilidad de un servicio en un momento determinado [8].

Si bien a este problema existente se encontró solución, a través de la implementación de un framework de ejecución de coreografías en ambientes ubicuos, tal como se plantea en [11], se necesita hacer pruebas y más investigaciones a partir de una implementación real.

Por lo antes expuesto se deducen las siguientes preguntas de investigación:

- ¿El framework de ejecución de coreografías propuesto, brinda soluciones en una huerta educacional?
- ¿Se necesitan mejoras para ser implementado en un ambiente real como el de la Huerta? ¿Cuáles?
- ¿Qué tipo de hardware se necesita para una implementación real, o alcanza con utilizar placas académicas de uso general como Arduino?
- ¿Los protocolos de comunicación seleccionados, son los correctos, o debe ampliarse la implementación y utilización de protocolos?

Por lo tanto, el proyecto plantea los siguientes objetivos:

- Realizar una implementación real del framework de ejecución de coreografías en ambientes ubicuos, en un caso de estudio particular y real.
- Realizar las mejoras y ajustes necesarios al framework para que pueda ser utilizado comercialmente.
- Poner a prueba la validez de la utilización de los protocolos de comunicación establecidos por SOA.
- Difundir los resultados obtenidos con publicaciones de los resultados parciales y finales.

La hipótesis principal es poner a prueba si los mecanismos principales y fundacionales de SOA son aplicables para la composición de coreografías en ambientes ubicuos, en un ambiente real. Se tienen fuertes indicios, por los estudios e investigaciones llevadas adelante en el proyecto de Investigación sobre el cual se basa el presente trabajo, que los mismos serán suficientes y que serán de utilidad en este tipo de ambientes.

Se espera, además logros en términos de formación en recursos humanos:

El Mg. Pablo M. García, trabaja en la propuesta de innovaciones en el ámbito del manejo de la memoria de las placas arduino, cuyos resultados deben reflejarse en una Tesis Doctoral.

La investigadora Sonia Ponteprimo se encuentra trabajando en la implementación del framework en la Huerta Tecnológica, donde se están haciendo los primeros análisis de cantidad de dispositivos a ubicar, lugares y qué actuadores se utilizarán para las tareas de automatización del riego.

Como parte de los resultados obtenidos hasta el momento podemos mencionar que se ha obtenido una precisa y clara situación actual o estado del arte respecto de los dispositivos ubicuos que formarán parte de la implementación real, así como los distintos avances que se han realizado en la teoría de microservicios y de coordinación a través de la utilización de coreografías de servicios. Además se ha logrado la implementación en el framework, de un nuevo protocolo de comunicación de los dispositivos como el caso de Bluetooth. También se ha logrado armar un prototipo real para la implementación en la Huerta Tecnológica, el cual sirve como modelo para su implementación real.

Formación de Recursos Humanos

Además de las metas propuestas en el punto 3, se espera como resultado en la formación de recursos humanos, la continuación de esta misma línea de proyecto como tesis doctoral del investigador Mg. Pablo García. La investigadora Sonia Ponteprimo ha defendido exitosamente su trabajo final de maestría sobre fines del año 2025, relacionada con las metodologías de enseñanza en espacios digitales sobre la aplicación en la Huerta Tecnológica Educativa, con calificación de sobresaliente. También se espera lograr una mayor interrelación con la Universidad Politécnica de Madrid para la aplicación de las metodologías aquí presentadas en los proyectos de Ingeniería de Software Empírica y realizando una estancia posdoctoral del Dr. Oscar A. Testa. Adicionalmente, se espera que otras tesis de Maestría surjan a partir de los logros obtenidos en la presente línea de investigación.

Bibliografía

[1] U. I. d. T. (UIT), "Unión Internacional de Telecomunicaciones." https://www.itu.int/net/pressoffice/pressreleases/2014/23-es.aspx, 10 2015.

[2] M. Weiser, "Hot topics-ubiquitous computing," Computer, vol. 26, pp. 71–72, Oct 1993.

[3] Q. Z. Sheng, X. Qiao, A. V. Vasilakos, C. Szabo, S. Bourne, and X. Xu, "Web services composition: A decade's overview," Information Sciences, vol. 280, no. 0, pp. 218–238, 2014.

[4] M. Viroli, "On competitive self-composition in pervasive services," Science of Computer Programming, vol. 78, no. 5, pp. 556–568, 2013. Special section: Principles and Practice of Programming in Java 2009/2010 & Special section: Self-Organizing Coordination.

[5] S. W. Loke, "Supporting ubiquitous sensor-cloudlets and context-cloudlets: Programming compositions of context-aware systems for mobile users," Future Generation Computer Systems, vol. 28, no. 4, pp. 619–632, 2012.

[6] F. Palmieri, "Scalable service discovery in ubiquitous and pervasive computing architectures: A percolation-driven approach," Future Generation Computer Systems, vol. 29, no. 3, pp. 693–703, 2013. Special Section: Recent Developments in High Performance Computing and Security.

[7] S. Najar, M. K. Pinheiro, and C. Souveyet, "A New Approach for Service Discovery and Prediction on Pervasive Information System," Procedia Computer Science, vol. 32, pp. 421–428, 2014. The 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014).

[8] G. Cassar, P. Barnaghi, W. Wang, S. De, and K. Moessner, "Composition of services in pervasive environments: A Divide and Conquer approach," in Computers and Communications (ISCC), 2013 IEEE Symposium on, pp. 000226–000232, July 2013.

[9] Wikipedia, "Industria 4.0 — Wikipedia, La enciclopedia libre," 2016. [Internet; descargado noviembre-2016].

[10] H.-I. Yang, R. Bose, A. (Sumi) Helal, J. Xia, and C. Chang, "Fault-Resilient Pervasive Service Composition," in Advanced Intelligent Environments (A. D. Kameas, V. Callagan, H. Hagras, M. Weber, and W. Minker, eds.), pp. 195–223, Springer US, 2009.

[11] Oscar A. Testa, Efraín R. Fonseca C., Germán Montejano and Oscar Dieste, "Coordination of Ubiquitous Devices in Pervasive Environments: A Proposal Based on WS-CDL" in 38th International Conference of the Chilean Computer Science Society. SCCC 2019.

AUTOMATIZACIÓN DE LA GESTIÓN DE ESTACIONAMIENTOS MEDIANTE VISIÓN POR COMPUTADORA E INTEGRACIÓN CON ODOO

Agustín Sánchez Luccini, Ezequiel Alejandro Alderete, Sebastian Musilli, Alejandro Sartorio

Facultad de Tecnología Informática Universidad Abierta Interamericana (UAI)
{agustin.sanchezluccini, ezequielalejandro.alderete, sebastian.musilli}@alumnos.uai.edu.ar,
{alejandro.sartorio}@uai.edu.ar.

RESUMEN

Este trabajo presenta el diseño, la implementación y la validación de Odo-Park, una solución para automatizar el control de acceso y la cobranza en estacionamientos vehiculares mediante la integración de reconocimiento automático de patentes (LPR), un middleware de baja latencia y el ERP Odoo[1]. La propuesta aborda la brecha entre la captura de eventos en tiempo real y su trazabilidad administrativa y contable. La arquitectura se organiza en tres capas: adquisición de eventos mediante cámara LPR, coordinación y orquestación a través de FastAPI con Odoo como núcleo transaccional, y ejecución de reglas de negocio con activación de actuadores y cobro por QR integrado con Mercado Pago. Como contribución, se presenta una hoja de ruta replicable para la integración IoT-ERP y un prototipo funcional y extensible. La validación experimental incluyó 30 casos de prueba sobre una flota de 15 vehículos. Los resultados muestran la automatización completa del ciclo de egreso y una reducción del tiempo promedio de operación desde 1 min 45 s hasta menos de 25 s (≈76%), lo que confirma la viabilidad y efectividad del esquema propuesto.

Palabras clave: Computer visión; Reconocimiento de Matrícula Vehicular; Sistema de Gestión de Estacionamiento; Integraciones; Odoo.

CONTEXTO

El proyecto Odo-Park se inscribe en un conjunto de proyectos de investigación aplicada enmarcados en la línea de investigación Ingeniería de Software del centro de investigación CAETI (UAI), orientados a la automatización de procesos operativos en infraestructuras urbanas mediante arquitecturas interoperables, extensibles y sostenibles.

Este trabajo define un marco de referencia experimental para abordar la integración entre la capa física de sensado y control (IoT, visión computacional/LPR) y la capa administrativa transaccional (ERP Odoo), con énfasis en trazabilidad de eventos, consistencia de datos y automatización de punta a punta. El desarrollo contó con soporte institucional del Laboratorio de Sistemas, que proveyó infraestructura de red, dispositivos de captura LPR y el entorno físico requerido para validar los prototipos.

Si bien existen soluciones comerciales —por ejemplo, Parklio[2] o SmartPark[3]— que automatizan parcialmente el control de acceso, en muchos escenarios operan como sistemas cerrados o débilmente integrados con los procesos administrativos y contables de las organizaciones. En consecuencia, la consolidación de la información operativa, su auditoría y su vinculación nativa con la facturación y la cobranza suelen requerir integraciones ad-hoc o componentes propietarios.

[1] Odoo es un ERP de código abierto con arquitectura modular que centraliza procesos administrativos y contables en una base de datos unificada.

[2] https://parklio.com/es/

[3] https://smartparkcr.com/

Odo-Park se propone cerrar esta brecha mediante una arquitectura basada en servicios y un middleware de orquestación que unifica eventos en tiempo real con transacciones ERP, evitando dependencias de conectores propietarios. De este modo, se busca disponer de una solución replicable, mantenible y escalable para escenarios de gestión urbana automatizada, habilitando su evolución hacia funcionalidades analíticas y adaptativas en etapas posteriores.

INTRODUCCIÓN

La gestión de estacionamientos ha transitado desde esquemas manuales y descentralizados hacia soluciones digitales, en las que la automatización del ciclo completo contribuye a reducir costos operativos, mitigar errores humanos y disminuir tiempos de espera. En este contexto, resulta crítica la integración consistente entre dispositivos de captura y control, sistemas de gestión transaccional y servicios de pago en línea.

En línea con esta necesidad, el presente trabajo propone una solución integral que articula un ERP empresarial (Odoo) con una capa de servicios de comunicación en tiempo real implementada sobre FastAPI [1]. Esta capa opera como middleware entre los sensores físicos y el backend administrativo, permitiendo procesar con baja latencia los eventos operativos críticos del ciclo (ingreso/egreso, emisión y actualización de tickets, validación de pagos), mientras que el ERP concentra la persistencia, la trazabilidad y el registro contable.

El diseño conceptual se apoya en principios de IoT Industrial (IIoT) —conectividad, control, digitalización y aumento— [2], operacionalizados mediante la integración de sensores inteligentes, actuadores y un sistema de gestión empresarial [3]. Desde el punto de vista arquitectónico, la propuesta adopta un enfoque basado en servicios, con componentes desplegados en contenedores, con el objetivo de favorecer mantenibilidad, evolución incremental y escalabilidad.

En términos operativos, el flujo se inicia con el ingreso del vehículo, cuando una cámara LPR identifica la patente y genera automáticamente en Odoo un ticket digital (modelo parking.ticket) en estado “Activo”. En la fase de egreso, el sistema reidentifica la matrícula y el middleware calcula en tiempo real la tarifa y la duración de la estadía, presentando al usuario un código QR de Mercado Pago en un panel con asistencia por voz. El ciclo se completa sin intervención humana tras la confirmación del pago mediante un webhook [4], a partir del cual Odoo recibe la notificación, emite el comprobante correspondiente (modelo parking.receipt) y realiza el cierre definitivo del ticket. De este modo, Odo-Park automatiza el ciclo de punta a punta, garantizando trazabilidad y registro contable dentro del ERP.

Si bien existen antecedentes internacionales —por ejemplo, Parklio [5], ParkVision [6] y ParkPlus [7]— que resuelven aspectos parciales del control de acceso y operación, en numerosos escenarios estas soluciones operan como sistemas aislados, sin integración nativa con plataformas ERP [8]. En este sentido, la contribución de Odo-Park se centra en la unificación del ciclo operativo con el ciclo administrativo y contable mediante su integración directa con Odoo.

En términos teóricos, el trabajo se propone definir una arquitectura de referencia para la integración entre sistemas de sensado físico, middleware de orquestación y plataformas ERP en escenarios de estacionamiento inteligente. En términos experimentales, el objetivo es validar la viabilidad técnica de dicha arquitectura mediante un prototipo funcional que permita automatizar el ciclo de ingreso y egreso, medir su latencia operativa y evaluar su impacto sobre los tiempos de procesamiento y la reducción de la intervención humana.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La arquitectura de Odo-Park se fundamenta en la premisa de desacoplar la velocidad de detección de la integridad transaccional para resolver la dicotomía entre el tiempo real y la consistencia de datos, un

desafío central en los sistemas ciberfísicos de estacionamiento inteligente [9]. El primer eje de desarrollo se centra en la interoperabilidad y gestión escalable, cuyo objetivo es lograr una plataforma integral y modular que permita extender las capacidades del sistema sin alterar su núcleo. Para su implementación se utiliza Odoo 19 como plataforma de desarrollo low-code, aprovechando una arquitectura que permite gestionar el ciclo de vida de la estadía a través de estados dinámicos, asegurando la integridad contable. La evidencia de este eje se halla en la implementación exitosa de la extensión del modelo detallada en la Figura 1, la cual se integra con soluciones de visión artificial para la asignación de espacios y gestión de registros [10].

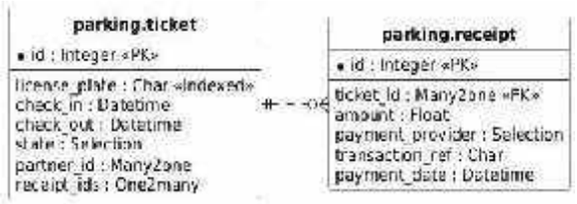


Figura 1: Modelo de Datos (Extensión de Odoo)

La Figura 1 muestra la extensión del modelo de datos en Odoo, incluyendo las entidades principales (parking.ticket y parking.receipt), sus atributos relevantes y las relaciones establecidas entre los distintos estados del ciclo de estacionamiento.

El segundo eje estratégico se orienta al procesamiento y respuesta en tiempo real, buscando garantizar una latencia mínima en la detección y apertura de barreras para manejar múltiples conexiones concurrentes. Esto se implementa mediante el uso de FastAPI y el framework ASGI Starlette para el manejo nativo de concurrencia [11], junto con Redis Streams como bus de mensajes de baja latencia para la comunicación entre sensores y el orquestador [12]. Las pruebas de entorno proporcionan la evidencia necesaria al validar un objetivo técnico de latencia inferior a 500 ms, permitiendo que el middleware procese la señal de la cámara y envíe la orden de apertura en milisegundos.

Como tercer eje, se desarrolla la línea de pagos, conciliación y automatización, que tiene como fin automatizar el flujo operativo integral desde el ingreso hasta el egreso mediante soluciones sostenibles en la nube [13]. La implementación combina el reconocimiento automático de patentes (LPR) con el uso de webhooks para asegurar la conciliación de los estados de pago en el ERP. El flujo cronológico de estas interacciones, que vincula eficazmente cámaras, middleware y el sistema de gestión, se presenta como evidencia en el diagrama de secuencia de la Figura 2.

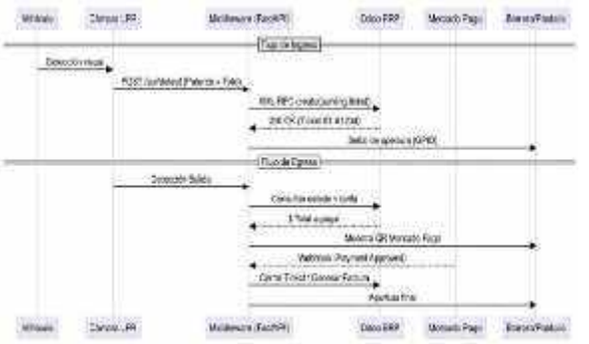


Figura 2: Diagrama de secuencia del ingreso y egreso de vehículos

La Figura 2 presenta el diagrama de secuencia del sistema, donde se detallan los componentes intervinientes (cámara LPR, middleware, ERP Odoo y pasarela de pagos), así como el orden temporal de las interacciones durante los procesos de ingreso y egreso.

El cuarto eje aborda la UX asistida y visualización inteligente para optimizar la experiencia del conductor y la toma de decisiones administrativa mediante interfaces dinámicas. Se implementan websockets para la actualización instantánea de pantallas y se despliega localmente Coqui-AI para servicios de Text-to-Speech (TTS), eliminando la latencia de servicios externos [14].

Tabla 1: Comparativa de Servicios Text-to-Speech (TTS).

Características	Google Cloud TTS	Coqui-AI (Open Source)
Modelo	WaveNet / Neural2 (Cloud)	VITS / Glow-TTS (Local)
Licencia	Comercial (Pago por Uso)	Open Source (MIT/Apache)
Latencia	Variable (Dep. de la Red)	Baja (< 200ms en Local)
Dependencia	Requiere Internet	Offline
Privacidad	Procesado en la Nube	Procesado Localmente

Como evidencia, la Tabla 1 demuestra que la implementación local garantiza una respuesta inferior a 200ms en la asistencia por voz, mientras que la arquitectura físico-lógica que soporta esta interacción se detalla en la Figura 3.

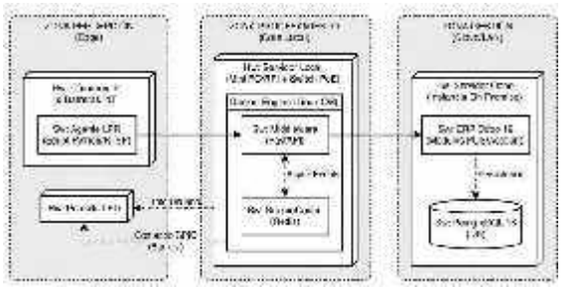


Figura 3: Arquitectura Físico-Lógica de Odo-Park

La Figura 3 describe la arquitectura físico-lógica del sistema, identificando los dispositivos de captura, los componentes de procesamiento intermedio, el núcleo transaccional ERP y los mecanismos de comunicación que articulan la infraestructura.

RESULTADOS OBTENIDOS Y ESPERADOS

Para la validación de la primera fase, se estableció un entorno de pruebas controlado, conformado por un servidor local de Odoo y una webcam configurada para simular la cámara LPR. Las pruebas se realizaron en el garage de la UAI.

Las condiciones experimentales incluyeron un tamaño de muestra n = 30 ciclos completos de ingreso y egreso, utilizando una flota heterogénea de 15 vehículos. Se definió como criterio de acierto o lectura correcta la identificación positiva y exacta del 100% de los caracteres de la patente en un único intento de captura. El ciclo de latencia se midió desde la captura del evento por la cámara LPR (on-device) hasta la confirmación de la transacción en el ERP Odoo, considerándose este intervalo como ciclo completo.

Los resultados técnicos, detallados en la Tabla 2 y en la Tabla 3, confirmaron la viabilidad del sistema, logrando una integración exitosa entre FastAPI y Odoo mediante protocolos RPC y autenticación JWT. Asimismo, se validó el cumplimiento de la latencia objetivo, registrándose una respuesta promedio de aproximadamente 450 ms.

Tabla 2: Comparativa de Tiempos de Operación

Métrica	Modo Manual	Odo-Park	Mejora
Tiempo de Procesamiento en Salida	1 minuto 45 seg	< 25 seg	~76%
Intervención Humana por Vehículo	100%	0%	100%

La Tabla 2 evidencia el impacto operativo del sistema. Los resultados muestran una reducción del tiempo de egreso de 1 minuto y 45 segundos (manual) a menos de 25 segundos (automatizado), validando una mejora de eficiencia del 76% y la eliminación exitosa de la ineficiencia administrativa.

Tabla 3: Precisión del Reconocimiento de Patentes (LPR) en Fase I

Métrica	Valor	Detalle
Total de Pruebas	30	Ciclos completos de ingreso/egreso
Vehículos Únicos	15	Flota de prueba heterogénea
Lecturas Correctas	29	Identificación positiva de caracteres
Lecturas Fallidas	1	Requiere reintento o validación manual
Tasa de Precisión	96.6%	(Correctas / Total) * 100
Condiciones de Luz	Artificial	Entorno controlado (Garage UAI)

La fiabilidad del módulo de visión artificial se resume en la Tabla 3. Tras 30 pruebas controladas, el sistema alcanzó una tasa de acierto del 96.6%, demostrando que la solución es lo suficientemente robusta para operar de forma autónoma con una intervención humana mínima

Tabla 4: Desglose de Latencia Técnica

En cuanto al análisis de latencia y la discusión de cuellos de botella expuestos en la Tabla 4, el

Componente	Operación	Tiempo Promedio	Desviación	% del Total
Cámara LPR	Captura + LPR (Ondevice)	120 ms	±15 ms	26.6%
Red (LAN)	Transmisión de Payload (JSON)	45 ms	±20 ms	10.0%
Middleware	Validación de lógica + Routing	35 ms	±5 ms	7.7%
Odoo (ERP)	Transacción XML-RPC	250 ms	±40 ms	55.5%
Total	Ciclo Completo	~450 ms	±80 ms	100%

desglose técnico permite identificar que la transacción en Odoo constituye el principal punto crítico del sistema, al concentrar más del 50% del tiempo total del ciclo. Esto implica que, ante un incremento en la concurrencia de vehículos, la comunicación sincrónica podría degradar la experiencia del usuario.

En consecuencia, resulta pertinente incorporar estrategias de mitigación tales como la optimización de llamadas RPC, el uso de colas de mensajes para desacoplar procesos y la implementación de

mecanismos de caching en el middleware. Estas mejoras permitirían que el hardware de acceso responda sin depender de la finalización inmediata de la escritura en la base de datos, preservando la coherencia transaccional sin comprometer la latencia operativa.

La proyección del sistema se vincula a objetivos medibles y a un alcance explícito mediante un roadmap que evita promesas no respaldadas estableciendo que para la segunda fase se busca reducir el tiempo de búsqueda de estacionamiento en un cuarenta por ciento mediante el despliegue de sensores iot ultrasónicos y modelos arima para predicción de ocupación mientras que la tercera fase se orienta a la escalabilidad para soportar mil transacciones simultáneas mediante la orquestación con kubernetes asegurando la trazabilidad y el realismo en cada etapa del desarrollo [15] [16]

FORMACIÓN DE RECURSOS HUMANOS

El proyecto se desarrolla en el marco del grupo de investigación en Sistemas de Información de la Universidad Abierta Interamericana. El equipo estuvo integrado por tres estudiantes de quinto año de la carrera Ingeniería en Sistemas Informáticos, en el marco de las asignaturas Seminario de Aplicación Profesional (SAP) y Trabajo Final Integrador (TFI), bajo la dirección de un docente de la Facultad de Tecnología Informática. A partir de estos avances, se continuará en 2026 con la incorporación de nuevos estudiantes de grado.

BIBLIOGRAFÍA

[1] Ramírez, S., & FastAPI Contributors. (2024). FastAPI: Modern, fast web framework for building APIs with Python.

[2] Association of Smart Home & Buildings. (2020). Four pillars of the Industrial IoT. ASHB Technical Report.

[3] Pradhan, G., Prusty, M. R., Negi, V. S., & Chinara, S. (2025). Advanced IoT-integrated parking systems with automated license plate recognition and payment management. Scientific Reports, 15.

[4] Mercado Pago Developers. (2024). API Reference: Webhooks & QR Payments.

[5] Parklio. (2024). Smart Parking Solutions & Management System.

[6] Ma, J., Arampath, A. M. S. Y., Illesingha, I. T. S., De Silva, K. T. S., & Wijayasekara, S. K. (2025). ParkVision: Smart way to automate the management of the parking system. International Research Journal of Innovations in Engineering and Technology.

[7] ParkPlus. (2024). ParkPlus System: Simplifying parking management.

[8] Pranaav, S., & Seshasayee, U. R. (2025). Automated parking management system using computer vision and parking optimization. IEEE ICCSP.

[9] Авсієвич, В., & Павлова, О. (2024). Security framework for cyber-physical smart parking systems. Cybersecurity: Education, Science, Technique.

[10] Chandrashekar, D. (2024). Vehicle number plate recognition and slot allocation in a cloud-based automatic parking management system. National College of Ireland.

[11] Encode. (2024). Starlette: The little ASGI framework.

[12] Redis Ltd. (2024). Redis Streams: Introduction and concepts.

[13] George, A. M. (2024). An intelligent and sustainable parking fee solution through cloud deployed computer vision. National College of Ireland.

[14] Coqui AI. (2025). Coqui TTS: A deep learning toolkit for Text-to-Speech.

[15] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th ed.). Wiley.

[16] Burns, B., Beda, J., Hightower, K., & Evenson, L. (2022). Kubernetes: Up and Running (3rd ed.). O'Reilly Media.

Circuito Paramétrico Cuántico para la Evaluación de Riesgo de Carteras de Inversión con 110 Qubits en Hardware Cuántico Real

Juan Pablo Braña, Alejandra M.J. Litterio, Alejandro Fernández, Lucas Jung, Rim Chebbaki
Centro de Altos Estudios en Tecnología Informática (CAETI) – Facultad de Tecnología Informática – UAI
Av. Montes de Oca 745 – (C1270AAH) Ciudad Autónoma de Buenos Aires, Argentina
{juan.brana, alejandra.litterio, alejandroa.fernandez}@uai.edu.ar, lucasemanuel.jung@alumnos.uai.edu.ar, rim@uni.minerva.edu

RESUMEN

Presentamos un circuito cuántico paramétrico original para la evaluación de riesgo de carteras de inversión, ejecutado en hardware cuántico real de IBM con 110 qubits —uno por cada activo financiero. El circuito codifica el riesgo individual de cada instrumento mediante compuertas de rotación RY parametrizadas y las dependencias entre activos mediante compuertas controladas CRY, cuyos ángulos se derivan de las correlaciones históricas. Al medir, la frecuencia de colapso al estado $|1\rangle$ de cada qubit refleja un índice de riesgo que integra tanto el componente individual como el sistémico. El circuito se ejecutó en un procesador IBM Quantum Eagle (127 qubits físicos) con datos reales de mercado. Para validar los resultados se implementó un modelo de referencia clásico basado en propagación matricial sobre el mismo grafo de correlaciones, comparando ambos enfoques mediante correlación de rangos (Spearman ρ, Kendall τ) y desplazamiento de posiciones. Se cuantificaron correlaciones conectadas entre qubits como evidencia de comportamiento no trivial del circuito. Se construyeron carteras ponderadas inversamente al riesgo y se realizó una prueba retrospectiva contra el índice SPY. Se discute el camino hacia la ventaja cuántica.

Palabras clave: *Computación Cuántica, Hardware Cuántico Real, IBM Quantum, Riesgo Financiero, Compuertas Paramétricas, Entrelazamiento, Ventaja Cuántica, Estimación de Amplitud, Portafolios de Inversión.*

CONTEXTO

El presente trabajo se enmarca en el proyecto "Inteligencia Artificial y Computación Cuántica en Finanzas", radicado en el CAETI de la UAI desde 2020. Da continuidad a investigaciones en análisis de sentimiento financiero [2, 3] y optimización de carteras mediante VQE e híbridos de computación cuántica e inteligencia artificial [4]. Es financiado por la Secretaría de Investigación de la Universidad y cuenta con docentes y alumnos de la Maestría en Tecnología Informática y la Diplomatura en Análisis de Datos.

En la fase anterior [4] se experimentó con 5 qubits y 5 instrumentos, donde la optimización clásica de Markowitz dominaba en el 92% de los escenarios. Se anticipó que al escalar qubits las aproximaciones cuánticas mejorarían. La presente fase materializa esa hipótesis utilizando procesadores IBM Quantum Eagle de 127 qubits físicos [18], e incorpora un modelo de referencia clásico formal y métricas cuantitativas, respondiendo a la pregunta: ¿qué aporta el circuito cuántico respecto a una propagación clásica equivalente?

1. INTRODUCCIÓN

La gestión cuantitativa de riesgo en carteras de inversión requiere modelar tanto el riesgo individual de cada activo como la propagación de riesgo sistémico a través de las correlaciones entre activos. Los métodos tradicionales estiman matrices de covarianza y propagan el riesgo linealmente, un enfoque que puede subestimar efectos no lineales de contagio financiero, especialmente en escenarios de crisis donde las correlaciones se amplifican de manera no proporcional [8]. La Computación Cuántica

ofrece un paradigma alternativo: codificar el riesgo en las amplitudes de probabilidad de los qubits y las dependencias entre activos en entrelazamiento cuántico. La técnica de estimación de amplitud cuántica ha demostrado ventajas cuadráticas teóricas para estimar el valor en riesgo [16], y se han propuesto aplicaciones en riesgo crediticio [5], valoración de opciones [14] y optimización dinámica de portafolios [10]. Herman y otros [7] ofrecen una revisión comprehensiva del estado del arte en computación cuántica para finanzas, mientras que Orus, Mugel y Lizaso [12] proporcionan un marco teórico general. Grant, Humble y Stump [17] realizaron una comparación rigurosa de optimización cuántica de portafolios mediante técnicas de recocido cuántico.

Proponemos un **circuito cuántico paramétrico de evaluación de riesgo** que codifica riesgo individual en compuertas RY y dependencias en compuertas controladas CRY. A diferencia de VQE/QAOA [1, 10], los parámetros se derivan de datos históricos, minimizando la profundidad del circuito para hardware con ruido [13]. Es importante precisar que se trata de un algoritmo cuántico paramétrico donde los parámetros se calculan clásicamente y la unidad cuántica realiza el muestreo del estado entrelazado resultante. No se demuestra ventaja cuántica computacional en esta fase, pero la arquitectura está diseñada como bloque extensible hacia protocolos con aceleración demostrable (Sección 3.B). La contribución es triple: (a) ejecución en hardware real a 110 qubits (22× la fase anterior); (b) **validación con modelo de referencia clásico**, métricas estadísticas, correlaciones conectadas entre qubits y prueba retrospectiva; y (c) **análisis del camino hacia la ventaja cuántica**.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Eje A – Algoritmo cuántico de evaluación de riesgo

Preprocesamiento de datos financieros.

Se descargan precios reales de cierre ajustados de 110 acciones demúltiples sectores Para cada activo j se calculan log-retornos diarios r(j,t) = ln(P(j,t)/P(j,t−1)) y se define un «indicador de día malo»: retorno inferior a una desviación estándar negativa. Se computan la probabilidad de día malo prob(j) = N(días malos)/N(total) y el impacto medio impact(j) = media(|r(j,t)| para días malos). Ambos se normalizan al rango [0,1] mediante mínimo-máximo, y su producto define el vector de riesgo compuesto: R(j) = prob_norm(j) × impact_norm(j). La matriz de correlación absoluta de retornos constituye la matriz de adyacencia A, donde A(i,j) = |corr(r_i, r_j)|.

Arquitectura del circuito cuántico.

Capa 1 – Superposición: Hadamard en cada qubit, P(|1⟩) = 0,5 uniforme.

Capa 2 – Riesgo individual (RY): Para cada activo j con riesgo normalizado Rj ∈ [0, 0,95], se aplica RY(θj) con θj = 2·arcoseno(√Rj). La parametrización mediante rotaciones RY introduce un sesgo probabilístico relativo respecto del estado de superposición inicial, calibrando el riesgo individual dentro del estado global del circuito. Activos más riesgosos reciben rotaciones mayores hacia |1⟩. El límite superior de 0,95 evita ángulos extremos que degradarían la fidelidad en hardware con ruido.

Capa 3 – Riesgo sistémico (CRY): Para cada activo i, se identifican k=3 vecinos con |corr| ≥ 0,6. Se aplica CRY(θij) con θij = 2·arcoseno(√|corr(i,j)|). La CRY actúa condicionalmente: si el qubit de control está en |1⟩, rota el objetivo. Un activo individualmente seguro pero correlacionado con activos riesgosos ve incrementada su frecuencia de |1⟩. La restricción a k vecinos con umbral limita la profundidad, crítico para hardware con ruido [13]. *Adicionalmente, el circuito soporta correlaciones con signo*: cuando la correlación es negativa, el ángulo CRY se invierte, modelando cobertura *(hedging)* donde un activo anticorrelacionado atenúa el riesgo del otro.

Ejecución en hardware cuántico real.

El circuito se ejecuta en un procesador IBM Quantum Eagle (127 qubits físicos) [18] mediante el muestreador SamplerV2 de Qiskit Runtime versión 2.1 [19]. El proceso incluye: (a)

selección automática del procesador con menor carga (b) transpilación del circuito abstracto al conjunto de instrucciones nativo del procesador mediante el optimizador de pases de Qiskit con nivel de optimización 1; y (c) ejecución con 4.000 muestras. Se extraen métricas de transpilación: profundidad del circuito, cantidad total de compuertas, distribución de operaciones (compuertas de uno y dos qubits), error promedio de lectura y error promedio de compuertas de dos qubits. Estas métricas permiten evaluar la calidad de la ejecución y estimar cuánto del resultado proviene de la señal cuántica frente al ruido del procesador. La frecuencia relativa de colapso a $|1\rangle$ de cada qubit constituye el índice de riesgo cuántico implícito del activo correspondiente.

Eje B – Modelo de referencia clásico

Se implementó un **modelo de referencia clásico que replica la lógica del circuito sin hardware cuántico**. Se construye exactamente el mismo grafo de correlaciones (k vecinos más correlacionados sobre umbral) y se normaliza como matriz de transición estocástica. El riesgo se propaga iterativamente (3 iteraciones) mediante una difusión ponderada: puntuación$_j$ = $(1-\alpha)\cdot R_j + \alpha\cdot\Sigma_i(T_{ij}\cdot$puntuación$_i)$, donde T es la matriz normalizada por filas y α=0,5 controla el equilibrio entre riesgo propio y propagado, análogo al algoritmo de clasificación de páginas [11]. Este modelo captura la propagación lineal de riesgo a través del grafo. En contraste, en el circuito cuántico la compuerta CRY produce una rotación condicional que solo actúa cuando el qubit de control está en $|1\rangle$, generando una dependencia multiplicativa entre amplitudes. Además, las CRY sucesivas crean entrelazamiento genuino entre qubits, introduciendo correlaciones sin análogo directo en la difusión matricial. Estas diferencias se manifiestan especialmente en activos que actúan como nodos centrales del grafo de correlaciones.

Eje C – Comparación estadística y prueba retrospectiva

Se comparan cuantitativamente los ordenamientos cuántico y clásico mediante cuatro métricas complementarias: (a) **Spearman ρ** (correlación de rangos), (b) **Kendall τ** (concordancia de pares), (c) solapamiento de las 10 posiciones de mayor riesgo (fracción de activos comunes), y (d) desplazamiento medio de posiciones en el ordenamiento. Se construyen dos carteras de inversión donde cada activo recibe un peso inversamente proporcional a su nivel de riesgo: $w_j = (1/(\text{riesgo}_j+\varepsilon)) / \Sigma(1/(\text{riesgo}_k+\varepsilon))$, de modo que los activos más seguros concentran la mayor participación. Se seleccionan los N=30 activos de menor riesgo según cada ordenamiento Se realiza una prueba de simulación histórica (backtesting) de tipo comprar y mantener con datos reales y se comparan indicadores: retorno total, retorno anualizado, ratio de Sharpe, volatilidad anualizada y máxima caída de la cartera cuántica, la cartera clásica y el índice de referencia SPY.

Eje D – Tablero interactivo

Para complementar nuestro estudio hemos desarrollado una plataforma web que integra todos los módulos del sistema: simulación de portafolio con configuración de parámetros, cálculo automático de indicadores financieros, gráficos comparativos interactivos (dispersión cuántico contra clásico, barras de divergencia, mapa de calor de correlaciones, curva de prueba de simulaciones), tablas de asignación de pesos y métricas de transpilación.

3. RESULTADOS OBTENIDOS / ESPERADOS

A. Resultados Obtenidos

Se ejecutó el circuito de 110 qubits en un procesador IBM Quantum Eagle con datos históricos de 110 instrumentos multisectoriales. El índice de riesgo cuántico captura patrones consistentes con el conocimiento financiero: activos de alta volatilidad como MSTR, SMCI y TSLA presentan índices de riesgo elevados; activos defensivos como PG, KO y UL muestran valores bajos.

Capa	Compuerta	Parámetro	Efecto
1. Superpos.	H	—	P(|1⟩)=0,5
2. Riesgo	RY(θj)	2·arcoseno(√Rj)	P(|1⟩) ∝ riesgo
3. Sistém.	CRY(θij)	±2·arcoseno(√|corr|)	Entrelaz./cobert.
4. Medición	Medida	4.000 muestras	Cadena → f(|1⟩)

Tabla 1. Arquitectura del circuito cuántico.

La comparación con el modelo de referencia clásico arrojó un coeficiente de Spearman ρ alto y significativo ($p < 0{,}001$), confirmando que ambos métodos identifican los mismos activos extremos. Sin embargo, el **desplazamiento medio de posiciones** y el **solapamiento parcial de las 10 posiciones superiores** revelan divergencias sistemáticas en la zona media del ordenamiento. Los activos con mayor divergencia son aquellos con *riesgo individual moderado pero alta correlación cruzada con múltiples activos*, donde las compuertas CRY introducen dependencias condicionales entre amplitudes que la difusión matricial clásica no replica.

Para cuantificar este comportamiento no trivial se calcularon **correlaciones conectadas** entre pares de qubits: $C_{ij} = \langle Z_i Z_j \rangle - \langle Z_i \rangle \langle Z_j \rangle$, donde $Z_i = +1$ si el qubit i colapsa a $|0\rangle$ y -1 si colapsa a $|1\rangle$. Si $C_{ij} \neq 0$, las mediciones de ambos qubits están correlacionadas más allá de lo que predicen sus marginales individuales. Se observaron valores C_{ij} no nulos en los pares donde actúan compuertas CRY, lo que constituye evidencia de que el circuito produce *correlaciones no separables en las estadísticas de medición*. Es importante aclarar que esta métrica no constituye una prueba formal de entrelazamiento (lo cual requeriría, por ejemplo, la violación de una desigualdad de Bell), sino evidencia de que la dependencia condicional de las CRY se manifiesta en la distribución conjunta de resultados.

La prueba retrospectiva mostró que la cartera construida a partir del ordenamiento cuántico (30 activos de menor riesgo, ponderados inversamente) exhibe un perfil de riesgo-retorno competitivo frente a la cartera clásica equivalente y frente al índice de referencia SPY. Las diferencias en composición entre ambas carteras se concentran en activos de riesgo medio donde el circuito cuántico asigna un riesgo implícito diferente al clásico, lo cual se traduce en distinta selección y ponderación de instrumentos. La estabilidad del ordenamiento fue verificada mediante múltiples ejecuciones independientes de 4.000 muestras, observándose consistencia entre corridas pese al ruido del procesador, lo que indica que la señal de riesgo supera el ruido del hardware NISQ.

B. Camino hacia la ventaja cuántica

Como se precisó en la Introducción, los parámetros (θ_j, θ_{ij}) se derivan clásicamente. El procesador cuántico muestrea el estado entrelazado pero no realiza optimización cuántica. Identificamos tres vías hacia la ventaja cuántica:

(i) Estimación de amplitud cuántica. Woerner y Egger [16] demostraron que esta técnica puede estimar el valor en riesgo y la pérdida esperada condicionada con aceleración cuadrática: $O(\sqrt{N})$ frente a $O(N)$ evaluaciones respecto a Monte Carlo clásico. Nuestro circuito puede servir como operador de carga de un protocolo de estimación de amplitud: las amplitudes codificadas por RY/CRY definen la distribución conjunta de pérdidas del portafolio, y la estimación de amplitud calcularía probabilidades de pérdida por encima de un umbral dado con significativamente menos mediciones. Esta extensión requiere hardware con menor tasa de error, pero la arquitectura del circuito ya está diseñada para servir como bloque base.

(ii) Optimización variacional. Una segunda vía consiste en reemplazar la derivación clásica de los ángulos θ por un ciclo de optimización variacional (VQE/QAOA) [1, 10]. Los ángulos se tratarían como parámetros libres optimizados mediante un optimizador clásico (COBYLA, SPSA) que minimiza una función de costo evaluada cuánticamente, por ejemplo la varianza condicional del portafolio. Este enfoque explotaría la capacidad del circuito de explorar

un espacio de Hilbert de dimensión $2^{110} \approx 10^{33}$, inaccesible para métodos clásicos de fuerza bruta.

(iii) Interferencia constructiva. La tercera vía involucra la adición de capas de interferencia al circuito. En la versión actual, el circuito aplica compuertas y mide sin explotar la interferencia entre caminos cuánticos. Añadiendo capas de amplificación tipo Grover [6] sobre el estado preparado, se podrían amplificar las amplitudes correspondientes a configuraciones de bajo riesgo del portafolio, aumentando la probabilidad de muestrear carteras óptimas. Esta técnica ha sido propuesta para optimización combinatoria en finanzas [12]. La presente implementación constituye un prototipo a escala real que valida la codificación y ejecución del problema. Las tres extensiones son pasos incrementales alcanzables con las mejoras de hardware proyectadas en las hojas de ruta de IBM Quantum [18].

C. Resultados Esperados

En próximas fases se espera: (a) implementar mitigación de errores cuánticos mediante extrapolación a ruido cero y cancelación probabilística de errores [15] para cuantificar el impacto del ruido en los resultados; (c) integrar el circuito como operador de carga dentro de un protocolo de estimación de amplitud cuántica para estimar el valor en riesgo con aceleración cuadrática; (d) incorporar optimización variacional de los ángulos; (e) integrar el módulo híbrido de computación cuántica con inteligencia artificial, incluyendo procesamiento de lenguaje natural y predicción con redes neuronales recurrentes; y (f) realizar una prueba retrospectiva rigurosa multianual comparando con estrategias de gestión activa profesional.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto consta de un equipo multidisciplinario que incluye docentes de la Maestría en Tecnología Informática, lingüistas, y expertos en Finanzas Cuantitativas. Recibió más de 10 alumnos de la UAI que realizaron pasantías profesionales y participaron de programas de incentivos a la investigación, siendo su primer contacto con la Computación Cuántica. Además, participaron activamente alumnos de universidades del exterior del país como Minerva.

5. BIBLIOGRAFÍA

[1] Alcazar, J., Leyton-Ortega, V. y Perdomo-Ortiz, A. (2020). "Classical versus Quantum Models in Machine Learning". Mach. Learn.: Sci. Technol., 1(3), 035003. DOI: 10.1088/2632-2153/ab9009.

[2] Braña, J.P., Litterio, A.M.J. y Fernández, A. (2018). "FSAL: Lexicón financiero de sentimiento en español rioplatense". Rev. Abierta Inform. Aplicada, 2(1), pp. 23-38. ISSN: 2591-5320.

[3] Braña, J.P., Litterio, A.M.J. y Fernández, A. (2020). "IA y Computación Cuántica en Finanzas". Actas del XXII WICC, UNPA, pp. 430-434.

[4] Braña, J.P., Litterio, A.M.J. y Fernández, A. (2021). "Optimización de Carteras: Benchmark Clásico, QC e Híbrido AI/QC". Actas del XXIII WICC.

[5] Egger, D. et al. (2021). "Credit Risk Analysis using Quantum Computers". IEEE Trans. Comput., 70(12), pp. 2136-2145. DOI: 10.1109/TC.2020.3038063.

[6] Grover, L. (1996). "A fast quantum mechanical algorithm for database search". Actas 28º STOC, pp. 212-219. DOI: 10.1145/237814.237866.

[7] Herman, D. et al. (2023). "Quantum computing for finance". Nature Rev. Phys., 5, pp. 450-465. DOI: 10.1038/s42254-023-00603-1.

[8] Jorion, P. (2007). Value at Risk. 3ra ed., McGraw-Hill. ISBN: 978-0071464956.

[9] Markowitz, H. (1952). "Portfolio Selection". J. Finance, 7(1), pp. 77-91. DOI: 10.1111/j.1540-6261.1952.tb01525.x.

[10] Mugel, S. et al. (2022). "Dynamic Portfolio Optimization with Quantum Processors". Phys. Rev. Research, 4(1), 013006. DOI: 10.1103/PhysRevResearch.4.013006.

[11] Page, L. et al. (1999). "The PageRank Citation Ranking". Informe Técnico, Stanford, 1999-66.

[12] Orus, R. et al. (2019). "Quantum computing for finance: overview". Rev. in Phys., 4, 100028. DOI: 10.1016/j.revip.2019.100028.

[13] Preskill, J. (2018). "Quantum Computing in the NISQ era". Quantum, 2, 79. DOI: 10.22331/q-2018-08-06-79.

[14] Stamatopoulos, N. et al. (2020). "Option Pricing using Quantum Computers". Quantum, 4, 291. DOI: 10.22331/q-2020-07-06-291.

[15] Temme, K. et al. (2017). "Error Mitigation for Short-Depth Circuits". Phys. Rev. Lett., 119(18), 180509. DOI: 10.1103/PhysRevLett.119.180509.

[16] Woerner, S. y Egger, D.J. (2019). "Quantum risk analysis". NPJ Quantum Inf., 5, art. 15. DOI: 10.1038/s41534-019-0130-6.

[17] Grant, E. et al. (2021). "Benchmarking Quantum Annealing with Portfolio Opt.". Phys. Rev. Applied, 15(1), 014012. DOI: 10.1103/PhysRevApplied.15.014012.

[18] IBM Quantum (2024). IBM Quantum System Technology. https://www.ibm.com/quantum/technology [Consultado: feb. 2026].

[19] Equipo de desarrollo de Qiskit (2024). "Qiskit: Open-Source Framework for QC". DOI: 10.5281/zenodo.2573505.

Colaboración de los LLMs para construir modelos conceptuales

Alberto Sebastián[1,2], Juan R. Lestani[1], Graciela D.S. Hadad[1,3]

[1] Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano
[2] Facultad de Tecnología Informática, Universidad Abierta Interamericana
[3] Instituto de Ingenierías y Nuevas Tecnologías, Universidad Nacional del Oeste
{alberto.sebastian, juan.lestani, graciela.hadad} @ comunidad.ub.edu.ar

RESUMEN

Para definir requisitos del software es necesario comprender el contexto en el que dicho software se usará. Esta actividad inicial de comprensión del problema precisa recolectar información de diversas fuentes, a veces contradictorios, y abstraerla en un modelo conceptual consistente, lo que insume tiempo y recursos humanos. Se propone transformar esa información capturada heterogéneas y no estructuradas, tales como transcripciones de entrevistas y documentos del dominio, en representaciones semi-formales escritas en lenguaje natural, que capturen los conceptos clave y sus relaciones en el contexto de aplicación. Para ello, se elaborará una solución basada en Grandes Modelos de Lenguaje, para que integre el conocimiento elicitado con base en las reglas sintácticas y semánticas del meta-modelo conceptual a utilizar. La solución de IA generará contenido preliminar para el modelo conceptual Léxico Extendido del Lenguaje. Este modelo facilita la comunicación entre los involucrados y, por ende, permite una mejor comprensión del contexto de aplicación. La contribución principal radica en agilizar el proceso de modelado mediante la integración automática de múltiples puntos de vista, incorporando la detección preliminar de defectos en la terminología, originados en las propias fuentes de información, con una solución adaptable a otros modelos.

Palabras clave: Ingeniería de Requisitos, Inteligencia Artificial Generativa, Modelo Conceptual, Léxico Extendido del Lenguaje.

CONTEXTO

La línea de investigación que se presenta es parte del proyecto de investigación "Asistencia inteligente para el modelado conceptual en un dominio" que se desarrolla en la Universidad de Belgrano.

1. INTRODUCCIÓN

La Ingeniería de Requisitos (IR) es una disciplina intensiva en conocimiento y comunicación, donde la brecha entre las necesidades informales de los usuarios y las especificaciones del sistema representa un desafío histórico (Hadad et al., 2025, cap.3). En el proceso de definición de requisitos se suelen construir modelos conceptuales escritos en lenguaje natural para tener un mayor acercamiento con los usuarios (Seyff et al., 2009) (Antonelli et al., 2012) (Mighetti & Hadad, 2016) (Elizalde et al., 2021), facilitando la comunicación entre los involucrados, tanto para validar los modelos como para negociar las propuestas de requisitos.

Para construir los modelos es indispensable elicitar información del contexto, la que proviene de distintas fuentes de información, ya sean personas, documentos o software existente. Diversos autores han expresado las varias dificultades para construir un modelo conceptual a partir de la información recogida en la actividad de elicitación (Mohamed et al., 2017) (Doorn et al., 2022). El modelo conceptual a construir será el Léxico Extendido del Lenguaje (LEL), que es un glosario con el vocabulario del contexto de aplicación que brinda una visión estática del dominio (Elizalde et al., 2021). En el trabajo de (Hadad et al., 2025b) se muestran las dificultades en la construcción del modelo LEL, aplicables a otros modelos conceptuales. Se ha elegido el modelo Léxico Extendido del

Lenguaje (LEL) para ser construido a través de un LLM, por ser un modelo conceptual que no solo ha sido objeto de estudio y perfeccionando en proyectos de investigación previos, sino también por su uso en la industria (Mira et al., 2018) (Andrianjaka et al., 2020) (Antonelli et al, 2024).

En las últimas décadas se avanzó en el empleo de técnicas de Procesamiento de Lenguaje Natural clásicas en tareas de síntesis, clasificación y generación de artefactos en la IR (Thakur & Gupta,2016) (Jugran et al., 2021) (Maltempo et al., 2024). Sin embargo, la irrupción de la Inteligencia Artificial Generativa, y específicamente de los Grandes Modelos de Lenguaje (LLMs), ha provocado cambios en la automatización de tareas de Ingeniería de Software, aunque su integración sistémica en las actividades de la IR aún requiere validación empírica rigurosa (Fan et al., 2023) (Hou et al., 2024).

Los LLMs son sistemas de inteligencia artificial basados en arquitecturas de redes neuronales (generalmente Transformers) que han sido entrenados con volúmenes masivos de datos para comprender y generar texto con una fluidez similar a la humana. Hou et al. (2024) describe a los LLMs como modelos que han demostrado capacidades superiores a las técnicas de Procesamiento de Lenguaje Natural tradicionales en tareas de síntesis, clasificación y generación de artefactos de software. Estos modelos no solo "predicen la siguiente palabra", sino que capturan estructuras sintácticas y semánticas complejas que permiten realizar razonamientos sobre el lenguaje (Vaswani et al., 2017).

La capacidad de los LLMs para realizar tareas de texto a modelo ha sido objeto de estudio reciente. Investigaciones preliminares han demostrado que modelos de la IA Generativa pueden generar código, diagramas de clase UML, historias de usuario o modelos de datos, con una precisión sintáctica aceptable a partir de descripciones textuales simples (Fan et al., 2023) (Spijkman et al., 2025). Sin embargo, la literatura señala que, si bien la sintaxis suele ser correcta, la precisión semántica depende fuertemente de la calidad del prompting y del contexto proporcionado (Arora et al., 2024) (Ronanki et al, 2024) (Binkhonain & Alfayez, 2025). De acuerdo con la revisión de la literatura en (Ronanki et al, 2026), se señala la escasez de guías de ingeniería de prompting para actividades específicas del dominio en la IR. En este sentido, trabajos recientes destacan que el enfoque zero-shot suele ser insuficiente para dominios complejos, sugiriendo la necesidad de estrategias más sofisticadas como few-shot learning o chain-of-thought para mejorar la fidelidad de los modelos generados (Ali, 2025) (Binkhonain & Alfayez, 2025).

La detección de ambigüedades e inconsistencias es otra área donde la Inteligencia Artificial Generativa promete avances. A diferencia de las herramientas basadas en reglas o diccionarios estáticos, los LLMs poseen una comprensión semántica que les permite identificar inconsistencias pragmáticas en textos (Xu et al., 2024). Estudios recientes comparativos indican que los LLMs pueden superar a los ingenieros de requisitos junior en la identificación de conflictos entre requisitos, aunque todavía requieren supervisión humana para validar falsos positivos (Arora et al., 2024). El desafío actual no es solo detectar defectos en un documento aislado, sino identificar contradicciones a través de fuentes heterogéneas (entrevistas y documentos), un área donde la integración de múltiples puntos de vista mediante IA es una línea de investigación activa y necesaria.

A pesar de estos avances, existe una carencia de métodos integrales que combinen la extracción de información desde fuentes no estructuradas y heterogéneas, con la construcción de modelos conceptuales consistentes que unifiquen diferentes puntos de vista. Este proyecto busca cómo cubrir este vacío de manera exploratoria mediante el uso de LLMs.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para lograr una buena comprensión del contexto de aplicación del software, se necesita adquirir conocimiento de diversas fuentes de información, para construir un modelo del dominio que luego permita propuesta de

solución de software. En ciertos contextos se suele tener múltiples usuarios e incluso abundante documentación, pero además estas fuentes pueden proveer información contradictoria, obsoleta o ambigua, e incluso irrelevante o fuera del alcance. Es por ello que se propone explorar la construcción automatizada de modelos conceptuales en Ingeniería de Requisitos, escritos en lenguaje natural, utilizando herramientas de la Inteligencia Artificial Generativa, reduciendo la intervención manual necesaria, solo para ajustes menores.

El proyecto que se inicia es de carácter exploratorio, e integra las disciplinas de Ingeniería de Requisitos con Inteligencia Artificial, dados los avances recientes en este campo y particularmente con LLMs.

Se formalizará el meta-modelo del LEL con sus reglas sintácticas y semánticas. Se conformará un corpus de entrenamiento y prueba con información elicitada de entrevistas a usuarios y documentación de dominio, provenientes de proyectos de variada envergadura y diversos dominios. Para algunos proyectos, esa información de entrada se asociará con el modelo LEL construido oportunamente por ingenieros de requisitos. Estos LEL se almacenarán separadamente de la colección de información elicitada y podrán servir luego de base para contrastar resultados respecto a los modelos construidos por la solución de IA.

Se seleccionarán dos LLMs a utilizar en la experimentación en base a criterios de disponibilidad, consumo de recursos, capacidades y eficiencia. Se definirá en detalle el flujo de trabajo basado en LLM, donde cada fase de ejecución se encargará de una actividad específica de construcción del modelo LEL, con sus entradas y salidas, incluyendo el procesamiento de las fuentes de información (ver Figura 1). Esta estructuración de las fases facilitará su adaptación a otros modelos conceptuales.

Se diseñarán diferentes estrategias de prompting para guiar en la extracción, clasificación y descripción de téminos del LEL, según el flujo de trabajo con LLM especificado. Se realizarán pruebas piloto con los prompts desarrollados y con el corpus de fuentes de información para comparar el desempeño de las distintas estrategias en los dos LLMs. En base a estas pruebas piloto, se desarrollarán pautas para integrar en el LLM información elicitada de múltiples fuentes, y detectar conflictos básicos de terminología, tales como sinónimos y polisemia. Asimismo, se definirá una fase del LLM para detectar inconsistencias, omisiones y ambigüedades mediante la comparación entre el modelo LEL generado y la información elicitada (formalizada en el corpus).

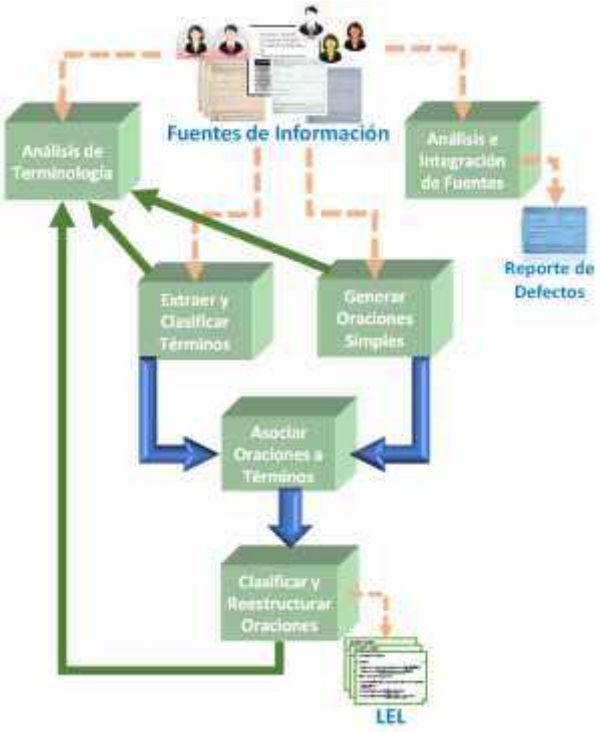


Figura 1. Esquema preliminar del flujo de trabajo basado en LLM

Finalmente, se hará una validación empírica de la solución de IA, integrando la secuencia de ejecución del LLM referidas a las actividades de construcción de un modelo LEL consistente. Para ello se llevará a cabo un experimento controlado utilizando dos casos, donde se construirán los modelos, por un lado, mediante la solución de IA usando los dos LLMs seleccionados y, por otro lado, de manera independiente por ingenieros de requisitos expertos. Se analizarán los resultados comparativamente en cuanto a tiempos y calidad de los modelos generados,

utilizando métricas de calidad previamente diseñadas.

En resumen, el objetivo general planteado consiste en desarrollar una guía metodológica basada en el uso de Grandes Modelos de Lenguaje para la generación automatizada de un modelo Léxico Extendido del Lenguaje en IR, a partir de información elicitada de diversas fuentes con diferentes puntos de vista, explorando su derivación a partir de escasa intervención humana.

Los objetivos específicos son:

- Establecer las estrategias de ingeniería de prompts y las fases de ejecución basadas en LLM necesarias para guiar la extracción de información relevante desde las fuentes no estructuradas (transcripciones de entrevistas y documentos del dominio) al modelo LEL.
- Definir guías para integrar los diversos puntos de vista provenientes de las fuentes de información en un modelo LEL unificado, resolviendo conflictos básicos de terminología.
- Incorporar capacidades de análisis estático y semántico en el LLM para la detección de defectos provenientes de las fuentes de información.
- Evaluar la calidad de los modelos LEL obtenidos por la solución de IA frente a los construidos por ingenieros de requisitos, mediante un estudio empírico comparativo.

Se asume que la generación automatizada de modelos conceptuales, específicamente glosarios, mediante LLMs debería reducir su tiempo de construcción en comparación con la elaboración manual, manteniendo una calidad semántica similar a la de un experto. Es esperable además que la incorporación de conocimiento del dominio y reglas del meta-modelo conceptual, utilizando técnicas de Ingeniería de Prompts en LLMs reduzca las alucinaciones y errores de sintaxis en los modelos generados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se dispone de una colección importante de información elicitada proveniente de proyectos de software en contextos reales de variada envergadura, lo que garantiza un input realista para las pruebas. Esta información requiere ser anonimizada y formalizada en el corpus, para un adecuado tratamiento por el LLM.

Se debe considerar que la dependencia de LLMs externos implica enfrentar el riesgo inherente de alucinaciones o imprecisiones semánticas que son difíciles de erradicar por completo. Para mitigar este problema, se harán sucesivos refinamientos de los prompts mediante pruebas piloto usando una variedad apropiada de entradas al LLM.

Como resultado de este proyecto, se espera obtener una solución inteligente basada en guías para el entrenamiento y uso de un LLM que permita construir modelos conceptuales a partir de narrativas obtenidas de transcripciones de entrevistas grabadas o provenientes de documentación que presenten los usuarios de un futuro sistema de software.

El uso de esta solución de IA será, en principio, en el ámbito académico, aunque se espera probarlo en algún proyecto en empresas de software, para colaborar con sus profesionales en la definición de requisitos de software.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto participan tres investigadores formados y un alumno de grado. En el marco del proyecto, se está desarrollando un trabajo final de grado y se está finalizando una tesis doctoral enfocada en el estudio de técnicas gráficas visuales que mejoren la comprensión del contexto de aplicación.

5. BIBLIOGRAFÍA

Ali, N. (2025). Prompting Strategies for Large Language Models in Requirements Elicitation. Tesis de Maestría. Tampere University, Finlandia.

Andrianjaka, R.M., Razafimahatratra, H., Mahatody, T., Ilie, M., Ilie, S., & Raft, R.N. (2020). Automatic generation of software components of the Praxeme methodology from ReLEL. Proceedings of the 24th International Conference on System Theory, Control and Computing, Sinaia, Romania, 843-849.

Antonelli, L., Rossi, G., Leite, J.C.S.P., & Oliveros, A. (2012). Deriving requirements specifications from the application domain language captured by Language Extended Lexicon. Proceedings of the 15th Workshop on Requirements Engineering (WER), Buenos Aires.

Antonelli. L., Lezoche. M., & Delle Ville. J. (2024). Knowledge Extraction from the Language Extended Lexicon Glossary Using Natural Language Processing. TecnoLógicas, 27(59).

Arora, C., Grundy, J., & Abdelrazek, M. (2024). Advancing requirements engineering through generative AI: Assessing the role of LLMs. En: Generative AI for Effective Software Development, 129-148. Cham: Springer Nature Switzerland.

Binkhonain, M., & Alfayez, R. (2025). Are prompts all you need? Evaluating prompt-based Large Language Models (LLM)s for software requirements classification. Requirements Engineering Journal, Elsevier, 30, 423–443.

Doorn, J.H., Hadad, G.D.S., Elizalde, M.C., Ridao, M.N., Casafuz, D., Sebastián, A., & Riera, G. (2022). Impacto del Proceso de las Entrevistas en la Calidad de los Modelos. Memoria de Trabajos de 10° Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI).

Elizalde, M.C., Hadad, G.D.S., & Doorn, J.H. (2021). Incorporación de heurísticas lingüístico-cognitivas en el Proceso de Requisitos. Memoria de Trabajos de 9° Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI), Mendoza, 312-320.

Fan, A. Gokkaya, B., Harman, M., Lyubarskiy, M., Sengupta, S., Yoo, S., & Zhang, J.M. (2023). Large Language Models for Software Engineering: Survey and Open Problems. Proceedings of 2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE), Melbourne, Australia, 31-53.

Hadad, G.D.S., Doorn, J.H., & Leite, J.C.S.P. (2025). Gestión de Requisitos para la Ingeniería de Software. Editorial Alfaomega. ISBN: 978-987-8983-90-5.

Hadad, G.D.S., Doorn, J.H., Elizalde, M.C., & Ridao, M.N. (2025b). Trayecto para Precisar Heurísticas en un Modelo Conceptual. Proceedings of the 28th Workshop on Requirements Engineering (WER), Agosto 2025, Rio de Janeiro, Brazil.

Hou, X., Zhao, Y., Liu, Y., Yang, Z., Wang, K., Li, L., Luo, X., Lo, D., Grundy, J. & Wang, H. (2024). Large Language Models for Software Engineering: A Systematic Literature Review. ACM Transactions on Software Engineering and Methodology, 33(8), article 220.

Jugran, S., Kumar, A., Tyagi, B.S., & Anand, V. (2021). Extractive Automatic Text Summarization using SpaCy in Python & NLP. Proceedings of 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 582-585.

Maltempo, G., Delle Ville, J., Cecconato, S., Pellegrino, F., Distante, D., & Antonelli, L. (2024). Multi-Word Entity Extraction and Rich Relationship Identification to Derive Conceptual Models from Natural Language Specifications. Proceedings of the 27th Workshop on Requirements Engineering (WER), Buenos Aires.

Mighetti, J.P., & Hadad, G.D.S. (2016). A Requirements Engineering Process Adapted to Global Software Development. CLEI Electronic Journal, 19(3), paper 7.

Mira, N.C., Boggio, M.A., Perez, S.B., & Salamon, A.G. (2018). Una propuesta de análisis de problemas utilizando LEL y Escenarios en entrenamientos de equipos que gestionan situaciones de crisis. Memorias del 6to Congreso Nacional de Ingeniería Informática – Sistemas de Información, Mar del Plata, 1041-1048.

Mohamed, K. A., Abd Ellatif, M., & Farhan, M. S. (2017). Using ontology-based concept maps for requirements engineering: A case study. Proceedings of International Computer Engineering Conference (ICENCO), 366–371.

Ronanki, K., Arvidsson, S., & Axell, J. (2026). Prompt Engineering Guidelines for Using Large Language Models in Requirements Engineering. En: Taibi, D., Smite, D. (eds) Software Engineering and Advanced Applications. SEAA 2025. Lecture Notes in Computer Science, vol 16083. Springer, Cham.

Ronanki, K., Cabrero-Daniel, B., Horkoff, J., & Berger, C. (2024). Requirements Engineering using Generative AI: Prompts and Prompting Patterns. En: Nguyen-Duc, A., Abrahamsson, P., Khomh, F. (eds) Generative AI for Effective Software Development. Springer.

Seyff, N., Maiden, N., Karlsen, K., Lockerbie, J., Grün bacher, P., Graf, F., & Ncube, C. (2009). Exploring how to use scenarios to discover requirements. Requirements Engineering Journal, Springer, 14(2), 91-111.

Spijkman, T., Molenkamp, B., Beudeker, S., Overbeek, S., & Dalpiaz, F. (2025). LLM-Assisted Requirements Engineering in Agile MDD: Industry Insights and Validation. Proceedings of the 2025 IEEE 33rd International Requirements Engineering Conference (RE), Valencia, Spain, 366-377.

Thakur, J.S., & Gupta, A. (2016). Anmodeler: a tool for generating domain models from textual specifications. Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering, ACM.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems, 5998-6008.

Xu, R., Qi, Z., Guo, Z., Wang, C., Wang, H., Zhang, Y., & Xu, W. (2024). Knowledge Conflicts for LLMs: A Survey. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 8541–8565, Association for Computational Linguistics.

Criterios y factores contextuales para la selección de recursos de *Design Thinking* en Ingeniería de Software: hacia una guía de buenas prácticas

Corso Cynthia, Paz Menvielle Alejandra, Guzmán Analía, Casatti Martín, Enamorado Constanza.

Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información
Departamento Ingeniería en Sistemas de Información
Centro de Investigación y Transferencia Acústica
Facultad Regional Córdoba. Universidad Tecnológica Nacional
Maestro M. López esq. Cruz Roja-Ciudad Universitaria-Córdoba
cynthia@bbs.frc.utn.edu.ar/mpaz@frc.utn.edu.ar/aguzman@frc.utn.edu.ar/mcasatti@frc.utn.edu.ar/
enamoradoconstanza0@gmail.com

RESUMEN

El desarrollo de software enfrenta desafíos crecientes relacionados con la comprensión profunda de las necesidades de los usuarios y la generación de soluciones innovadoras en contextos cambiantes. En este escenario, *Design Thinking* (DT) ha sido adoptado como un enfoque centrado en las personas que aporta valor especialmente en las etapas iniciales del ciclo de desarrollo. Sin embargo, la literatura evidencia que, aunque se documenta qué recursos (métodos) de DT se utilizan, existe una limitada exploración respecto a los criterios y factores contextuales que orientan su selección. Esta brecha genera dificultades para su integración estratégica en proyectos reales de software. El presente trabajo plantea un abordaje metodológico para la construcción y validación de una guía de buenas prácticas orientada a la selección e integración contextualizada de recursos (métodos) de DT en Ingeniería de Software.

Palabras claves: *Design Thinking, Ingeniería de Software, recursos (métodos), guía de buenas prácticas, diseño centrado en el usuario.*

CONTEXTO

Este trabajo se enmarca en el proyecto de investigación "Innovación Centrada en el Usuario: Explorando *Design Thinking* en el desarrollo de software" (PID COSITC770), aprobado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional. El contexto de desarrollo de la presente investigación es el Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS) radicado en la UTN – Facultad Regional Córdoba, y se alinea con una línea sostenida de investigaciones y publicaciones previas del equipo en torno a la aplicación de *Design Thinking* y Diseño Participativo en el ámbito del desarrollo de software.

1. INTRODUCCIÓN

El desarrollo de software enfrenta desafíos crecientes vinculados con la identificación precisa de las necesidades reales de los usuarios, la adaptación a contextos dinámicos y la generación de

soluciones innovadoras que respondan efectivamente a situaciones de uso complejas. Estos desafíos se vuelven especialmente críticos en las etapas tempranas del diseño, donde las decisiones adoptadas condicionan de manera significativa la calidad y pertinencia de los requerimientos que posteriormente guiarán el desarrollo.

En este escenario, se vuelve fundamental incorporar enfoques que permitan abordar la complejidad de los problemas desde una perspectiva centrada en las personas, favoreciendo la empatía, la revisión de supuestos y la exploración creativa de alternativas antes de formalizar especificaciones.

En los últimos años, *Design Thinking* (DT) ha cobrado relevancia como un enfoque que promueve este tipo de abordaje. Brown lo describe como un proceso iterativo orientado a comprender al usuario, cuestionar supuestos y redefinir problemas para identificar soluciones no evidentes [1]. Diversos frameworks conceptuales, como los propuestos por Stanford d.school, IDEO e IBM, coinciden en estructurar este proceso en etapas de empatía, ideación y prototipado, proponiendo dinámicas que favorecen la generación de conocimiento contextualizado y la exploración de soluciones centradas en las personas [2][3][4]. Su aplicación se ha extendido a múltiples dominios como la salud y el urbanismo, evidenciando su versatilidad para abordar problemas complejos donde la comprensión del contexto de uso resulta determinante [5][6]. Esta expansión demuestra que el valor de DT radica no solo en las etapas que propone, sino en el conjunto de recursos (métodos) que habilitan la comprensión profunda de las necesidades y la experimentación iterativa de soluciones.

En el marco de la *Ingeniería de Software*, distintos trabajos evidencian que DT se integra principalmente en etapas previas al desarrollo formal, actuando como una antesala al ciclo tradicional y enriqueciendo la definición de requerimientos desde una perspectiva centrada en el usuario [7][8]. Esta integración permite explorar problemáticas no evidentes, comprender mejor el contexto de uso y generar una base más sólida para la especificación posterior.

El mapeo sistemático realizado por Parizi et al. muestra la diversidad de recursos (métodos) utilizados y la variedad de actividades del desarrollo de software en las que se aplican [9]. Sin embargo, estos trabajos coinciden en un aspecto central: el foco se concentra en describir qué recursos (métodos) se utilizan, sin profundizar en cómo se seleccionan ni en qué condiciones contextuales resultan más adecuados. La literatura documenta el uso de entrevistas, brainstorming, prototipado y observación contextual, entre otros, pero ofrece escasa orientación respecto a los criterios que guían su elección.

Esta situación genera dificultades prácticas en la adopción del enfoque, especialmente para equipos que buscan trasladar DT a proyectos concretos de software. La abundancia de frameworks y recursos disponibles, sumada a la ausencia de lineamientos claros de selección, conduce frecuentemente a la aplicación aislada de dinámicas sin una comprensión integrada del enfoque. A su vez, se evidencia la falta de una guía estructurada que oriente su integración estratégica en el ciclo de desarrollo de software [8][10], lo que pone de manifiesto una brecha en la literatura respecto a la toma de decisiones sobre la

selección e integración de recursos (métodos) de DT en el diseño de software.

A partir de esta problemática, surge la necesidad de generar conocimiento que permita orientar de manera estructurada la planificación de experiencias de DT en el ámbito de la Ingeniería de Software, aportando criterios claros que vinculen el objetivo de diseño, el contexto del proyecto y las características de los participantes con los recursos más adecuados.

En este sentido, el objetivo general del proyecto en el que se enmarca este trabajo es proponer una guía de buenas prácticas para la selección e integración de recursos (métodos) de *Design Thinking* en el diseño de software, orientada a favorecer una adopción estratégica, contextualizada y coherente con los desafíos propios del dominio.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

El presente trabajo se desarrolla en el marco de un proyecto de investigación cuyo objetivo general es proponer una guía de buenas prácticas para la selección e integración de recursos (métodos) de *Design Thinking* en el diseño de software. Dado que el proyecto se encuentra en su primer año de ejecución, a continuación, se describen los ejes de investigación y desarrollo que estructuran el trabajo previsto para alcanzar dicho objetivo.

- Análisis sistemático de la literatura científica, orientado a identificar los recursos (métodos) de *Design Thinking* utilizados en el ámbito del desarrollo de software, sus fortalezas, limitaciones y condiciones de uso reportadas.
- Relevamiento de prácticas actuales en contextos profesionales, mediante entrevistas a perfiles clave del desarrollo de software (UX, desarrolladores, líderes técnicos, product managers), con el fin de comprender cómo se aplican estos recursos (métodos) en situaciones reales de diseño.
- Identificación de criterios y factores contextuales que influyen en la selección de recursos (métodos) de *Design Thinking*, a partir del análisis comparativo entre la evidencia reportada en la literatura y la práctica profesional relevada.
- Diseño de una guía de buenas prácticas, basada en la sistematización de los hallazgos obtenidos, que vincule objetivos de diseño, contexto del proyecto y características de los participantes con los recursos (métodos) más adecuados.
- Validación de la guía con expertos en *Ingeniería de Software*, mediante encuestas y entrevistas que permitan evaluar su claridad, pertinencia y aplicabilidad.
- Validación aplicada de la guía, solicitando a expertos la planificación de una experiencia de *Design Thinking* a partir de un caso ficticio, con el fin de analizar cómo interpretan y utilizan los lineamientos propuestos.

En conjunto, estos ejes definen el plan de trabajo y el abordaje metodológico para la elaboración de la guía propuesta.

3. RESULTADOS OBTENIDOS/ESPERADOS

Dado que el proyecto se encuentra en su etapa inicial de ejecución, los resultados esperados se vinculan con las evidencias científicas y aportes que se prevé generar a partir del desarrollo de las líneas de investigación planteadas.

En primer lugar, se proyecta una sistematización estructurada de los recursos (métodos) reportados en la literatura científica, ampliando y profundizando el relevamiento previo realizado por Parizi et al. (2022). Esta actualización permitirá incorporar evidencia reciente y analizar los recursos (métodos) no solo en términos de uso, sino también considerando sus propósitos, condiciones de aplicación, fortalezas y limitaciones, con el fin de identificar patrones y criterios que orienten su selección en distintos contextos.

En segundo lugar, se prevé construir una caracterización actualizada de las prácticas profesionales vinculadas al uso de estos recursos (métodos), a partir del relevamiento empírico en contextos reales de diseño de software. Esta caracterización contemplará criterios y factores contextuales tales como el tipo de organización, los perfiles involucrados, el momento del ciclo en que se aplican los métodos, el nivel de formalización del enfoque, las restricciones organizacionales y los criterios explícitos o implícitos que influyen en su selección.

A partir de la integración entre la evidencia relevada en la literatura y la caracterización de las prácticas profesionales, se espera delimitar un conjunto sistemático de criterios y factores contextuales que inciden en la selección de recursos (métodos) de *Design Thinking*. Estos criterios apoyarán la toma de decisiones y permitan vincular condiciones del contexto con opciones metodológicas pertinentes.

Como resultado central, se proyecta el diseño de una guía de buenas prácticas concebida como un instrumento de apoyo a la planificación de experiencias de *Design Thinking* en el ámbito del desarrollo de software. La guía articulará objetivos de diseño, características del proyecto (nivel de incertidumbre, restricciones organizacionales, momento del ciclo de desarrollo) y perfiles de los participantes con recursos (métodos) de DT pertinentes. Asimismo, establecerá correspondencias explícitas entre los criterios y factores contextuales identificados y los recursos (métodos) sugeridos, incluyendo orientaciones para su aplicación e integración en el proceso de desarrollo. De este modo, la guía no se limitará a describir recursos disponibles, sino que ofrecerá lineamientos estructurados para su selección contextualizada y su uso estratégico.

Asimismo, se prevé obtener evidencia sobre la validez conceptual de la guía mediante un proceso de validación con expertos en Ingeniería de Software y *Design Thinking*. Esta validación incluirá instancias de encuestas y entrevistas orientadas a evaluar la coherencia interna de los criterios propuestos, la claridad y comprensibilidad de su formulación, y la utilidad percibida del instrumento para el diseño de experiencias de *Design Thinking* en el contexto de estudio.

Complementariamente, se validará su aplicación en un caso de diseño ficticio con expertos previamente involucrados y nuevos participantes, con el fin de observar cómo interpretan y utilizan los criterios propuestos por la guía. Estos resultados permitirán generar aportes tanto al ámbito académico, al ofrecer una sistematización fundamentada de criterios y factores contextuales para la selección de recursos (métodos) de DT, como al ámbito profesional, al proporcionar un instrumento de apoyo a la toma de decisiones que facilite su integración

estratégica en procesos reales de diseño de software.

Se espera que los avances y evidencias generadas en el marco del proyecto favorezcan instancias de difusión y transferencia hacia otros ámbitos académicos y profesionales. En este sentido, se prevé la realización de seminarios, charlas y espacios de intercambio con docentes, estudiantes y profesionales del área, así como la transferencia de los resultados a cátedras vinculadas al desarrollo de software.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto es llevado adelante por un equipo de investigadores pertenecientes al Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS) de la UTN – Facultad Regional Córdoba, con perfiles complementarios que aportan distintas miradas al abordaje de la problemática.

Por un lado, parte del equipo cuenta con trayectoria previa en investigaciones vinculadas a *Design Thinking* y Diseño Participativo aplicados al ámbito del desarrollo de software. Por otro lado, otros integrantes poseen experiencia consolidada en construcción y desarrollo de software, aportando una visión práctica y contextualizada del dominio en el cual se busca aplicar la propuesta.

Esta combinación de perfiles permite articular el análisis conceptual con la comprensión práctica del desarrollo de software.

Asimismo, el proyecto contempla la participación de jóvenes investigadores y estudiantes en distintas etapas del trabajo, favoreciendo la formación en metodologías de investigación y en enfoques centrados en el usuario.

La articulación con el Laboratorio LIFIA (UNLP), a través del intercambio con profesionales del área temática, fortalece el enfoque metodológico y el desarrollo del proyecto.

5. REFERENCIAS

[1] T. Brown, Change by Design. Harper Business, 2009.

[2] Stanford d.school (Hasso Plattner Institute of Design). (2026). *Design Thinking – d.school, Stanford University*. Recuperado de https://dschool.stanford.edu/

[3] IDEO. (2026). *Design thinking – IDEO*. Recuperado de https://designthinking.ideo.com/

[4] IBM. (2026). *Enterprise Design Thinking: Framework – IBM*. Recuperado de https://www.ibm.com/training/enterprise-design-thinking/framework

[5] J. E. McLaughlin et al., "A qualitative review of the design thinking framework in health professions education," BMC Medical Education, 2019.

[6] G. Bermejo-Martín and C. Rodríguez-Monroy, "Design thinking methodology to achieve household engagement in urban water sustainability," Water, 2020.

[7] E. D. Canedo et al., "Design Thinking Use in Agile Software Projects," ICEIS, 2020.

[8] J. Hehn and D. Méndez, Design Thinking for Software Engineering, 2022.

[9] R. Parizi et al., "How has design thinking been used and integrated into software development activities? A systematic mapping," Journal of Systems and Software, 2022.

[10] C. Challiol et al., "Co-diseño in-situ de juegos móviles usando recursos de Design Thinking," 2020.

DE LA WEB COLABORATIVA A LA WEB SEMÁNTICA: RECUPERACIÓN Y REUSO INTELIGENTE DE INFORMACIÓN EN COMUNIDADES VIRTUALES

Valeria Zoratto[1,2], Gabriela Aranda[1,2], Nadina Martinez Carod[1,2], Lucas Cavaliere[1], Natalia Baeza[1], Sandra Lucero[1], Nahuel Mamani[1]

[1] Facultad de Informática. Universidad Nacional del Comahue
Buenos Aires 1400, (8300) Neuquén
[2] GIISCo, Grupo de Investigación en Ingeniería de Software del Comahue
{vzoratto, gabriela.aranda}@fi.uncoma.edu.ar

RESUMEN

En la actualidad, la Web constituye un ecosistema dinámico de producción y circulación de conocimiento generado por los propios usuarios. La consolidación de la Web 2.0 impulsó la transición desde un modelo centrado en la publicación de información hacia entornos participativos y colaborativos, donde plataformas como redes sociales, blogs, wikis y comunidades virtuales favorecen el intercambio de experiencias y la construcción colectiva de contenido. Este crecimiento exponencial de información producida por los usuarios plantea nuevos desafíos asociados a la organización, recuperación y aprovechamiento efectivo del conocimiento disponible.

En este contexto, la evolución hacia la Web 3.0 o Web semántica propone incorporar mecanismos de interpretación computacional capaces de estructurar y comprender la información publicada en la red, lo que facilita experiencias más inteligentes, personalizadas y orientadas al usuario. Bajo este paradigma, las comunidades virtuales de preguntas y respuestas (Community Question Answering, CQA) emergen como espacios relevantes para la adquisición y reutilización del conocimiento colectivo.

El presente proyecto se orienta al desarrollo de modelos y herramientas destinadas a la recuperación, análisis, clasificación y reuso de contenido proveniente de dichas comunidades virtuales. El enfoque propuesto integra técnicas de aprendizaje automático, detección de información relevante y algoritmos de recomendación, con el objetivo de asistir a los usuarios mediante la reutilización de respuestas previas y la integración de múltiples fuentes de información. De esta manera, se busca avanzar hacia sistemas de software capaces de transformar la interacción colaborativa de la Web 2.0 en conocimiento estructurado y reutilizable propio de la Web 3.0.

Palabras clave: Recuperación de información, comunidades virtuales, CQA, Web 3.0, reuso de información

CONTEXTO

El presente trabajo se enmarca en el proyecto de investigación "Reuso de información en comunidades virtuales" (04/F018), desarrollado durante el período 2022–2026 en la Facultad de Informática de la Universidad Nacional del Comahue. Este proyecto constituye la continuación de una de las líneas del Programa 04/F001, "Desarrollo de Software basado en Reuso", ejecutado entre los años 2013 y 2021, cuyo foco principal estuvo orientado al estudio del reuso de información proveniente de foros de discusión

técnicos.

A partir de los resultados obtenidos previamente, la investigación actual amplía su alcance hacia distintos tipos de comunidades virtuales y nuevos dominios de aplicación, incluyendo áreas de la informática, la educación y ámbitos sociales. El objetivo principal es profundizar el conocimiento adquirido mediante el desarrollo de modelos y herramientas que permitan recuperar, procesar y reutilizar información disponible en la web, contribuyendo a la toma de decisiones basada en información generada en comunidades virtuales.

1. INTRODUCCIÓN

La búsqueda de información en la Web ha experimentado una transformación significativa como consecuencia de la generación continua de contenido producida por la interacción de los usuarios en línea. Este fenómeno ha favorecido la consolidación de las denominadas comunidades virtuales, espacios digitales donde los usuarios crean, comparten y reutilizan ese conocimiento de manera colaborativa [13]. Dentro de este ecosistema, las Comunidades de Preguntas y Respuestas (Community Question Answering, CQA) se destacan como uno de los entornos más utilizados para la recuperación de información.

Desde la aparición de la Web 2.0, múltiples investigaciones han analizado el potencial de estas plataformas para la adquisición y construcción colectiva de conocimiento [1]. Sin embargo, el crecimiento masivo de información generada por los usuarios introduce desafíos asociados a la calidad, veracidad y relevancia de las respuestas proporcionadas, motivando diversas líneas de investigación orientadas al análisis de usuarios, sus roles e interacciones [14], la evaluación de la calidad de las respuestas [2, 5, 9], la detección de información confiable [12], la gestión y organización del contenido disponible [18], así como el estudio de las dinámicas sociales que permiten la conformación de comunidades especializadas [7].

El análisis de estas comunidades suele realizarse a partir de conjuntos de datos obtenidos directamente de plataformas CQA, incluyendo preguntas, respuestas, comentarios e información asociada a los usuarios. Dichos conjuntos pueden derivarse de corpus académicos existentes o generarse mediante herramientas de procesamiento del lenguaje natural aplicadas a datos extraídos de plataformas como StackOverflow y otros sitios del ecosistema StackExchange [8, 17].

En los últimos años, los avances en aprendizaje automático y aprendizaje profundo han impulsado nuevos enfoques para mejorar la recuperación y clasificación de información en comunidades virtuales, incluyendo modelos recientes orientados a la predicción de relevancia entre preguntas en distintas plataformas colaborativas [28].

Paralelamente, diversas investigaciones han explorado el uso de la información generada en comunidades virtuales como soporte para procesos de participación ciudadana y toma de decisiones públicas. Estudios recientes analizan los fundamentos y desafíos de la e-participación digital [27], así como el desarrollo de plataformas tecnológicas destinadas a promover la participación ciudadana mediada por tecnologías [26].

En este contexto, el análisis de la información generada por los usuarios presenta desafíos similares a los abordados en comunidades CQA, permitiendo reutilizar técnicas previamente desarrolladas para el análisis de contenido colaborativo. En consecuencia, el presente proyecto propone aplicar estrategias de recuperación, reuso e integración de información (Information Reuse and

Integration, IRI) para obtener, integrar y validar conocimiento proveniente de comunidades virtuales, orientándose al soporte de procesos de toma de decisiones en distintos dominios de aplicación [6].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto cuenta con la participación de docentes y estudiantes del Grupo de Investigación en Ingeniería de Software del Comahue (GIISCo), perteneciente a la Facultad de Informática de la Universidad Nacional del Comahue (UNCo). Asimismo, incluye la colaboración de asesores externos de la Facultad de Humanidades de la misma universidad y de la Facultad de Ciencias Exactas de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA).

El enfoque principal del proyecto se orienta al desarrollo de técnicas y algoritmos de recomendación destinados a brindar asistencia inteligente a los usuarios de comunidades virtuales en la búsqueda de información relevante. Los docentes participantes poseen formación en diversas áreas, entre ellas Programación, Ingeniería en Sistemas y Teoría de la Computación, lo que aporta una perspectiva multidisciplinaria al desarrollo de la investigación.

La diversidad de perfiles académicos permite abordar los objetivos del proyecto desde distintos enfoques, favoreciendo la implementación de mejoras continuas y promoviendo el trabajo colaborativo dentro del equipo de investigación.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Desde 2013, nuestro equipo de investigación trabaja en temáticas vinculadas al análisis de comunidades virtuales. En una primera etapa, se propuso un modelo de calidad para publicaciones en foros de discusión técnicos, basado en criterios orientados a evaluar la información contenida en los hilos de discusión [3]. Posteriormente, el trabajo se centró en el estudio de los foros como un tipo específico de comunidad virtual, abordando tanto la clasificación de roles de usuarios según su nivel de conocimiento y experticia [11], como la definición de métricas de calidad aplicadas al contenido de los hilos de discusión [10].

Asimismo, se incorporaron herramientas gramaticales y recursos léxicos para la clasificación del contenido de foros de discusión [4], junto con el análisis de características mínimas para clasificar publicaciones según su calidad mediante distintos clasificadores [19]. A partir de estos avances, se desarrollaron modelos de aprendizaje automático orientados al análisis de la información en foros, incluyendo la clasificación de mensajes según su función dentro de los hilos [21] y la caracterización de usuarios [22].
Posteriormente, dichos modelos fueron ampliados mediante la incorporación de técnicas de selección de características basadas en Chi-cuadrado, para encontrar conjuntos más pequeños que permitieron optimizar el desempeño e interpretabilidad del proceso de clasificación [23].

Finalmente, los resultados obtenidos comenzaron a proyectarse hacia aplicaciones vinculadas a la participación ciudadana, incluyendo una revisión sistemática de herramientas tecnológicas de e-participación [24] y el desarrollo de una plataforma web open source orientada a la gestión de encuestas y votaciones urbanas como soporte para la toma de decisiones en contextos sociales locales [25].

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto cuenta con la participación de profesionales, asesores e investigadores del CONICET, docentes del área de informática y estudiantes de grado y posgrado que desarrollan sus trabajos finales y tesis en el marco de la investigación, lo que fortalece su carácter multidisciplinario. Entre los integrantes de la Universidad Nacional del Comahue se incluyen *tres docentes del Departamento de Programación con dedicación exclusiva* y *Doctorado en Ciencias de la Computación*; *una docente* del mismo departamento con dedicación parcial que se encuentra realizando su tesis de Maestría en Ciencias de la Computación; *un docente del Departamento de Ingeniería de Computadoras* con dedicación simple; y *una docente* colaboradora del Departamento de Teoría de la Computación con dedicación exclusiva.

Asimismo, el proyecto cuenta con la asesoría de una investigadora del Instituto Superior de Ingeniería del Software (ISISTAN), de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), especialista en Sistemas de Recuperación de Información y doctora en Ciencias de la Computación. También participan docentes de la Facultad de Humanidades y Facultad de Economía y Administración del Instituto Patagónico de Estudios de Humanidades y Ciencias Sociales.

En relación con la formación académica, el proyecto incluye actualmente una estudiante de Maestría en Ciencias de la Computación y previendo la incorporación de nuevos tesistas durante el presente año.

5. BIBLIOGRAFÍA

[1] Al-ghamdi, H. A. K., & Al-ghamdi, A. A. K. (2015). The role of virtual communities of practice in knowledge management using Web 2.0. *Procedia Computer Science*, *65*, 406-411.

[2] Amancio L., Dorneles C., and Dalip D.. Recency and quality-based ranking question in CQAs: A stack overflow case study. Information Processing Management, 58(4), 2021.

[3] Aranda, G. N., Martínez Carod, N., Faraci, P., & Cechich, A. (2013). Hacia un framework de evaluación de calidad de información en foros de discusión técnicos. In *XIV Simposio Argentino de Ingeniería de Software (ASSE)-JAIIO 42 (2013)*.

[4] Aranda G, Zoratto V., Martínez Carod N., Roger S, Otermin F., Cechich A. . Clasificación de contenido de hilos de discusión mediante análisis sintáctico y morfológico. CICCSI 2018. ISBN 9789874568366. Mendoza, 2018, pp. 35-44.

[5] Braslavski, P., Savenkov, D., Agichtein, E., & Dubatovka, A. (2017, March). What do you mean exactly? Analyzing clarification questions in CQA. In *Proceedings of the 2017 conference on conference human information interaction and retrieval* (pp. 345-348).

[6] Day, M. Y., Ong, C. S., & Hsu, W. L. (2010). An analysis of research on information reuse and integration (2003-2008). *International Transactions on Systems Science and Applications*, *6*(2/3), 146-157.

[7] Fu, C. (2020). User intimacy model for question recommendation in community question answering. *Knowledge-Based Systems*, *188*, 104844.

[8] Hazoom, M., Malik, V., & Bogin, B. (2021). Text-to-SQL in the wild: a naturally-occurring dataset based on Stack exchange data. *arXiv preprint arXiv:2106.05006*.

[9] Lu, H., & Kong, M. (2017, February). Community-based question answering via contextual ranking metric network learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 31, No. 1).

[10] Martínez Carod, N. M., Zoratto, V., Aranda, G., & Faraci, P. (2018). Aplicación de Métricas de calidad a hilos de discusión. CICCSI 2018. ISBN 9789874568366. Mendoza, 2018, pp. 151-160.

[11] Martínez Carod, N. , Aranda, G., Zoratto, V., & Murray, C.. Una propuesta para clasificación de roles de usuarios en foros de discusión técnicos. CACIC (2019).

[12] Mihaylova, T., Karadjov, G., Atanasova, P., Baly, R., Mohtarami, M., and Nakov, P. (2019). SemEval-2019 task 8: Fact checking in community question answering forums. In Proceedings of the 13th International Workshop on Semantic Evaluation, pages 860–869, Minneapolis, Minnesota, USA. Association for Computational Linguistics.

[13] Ogbamichael, H. B., & Warden, S. (2018). Information and knowledge sharing within virtual communities of practice. South African Journal of Information Management, 20(1), 1-11.

[14] Rivera Rásury, M. X. (2019). Desarrollo de una herramienta de soporte metodológico a los procesos de e-participación.

[15] Roy, P. K., Saumya, S., Singh, J. P., Banerjee, S., & Gutub, A. (2023). Analysis of community question-answering issues via machine learning and deep learning: State-of-the-art review. CAAI Transactions on Intelligence Technology, 8(1), 95-117.

[16] Santamaría-Philco, A., Cerdá, J. H. C., & Gramaje, M. C. P. (2019). Advances in e-Participation: A perspective of Last Years. *Ieee Access*, *7*, 155894-155916.

[17] Wang, S., Lo, D., & Jiang, L. (2013, March). An empirical study on developer interactions in stackoverflow. In *Proceedings of the 28th annual ACM symposium on applied computing* (pp. 1019-1024).

[18] Zhang, Y., Lo, D., Xia, X., & Sun, J. L. (2015). Multi-factor duplicate question detection in stack overflow. *Journal of Computer Science and Technology*, *30*, 981-997.

[19] Zoratto V., Aranda G, Martinez Carod N., Otermin F.. Evaluación de estrategias para clasificar hilos de foros de discusión según su contenido, ASSE-JAIIO 2021, Argentine Symposium on Software Engineering, Argentina, 2021.

[20] Zoratto, V., Godoy, D., & Aranda, G. N. (2023). A Study on Influential Features for Predicting Best Answers in Community Question-Answering Forums. Information, 14(9), 496.

[21] Zoratto, V., Aranda, G. N., Saldivia Urra, J. (2024) Clasificación de posts según su rol en hilos de foros de discusión, CACIC 2024: XXX Congreso Argentino de Ciencias de la Computación, Oct 2024, La Plata, ISBN: 978-950-34-2428-5, pp. 678-687

[22] Aranda, G., Zoratto, V., Cárdena, A. (2024): Análisis del comportamiento de los usuarios que preguntan en CQA, CACIC 2024: XXX Congreso Argentino de Ciencias de la Computación, Oct 2024, La Plata, ISBN: 978-950-34-2428-5, pp. 629-638

[23] V. Zoratto, G. Aranda, (2025): Evaluating Feature Importance in Post Classification Using Chi-Square Analysis. Communications in Computer and Information Science, vol 2520. Springer, Cham. pp. 167-183.

[24] N. Mamani, G. Aranda (2025), Hacía una herramienta de participación ciudadana: primeros pasos en el diseño de una aplicación para opinar sobre asuntos urbanos. CoNaIISI 2025. 13º Congreso Nacional de Ingeniería Informática/ Sistemas de Información. Nov. 2025.

[25] S. Lucero, G. Aranda. (2025): Herramientas para la e-participación: Una revisión sistemática de la literatura. CACIC 2025: XXXI Congreso Argentino de Ciencias de la Computación, Oct 2025, Universidad Nacional de Río Negro, Argentina. ISBN: 978-987-8258-99-7.

[26] López, F., Abitú, V. A., & Ordinez, L. (2024). Plataformas virtuales de participación ciudadana. Revista Electrónica De Divulgación De Metodologías Emergentes En El Desarrollo De Las STEM, 6(1), 43–67.

[27] Mariani, I., Mortati, M., Rizzo, F., Deserti, A. (2025). The Theoretical Background on E-Participation. In: Design Thinking as a Strategic Approach to E-Participation. SpringerBriefs in Applied Sciences and Technology(). Springer, Cham.

[28] Song, Y., Jiang, B., Zhang, D., Liao, Z.: Cross-community question relevance prediction for stack overflow and github. Journal of Universal Computer Science 31(1), 52 (2025)

Desafíos en el Diseño y la Comunicación entre Microfrontend

Franco Pappatella, Natalia Cirillo, Nicolas Battaglia, Gustavo Rossi

CAETI. Facultad de Tecnología Informática, Universidad Abierta Interamericana (UAI)
pappatella.francoariel@gmail.com, nataliacirilloo@gmail.com, nicolas.battaglia@uai.edu.ar, gustavo.rossi@uai.edu.ar

RESUMEN

Los Microfrontend se consolidaron como un enfoque arquitectónico para aplicaciones interactivas distribuidas al dividir la interfaz en componentes independientes desarrollados por distintos equipos y tecnologías. Esta autonomía mejora la escalabilidad, pero introduce el desafío de la comunicación entre módulos.

Este proyecto busca analizar y clasificar las estrategias de comunicación existentes entre Microfrontend y sus implicancias en desacoplamiento, autonomía y complejidad, aportando criterios para su selección en arquitecturas interactivas distribuidas.

Palabras clave: *Microfrontends, arquitecturas distribuidas, comunicación entre componentes, desacoplamiento, front-end modular*

CONTEXTO

Esta investigación representa una línea temática dentro del proyecto: Aspectos de Arquitectura y diseño en aplicaciones distribuidas (Web, Móviles, IoT, etc) con foco en Microservicios (Battaglia et al 2025ª)

La investigación se enmarca en el área de Arquitecturas de Software y Sistemas Distribuidos, enfocada en patrones aplicados al front-end de aplicaciones distribuidas complejas.

La necesidad de escalabilidad y despliegue independiente impulsó la adopción de Microfrontend. Este tipo de descomposición plantea desafíos en la gestión del estado compartido y la coordinación entre módulos, para mantener la independencia sin afectar la coherencia del sistema. Asimismo los Microfrontend plantean otros desafíos de diseño interesantes respecto al diseño de la interfaz y la interacción

1. INTRODUCCIÓN

Las arquitecturas de Microservicios representaron un avance importante en cuanto a la obtención de modularidad en aplicaciones complejas. Descomponer un dominio de aplicación en componentes independientes (los microservicios) que se comunican entre sí permite (cuando los componentes han sido diseñados apropiadamente) maximizar la modularidad, garantizando mayor eficiencia no solamente en el proceso de evolución del software sino también en su despliegue (Jamshidi et al. 2018, Kapferer et al. 2020, Newman 2021, Oumoussa et al. 2024).

Hasta hace unos años este tipo de descomposición se aplicaba fundamentalmente a los aspectos del dominio de aplicación (por ejemplo descomponiendo en función de “áreas” del negocio), pero no a los aspectos de presentación e interacción.

Los Microfrontend también proponen descomponer un front-end monolítico en

partes desarrollables y desplegables de forma independiente, habilitando autonomía de equipos y flexibilidad tecnológica, aunque con costos como mayor complejidad operativa, desafíos de gestión de estado y duplicación de código (Silva et al., 2025).

Los Microfrontend aplican los mismos conceptos de descomposición en componentes, pero en el dominio de la interfaz y la interacción, dividiendo la interfaz en módulos autónomos (por ejemplo, por dominio funcional), que, de forma semejante a los microservicios, podrán ser desarrollados y desplegados de forma independiente.

Si bien es esperable obtener las mismas ventajas que se han obtenido en el uso de microservicios, existen algunos problemas que aun deben ser caracterizados y resueltos. Por un lado la descomposición de interfases en módulos independientes durante el proceso de diseño no necesariamente resulta trivial dado que los diseñadores (de interfaz) suelen ver a la misma como un desarrollo holístico, que debe resultar en interfaces consistentes. Con este análisis, entonces, la división en módulos propuesta por los Microfrontend (eventualmente dividiendo también los aspectos estéticos en equipos de desarrollo diferentes), podría no ser la mejor.

Por otro lado, la independencia de estas componentes requiere el establecimiento de mecanismos de comunicación, que exceden el tipo de mecanismos utilizados por los microservicios, dada la propia naturaleza de los Microfrontend. Mientras que en arquitecturas monolíticas tradicionales, la relación entre la interfaz y el back-end suele basarse en dependencias directas y estados compartidos (por ejemplo siguiendo el estilo del patrón Model-View-Controller), en entornos distribuidos pueden generar acoplamiento y reproducir limitaciones típicas de los monolitos.

Este problema de acoplamiento no es sólo técnico, sino también organizacional. La autonomía de los Microfrontend no se limita al despliegue independiente, sino que también comprende la capacidad de ser desarrollados y mantenidos por equipos distintos, con control sobre sus decisiones técnicas.

El desacoplamiento del front-end favorece equipos autónomos organizados por rebanadas verticales de funcionalidad, con responsabilidad end-to-end sobre una parte del producto (Jackson, 2019)

Para que esta autonomía sea efectiva, la comunicación entre componentes no puede basarse en el conocimiento de las implementaciones internas de otros Microfrontends. Cuando esto ocurre, se introducen dependencias que obligan a una coordinación constante y reducen la capacidad de evolución independiente.

Por otra parte, no se debe confundir el concepto de Microfrontend con que cada módulo debe obtener todos sus datos de forma independiente, ya que podríamos tener datos duplicados entre los Microfrontend, y eso podría conllevar un derroche innecesario de memoria del navegador. Para evitar esto, es necesario diseñar un flujo de datos eficiente para garantizar una comunicación performante y coherente, formando una estructura de datos capaz de almacenar el estado que tienen en común los Microfrontend de nuestra aplicación.

Ante estos desafíos, la elección de una buena estrategia de comunicación se convierte en un factor clave para implementar esta arquitectura. Evitar el acoplamiento innecesario, maximizar la autonomía de los

módulos y garantizar la eficiencia del flujo de datos requiere emplear mecanismos diseñados específicamente para entornos distribuidos y heterogéneos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta línea de investigación buscamos profundizar diversos aspectos aún abiertos en la tecnología de Microfrontend.

Por un lado, nos interesa discutir los problemas de diseño mencionados en la sección anterior, específicamente cómo hacer compatible la naturaleza distribuida del diseño, desarrollo y despliegue de los Microfrontend con el diseño en general centralizado de los estilos de interfaz e interacción en las aplicaciones modernas. Observamos en este punto que existe una tensión en este aspecto entre la modularidad, típica de las aplicaciones distribuidas desarrolladas con microservicios, y la usabilidad deseable en cualquier aplicación moderna y que, según el estado del arte en diseño, suele alcanzarse con interfaces consistentes y desarrolladas como un "todo".

Esta tensión puede derivar en anti-patrones de descomposición: nano frontends (granularidad excesiva y mayor costo/acoplamiento) o mega frontends (granularidad insuficiente y "efecto monolito") (Silva et al., 2025).

En los aspectos de comunicación investigaremos sobre las diferentes estrategias de comunicación entre Microfrontend y su impacto en el desacoplamiento y la evolución independiente dado que la elección del mecanismo de comunicación debe responder a criterios técnicos y organizacionales, priorizando la evolución independiente y la coherencia arquitectónica.

3. RESULTADOS OBTENIDOS / ESPERADOS

En un análisis inicial hemos observado lo siguiente respecto a algunos de los mecanismos de comunicación más comunes, expresados por los patrones: Event Bus, Local Storage, Backend for Front-end y Callbacks.

El patrón **Event Bus** centraliza la comunicación mediante eventos compartidos. Los módulos emiten y escuchan sin dependencia directa, favoreciendo el desacoplamiento. Sin embargo, los eventos no persisten y requieren convenciones.

El uso de **Local Storage** permite persistencia y sincronización entre pestañas, útil para información global, aunque con limitaciones en reactividad.

El patrón **Backend For Front-end** (BFF) introduce una capa intermedia entre Microfrontend y backend. Cada módulo interactúa con un backend propio, reduciendo la complejidad en el cliente y favoreciendo la autonomía, aunque incrementa la complejidad general.

Los **callbacks** orquestados por un contenedor centralizan el estado y el flujo de información. Es un enfoque simple, pero limita la independencia tecnológica.

En estos momentos nos encontramos analizando un repertorio más completo de estilos existentes y las características de cada uno.

En la tradición de los patrones de diseño y arquitectural esperamos desarrollar una guía exhaustiva con recomendaciones de cuándo usar cada tipo de mecanismo y casos de éxito, junto con discusiones respecto a trade-offs en cada caso.

Otro resultado esperado es un conjunto de guidelines de diseño para buscar la compatibilidad entre desacoplamiento del diseño y usabilidad de la interfaz.

4. FORMACIÓN DE RECURSOS HUMANOS

Este trabajo está originado en la participación de estudiantes de grado de la Facultad de Tecnología Informática de la UAI en el proyecto de investigación Aspectos de Arquitectura y diseño en aplicaciones distribuidas (Web, Móviles, IoT, etc) con foco en Microservicios. Se espera que parte de esta investigación termine en trabajos de fin de carrera. Al mismo tiempo se están extendiendo dos tesis doctorales, sobre diseño de microservicios usando métodos ágiles (Battaglia et al 2024, Battaglia 2025, Battaglia et al 2025b) e inteligencia artificial generativa (Narvaez 2025) de manera de definir proyectos finales de maestría y tesis doctorales alrededor de esta nueva temática.

5. BIBLIOGRAFÍA

Battaglia, N., Garcia, A. N., & Congiusti, A. (2024). Descubrimiento de Microservicios en Metodologías Ágiles: un mapeo sistemático de la literatura. XXX Congreso Argentino de Ciencias de La Computación, 1506–1510.

Battaglia, N. (2025). Identificación Ágil de Microservicios Utilizando DDD y BDD. XXVIII Ibero-American Conference on Software Engineering (CIbSE 2025 Doctoral Symposium).

Battaglia, N., Narváez, D., & Rossi, G. H. (2025a). Aspectos de arquitectura y diseño en aplicaciones distribuidas (web, móviles, Iot, etc) con foco en microservicios. In XXVII Workshop de Investigadores en Ciencias de la Computación (Mendoza, 10 y 11 de abril de 2025)

Battaglia, N., Rossi, G., & Fernández., A. (2025b). Behavior-Driven Microservice Architecture: Un Marco Metodológico para la Identificación Iterativa de Microservicios en Proyectos Ágiles Greenfield. Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería Del Conocimiento JIISIC 2025.

Fernandez, T. (2022, 29 de Septiembre). *Microfrontend: Microservices for the Front-end*. Semaphore. https://semaphore.io/blog/Microfrontend

Jackson, C. (2019, 19 de Junio). *Micro front-ends*. Martin Fowler. https://martinfowler.com/articles/micro-front-ends.html

Jamshidi, P., Pahl, C., Mendonça Nane Kratzke, & others. (2018). Microservices: The Journey So Far and Challenges Ahead. IEEE Software, 35(3), 24–35.

Kapferer, S., & Zimmermann, O. (2020). Domain-specific Language and Tools for Strategic Domain-driven Design, Context Mapping and Bounded Context Modeling. MODELSWARD, 299–306.

Narvaez D. Explorando el Uso de Inteligencia Artificial en el Descubrimiento y Diseño de Microservicios. In: Congresso Ibero-Americano em Engenharia de Software (CIbSE). SBC. 2025:224–31.

Newman, S. (2021). Building Microservices (2nd ed.). O'Reilly Media.

Oumoussa, I., & Saidi, R. (2024). Evolution of microservices identification in monolith

decomposition: A systematic review. IEEE Access, 12, 23389–23405.

Silva, N., Rodrigues, E., & Conte, T. (2025, 27 de Marzo). *A Catalog of Micro Front-ends Anti-patterns*. Federal University of Amazonas. https://arxiv.org/abs/2411.19472

DISEÑO Y VALIDACIÓN DE UN MODELO DE CALIDAD BASADO EN CMMI E ISO/IEC 25010 PARA GESTIÓN DE RIESGOS DE CALIDAD DERIVADA DE LA ROTACIÓN DE PERSONAL EN LOS EQUIPOS DE DESARROLLO

David Oghenerume Brughoro, Alberto Sanchez, Carlos Salgado, Mario Peralta

Departamento de Informática, Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis, Ejército de los Andes 950, San Luis
dbrughoro@gmail.com, {alfanego, csalgado, mperalta}@unsl.edu.ar

RESUMEN

En el ámbito del desarrollo de software, un problema/fenómeno que provoca pérdida de conocimiento crítico y, en consecuencia, un impacto directo en la calidad del producto, es la rotación del personal en equipos de desarrollo. Si bien existen modelos consolidados para la gestión de procesos y la evaluación de la calidad, se identifica un vacío relevante: no se dispone de un enfoque que integre de manera sistemática los lineamientos de CMMI, las categorías de calidad de ISO 25010 y los factores asociados a la rotación del personal. Para cubrir esta brecha, en esta línea de investigación se propone definir un modelo conceptual sustentado en una matriz de relaciones que permite mapear los efectos de la rotación sobre atributos específicos de calidad y prácticas de madurez de procesos. La metodología utilizada combina una revisión de la literatura sobre calidad, gestión del conocimiento y rotación, junto con la formulación de una propuesta preliminar del modelo. Como aporte principal, se ofrece un marco inicial que facilita la identificación, análisis y gestión de riesgos derivados de la pérdida de conocimiento en entornos de desarrollo, sirviendo como base para futuras validaciones empíricas y refinamientos metodológicos.

Palabras Clave: CMMI, ISO/IEC 25010, rotación de personal, calidad de software, gestión de riesgos.

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto Nº P-03-2020. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

La rotación de personal es un fenómeno persistente en la industria del software y afecta tanto a organizaciones grandes como pequeñas. En un entorno caracterizado por alta demanda de talento, ciclos de desarrollo acelerados y equipos distribuidos, la salida frecuente de integrantes clave implica la pérdida de conocimiento tácito y explícito acumulado durante el proyecto. Esta pérdida repercute directamente en la continuidad del trabajo, la capacidad de toma de decisiones y, en última instancia, en la calidad del producto entregado. Diversos estudios han documentado cómo la rotación incrementa la probabilidad de defectos, retrabajos y desviaciones en tiempos y costos, convirtiéndose en un riesgo crítico para la sostenibilidad de los equipos.

Analizar este problema es relevante porque, aunque existen marcos consolidados como CMMI [1] para la gestión de procesos y la norma ISO 25010 [2] para la evaluación de la calidad del producto, no se dispone de un modelo que conecte explícitamente estas perspectivas con los efectos de la rotación. Esta

ausencia limita la capacidad de las organizaciones para anticipar el impacto de la pérdida de conocimiento y diseñar estrategias integrales de mitigación.

El presente artículo aporta una primera aproximación conceptual, describiendo la idea central de un modelo de calidad que relaciona prácticas de madurez, atributos de calidad y factores asociados a la rotación.

Así en esta etapa temprana de la investigación, se ha iniciado un estudio del estado del arte sobre gestión del conocimiento, calidad del software y rotación del personal.

En este sentido, la literatura sobre rotación de personal en equipos de desarrollo de software ha destacado consistentemente sus efectos adversos en la continuidad operativa y en los resultados de los proyectos [3]. La rotación implica la pérdida de conocimiento tácito, aquel que los desarrolladores adquieren mediante la experiencia diaria, la interacción con el código y el entendimiento de las decisiones arquitectónicas. Este tipo de conocimiento es especialmente difícil de documentar y transferir, por lo que su desaparición genera brechas cognitivas dentro del equipo. Diversos estudios, muestran que, cuando este conocimiento se pierde, disminuye la productividad debido al tiempo requerido para capacitar a nuevos integrantes, y aumenta la probabilidad de errores, fallos de comunicación y demoras en el ciclo de desarrollo. En consecuencia, la calidad del producto también se ve afectada, manifestándose en defectos más frecuentes, menor coherencia en el diseño y degradación de atributos clave.

En el ámbito de los modelos de madurez, CMMI aporta prácticas que indirectamente contribuyen a mitigar la pérdida de conocimiento, al promover la estandarización de procesos, la gestión sistemática de la capacitación y el aseguramiento de la calidad. Áreas de proceso como *Training, Quality Assurance y Causal Analysis and Resolution* (CAR [1]) ofrecen lineamientos que facilitan la transferencia y retención del conocimiento, ya sea mediante prácticas formales de aprendizaje, actividades de monitoreo de calidad o análisis estructurado de causas raíz.

Por su parte, el modelo ISO/IEC 25010 define atributos de calidad del producto particularmente sensibles a la rotación del personal. La mantenibilidad puede verse comprometida si la salida de desarrolladores clave dificulta comprender, modificar o extender el sistema. La confiabilidad puede afectarse cuando se pierde el entendimiento de los mecanismos internos que garantizan funcionamiento estable ante fallos. La idoneidad funcional también puede degradarse si no se preserva el conocimiento relacionado con requisitos, contexto de uso y decisiones funcionales tomadas durante el proyecto. En conjunto, estos atributos dependen fuertemente de la continuidad del conocimiento dentro del equipo, lo que hace que la rotación sea un factor de riesgo crítico.

2. PROPUESTA

Como se mencionó previamente, en el este trabajo, se propone definir un modelo de calidad que relaciona prácticas de madurez, atributos de calidad y factores asociados a la rotación de personal en equipos de desarrollo de software. El objetivo del modelo es ofrecer un mecanismo sistemático para identificar, analizar y mitigar los riesgos de calidad derivados de la rotación del personal en equipos de desarrollo de software. No se trata de un software ni de un prototipo operativo, sino de un marco conceptual que articula riesgos asociados a la salida de personal con prácticas de proceso y atributos de calidad establecidos en marcos ampliamente aceptados. El modelo busca servir como una guía para organizaciones que desean anticipar el impacto de la pérdida de conocimiento y diseñar estrategias de mitigación basadas en estándares reconocidos.

2.1. Componentes del modelo

El modelo integra tres componentes fundamentales:

1. Riesgos por rotación: Incluyen la pérdida de conocimiento tácito, interrupción de la

continuidad técnica, incorporación de errores por desconocimiento del diseño previo, y disminución de la mantenibilidad del código.

2. Procesos de CMMI: Se consideran áreas relevantes relacionadas con transferencia y preservación del conocimiento, tales como capacitación, aseguramiento de calidad y análisis y resolución de causas. Estas prácticas permiten establecer mecanismos formales de documentación, monitoreo y aprendizaje organizacional.
3. Atributos de calidad de ISO 25010: Se incorporan atributos sensibles a la rotación, como mantenibilidad, confiabilidad e idoneidad funcional. La matriz del modelo permite vincular cada riesgo con los atributos afectados y con las prácticas de CMMI que pueden contribuir a mitigarlo.

2.2.La matriz del modelo

El corazón de la propuesta es una matriz de relaciones que permite mapear riesgos, procesos y atributos. A modo de ejemplo ilustrativo, la Figura uno muestra la matriz simplificada.

Riesgo	CMMI (Proceso relevante)	ISO 25010 (Atributo afectado)	Posible mitigación
Pérdida de conocimiento tácito	Capacitación; CAR	Mantenibilidad	Programas de onboarding; lecciones aprendidas
Introducción de defectos	Aseguramiento de calidad	Confiabilidad	Revisiones técnicas; estándares de codificación
Desalineación funcional	Gestión de requisitos	Idoneidad funcional	Trazabilidad; documentación estructurada

Figura 1. Matriz del Modelo Simplificada

Esta matriz puede ampliarse con mayor detalle según la organización y los tipos de proyecto. Ejemplo de aplicación

En una organización de desarrollo de software, el modelo se aplicaría iniciando con la identificación de roles críticos y la evaluación de riesgos asociados a su potencial rotación. Luego, el equipo de procesos mapearía cada riesgo con las prácticas de CMMI disponibles internamente y con los atributos de calidad del producto que podrían verse afectados. A

partir de esta vinculación, la organización seleccionaría estrategias de mitigación concretas, tales como fortalecer los planes de capacitación, mejorar los mecanismos de aseguramiento de calidad o reforzar la documentación arquitectónica.

Durante la ejecución del proyecto, la matriz serviría como herramienta de monitoreo: permitiría evaluar si la salida de un miembro del equipo activa alguno de los riesgos catalogados y, en consecuencia, ejecutar las medidas de mitigación previamente definidas. De esta forma, el modelo actúa como un marco preventivo y adaptable, ofreciendo una perspectiva integrada que combina procesos, calidad del producto y gestión del conocimiento ante la rotación del personal.

2.3.Metodología

La metodología propuesta describe el plan de trabajo para desarrollar y consolidar el modelo conceptual, sin requerir que todas las etapas hayan sido ejecutadas al momento de la redacción. El proceso se compone de tres fases principales.

En primer lugar, se realizará una revisión bibliográfica sistemática centrada en tres ejes: estudios sobre rotación de personal y pérdida de conocimiento en equipos de software, marcos de madurez organizacional y procesos relacionados con la gestión del conocimiento, y atributos de calidad del software definidos en estándares internacionales. Esta revisión permitirá identificar conceptos clave, relaciones teóricas y vacíos existentes en la literatura, proporcionando la base conceptual para la propuesta.

En segundo lugar, se llevará a cabo la construcción de la matriz de relaciones, que constituye el componente central del modelo.

Esta matriz se formulará mapeando riesgos asociados a la rotación con prácticas relevantes de gestión de procesos y con atributos de calidad susceptibles de afectarse. El objetivo es generar un instrumento estructurado que permita visualizar vínculos causales y orientar estrategias de mitigación.

Finalmente, se contempla una validación por expertos, aún en etapa planificada. Esta consistirá en recabar la opinión de especialistas en calidad de software, modelos de procesos y gestión del conocimiento, quienes evaluarán la coherencia del modelo, la pertinencia de las relaciones propuestas y la utilidad práctica de la matriz. Esta etapa permitirá refinar y fortalecer el modelo antes de su futura aplicación en estudios empíricos o en entornos organizacionales reales.

3. RESULTADOS OBTENIDOS Y OBJETIVOS

El desarrollo de este trabajo abre varias líneas de acción que permitirán consolidar y fortalecer la propuesta. En primer lugar, será necesario finalizar el modelo conceptual, ampliando la matriz de relaciones con un conjunto más completo de riesgos, prácticas de proceso y atributos de calidad. Esto incluye refinar definiciones, ajustar categorías y profundizar en los vínculos causales identificados preliminarmente.

En segundo lugar, se llevará a cabo la validación con expertos, un paso crítico para evaluar la solidez y aplicabilidad del modelo. Esta validación permitirá obtener retroalimentación sobre la pertinencia de las relaciones propuestas, la claridad del enfoque y el potencial de uso en contextos reales. La opinión de especialistas en calidad del software, modelos de madurez y gestión del conocimiento será fundamental para ajustar la propuesta antes de su implementación práctica.

Finalmente, se proyecta la aplicación del modelo en un caso de estudio, idealmente en un entorno organizacional que experimente niveles de rotación representativos. Este ejercicio permitirá observar cómo la matriz orienta la identificación de riesgos, qué tan efectivas resultan las estrategias de mitigación seleccionadas y qué ajustes adicionales son necesarios para mejorar su utilidad operativa. Esto constituirá la base para la evolución hacia un modelo más robusto y validado empíricamente.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del proyecto, donde se encuentra inserta esta línea de investigación, se están llevando a cabo diferentes tesis de grado y posgrado. Éstas últimas correspondientes a la Especialización en Ingeniería de Software, Maestría en Ingeniería de Software, Maestría en Calidad del Software y al Doctorado en Ingeniería Informática. En particular, la línea de investigación forma parte de una tesis de la Maestría en Calidad del Software.

BIBLIOGRAFÍA

[1] C. Institute, "CMMI for Development, Version 2.0. ," 2018.

[2] ISO, "ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models," ed, 2011.

[3] S. Delgado Llamas, V. G. López Torres, and L. R. Moreno Moreno, "Impacto del Clima Organizacional en la Rotación del Personal: Evidencia en Sector Desarrollo de Software," *Revista Internacional Administración & Finanzas,* vol. 10, pp. 19-61, 2017.

El Paradigma de la Computación Cuántica: Estrategias de Desarrollo, Seguridad y Formación ante la Próxima Revolución Tecnológica

Leonardo Corbalán, Pablo Thomas, Juan Fernández Sosa, Luciano Marrero, Verena Olsowy, Fernando Tesone, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

526 e/ 10 y 11 La Plata Buenos Aires

{corbalan, pthomas, jfernandez, lmarrero, volsowy, ftesone, ppesado}@lidi.info.unlp.edu.ar

RESUMEN

Esta línea de investigación aborda el estudio y desarrollo de la Computación Cuántica desde una perspectiva integral de Ingeniería de Software, enmarcada en el tránsito actual desde la era NISQ hacia la era de tolerancia a fallos temprana. El proyecto se centra en seis ejes estratégicos: metodologías para el ciclo de vida del software cuántico (QSE), seguridad y coexistencia de sistemas criptográficos híbridos y blockchain, educación superior en programación cuántica, algoritmos variacionales híbridos, arquitecturas para laboratorios remotos y la experimentación directa con hardware de Resonancia Magnética Nuclear (RMN).

A través de la colaboración interuniversitaria y la integración de hardware real (como el procesador SpinQ Triangulum), se busca no solo la producción científica de impacto —con resultados ya publicados en foros internacionales como Springer CCIS—, sino también la formación de recursos humanos capaces de enfrentar la "amenaza" cuántica y de aprovechar las oportunidades de la utilidad cuántica en optimización y simulación. Esta línea de investigación combina el rigor teórico con la validación experimental, para desarrollar las competencias necesarias y hacer frente a los desafíos que plantea la próxima revolución computacional.

Palabras claves: Computación Cuántica, Ingeniería de Software Cuántico, Era NISQ, Tolerancia a Fallos, Criptografía Post-Cuántica, Educación en Informática.

CONTEXTO

Esta línea de Investigación forma parte del Proyecto "*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital*" del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa y Latinoamérica.

1. INTRODUCCIÓN

En 1980, Paul Benioff presentó el primer modelo cuántico de una máquina de Turing, sentando las bases teóricas para la computación cuántica al mostrar que la computación podía realizarse bajo las leyes de la mecánica cuántica [1]. Poco después, en 1982, Richard Feynman propuso explícitamente utilizar principios cuánticos para construir computadoras que simularan eficientemente problemas de física cuántica, superando las limitaciones de las computadoras clásicas [2].

La década de 1990 marcó un punto de inflexión en la computación cuántica con el surgimiento de algoritmos que demostraron ventajas teóricas cruciales frente a sus contrapartes clásicas. Si bien el algoritmo de Deutsch (1985) [3] estableció los cimientos al determinar si una función booleana es constante o balanceada con una única consulta, fue en los años subsiguientes donde se

consolidó el verdadero potencial del campo. Un avance fundamental fue la generalización propuesta por Deutsch y Jozsa en 1992 [4], la cual evidenció la potencia del paralelismo cuántico al resolver el problema para **n** bits con una sola consulta, frente a las **2^(n-1) + 1** necesarias en el peor caso clásico.

Posteriormente, 1994 se convirtió en un año decisivo. El algoritmo de Simon [5,6] ofreció una solución con ventaja exponencial para el problema del subgrupo oculto, sirviendo de inspiración conceptual directa para el hito más disruptivo de la década: el algoritmo de Shor [7,8]. Presentada ese mismo año, la propuesta de Shor revolucionó la criptografía al demostrar que la factorización de enteros y el logaritmo discreto pertenecen a la clase de complejidad BQP (tiempo polinomial cuántico). Esto evidenció una amenaza latente para la infraestructura de clave pública actual, cuya seguridad depende de la intratabilidad clásica de estos problemas.

Finalmente, en 1996, el algoritmo de Grover [9] introdujo una mejora cuadrática en la búsqueda de bases de datos no estructuradas, reduciendo la complejidad temporal de $O(N)$ a $O(sqrt(N))$. Este conjunto de desarrollos no sólo teorizó la supremacía cuántica en dominios específicos de álgebra y optimización, sino que impulsó la carrera global por el desarrollo de hardware cuántico funcional capaz de ejecutar estas rutinas.

En 1998 se logró un hito fundamental en la computación cuántica: la primera demostración experimental de un algoritmo cuántico en un computador funcional de 2 qubits, utilizando tecnología de resonancia magnética nuclear (RMN) en estado líquido. Las implementaciones pioneras incluyeron el algoritmo de Deutsch (Jones y Mosca, Oxford) [10] y una versión del algoritmo de Grover (Chuang, Gershenfeld y Kubinec, colaboración IBM/MIT/Berkeley) [11], marcando el inicio de la era experimental del campo.

A partir del siglo XXI, la competencia global por liderar el desarrollo de hardware cuántico ha trasladado el foco de las demostraciones teóricas a hitos prácticos de ingeniería. En 2019, Google anunció haber alcanzado la "supremacía cuántica" (*quantum supremacy*) con su procesador Sycamore [12], al resolver una tarea de muestreo de circuitos aleatorios en 200 segundos. Aunque este hito generó un intenso debate sobre la utilidad de tareas artificiales frente a algoritmos clásicos optimizados, marcó el inicio de una fase de experimentación intensiva a escala intermedia.

En la actualidad, el paradigma ha evolucionado desde la búsqueda de la supremacía hacia la utilidad cuántica [13]. Este cambio de enfoque busca extraer valor científico o comercial tangible de los dispositivos de la era NISQ (*Noisy Intermediate-Scale Quantum*) [14], mientras simultáneamente se acelera la transición hacia la era de tolerancia a fallos temprana (*Early Fault-Tolerant*). El desafío central hoy es la implementación de la corrección de errores cuánticos (QEC) mediante la creación de qubits lógicos [15], los cuales permiten preservar la información frente al ruido ambiental y son el pilar necesario para alcanzar algoritmos de gran escala.

Líderes del sector como IBM, con su hoja de ruta hacia sistemas modulares, y Google Quantum AI, con los avances en su procesador Willow, han demostrado recientemente que es posible reducir las tasas de error por debajo del umbral crítico. Estos avances señalan una transición fundamental: el paso de procesadores ruidosos a la era de tolerancia a fallos temprana (*Early Fault-Tolerant*) [16], donde aplicaciones en química, optimización y criptografía comienzan a mostrar un valor científico tangible.

En este contexto de maduración tecnológica y transición hacia sistemas robustos, se proponen las siguientes líneas de trabajo:

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Ingeniería de Software Cuántico (Quantum Software Engineering - QSE): Definición y validación de metodologías para el ciclo de vida del software cuántico. Esto incluye el estudio de técnicas de *testing* y *debugging* de

circuitos sobre simuladores, el uso de aserciones cuánticas, así como el análisis comparativo (benchmarking) de procesos de transpilación y optimización de circuitos.

Seguridad y Criptografía en la Era Post-Cuántica: Estudio del impacto de la computación cuántica en la seguridad de sistemas distribuidos y Blockchain Desarrollo de marcos conceptuales y arquitecturas de transición para la coexistencia de criptografía clásica, post-cuántica (PQC) y cuántica (QKD), analizando la interoperabilidad y el rendimiento en redes híbridas mediante simulación.

Educación en Computación Cuántica: Investigación sobre estrategias pedagógicas y curvas de aprendizaje para la enseñanza de la programación cuántica en carreras de Informática. Evaluación de herramientas de visualización, lenguajes de alto nivel y diseño de material didáctico interactivo (Cuadernos Jupyter) para la formación de recursos humanos en el área.

Algoritmos Híbridos Clásicos-Cuánticos: Investigación en la orquestación de sistemas que combinan procesamiento clásico (CPU/GPU) con unidades de procesamiento cuántico (QPU) simuladas. Exploración de algoritmos variacionales (VQE, QAOA).

Arquitecturas para Laboratorios Cuánticos Remotos: Diseño de arquitecturas de software para la gestión y exposición de recursos cuánticos como servicio. Esto abarca el desarrollo de *middleware* para la gestión de colas, APIs de control y la integración de hardware real (procesador SpinQ Triangulum) con plataformas de educación web.

Caracterización y Mitigación de Errores en Hardware NISQ: Experimentación práctica con procesadores cuánticos de escritorio basados en Resonancia Magnética Nuclear (RMN). Implementación y análisis de códigos de corrección de errores cuánticos (como el código de repetición de *bit-flip*) y técnicas de mitigación de ruido (*Zero Noise Extrapolation*) en entornos ruidosos reales.

3. RESULTADOS OBTENIDOS/ESPERADOS

Estado del Arte en Herramientas de Desarrollo: Se ha realizado un relevamiento exhaustivo y clasificación taxonómica de los lenguajes de programación, SDKs y simuladores cuánticos actuales. Este trabajo ha permitido identificar las herramientas más aptas para la inserción de la temática en la currícula académica de grado y posgrado. Este resultado se materializó inicialmente en el *Congreso CACIC 2024* y, tras una selección, su versión extendida fue publicada en la serie global *Communications in Computer and Information Science* de Springer: "*Tools for Quantum Programming: A Review*" [17].

Marco de Transición para Blockchain Cuántica: Se ha desarrollado un marco conceptual integral para abordar la amenaza cuántica en sistemas Blockchain. El modelo propone una estrategia de tres fases para la migración desde criptografía clásica hacia esquemas post-cuánticos y nativos cuánticos, estableciendo protocolos de interoperabilidad y gobernanza. Este resultado se materializó en la publicación: *"Desarrollo de un Marco Conceptual para la Transición y Coexistencia Segura entre Sistemas Blockchain Clásicos, Post-Cuánticos y Cuánticos"* (CACIC 2025) [18].

Se espera alcanzar resultados en:

Capacitación de Recursos Humanos: Consolidación de un núcleo de investigación con formación teórica en álgebra lineal aplicada a computación cuántica, formalismo de Dirac y programación en Qiskit.

Metodologías de Verificación y Validación para Circuitos Cuánticos: Investigación y propuesta de estrategias de *testing* de software cuántico, enfocándose en el uso de aserciones cuánticas y testing metamórfico para la depuración de algoritmos sobre simuladores.

Evaluación de Rendimiento de Criptografía Post-Cuántica: Simulación y medición del impacto en el rendimiento (latencia y tamaño de bloque) al integrar firmas digitales post-cuánticas (como CRYSTALS-Dilithium) en entornos de Blockchain emulados.

Simulación de Protocolos de Comunicación Cuántica: Modelado y análisis del protocolo BB84 bajo distintas condiciones de ruido simulado para determinar umbrales de seguridad en la distribución de claves (QKD).

Producción de Material Didáctico: Generación de una suite de *Cuadernos Jupyter* y guías de laboratorio interactivas orientadas a la enseñanza de la programación cuántica en carreras de Informática.

Validación Experimental en Hardware RMN: Ejecución comparativa de algoritmos fundamentales en hardware real vs. simuladores para caracterizar el ruido físico y la decoherencia en tecnología de Resonancia Magnética Nuclear.

Despliegue de Infraestructura Experimental: Puesta en marcha y calibración operativa del computador cuántico SpinQ Triangulum, integrándolo como un recurso de "Backend Real".

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con este proyecto. Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

5. BIBLIOGRAFÍA

1. Benioff, P. (1980). The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of statistical physics*, *22*(5), 563-591.

2. Feynman, R. P. (1982). Simulating physics with computers. International Journal of Theoretical Physics, 21(6/7), 467–488. DOI: https://doi.org/10.1007/BF02650179

3. Deutsch, D. (1985). Quantum theory, the Church–Turing principle and the universal quantum computer. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 400(1818), 97–117. https://doi.org/10.1098/rspa.1985.0070

4. Deutsch, D., & Jozsa, R. (1992). Rapid solution of problems by quantum computation. Proceedings of the Royal Society of London. Series A, 439(1907), 553–558. DOI: 10.1098/rspa.1992.0167

5. Simon, D. R. (1994). On the power of quantum computation. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science (pp. 116–123). IEEE Computer Society. https://doi.org/10.1109/SFCS.1994.365701

6. Simon, D. R. (1997). On the power of quantum computation. SIAM Journal on Computing, 26(5), 1474–1483. https://doi.org/10.1137/S0097539796298637

7. Shor, P. W. (1994). Algorithms for quantum computation: Discrete logarithms and factoring. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science (pp. 124–134). IEEE Computer Society. https://doi.org/10.1109/SFCS.1994.365700

8. Shor, P. W. (1997). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Journal on Computing, 26(5), 1484–1509. https://doi.org/10.1137/S0097539795293172

9. Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In Proceedings of the twenty-eighth annual ACM symposium on Theory of computing (pp. 212–219). Association for Computing Machinery. https://doi.org/10.1145/237814.237866

10. Jones, J. A., & Mosca, M. (1998). Implementation of a quantum algorithm on a nuclear magnetic resonance quantum computer. The Journal of Chemical Physics, 109(5), 1648–1653. https://doi.org/10.1063/1.476739

11. Chuang, I. L., Gershenfeld, N., & Kubinec, M. (1998). Experimental implementation of fast quantum searching. Physical

Review Letters, 80(15), 3408–3411. https://doi.org/10.1103/PhysRevLett.80.3408

12. Arute, F. et al. (2019). Quantum supremacy using a programmable superconducting processor. Nature, 574(7779), 505–510. https://doi.org/10.1038/s41586-019-1666-5

13. Kim, Y. et al. (2023). Evidence for the utility of quantum computing before fault tolerance. Nature, 618(7965), 500–505. https://doi.org/10.1038/s41586-023-06096-3

14. Preskill, J. (2018). Quantum computing in the NISQ era and beyond. Quantum, 2, Article 79. https://doi.org/10.22331/q-2018-08-06-79 (Definición clave de la era NISQ: dispositivos ruidosos de escala intermedia, sin tolerancia a fallos completa).

15. Acharya, R. et al. (2023). Suppressing quantum errors by scaling a surface code logical qubit. Nature, 614(7949), 676–681. https://doi.org/10.1038/s41586-022-05434-1

16. Bluvstein, D. et al. (2024). Logical quantum processor based on reconfigurable atom arrays. Nature, 626(7997), 58–65. https://doi.org/10.1038/s41586-023-06927-3

17. Marrero, L.; Olsowy, V.; Fernández Sosa, J.; Tesone, F.; Corbalán, L.; Thomas, P. J. & Pesado, P. (2024). "Una revisión de herramientas para programación cuántica". En *XXX Congreso Argentino de Ciencias de la Computación (CACIC)*, La Plata (pp. 802-813). Versión extendida seleccionada y publicada internacionalmente como: "Tools for Quantum Programming: A Review". En: Pesado P., Thomas P. (eds) *Computer Science – CACIC 2024*. Communications in Computer and Information Science (CCIS), vol. 2520, pp. 197–211. Springer, Cham.

18. Corbalán, L.; Fernández Sosa, J.; Marrero, L.; Olsowy, V.; Thomas, P. J. & Pesado, P. (2025). "Desarrollo de un Marco Conceptual para la Transición y Coexistencia Segura entre Sistemas Blockchain Clásicos, Post-Cuánticos y Cuánticos". En *XXXI Congreso Argentino de Ciencias de la Computación (CACIC)*. Universidad Nacional de Río Negro (Sede Atlántica), 6 al 10 de octubre.

El uso de NLP e IA para mejorar la calidad y extraer información de especificaciones en lenguaje natural

Marcela Samela[1], Gustavo Eduardo Gonzalez[1], Mariana Otero[1], Lovena Laurent[1], Guillermo Omar Pizarro Vasquez[12], Adrian Federico Gil[1], Dario Cardacci[1], Leandro Antonelli[13]

[1] CAETI, Facultad de Tecnología Informática, Universidad Abierta Interamericana, CABA, Argentina. Universidad Politécnica Salesiana, Ecuador.
[3]Lifia, Facultad. de Informática, Universidad Nacional de La Plata, Calle 50 y 120 - La Plata - Bs. As. – Argentina.

{marcela.samela, gustavoernesto.gonzales, mariana.otero, lovena.laurent, guillermoomar.pizarrovasquez, adrianfederico.gil, dario.cardacci, leandro.antonelli}@uai.edu.ar

RESUMEN

Se conoce como revolución industrial al proceso en el cual con la invención de la máquina a vapor se lograron automatizar procesos en los que intervenía la fuerza humana, logrando resultados más rápidos y económicos. Algunos autores identifican 4 revoluciones industriales, en donde la tercera y cuarta etapa dependen fuertemente de los sistemas informáticos. Lo cierto es que actualmente los sistemas informáticos están en todos los aspectos de nuestra vida, y la cantidad de información que se produce y consume es muy grande. Hay información que está codificada y que sólo entienden los sistemas informáticos, pero también hay mucha información escrita en lenguaje natural que generan y consumen los humanos. Con el fin de optimizar el tratamiento de esta información en lenguaje natural, este proyecto tiene por objetivo desarrollar y evaluar herramientas basadas en Procesamiento de Lenguaje Natural e Inteligencia Artificial para el tratamiento de especificaciones técnicas y no técnicas con el fin de mejorar la calidad, extraer información y consolidar el conocimiento.

Palabras clave: Lenguaje Natural, especificaciones, procesamiento de lenguaje natural, inteligencia artificial.

CONTEXTO

El proyecto de I/D que se presenta en este artículo lo realiza un grupo de docentes e investigadores de la Universidad Abierta Interamericana vinculados a un proyecto de investigación del centro CAETI de la misma Universidad.

1. INTRODUCCIÓN

Los sistemas informáticos tienen varias décadas de funcionamiento y en su evolución se ha consolidado mucha información en forma de bases de datos. Sin embargo, hoy día hay muchísima información que se encuentra en documentos escritos en lenguaje natural. Incluso, las aplicaciones han hecho que los documentos proliferen aún más debido a las facilidades de edición (por ejemplo, al tomar documentos previos y editarlos para obtener nuevos documentos). Es por eso que la cantidad de información que existe hoy día en documentos digitales escritos en lenguaje natural es muy grande y muy variada. Hay documentos técnicos como por ejemplo escritos jurídicos en muchas formas (presentaciones, informes, pericias, fallos, dictámenes y leyes entre otras). O también documentos médicos (diagnósticos, tratamientos, informes e historias clínicas, entre otros). Más aún, hay una cantidad muy variada de documentos no técnicos: noticias, reseñas de películas y productos, conversaciones en mensajería instantánea y en redes sociales.

En los últimos años, el procesamiento de lenguaje natural (NLP por sus siglas en inglés), el cual es una disciplina dentro de la Inteligencia Artificial (AI por sus siglas en

inglés) ha logrado avances muy grandes [1] [2] [3]. Han aparecido muchas herramientas que brindan soluciones como: extracción de texto (text minining) [4], agrupamiento de documentos (clustering), comparación de similaridad (a través de una cantidad muy grande de técnicas), y análisis de sentimiento.

Estas herramientas permiten automatizar (es decir, proveen mecanismos más eficientes y con menor probabilidad de error) ciertas tareas comunes que realizan los seres humanos. Por ejemplo, las herramientas de NLP e IA permiten analizar el estilo de redacción y sugerir mejoras en un documento [5]. Sea un fallo judicial en donde se debe indicar explícitamente que deben realizar el actor y el demandado, es necesario que el fallo indique explícitamente el sujeto (actor o demandado) y las acciones que deben realizar. Es posible implementar herramientas basadas en NLP e IA para que identifique si las oraciones son completas (con sujeto gramatical y predicado conteniendo una acción). Además, es posible identificar si las oraciones están escritas en voz activa (es decir, si el sujeto gramatical de la oración, es realmente el sujeto que realiza la acción).

Las herramientas de NLP también permiten extraer información [6] [7]. Considerando textos escritos de una forma homogénea (a partir de técnicas del párrafo previo), es posible identificar el sujeto de la oración, el núcleo del predicado, y el objeto directo para establecer que cierto sujeto, realiza cierta acción, sobre cierto objeto.

Las herramientas de NLP e IA también permiten identificar tópicos, agrupar documentos e identificar cuan similares son los documentos [8]. Considerando el ejemplo jurídico, es muy importante buscar antecedentes sobre cierto tema, ya sea para diseñar la estrategia de una defensa, preparar una acusación o dictar un fallo.

Si bien los ejemplos previos son propios del ámbito jurídico, las mismas situaciones se presentan en otros ámbitos (dominios). En la ingeniería de requerimientos y en la administración de proyectos son problemas recurrentes. También lo son en ámbitos médicos, de arquitectura e ingeniería. Y también lo son en ámbitos donde las especificaciones (los documentos) no son tan técnicos ni estructurados, como puede ser en periodismo.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La estrategia general para llevar a cabo este proyecto es un método iterativo e incremental, el cual comienza con un relevamiento del estado del arte, para seleccionar una o varias técnicas, y proponer una contribución al estado del arte en algún área o dominio específico. Se buscara experimentar en el mismo con el fin de validar la contribución y luego evaluar y analizar la posibilidad de llevar la contribución a un dominio diferente, con la eventual generalización de la contribución.

Básicamente, el proyecto buscará analizar estilos de redacción con el fin de sugerir mejoras a la vez de estandarizar la estructura. De esta forma, será más sencillo el comparar distintos textos para establecer similaridades. Con esto se podrán agrupar textos relacionados y poder extraer información de los mismos. Todas estas acciones apuntan a consolidar la información permitiendo que sea procesada de una forma más eficiente.

3. RESULTADOS ESPERADOS

El objetivo general del proyecto es el de diseñar, implementar y testear herramientas basadas en NLP e IA para el tratamiento de documentos (especificaciones) técnicas y no técnicas, con el fin de mejorar la calidad y consolidar la información que proveen.

Los objetivos específicos se basan en explorar distintos matices: (i) analizar la calidad de las especificaciones con el fin de sugerir mejoras, (ii) agrupar y categorizar documentos, (iii) comprar nivel de similaridad, (iv) identificar conceptos relevantes y sus relaciones, y (v) consolidar la información extraía con el fin de darle un tratamiento automatizado.

4. FORMACIÓN DE RECURSOS HUMANOS

Se espera formar a jóvenes docentes investigadores, introducir a recientes graduados en el mundo académico de la investigación y finalmente a brindar a estudiantes avanzados la posibilidad de que un marco para llevar un trabajo final de carrera.

5. REFERENCIAS

[1] Kulkarni, A., Shivananda, A.: Natural Language Processing Recipes: Unlocking Text Data with Machine Learning and Deep Learning using Python. Apress; 1st ed. edición (1 Enero 2019).

[2] Zhao, L., Alhoshan, W., Ferrari, A., Letsholo, K. J., Muideen A. A., Chioasca, E. V., Batista-Navarro, R.T., Natural Language Processing for Requirements Engineering: A Systematic Mapping Study, ACM Comput. Surv. 54(3): 55:1-55:41 (2022)

[3] Ferrari, A., Zhao, L., Alhoshan, W.: NLP for Requirements Engineering: Tasks, Techniques, Tools, and Technologies ICSE (Companion Volume) pp 322-323 (2021)

[4] Charu C. Aggarwal, C., Xiang Z.: Mining Text Data, Springer (2012)

[5] Antonelli, L., Fernandez, A., Ruffolo, N., Sansone, E., Torres, D., "A Collaborative Approach to specify Kernel Sentences using Natural Language", Workshop in Requirements Engineering (WER), Natal, Brasil, Agosto 23 – 26 (2022).

[6] Antonelli, L. Delle Ville, J., Adorno, M. A., Ballestero, L. P., Cecconato, S. A., Fernandez, A., Maclen, G., Maltempo, G., Mattei, J.E., Tanevitch, L., Torres, D., "An approach to extract a conceptual model from natural language specifications", Workshop in Requirements Engineering (WER), Porto Alegre, Brasil, Agosto 15 – 17 (2023).

[7] Antonelli, L., Lezoche, M., Delle Ville, J., "A Method to obtain a Knowledge Representation from a Natural Language Specification of the Domain using the Glossary LEL" Decisioning 2023, Popayan, Colombia (2023).

[8] Delle Ville, J., Torres, D., Fernández, A., Antonelli, L., "An Approach to Cluster Scenarios According to their Similarity using Natural Language Processing", IX Jornadas Iberoamericanas de Interacción Humano – Computadora (JIHCI 2023), 13 al 15 de Septiembre, Universidad de La Matanza, Argentina, (2023)

Elicitación asistida por IA en Ingeniería de Requisitos

María Inés Bertín[1], Gerardo Riera[1], Gladys Kaplan[1], Gabriel Blanco[1]

[1]Departamento de Ingeniería e Investigaciones Tecnológicas,
Universidad Nacional de La Matanza. San Justo, Buenos Aires
{mbertin,griera,gkaplan,gblanco}@unlam.edu.ar

RESUMEN

La elicitación en Ingeniería de Requisitos constituye una actividad crítica, ya que la calidad de los artefactos depende directamente de la información capturada en las etapas tempranas. En el Proceso de Requisitos Basado en Escenarios, la construcción del Léxico Extendido del Lenguaje (LEL) se basa en el análisis de desgrabaciones de entrevistas iniciales no estructuradas, lo que implica un esfuerzo cognitivo elevado y una fuerte dependencia de la experiencia del ingeniero de requisitos.

Estas entrevistas suelen generar documentos incompletos o inconsistentes que impactan negativamente en la calidad de los modelos derivados. Este trabajo analiza el efecto de mejorar el documento de elicitación mediante asistencia basada en inteligencia artificial. Se propone un proceso iterativo compuesto por una entrevista abierta inicial, su análisis semántico mediante IA para detectar defectos y vacíos informacionales, la generación de una guía estructurada de profundización y una segunda entrevista validada e integrada por el ingeniero de requisitos.

Resultados preliminares en dos casos de estudio muestran que la asistencia en elicitación reduce vacíos e inconsistencias del insumo, y que esto se refleja en un LEL con mayor completitud, coherencia y precisión semántica.

Palabras clave: Ingeniería de Requisitos; elicitación; Léxico Extendido del Lenguaje; Defectos de elicitación; Modelos de lenguaje.

CONTEXTO

Este trabajo se enmarca dentro del proyecto "Impacto de la inteligencia artificial en la Ingeniería de Requisitos" como parte de la línea de investigación "Ingeniería de requisitos: Un enfoque a los negocios" que se realiza en la Universidad Nacional de La Matanza en el Departamento de Ingeniería e Investigaciones Tecnológicas.

1. INTRODUCCIÓN

La Ingeniería de Requisitos [1] constituye una disciplina crítica dentro del desarrollo de software, ya que los defectos introducidos en las etapas tempranas del proceso tienden a propagarse y amplificarse a lo largo del ciclo de vida del sistema. Aunque muchos de estos defectos pueden ser detectados en fases posteriores, el costo asociado al retrabajo es elevado y, en determinados casos, los errores no identificados repercuten directamente en la calidad del producto final. Diversos estudios han analizado la relevancia de la etapa de elicitación y sus desafíos metodológicos [2][3][4], particularmente cuando la información obtenida resulta incompleta, ambigua o inconsistente.

En el marco del Proceso de Requisitos Basado en Escenarios [5], la construcción del Léxico Extendido del Lenguaje (LEL) constituye el primer artefacto formal que organiza el conocimiento del dominio a partir del discurso de los stakeholders. Este artefacto cumple un rol estructural, ya que establece el soporte semántico para la construcción de modelos posteriores. Sin embargo, la calidad del LEL depende directamente de la calidad de los documentos de elicitación que le sirven de insumo. Cuando en las entrevistas aparecen DEO (Defectos, Errores y Omisiones), entendidos como inconsistencias, vacíos informacionales, ambigüedades y fallas de precisión presentes en el discurso del cliente, estos afectan la calidad del documento de

elicitación [6], que constituye el insumo del proceso. En consecuencia, dichos defectos tienden a trasladarse al LEL y a los modelos derivados, comprometiendo su coherencia conceptual y su utilidad analítica [7].

En este sentido, el problema no radica únicamente en la calidad del artefacto construido, sino en la calidad de la información fuente que alimenta el proceso. Tradicionalmente, la depuración y estructuración de dicha información depende en gran medida de la experiencia del ingeniero de requisitos, lo que introduce variabilidad y limita la reproducibilidad metodológica. En consecuencia, resulta necesario explorar mecanismos sistemáticos que permitan fortalecer la calidad del documento de elicitación antes de la construcción formal de modelos.

El reciente avance de los modelos de lenguaje de gran escala [8] ha demostrado capacidades relevantes en el análisis y procesamiento de texto en lenguaje natural, habilitando nuevas posibilidades para asistir tareas cognitivamente demandantes. En lugar de concebir a la Inteligencia Artificial como un sustituto del ingeniero de requisitos o como un generador automático de artefactos, este trabajo la posiciona como una herramienta de apoyo orientada a mejorar la calidad de la información elicitada.

En este contexto, se propone evaluar el uso de Inteligencia Artificial para analizar desgrabaciones de entrevistas iniciales, identificar posibles DEO y generar un conjunto estructurado de preguntas orientadas a una segunda instancia de elicitación.

Objetivo y alcance. El objetivo de este trabajo es asistir sistemáticamente la elicitación mediante el análisis de entrevistas y la generación de guías de profundización orientadas a reducir DEO en el documento fuente. La mejora del LEL se analiza como efecto derivado de disponer de un insumo de elicitación más completo y consistente, y no como el foco de automatización del modelado.

Este trabajo realiza las siguientes contribuciones:

- Propone un proceso de elicitación asistido por IA para reducir DEO en entrevistas abiertas.
- Introduce un esquema de complementariedad Humano–IA orientado a la validación e integración del insumo.
- Presenta evidencia preliminar de que la mejora del documento de elicitación se refleja en un LEL más completo y consistente.
- Discute indicadores iniciales para observar el impacto (DEO y cantidad de símbolos LEL) y sienta bases para la formalización futura de métricas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El LEL se construye a partir del análisis manual de entrevistas no estructuradas realizadas a los stakeholders. Si bien este tipo de entrevistas favorece la comprensión contextual del dominio, la experiencia profesional muestra que esta etapa presenta dificultades recurrentes que impactan directamente en la calidad de los artefactos resultantes.

El entrevistado suele describir procesos desde su perspectiva operativa inmediata, enfatizando aquello que considera problemático o central, mientras otros aspectos permanecen implícitos o ausentes.

Entre los problemas más frecuentes se encuentran:

- Omisión de estados formales del proceso.
- Reglas de negocio no explicitadas.
- Criterios de validación entre actividades o módulos insuficientemente definidos.
- Manejo de excepciones o situaciones no habituales no documentado.
- Responsabilidades organizacionales ambiguas o incompletas.
- Objetivos estratégicos del sistema no claramente formulados.

Asimismo, el análisis manual depende fuertemente de la experiencia del ingeniero de requisitos. Diferentes analistas pueden advertir, o no, determinados problemas a partir del mismo material de entrada, lo que introduce variabilidad en el proceso y limita su

reproducibilidad metodológica. Frente a este escenario, surge la necesidad de incorporar mecanismos que permitan fortalecer la calidad del documento de elicitación antes de proceder a la formalización del dominio.

Uso de Inteligencia Artificial como herramienta de asistencia:

En este proyecto se propone utilizar un modelo de Inteligencia Artificial, específicamente ChatGPT de OpenAI, como herramienta de apoyo durante la etapa de análisis de entrevistas abiertas. El enfoque adoptado no se orienta a la generación automática del LEL, sino a mejorar la calidad del insumo previo al modelado.
En particular, se explora la capacidad de la herramienta para:

1. Analizar desgrabaciones completas de entrevistas.
2. Detectar lagunas informacionales, ambigüedades y posibles inconsistencias.
3. Identificar áreas conceptuales insuficientemente desarrolladas (actores, procesos, estados, reglas, excepciones, integraciones).
4. Organizar los DEO detectados en categorías relevantes para Ingeniería de Requisitos.
5. Generar una guía estructurada de preguntas orientada a una entrevista de profundización.

De este modo, la IA se posiciona como un asistente metodológico que actúa como instancia de revisión semántica del discurso inicial. Su función no es reemplazar el juicio del ingeniero de requisitos ni automatizar la construcción del modelo, sino ampliar la capacidad de detección temprana de omisiones y sistematizar la planificación de una segunda instancia de elicitación.

Metodología:

La investigación se desarrolló sobre dos casos de estudio pertenecientes a diferentes contextos organizacionales: uno correspondiente al ámbito industrial (Petroquímica) y otro vinculado al sector de servicios técnicos (Remitos en línea para Servicios de Autoelevadores).

En ambos casos se partió de desgrabaciones completas de entrevistas abiertas realizadas a actores clave del dominio. El procedimiento adoptado consistió en:

- Someter la desgrabación al análisis mediante un modelo de lenguaje, solicitando explícitamente la identificación de lagunas de información, omisiones y ambigüedades.
- Clasificar los vacíos detectados en categorías conceptuales alineadas con la Ingeniería de Requisitos (procesos, actores, estados, reglas de negocio, integraciones, trazabilidad, entre otras).
- Generar, a partir de dichas categorías, un plan de entrevista estructurada de duración acotada (aproximadamente una hora), orientado a profundizar los aspectos incompletos.
- Evaluar cualitativamente la coherencia, cobertura y pertinencia del plan generado.

El alcance del presente trabajo se centra en evaluar la capacidad de la IA para asistir sistemáticamente en la detección temprana de defectos de elicitación y en la planificación estructurada de una instancia de profundización.
El proceso propuesto se estructura en seis etapas: (1) realización de una entrevista abierta inicial; (2) análisis semántico de la desgrabación mediante IA; (3) identificación preliminar de defectos en el documento de elicitación; (4) generación de una guía estructurada de profundización; (5) realización de una segunda entrevista focalizada; y (6) integración de la nueva información por parte del ingeniero de requisitos, quien valida, consolida y refina el documento de elicitación original.
La IA interviene exclusivamente como herramienta de apoyo en el análisis preliminar y en la generación de la guía de profundización, mientras que la validación contextual y la integración final permanecen bajo responsabilidad del ingeniero de requisitos.

3. RESULTADOS OBTENIDOS / ESPERADOS

Durante el análisis comparativo se observó una diferencia significativa entre ambos casos: la entrevista correspondiente al caso 1 (Petroquímica) presentaba mayor nivel de detalle, explicitación de procesos y definición de actores, lo que derivó en un LEL inicial con mayor cantidad de símbolos identificados, como se puede ver en la Tabla 1. También se observó un mayor detalle en la definición de cada símbolo. En contraste, la entrevista del caso 2 ("Remitos en línea para Servicios de Autoelevadores") resultó más general y con muchas omisiones, lo que se reflejó en un LEL inicial con menor cobertura de léxico del dominio y mayor necesidad de refinamiento posterior.

Caso	Símbolos LEL	DEO detectados
Petroquímica	35	15
Remitos en línea	26	23

Tabla 1. Indicadores observables: entrevista (DEO) y LEL resultante

A partir de esta observación, se procedió a aplicar el modelo de lenguaje para analizar las desgrabaciones iniciales de ambos casos, solicitando específicamente la identificación de posibles DEO en la información relevada, la cual arrojó los resultados que se evidencian en la **Tabla 2**.

Tipo de DEO	Petroquímica	Remitos en línea	Total
Inconsistencias	3	6	**9**
Omisiones	7	10	**17**
Errores conceptuales	3	4	**7**
Fallas de precisión	2	3	**5**
Total DEO detectados	**15**	**23**	**38**

Tabla 2. DEO identificados en el insumo de elicitación (entrevista inicial)

Las preguntas generadas fueron revisadas por el equipo de investigación y utilizadas como base para realizar una segunda entrevista en cada caso. Posteriormente, se reconstruyó el LEL considerando la información adicional obtenida. En ambos casos se verificó una mejora en la calidad del léxico resultante, evidenciada en una mayor cantidad de símbolos correctamente definidos (**Figura 1**), mejor delimitación de responsabilidades y reducción de ambigüedades detectadas en la versión inicial.

Los resultados también evidencian que la intervención del ingeniero de requisitos continúa siendo indispensable. Mientras la IA opera sobre patrones lingüísticos, coherencia textual y relaciones semánticas explícitas en la desgrabación, el análisis humano incorpora interpretación contextual, comprensión situacional del dominio y valoración de matices implícitos en el discurso del stakeholder. La identificación de DEO no constituye un proceso puramente mecánico, sino que requiere juicio profesional para validar la pertinencia de las observaciones realizadas por la herramienta. En consecuencia, más que un reemplazo, se configura un esquema de complementariedad en el que la IA asiste el análisis preliminar y el juicio experto garantiza la relevancia, coherencia y adecuación del proceso de elicitación mejorado.

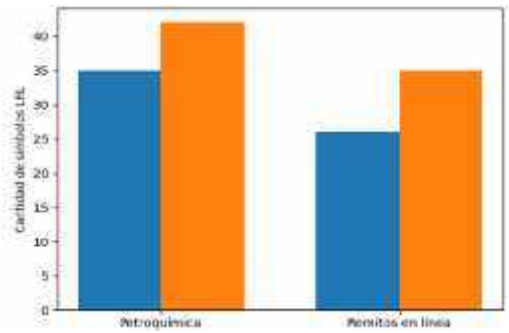


Fig. 1 – Incremento de símbolos del LEL después de mejorar la entrevista

Como resultado esperado en etapas futuras, se propone formalizar métricas que permitan cuantificar el impacto de la segunda entrevista en términos de incremento de símbolos, reducción de ambigüedades y mejora en la consistencia semántica del LEL, con el fin de validar de manera más sistemática la hipótesis de mejora iterativa asistida por IA.

Estos resultados sugieren que la mejora sistemática del documento de elicitación constituye un punto de apalancamiento clave

para aumentar la calidad semántica de los artefactos tempranos del proceso de requisitos. Los DEO reportados corresponden al documento de elicitación (entrevista desgrabada). La mejora observada en el LEL se interpreta como consecuencia de disponer de un insumo más completo y consistente, y no como resultado de automatizar la construcción del léxico.

3.1 Amenazas a la validez

Si bien los resultados preliminares muestran una mejora en la calidad del LEL tras la aplicación del proceso asistido por IA, el estudio presenta ciertas limitaciones que deben considerarse. En primer lugar, el análisis se realizó sobre dos casos de estudio, lo que restringe la generalización estadística de los resultados. No obstante, los casos fueron seleccionados deliberadamente por su heterogeneidad organizacional (industrial y servicios), buscando representatividad en términos de contexto y complejidad discursiva. Además del análisis cualitativo del plan de entrevista, se consideraron indicadores observables del LEL reconstruido, como cantidad de símbolos identificados, reducción de defectos y mayor precisión semántica, pero aún no se han formalizado métricas cuantitativas estandarizadas que permitan medir el impacto con mayor rigor experimental.

Finalmente, aunque el principio metodológico según el cual la calidad de los artefactos depende de la calidad de los insumos es ampliamente aceptado en Ingeniería de Software, la validación empírica sistemática de este proceso asistido requiere estudios adicionales con muestras más amplias y protocolos controlados. En este sentido, los resultados presentados deben interpretarse como evidencia preliminar que sustenta la viabilidad del enfoque, más que como una validación concluyente del enfoque propuesto.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación presentada es parte de la formación inicial en investigación de la Ingeniera María Inés Bertín y del Ingeniero Gerardo Agustín Riera.

5. BIBLIOGRAFÍA

[1] Sommerville, I. (2016). *Software Engineering* (10th ed.). Pearson.

[2] Ramesh, B., & Prithi, S. (2018). Requirements Engineering and Artificial Intelligence. *Journal of Systems and Software*, 137, 42-54.

[3] Kassab, M. (2019). The role of artificial intelligence in requirements engineering: A review of the state of the art. Requirements Engineering, 24(4), 1-19. https://doi.org/10.1007/s00766-019-00313-8

[4] Khlood Ahmad, Muneera Bano, Mohamed Abdelrazek, Chetan Arora, John Grundy (2021). *What's up with Requirements Engineering for Artificial Intelligence Systems?* IEEE 29th International Requirements Engineering Conference (RE), USA, 2021, pp. 1-12, doi: 10.1109/RE51729.2021.00008.

[5] Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., Hadad, G.D.S., Ridao, M.N.: *Defining system context using scenarios*. In: Cheng, B.H.C. et al. (eds.) *Perspectives on Software Requirements*, pp. 169–199. Kluwer Academic Publishers, Boston (2004)

[6] Berry, D. M., Kamsties, E., & Krieger, M. M. (2003). From contract drafting to software specification: Linguistic sources of ambiguity. Automated Software Engineering, 10(4), 299–322. https://doi.org/10.1023/A:1024486823503

[7] Nuseibeh, B., & Easterbrook, S. (2000). Requirements engineering: A roadmap. In *Proceedings of the Conference on the Future of Software Engineering (pp. 35–46). ACM.* https://doi.org/10.1145/336512.336523

[8] OpenAI. GPT-4 (2023). *GPT-4 Technical Report*.

Enseñanza de la Gestión de Proyectos con Metodología BIM en la Industria del Petróleo y Gas

Leandro Antonelli[1,3], Marcelo De Vincenzi[1], Marcelo Duarte dos Santos [2], Jorge Scali[1], Nicolás Battaglia[1], Claudia Pons[1,3], Susana Darin[1]

[1] CAETI, Facultad de Tecnología Informática, Universidad Abierta Interamericana.
[2] Advisor en Proyectos de Capital, Transformación Digital y Gobernanza
[3]Lifia, Facultad. de Informática, Universidad Nacional de La Plata.

leandro.antonell@uai.edu.ar, medevincenzi@uai.edu.ar, duartedossantosmarcelo@gmail.com, jorge.scali@uai.edu.ar, nicolas.battaglia@uai.edu.ar, claudia.pons@uai.edu.ar, susana.darin@uai.edu.ar

RESUMEN

Este artículo describe el proyecto de investigación llevado a cabo en el marco de la Universidad Abierta Interamericana con el fin de diseñar un curso de gestión de proyectos utilizando la metodología BIM (*Building Information Modeling)* para dirigir proyectos específicos de la industria del petróleo y gas. La dirección de proyectos tiene una larga trayectoria y se podría identificar un hito significativo al final de la década del 60 donde se formalizó la disciplina. A pesar de toda la historia, la disciplina tiene que seguir desarrollándose para adaptar sus métodos, técnicas y herramientas a las necesidades actuales. Particularmente, la industria de petróleo y gas, presenta ciertos desafíos provenientes de la criticidad de los productos que maneja, los materiales involucrados y el contexto en donde se realizan los proyectos. BIM es una metodología de trabajo colaborativa basada en modelos digitales inteligentes que integran geometría tridimensional (3D) con información estructurada del activo. Estos modelos permiten centralizar, gestionar y actualizar datos confiables a lo largo de todo el ciclo de vida, desde el diseño y la construcción hasta la operación, habilitando una toma de decisiones basada en evidencia, trazabilidad y enfoque predictivo. Este proyecto tiene por objetivo diseñar un curso de gestión de proyectos incorporando la metodología BIM para la industria de petróleo y gas.

Palabras clave: gestión de proyectos, seminario de actualización, Building information modeling.

CONTEXTO

El proyecto de I/D que se presenta en este artículo lo lleva a cabo un grupo de profesionales con experiencia en educación, tanto en docencia e investigación, como en la gestión educativa. Los participantes de este proyecto también poseen experiencia en la industria, ya sea en la disciplina de gestión de proyectos como en el sector del petróleo y gas. Se busca diseñar un curso para impartirlo en el marco de la Universidad Abierta Interamericana.

1. INTRODUCCIÓN

Un proyecto es un esfuerzo temporal emprendido para crear un producto, brindar un servicio o proveer un resultado único. Se caracteriza por tener un inicio y un final bien definidos, objetivos específicos, recursos limitados y restricciones de alcance, tiempo y costo [1].

La disciplina de gestión de proyectos es la que se encarga de planificar, ejecutar y controlar el desarrollo de cualquier proyecto. Dada las limitaciones y restricciones de las variables que incluso se interrelacionan compitiendo entre ellas, es muy difícil cumplir con objetivos, tiempos y costos. La historia de proyectos que fracasan es muy grande [2]. Y se sostiene que "un proyecto bien gestionado puede fracasar. Mientras que un proyecto mal o no gestionado seguro fracasará" [3].

La historia y evolución de la disciplina ha sido muy grande. Se puede atribuir sus orígenes miles de años atrás, en la construcción de las pirámides de Egipto, que al día de hoy genera mucha curiosidad sobre cómo se han construido en tan poco tiempo las pirámides de la meseta de Guiza: Keops, Kefrén y Micerino. Particularmente, la Gran Pirámide de Keops posee aproximadamente 2,3 millones de bloques, se habría construido en unos 20 años y ello implica colocar, en promedio, un bloque cada dos minuto [4]. La ciencia no tiene respuestas sobre la técnica de construcción, es decir, como se cortaron, trasladaron y ubicaron los bloques en una época en donde si bien había recursos y mano de obra de sobra, era necesario organizar todo ese esfuerzo.

Sin embargo, se estima que en Egipto existen entre 120 y 140 pirámides, las cuales evolucionaron desde mastabas (bases planas rectangular), pasando por pirámides escalonadas (varias mastabas superpuestas), hasta llegar a pirámides de caras lisas (en donde se fue ajustando el grado de inclinación). Toda esta evolución supone que permitió ajustar el proceso de construcción (organización y gestión) por lo cual, permitió llegar a la construcción de la gran pirámide de Keops con los desafíos que implica.

Durante la década del 60 se formalizó la disciplina (durante la carrera espacial) cuando en 1969 cinco profesionales interesados en formalizar la profesión fundan el *Project Management Institute* quien edita el PMBoK [1] *Project Management Body of Knowledge*, es decir, el cuerpo de conocimiento de la gestión de proyectos. Esta guía incluye las prácticas consolidadas de profesionales de todo el mundo de variados dominios y negocios.

Sin embargo, los diferentes productos poseen diferentes características y hacen que la gestión de su construcción o desarrollo sea particular. El software posee 4 características esenciales: es complejo por naturaleza, intangible, maleable y se tiene que adaptar a un contexto. Esto sumado a que posee un alto contenido intelectual, es decir, es básicamente conocimiento, determina que su gestión es diferente de la gestión para construir una pirámide o un puente romano [5]. El esfuerzo para capturar el conocimiento y difundirlo para que un equipo construya un producto intangible que se comunica a través de diferentes representaciones complementarias, es muy diferente del esfuerzo de obtener materiales y organizar la logística para su traslado y ubicación.

La industria del petróleo y gas posee desafíos muy particulares: complejidad técnica y operativa (trabajo en zonas remotas), riesgos financieros y altos costos de inversión, regulaciones y cumplimiento normativo, impacto ambiental y sostenibilidad, seguridad y gestión de riesgos, logística y cadena de suministros entre otros factores [6].El *Building Information Modeling* (BIM) es una metodología colaborativa que integra información digital multidimensional para diseñar, construir y operar infraestructuras. Permite centralizar datos geométricos, técnicos y de gestión a lo largo del ciclo de vida del proyecto. Su uso mejora la coordinación interdisciplinaria, reduce errores y optimiza costos, tiempos y mantenimiento [7].

La integración entre los estándares de gestión de proyectos (PMBOK 7) y la familia de normas ISO 19650 sobre gestión de la información en entornos BIM establece un marco convergente que permite una gestión basada en datos, trazabilidad documental y control colaborativo del ciclo de vida. Esta convergencia habilita nuevas prácticas de interoperabilidad especialmente útiles en megaproyectos del sector petróleo y gas.Es así que la combinación de prácticas de gestión de proyectos tradicionales con el método BIM ofrece una oportunidad para enriquecer la disciplina de gestión de proyectos. Ahora bien, la enseñanza de la gestión de proyectos con BIM presenta ciertos desafíos. Los alumnos necesitan conocer la complejidad de la industria de petróleo y gas para sacar provecho de los métodos, técnicas y herramientas propios de la disciplina de gestión de proyectos general, además de las del método BIM. Por lo tanto, diseñar un curso y llevarlo a la práctica posee los

desafíos de poder transmitir el contexto, la motivación y los métodos, técnicas y herramientas. Sobre todo, es extremadamente difícil realizar actividades realistas, dado lo complejo del contexto y la dimensión que podrían tener las actividades.

Este proyecto, tiene por objetivo diseñar un curso de gestión de proyectos utilizando la metodología BIM en la industria del petróleo y gas.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La estrategia general es la siguiente. En primer lugar de plantea tomar los conceptos relevantes de la gestión de proyectos tradicional y explorar los de BIM, para lograr una currícula unificada. Luego se pretende armar 3 líneas: motivación, ejercitación y conceptos teóricos. Se pretende realizar capsulas educativas articulando con microcredenciales formativas de manera de articular conceptos relevantes con su motivación y actividad práctica.

3. RESULTADOS ESPERADOS

Los resultados esperados son varios. Por un lado, a nivel pedagógico, se pretende diseñar un curso a través de píldoras educativas y microcredenciales. Por otro lado, a nivel de la disciplina, se espera enriquecer la misma con la articulación con BIM. Finalmente, a nivel formativo se espera contribuir con el sector del petróleo y gas brindando profesionales mejor preparados que puedan mejorar el desarrollo de esta área.

4. FORMACIÓN DE RECURSOS HUMANOS

Como se indica en la sección previa, se espera mejorar la formación, capacidades y habilidades de los profesionales del sector del petróleo y gas, con el fin de mejorar el sector y poder hacer frente a todas las problemáticas que se presentan: optimización, uso eficiente, cumplimiento regulatorio, ecología y sostenibilidad. Todo en un marco de ubicación geográfica distante y extrema.

5. REFERENCIAS

[1] Project Management Institute: A Guide to the Project Management Body of Knowledge (PMBOK® Guide), 7th edition, (2021).

[2] Why Do Projects Fail?, accedido 2026/02/16 https://calleam.com/WTPF/?page_id=3.

[3] Kerzner, H.: Project Management: A Systems Approach to Planning, Scheduling, and Controlling, 12th edition, Wiley (2017).

[4] Hawass, Z.: The Pyramids of Ancient Egypt. University of Pennsylvania Press (1990).

[5] Brooks, F. P.: No Silver Bullet: Essence and Accidents of Software Engineering. IEEE Computer, 20(4), 10–19 (1987).

[6] Merrow, E. W.: Industrial Megaprojects: Concepts, Strategies, and Practices for Success. Wiley. ISBN: 978-1-119-89317-2 (2024).

[7] Sacks, R., Eastman, C., Lee, C., Teicholz, P.: BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers (3rd ed.). Wiley. (2018).

Estudio de la actividad de instalación de sistemas de software en PyMES de Argentina

Rodrigo Maestre[1][0009-0001-6851-6463], Marisa Panizzi[1,2][0000-0002-9699-1544], Rodolfo Bertone[3][0000-0003-0609-0310]

[1] Programa de Maestría en Ingeniería en Sistemas de Información, Escuela de Posgrado, Universidad Tecnológica Nacional - Facultad Regional Buenos Aires (UTN - FRBA) - Dirección: Medrano 951. (C1179AAQ). CABA. Argentina.

[2] Instituto de Tecnología e Ingeniería, Universidad Nacional de Hurlingham, Av. Vergara 2222 (B1688GEZ) - Villa Tesei - Bs. As. Argentina.

[3] Facultad de Informática -Instituto de Investigación en Informática LIDI (III-LIDI). Universidad Nacional de La Plata, 50 y 120. La Plata, Argentina.

rmaestre@frba.utn.edu.ar, marisapanizzi@outlook.com, pbertone@lidi.unlp.edu.ar

Resumen

La línea I+D+I presentada se enmarca en el fortalecimiento de la actividad de instalación de sistemas de software dentro del proceso de despliegue de sistemas de software en Pequeñas y Medianas Empresas (PyMES) de Argentina.

Las PyMES suelen enfrentar dificultades para ejecutar la actividad de instalación de sistemas de software de manera consistente, debido a la falta de procesos definidos y herramientas adecuadas. Por lo tanto, es necesario avanzar en el estudio de esta actividad con el objetivo de identificar problemáticas recurrentes y oportunidades de mejora que permitan reforzar su ejecución, y de esta forma mejorar la calidad de los sistemas de software y la competitividad de estas empresas en la industria, tanto a nivel nacional como internacional.

En este artículo se presenta el proceso investigativo para la definición de un proceso de instalación de sistemas de software orientado a apoyar a las PyMES de Argentina.

Palabras clave: instalación, proceso de despliegue, sistemas de software, PyMES de Argentina.

Contexto

La línea I+D+I que se reporta en este artículo forma parte de un proyecto de investigación titulado "Estudio del proceso de implantación de sistemas informáticos en el contexto industrial de la República Argentina" (Código SIUTNBA0006576) de la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional.

Cuenta con el asesoramiento científico del grupo de investigación de Ingeniería de Software del Instituto de Investigación en Informática (III-LIDI) de la UNLP.

Introducción

La instalación de sistemas de software es una actividad fundamental en el ciclo de vida del software, donde un sistema de software se incorpora al ambiente productivo y se lo deja operativo para los usuarios finales. Una instalación mal ejecutada puede comprometer aspectos críticos como la disponibilidad del sistema, la continuidad del servicio y aumentar los riesgos de fallas en el ambiente de producción. Cuando las organizaciones descuidan esta actividad, suelen enfrentarse a problemas costosos y complejos, producto de errores introducidos durante la instalación [1].

Según el estándar internacional ISO/IEC/IEEE 12207:2017 [2], la instalación no se limita a la simple transferencia de archivos de un servidor a otro, sino que comprende un conjunto de tareas técnicas orientadas a integrar el sistema de software en su ambiente de destino,

incluyendo su configuración, la adecuación del ambiente, la creación y configuración de bases de datos y, si corresponde, la migración de datos. Esto evidencia la necesidad de abordar la instalación como una actividad crítica en el ciclo de vida del software.

La actividad de instalación de sistemas de software pertenece a un proceso más amplio denominado despliegue de sistemas de software. El estándar ISO/IEC/IEEE 12207:2017 ubica a la actividad de instalación dentro de dicho proceso, cuyo objetivo es poner el sistema de software en condiciones operativas y asegurar que sea funcional y compatible en el ambiente productivo. Por su parte, ITIL 4 [3], en la práctica de *Deployment Management*, aborda la instalación como parte de un conjunto de actividades orientadas a implementar cambios y nuevas versiones de software de forma controlada, minimizando impactos negativos en los servicios existentes. Aunque los estándares consideran a la instalación como una actividad integrada dentro del despliegue, su alcance y complejidad técnica justifican un análisis específico, ya que es el momento exacto donde el sistema de software desarrollado y el ambiente productivo se integran.

En la práctica, la instalación de sistemas de software presenta una serie de dificultades técnicas y organizacionales que afectan directamente la puesta en producción del sistema. Entre los problemas más frecuentes se encuentran la preparación inadecuada del ambiente de destino [4], configuraciones incorrectas del sistema o de su infraestructura, incompatibilidades con componentes existentes y otras condiciones técnicas y organizacionales que dificultan la actividad [5]. Estas situaciones suelen generar fallas durante la ejecución, interrupciones del servicio, necesidad de retrabajos y aumento de los costos y plazos originalmente previstos para el proyecto.

A partir de esta problemática, y con el objetivo de conocer el estado del arte en relación con la actividad de instalación de sistemas de software, se realizó un mapeo sistemático de la literatura (en inglés, *Systematic Mapping Study* o SMS). Para su desarrollo se siguieron las directrices propuestas en investigaciones académicas ampliamente aceptadas en el área de Ingeniería de Software basada en la evidencia [6], [7], [8], [9]. Este estudio permitió analizar las contribuciones existentes en la literatura científica vinculadas a la actividad de instalación de sistemas de software, identificando enfoques, herramientas y tareas asociadas, así como los tipos de investigación predominantes.

El objetivo del SMS fue dar respuesta a la pregunta de investigación (PI): *¿Cuál es el estado del arte en cuanto a la actividad de instalación de sistemas de software?* Se logró responder las 4 preguntas de investigación definidas, las cuales se sintetizan a continuación [10]:

PI1: ¿Qué contribuciones existen en la actividad de instalación de sistemas de software?

Para esta pregunta de investigación (PI1) los resultados arrojan que predominan los modelos (25%), que explican cómo encarar la instalación, seguidos por métodos (17%) y herramientas (17%). También aparecen guías (13%), métricas (13%) y en menor medida aportes sobre seguridad (10%) y riesgos (5%), estos dos últimos relacionados principalmente a entornos cloud.

PI2: ¿Qué herramientas tecnológicas se utilizan para la actividad de Instalación de sistemas de software?

Por otra parte, la segunda pregunta de investigación (PI2), respecto a las herramientas tecnológicas que se utilizan en la actividad de instalación de sistemas de software se focalizan principalmente en la automatización de CI/CD (29%), las cuales permiten instalar versiones de software de forma rápida y confiable, se destacan herramientas como Jenkins y GitLabCI. El uso de herramientas de empaquetamiento (21%), como RPM o DEB, que se utilizan para distribuir e instalar componentes de software, y las herramientas Cloud (21%) como AWS se posicionan en segundo lugar. Luego se destaca el uso de scripting (13%) en Bash o PowerShell. Las herramientas menos mencionadas en los artículos fueron los contenedores y orquestadores (8%) como Docker y Kubernetes, y los gestores de configuración (8%) como Ansible y Mender, que permiten asegurar que los entornos siempre estén configurados de la misma manera.

PI3: ¿Qué tareas de la actividad de instalación de sistemas de software se mencionan en las contribuciones?

La tercera pregunta de investigación (PI3) respecto a las tareas más mencionadas corresponden a la instalación de sistemas de software en el entorno productivo (29%). Le sigue la tarea de verificación y validación (26%) que permite comprobar que la instalación fue correcta. En tercer lugar, se encuentra la preparación del entorno (24%) vinculada a la definición de infraestructura, configuraciones y dependencias previas. La configuración del sistema (17%) es mencionada en varios artículos, en esta tarea se configura los permisos, las variables de entorno y otras opciones. La instalación de base de datos (2%) y migración de datos (2%) prácticamente no se mencionan. Y por último la tarea de desinstalación no se menciona en ninguno de los artículos analizados.

PI4: ¿Qué tipos de investigación se utilizan?

Finalmente, respecto a la última pregunta de investigación (PI4), se identificaron los tipos de investigación con la taxonomía propuesta por Wieringa *et al*. [11]. Se encontró que la mayoría de los artículos tienen como propósito realizar una propuesta de solución (66%), es decir una forma de introducir y justificar nuevos enfoques para la actividad de instalación. Otros artículos presentan evaluaciones de las investigaciones realizadas (27%), en donde se analiza la efectividad de diferentes soluciones. Por último, se observó un único artículo de opinión (7%) en el cual el autor brinda una comparación de diferentes modelos de instalación en entornos *cloud,* mencionando sus riesgos y beneficios.

El SMS permitió evidenciar la importancia de la actividad de instalación de sistemas de software, identificando sus principales desafíos y dejando en claro la necesidad de continuar investigando y desarrollando enfoques que la refuercen como una actividad crítica en el ciclo de vida del software.

Como continuación del proceso investigativo y previo a iniciar el diseño del proceso de instalación de sistemas de software para PyMES, se decidió llevar a cabo un proceso exploratorio mediante una encuesta [9], [12], [13] con el propósito de relevar el estado actual de la práctica de instalación de sistemas de software en la industria de software de Argentina.

Líneas de Investigación, Desarrollo e Innovación

La línea de I+D+I presentada consiste en la definición de un proceso de instalación de sistemas de software para apoyar a las PyMES

de Argentina. Este proceso propone ser una guía para la ejecución de la actividad de instalación de forma ordenada, considerando las particularidades y limitaciones propias de este tipo de empresas. Una vez diseñado el proceso se validará en la industria mediante un conjunto de estudios de caso.

Resultados y Objetivos

En esta línea de I+D+I en progreso respecto al fortalecimiento del proceso de despliegue de sistemas de software, se han logrado una serie de resultados que se detallan a continuación:

a) Académicos, se han concluido cuatro trabajos de especialidad en Ingeniería en Sistemas de Información, tres tesis de Maestría en Ingeniería en Sistemas de Información y dos tesis de Maestría en Ingeniería en Sistemas de Información en curso. Además, se ha defendido una tesis Doctoral en Ciencias Informáticas.

b) Producción Científica: se ha presentado la línea de I+D+I en eventos científicos de alcance nacional (WICC[1] 2017, WICC 2019, WICC 2020, WICC 2021, WICC 2022, CACIC[2] 2016, CACIC 2018, CACIC 2019, CACIC 2020, CACIC 2021, CACIC 2022, CACIC 2023, CACIC 2024, CACIC 2025, CONAIISI 2022[3], y en el ámbito internacional, CIACA[4] 2017, SEKE[5] 2017, CIbSE[6] 2019, CIbSE 2020, CIbSE 2021, CIbSE 2023, CIbSE 2024 y ICAETT[7] 2019. Además, se lograron tres publicaciones en Springer, una publicación en la revista *Brazilian Journal of Development* y otra en la revista *Journal of Computer Science and Technology* (JCS&T).

c) Proyectos de Investigación: los proyectos finalizados UTNBA4347 titulado: ¨Impacto del factor peopleware en el proceso de implantación de sistemas informáticos¨ y UTNBA6576 titulado: ¨Estudio del proceso de implantación de sistemas informáticos en el contexto industrial de la República Argentina¨. El proyecto a ejecutarse a partir del 1 de abril PID SIECBA0008576 titulado ¨Herramientas para fortalecer el proceso de despliegue de sistemas de software¨.

d) Formación en investigación: proceso de aprendizaje constante de métodos de investigación de ingeniería de software basada en evidencia [9].

Formación de Recursos Humanos

El grupo se encuentra conformado por un director, dos tesistas de Maestría, y un asesor científico-tecnológico. Se estima la formación de un Magíster en Ingeniería en Sistemas de Información, que actualmente se encuentra desarrollando como primera instancia su trabajo de especialidad.

Referencias

[1] - Alghamdi, A., & Niazi, M. (2023). Toward Successful Secure Software Deployment: An Empirical Study. In Proceedings of the 27th International

[1] Workshop de Investigadores en Ciencias de la Computación

[2] Congreso Argentino de Ciencias de la Computación.

[3] Congreso Nacional de Ingeniería Informática / Sistemas de Información

[4] Conferencia Iberoamericana de Computación Aplicada.

[5] International Conference on Software Engineering and Knowledge Engineering.

[6] Congreso Iberoamericano en Ingeniería de Software.

[7] Congreso Internacional sobre Avances en Nuevas Tendencias y Tecnologías.

Conference on Evaluation and Assessment in Software Engineering (EASE '23), 487–492. https://doi.org/10.1145/3593434.3593966

[2] - ISO/IEC/IEEE 12207:2017. Systems and software engineering - Software life cycle processes (2017).

[3] - AXELOS. ITIL® 4: Create, deliver and support: Reference and study guide. TSO (The Stationery Office). ISBN 9780113317998 (2021).

[4] - Gao, H., Treude, C., & Zahedi, M. (2025). Adapting installation instructions in rapidly evolving software ecosystems. IEEE Transactions on Software Engineering, 51(4), 1334-1357. https://doi.org/10.1109/TSE.2025.3552614

[5] - Chen, Z., Cao, Y., Liu, Y., Wang, H., Xie, T., & Liu, X. (2020). A comprehensive study on challenges in deploying deep learning based software. Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), 750–762, https://doi.org/10.1145/3368089.3409759

[6] Kitchenham, B., Budgen, D., & Brereton, P. (2015). Evidence-Based Software Engineering and Systematic Reviews (1st ed.). Chapman and Hall/CRC.

[7] Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), 68–77.

[8] Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, 64, 1–18.

[9] Genero, M., Piattini, M., & Cruz Lemus, J. A. (2014), Métodos de investigación en Ingeniería del Software. Madrid: Ra-Ma Editorial y Publicaciones.

[10] Maestre R., Panizzi M., Bertone R. (2025). Análisis del estado del arte sobre la actividad de instalación de sistemas de software. Publicado en el XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025).

[11] Wieringa, R., Maiden, N., Mead, N., Rolland, C. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. Requirements Engineering, 11(1), pp. 102-107.

[12] Kitchenman, B. A.y Pfleeger, S. L. (2008). Personal opinion survey. Capítulo 3 del Libro Guide to Advanced Empirical Software Engineering. Shull F., Janice Singer J., Sjøberg D.I.K. Editors. Springer.

[13] Jefferson Seide Molléri, Kai Petersen, Emilia Mendes (2020). An empirically evaluated checklist for surveys in software engineering. Information and Software Technology 119, 106240.

Evaluación de Prácticas de Ingeniería del Software en Organizaciones Regionales: Métricas, Calidad y Validación Empírica

María A. Barrera[1], Gabriel D. Vilallonga[1], Carolina I. Chayle[1], Claudia M. Herrera[1], Ana C. Figueroa Vicario[1], Manuel Baquinzay [1], Nancy del V. Lamas[1], Juan C. Leguizamón Almendra[1], Juan C. Daneri[1]
[1] Departamento de Informática- Facultad de Tecnología y Ciencias Aplicadas - Universidad Nacional de Catamarca
mbarrera@tecno.unca.edu.ar gvilallo@tecno.unca.edu.ar cchayle@tecno.unca.edu.ar
cm_herrera@tecno.unca.edu.ar

RESUMEN

En las organizaciones de Catamarca, la brecha entre los estándares internacionales y la práctica diaria de desarrollo dificulta la mejora continua y frecuentemente presentan limitaciones en la adopción de modelos de referencia como CMMI o ISO/IEC 25010. Este trabajo presenta una línea de investigación orientada al desarrollo y validación de un modelo integrado para la evaluación de prácticas de ingeniería del software, que combina métricas de proceso, calidad de producto y niveles de madurez, adaptado a las características de éstas organizaciones. El modelo propuesto se estructura en dimensiones, indicadores y métricas operativas que permiten caracterizar el estado de las prácticas de desarrollo y detectar brechas respecto a estándares reconocidos. La metodología contempla un enfoque empírico basado en estudios de caso múltiples y validación mediante método Delphi con expertos. Los resultados esperados incluyen la definición de un instrumento de evaluación aplicable en contextos regionales, así como lineamientos para la mejora continua.

Palabras clave: Ingeniería de Software, Organizaciones Regionales, Métricas, Calidad, Validación Empírica.

CONTEXTO

Este trabajo presenta las líneas de investigación del proyecto "Evaluación de Prácticas de Ingeniería del Software en Organizaciones Regionales: Métricas, Calidad y Validación Empírica", aprobado por la Secretaría de Investigación y Posgrado de la Universidad Nacional de Catamarca mediante Resolución Rectoral N° 774/2025 (Código SIGEVA: 80020250100069CA). La investigación se encuentra radicada en el Departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA. Nuestro equipo capitaliza la trayectoria y los resultados de proyectos anteriores dirigidos por el equipo, tales como: “Desafíos y Oportunidades de la IA Generativa: Análisis Profundo en la Transformación Digital y la Ingeniería de Software”, “Ingeniería de Software en la Era de las Industrias 4.0” y “Las TIC al servicio del Dato Abierto: Situación actual, conceptualización e iniciativas de apertura de información pública”. Esto permite abordar la problemática de la calidad del software no solo desde una perspectiva de código, sino también considerando la interoperabilidad y la reutilización de activos digitales en organizaciones de la provincia de Catamarca.

1. INTRODUCCIÓN

La calidad en el desarrollo de software constituye un factor crítico para el éxito de los sistemas informáticos, ya que impacta directamente en su confiabilidad, mantenibilidad y capacidad de evolución. En este sentido, la Ingeniería del Software ha desarrollado diversos modelos, estándares y enfoques orientados a la evaluación y mejora de los procesos y productos software. Entre los más relevantes se encuentra el modelo de calidad definido por la norma ISO/IEC 25010, elaborada por la International Organization for

Standardization, que establece un marco conceptual para la evaluación de características de calidad del producto software (ISO/IEC, 2011). Asimismo, los modelos de mejora de procesos, como CMMI, promovido por el CMMI Institute, proporcionan estructuras organizadas para evaluar y mejorar la capacidad de los procesos de desarrollo (Chrissis, Konrad & Shrum, 2011). Complementariamente, el uso de métricas de software constituye un elemento central para la evaluación objetiva, permitiendo cuantificar atributos de los procesos y productos, y facilitando la toma de decisiones basada en evidencia (Fenton & Bieman, 2014).

Sin embargo, diversos estudios han señalado que la adopción efectiva de estos modelos en organizaciones pequeñas y medianas presenta dificultades significativas, asociadas a limitaciones de recursos, falta de formalización de procesos y características propias del contexto organizacional (Laporte, Alexandre & Renault, 2008; Kuhrmann et al., 2018). Estas limitaciones son particularmente relevantes en organizaciones de la región, donde los modelos generalistas no siempre se ajustan adecuadamente a las condiciones reales de desarrollo. En este sentido, la literatura reciente en Ingeniería del Software empírica destaca la necesidad de desarrollar enfoques de evaluación adaptados al contexto organizacional específico, que permitan obtener diagnósticos realistas y orientar iniciativas de mejora de manera contextualizada (Wohlin et al., 2012).

En este contexto, este trabajo presenta una línea de investigación orientada al desarrollo de un modelo integrado de evaluación de prácticas de ingeniería del software, que combina métricas de proceso, calidad de producto y niveles de madurez, adaptado a las características de organizaciones públicas y privadas del medio. El modelo propuesto se estructura en dimensiones, indicadores y métricas operativas que permitirán evaluar sistemáticamente el estado de las prácticas de desarrollo e identificar oportunidades de mejora.

Dado que el proyecto de investigación en el cual se enmarca este trabajo ha iniciado recientemente, la investigación se encuentra en la etapa de diseño y definición del modelo conceptual y metodológico. Su validación empírica se realizará mediante un enfoque basado en estudios de caso múltiples en organizaciones reales y la aplicación del método Delphi con expertos, conforme a las recomendaciones metodológicas propuestas en la literatura especializada (Wohlin et al., 2012). Se espera que los resultados contribuyan al desarrollo de un instrumento de evaluación contextualizado, así como a la generación de conocimiento sobre la evaluación y mejora de prácticas de ingeniería del software en el ámbito provincial.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto tiene como contribución científica el desarrollo y validación de un modelo integrado de evaluación de prácticas de ingeniería del software adaptado a organizaciones del medio. Este modelo combina métricas de proceso, calidad de producto, niveles de madurez y prácticas de gobernanza de datos, con el propósito de generar un instrumento de evaluación contextualizado y empíricamente validado. En el presente trabajo se presenta la etapa inicial del proyecto acreditado, estructurando su desarrollo en dos ejes de investigación complementarios. Si bien el Eje 1 se centra en la métrica pura, su éxito depende de la gobernanza de datos que planteamos en el Eje 2, especialmente al integrar sistemas heterogéneos en el ámbito público provincial.

- *Eje 1Métricas de Software y Evaluación de Calidad Situada:* Las métricas constituyen un mecanismo fundamental para evaluar atributos de calidad (Fenton & Bieman, 2014). Este eje se centra en el diseño, adaptación y validación de métricas orientadas a evaluar la calidad de procesos y productos software en organizaciones con diferentes niveles de formalización en la región.

Adaptación de estándares: Se utilizan como referencia modelos reconocidos como ISO/IEC 25010 (ISO, 2011) e ISO/IEC 33001

(ISO, 2015), ampliamente utilizados para la evaluación de calidad y procesos software.
Validación empírica: Se aplica una estrategia metodológica basada en triangulación y método Delphi, ampliamente utilizado en investigación en Ingeniería del Software (Wohlin et al., 2012).
Evaluación de atributos de calidad: Se analizan atributos como mantenibilidad, eficiencia y confiabilidad, considerando su relación con factores técnicos y organizacionales.

- *Eje 2 Gobernanza de Datos e Interoperabilidad en Sistemas Software:* Se analizan marcos como ISO/IEC 38505-1 (ISO, 2017), orientados a la gobernanza de datos.

Gobernanza de datos: Se analizan marcos como ISO/IEC 38505-1 para definir criterios que permitan evaluar la gestión y reutilización de los datos como activos estratégicos.
Interoperabilidad: Se estudia el impacto de arquitecturas orientadas a servicios y el uso de estándares en la capacidad de integración y evolución de los sistemas software.
Reutilización: Se definen lineamientos que permitan evaluar el nivel de reutilización de componentes y datos, como indicador de madurez y sostenibilidad de los sistemas desarrollados.
Estos ejes constituyen la base conceptual y metodológica para el desarrollo y validación del modelo de evaluación propuesto, cuya aplicación empírica permitirá generar conocimiento científico y herramientas prácticas para la mejora de la ingeniería del software en el contexto regional.

3. RESULTADOS ESPERADOS

Considerando que el proyecto se encuentra en su fase inicial, los resultados obtenidos a la fecha corresponden a la consolidación del marco teórico y el diseño del protocolo de investigación. Se ha avanzado en la selección de los atributos de calidad de las normas ISO/IEC 25010 y 33001 que conformarán el núcleo del modelo de evaluación.
Como resultados esperados, se prevé alcanzar:
Un Modelo de Métricas Situadas: Un catálogo de indicadores validados empíricamente que permitan medir la mantenibilidad y la interoperabilidad en organizaciones con procesos de desarrollo heterogéneos.
Un Instrumento de Diagnóstico: Una herramienta basada en el método Delphi para que las organizaciones identifiquen brechas críticas respecto a estándares internacionales.
Marco de Gobernanza de Datos: Un conjunto de lineamientos técnicos para integrar la gestión de datos como un activo estratégico dentro del ciclo de vida del desarrollo de software (basado en ISO/IEC 38505-1).
Validación de Hipótesis: Evidencia empírica sobre la correlación entre la baja densidad tecnológica del entorno y las dificultades específicas en la adopción de modelos de madurez como CMMI.
La transferencia de estos resultados no se limitará al ámbito académico; buscamos impactar directamente en el medio regional mediante talleres y la formación de los futuros ingenieros de la UNCA

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo integra a docentes investigadores, tesistas y estudiantes de la carrera de Ingeniería en Informática de la UNCA. Actualmente, la línea de investigación vincula de manera directa el desarrollo de una (1) tesis de Doctorado y dos (2) tesis de Maestría radicadas en la Universidad Nacional de San Luis (UNSL), enfocadas en métricas de software y calidad de procesos. Asimismo, se prevé la dirección de al menos tres (3) tesinas de grado anuales y la incorporación de alumnos becarios para tareas de relevamiento y análisis técnico, fortaleciendo la formación de recursos humanos en investigación aplicada dentro del ámbito regional.

5. BIBLIOGRAFÍA

ISO/IEC. (2011). ISO/IEC 25010:2011 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models. International Organization for Standardization.
ISO/IEC. (2015). ISO/IEC 33001:2015 Information technology — Process assessment — Concepts and terminology.

International Organization for Standardization.
ISO/IEC. (2017). ISO/IEC 38505-1:2017 Governance of data — Part 1. International Organization for Standardization.
Chrissis, MB, Konrad, M. y Shrum, S. (2011). CMMI para el desarrollo: directrices para la integración de procesos y la mejora de productos . Pearson Educación.
Fenton, N., & Bieman, J. (2014). Software metrics: a rigorous and practical approach. CRC press.
Kuhrmann, M., Diebold, P., Münch, J., Tell, P., Garousi, V., Felderer, M., ... & Prause, C. R. (2018). Hybrid Software and System Development in Practice. In Software Engineering und Software Management 2018 (pp. 103-105). Gesellschaft für Informatik.
Laporte, C. Y., Alexandre, S., & Renault, A. (2008). Developing international standards for very small enterprises. Computer, 41(3), 98-101.
Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering (Vol. 236, p. P3). Berlin: Springer.

Evaluación del cumplimiento de OWASP en Productos *Software*

Paula Angeleri, Jorge Ceballos, Gustavo Aldegani,
Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano, CABA
Facultad de Ingeniería, Universidad Nacional de la Matanza, San Justo.
{paula.angeleri; jorge.ceballos; gustavo.aldegani}@comunidad.ub.edu.ar;
{pangeleri; jceballos}@unlam.edu.ar;

RESUMEN

El objetivo de este artículo es presentar la aplicación en la academia y la industria de los avances en los dos últimos proyectos de la Línea de investigación MyFEPS *Metodologías y Framework para la Evaluación de Productos de Software*, que está activa desde hace más de 15 años. Asimismo, se presenta una nueva sub-línea de investigación en el campo de la seguridad informática, orientada al desarrollo de herramientas de análisis estático que permiten detectar vulnerabilidades de seguridad en código fuente.

Palabras clave: Calidad de software, Evaluación de calidad de producto *software*, Evaluación ágil de seguridad informática del producto software, Software seguro.

CONTEXTO

La Línea de Investigación en Evaluación de Productos Software de la Universidad de Belgrano[1], conocida como "Línea MyFEPS" toma su nombre del primer proyecto de investigación *Metodologías y Framework para la Evaluación de Productos Software (MyFEPS)*[2], llevado a cabo desde el año 2010 al 2014, con el fin de colaborar con el IRAM para ofrecer el servicio conjunto de certificación de productos software[3], habiendo cumplido UB e IRAM, este propósito inicial[4]. En este primer proyecto[5] se especificaron:

(a) la primera versión del **modelo de calidad QSAT** [6], más actualizado que los modelos publicados hasta ese momento,
(b) un **proceso de evaluación MyFEPS**[7], alineado a las series de normas ISO/IEC 14598[8] y la actual serie ISO/IEC 25000 conocida como (SQuaRE)[9], que describe su proceso en la norma ISO/IEC 25040[10], y
(c) un **conjunto (kit) de herramientas de apoyo** durante la evaluación.

Con estos 3 componentes más un set de bases teóricas se dio vida al **Framework MyFEPS** de evaluación de productos *software*[11]. A partir de la creación del framework, se llevaron a cabo numerosos proyectos de evaluación de software, con fines académicos (para validación y mejora del framework y de sus instrumentos) y de industria (con el propósito de certificación del software)[11][12][13][14][15][16].

1. INTRODUCCIÓN

Una vez creado el Framework MyFEPS, se realizaron otros proyectos para mejorarlo y mantenerlo actualizado, lo que es una ventaja respecto de las normas ISO/IEC, que, por su propio proceso de estudio, requieren un promedio de 5 años en ser actualizadas.

El primero de estos proyectos se hizo en conjunto con la empresa TSOFT S.A., denominado: "*Diseño y desarrollo de un framework metodológico e instrumental para asistir a la evaluación de software (MBCPS)",* desarrollado del 2011 al 2013, en forma conjunta con el proyecto

MyFEPS1, dándole al proyecto una impronta de la industria[17]. Luego UB apuesta por el proyecto MyFEPS2: *Metodologías y Framework para la Evaluación de Productos Software: Evaluación comparativa de productos software de propósito similar* (2016-2018)[18], que busca desarrollar un nuevo proceso de evaluación más flexible y liviano [19] y un tercer proceso, que permita realizar evaluaciones comparativas de más de un producto software de propósito similar[20][21]. A lo largo de estos años se también hubo una colaboración entre este grupo de investigación y el grupo perteneciente a la Línea de Investigación en Ingeniería de Software y Sistemas de la Pontificia Universidad Católica del Perú (PUCP)[22]. Asimismo, se ha colaborado en proyectos del Grupo de Investigación y Desarrollo en Ingeniería de Software (red internacional GIDIS)[23].

En UB, luego de más de una década participando en proyectos referidos a Calidad del producto software, Calidad de los servicios Informáticos, y Calidad en procesos de desarrollo y mantenimiento de software, se abre una sub-línea de investigación en Seguridad informática.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En 2023 inició en UB un proyecto de investigación de la sub-línea Seguridad informática: *Framework MyFEPS para evaluación de la característica básica de Seguridad Informática de un producto software (MyFEPS3)*[24]. El objetivo principal del proyecto es "establecer un Framework y metodologías ágiles que permitan realizar evaluaciones de conformidad del nivel de seguridad informática de aplicaciones software principalmente desarrolladas a medida (calidad interna). La mención a una evaluación "ágil" refiere a eficiencia en cuanto al tiempo y los recursos"[25][26].

Paralelamente, en la Universidad Nacional de la Matanza (UNLaM) se llevó a cabo (2024-2025) el proyecto de investigación *Protocolo de verificación del uso de buenas prácticas de "programación segura" en un producto software*[27].

A partir de la sub-línea de seguridad informática, surgen otras sub-líneas de investigación que se fueron desarrollando en forma paralela:

1. Extensión de QSAT para evaluar el Cumplimiento Legal y/o /normativo respecto de seguridad informática,
2. Desarrollo de herramientas de análisis estático que permiten detectar vulnerabilidades de seguridad en código fuente.

Entre futuras líneas de investigación, se puede seguir validando la adecuación del Framework MyFEPS, y profundizando en lo que respecta a la evaluación de otras características básicas (CB) del modelo QSAT [6] como ser: Calidad de los artefactos, Efectividad, Facilidad de mantenimiento, Manejo de fallas, Satisfacción subjetiva de los usuarios, Satisfacción de stakeholders no usuarios, Usabilidad Objetiva, Accesibilidad, Facilidad de Instalación, entre otras. También sería interesante explorar si se puede extender el Framework MyFEPS para evaluar no sólo software, sino un Ecosistema Digital de Integrabilidad (EDI)[28]. Asimismo, evaluar la calidad de los datos, ya sea en un entorno EDI, o en un sistema de Inteligencia Artificial (IA), por ejemplo.

Los resultados obtenidos desde el inicio de la línea de investigación MyFEPS fueron explicados en [1][2][10][11][12][13][14][15]. Los más representativos son: la creación del **Framework MyFEPS** para evaluar productos software, con un **modelo de calidad propio QSAT** o utilizando otros modelos, como el propuesto por ISO/IEC 25010[29], lo que fue validado en los

proyectos eCom[30][31], Morpho[4][16], CKids, y WebChat, entre otros. Además, se establecieron tres métodos de evaluación para adecuarse al propósito de distintos proyectos de evaluación:

1) **método de evaluación MyFEPS tradicional**, centrado en la documentación (utilizado en procesos de certificación por IRAM y por UB),
2) **método de evaluación MyFEPS Ágil** para proyectos más ágiles que no requieran tanta formalidad, ni tanta documentación[18].

Ambos métodos se utilizan para evaluar un único producto software por proyecto, y se han utilizado para evaluar la calidad interna, externa o en uso.

3) **método MyFEPS comparativo**, para evaluar más de un producto software de propósitos similares, en un mismo proyecto[20][21].

3 RESULTADOS OBTENIDOS

Como se explica en [1], "el objetivo principal de los últimos años de MyFEPS fue la continua experimentación y uso del Framework MyFEPs … en diferentes tipos de proyectos académicos… y de transferencia en la industria, … buscando validar y mejorar el Framework MyFEPS".

En particular, durante el proyecto MyFEPS3 se ha podido validar que el método de evaluación MyFEPS Ágil es adecuado para realizar evaluaciones de seguridad informática de un producto software (en proyectos de industria y de academia), y que las nuevas SubCaracterísticas, Sub-Sub-Características, Atributos y Métricas incorporadas al modelo QSAT son adecuadas para medir las CB *seguridad informática, cumplimiento legal/normativo de seguridad informática*, y SubCB *disponibilidad*. Es decir, se alcanzaron los objetivos propuestos, detallados en [24] y [27].

Respecto de la sub-línea: 2. Desarrollo de herramientas de análisis estático…se desarrolló un analizador semántico durante el proceso de compilación, con la capacidad de identificar patrones de vulnerabilidad en el código fuente, el cual se enfocó a la detección de inyecciones SQL, al diseño inseguro y configuración incorrecta, y los fallos de identificación y autenticación. Este analizador semántico que detecta las 3 vulnerabilidades de OWASP Top10 [32]. mencionadas fue evaluado a través de casos prácticos.

Se transfirieron los conocimientos adquiridos a la academia y a la industria (CESSI, CECRA, IRAM, TSOFT S.A., Estudio Biscardi, Levin Software S.A., Ahorrocoop de Chile, SyK S.A., y a Gobierno (Dirección General de Rentas de Corrientes) a través de presentaciones, trabajos de cátedra, trabajos finales de carrera, curso *in company,* conferencias/ videoconferencias y servicios de asistencia técnica.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de la UB participaron 5 investigadores formados y 3 docentes colaboradores que han dirigido tesistas de grado. Además, participaron 3 estudiantes que finalizaron su trabajo final de carrera, y hay 8 trabajos finales en curso.

En el proyecto de UNLaM participaron 2 investigadores formados y 2 en formación.

En ambas universidades participaron estudiantes en el marco de las cátedras: Ingeniería de Software IV: Testing y Calidad, Ingeniería de Software IV: Integrabilidad (UB); Modelos de Calidad y Gestión de la Calidad de proyectos informáticos (UNLaM).

5. BIBLIOGRAFIA

[1] Angeleri, Paula M.; Titiosky, Rolando; Ceballos, Jorge L; Aguilera, Sergio O.; *Línea de investigación*

en evaluación de productos software, XX Workshop de Investigadores en Ciencias de la Computación WICC, 2018, URL: https://sedici.unlp.edu.ar/handle/10915/67914 consultada el 06/02/2026 a las 14:28 hs.

[2] Angeleri, Paula M.; Titiosky, Rolando; Ceballos, Jorge L.; *Línea de investigación en evaluación de productos software: dificultades encontradas*, XXI Workshop de Investigadores en Ciencias de la Computación WICC, 2019, URL: https://sedici.unlp.edu.ar/handle/10915/77122 consultada el 06/02/2026 a las 15:35hs.

[3] Angeleri, P.: Premios Sadosky 2008, IRAM Newsletter, Año 2 Nro.6, 2009.

[4] Ceballos, J.: Caso de Éxito de certificación de producto software según las normas serie ISO/IEC 25000, IRAM Newsletter, AÑO 9, Nº 72, 2015.

[5] Sorgen, A., Angeleri, P.: *El Modelo de Evaluación del Proyecto MyFEPS*, 40JAIIO -ASSE 2011 - ISSN: 1850-2792, 180-191, 2011, URL: https://sedici.unlp.edu.ar/handle/10915/125495 consultada el 06/02/2026 a las 15:52 hs.

[6] Angeleri, P.; Oliveros, A; Sorgen, A.; Titiosky, R.; Wuille Bille J.: Modelo de calidad de productos de software, Libro de Actas del 2do. Congreso Nacional de Ingeniería Informática y Sistemas de Información CONAIISI 2014, ISSN: 2346-9927 - Página 1043. – ISSN: 2346-9927 http://conaiisi.unsl.edu.ar/ProceedingsCoNaIISI2014.pdf consultada el 11/92/2926 a las 12:12 hs.

[7] Angeleri, P.; Titiosky, R.; Santi, M..; Davila, A.: Proceso de Evaluación de Productos *Software*, Memorias del 3er Congreso Nacional de Ingeniería Informática / Sistemas de Información, CONAIISI 2015, ISBN: 978-987-1896-47-9.

[8] serie de normas IRAM-ISO-IEC 14598: Tecnología de la información. Ingeniería de software. Evaluación del producto de software. Information technology. Software engineering. Software product evaluation. Parte 1, 2006, Partes 2,3, 2007, Partes 4,5, 2009, Parte 6, Buenos Aires, 2012.

[9] ISO/IEC 25000 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) – Overview, ISO, Ginebra.

[10] ISO/IEC 25040 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE)— Evaluation process. ISO, Ginebra, 2012.

[11] Angeleri, P., Titiosky, R., Ceballos J.: Framework de Evaluación de Productos Software, XVIII Workshop de Investigadores en Ciencias de la Computación WICC 2016, Entre Ríos, Argentina, ISBN: 978-950-698-377-2, p. 657-661, URL: https://sedici.unlp.edu.ar/bitstream/handle/10915/54072/Documento_completo.pdf-PDFA.pdf?sequence=1 consultada el 12/02/2026 a las 12:14 hs.

[12] Dávila, Abraham; Angeleri, Paula; "Software Product Evaluation - MyFEPS: A Past and Future Reflective Analysis," 2022 11th International Conference On Software Process Improvement (CIMPS), Acapulco, Guerrero, Mexico, 2022, pp. 37-45, doi: 10.1109/CIMPS57786.2022.10035711.

[13] Sorgen, A., Angeleri, P.: Teoría y Práctica en la Evaluación de Productos de Software, XIV Workshop de Investigadores en Ciencias de la Computación WICC 2012. ISBN: 978-950-766-082-5, Páginas: 552-556, URL: http://sedici.unlp.edu.ar/handle/10915/18972 consultada el 13/02/2026 a las 12:39 hs.

[14] Titiosky, R., Angeleri, P., Sorgen, A., Wuille Bille, J.: "Proyectos de Evaluación de Productos de Software con un nuevo Framework de Calidad", XV Workshop de Investigadores en Ciencias de la Computación WICC 2013, ISBN: 9789872817961. URL: http://sedici.unlp.edu.ar/handle/10915/27225 hs. consultada el 12/02/2026 a las 13:33.

[15] Angeleri, P. Páginas: 500-504, Titiosky, R., Sorgen, A., Wuille Bille, J., Oliveros, A.: Ajustes al Framework de Evaluación de Productos de Software MyFEPS, XVI Workshop de Investigadores en Ciencias de la Computación WICC 2014. Páginas: 565-569, URL: http://sedici.unlp.edu.ar/handle/10915/41583 consultada el 12/02/2025 a las 13:16 hs.

[16] Angeleri, P., Titiosky, R., Ceballos J. Maspero, Ch., Sánchez A., Menal M., Vinjoy M., SOFTWARE PRODUCT CERTIFICATION: AN ACADEMIC-INDUSTRY CERTIFICATION PROJECT, 2016 IEEE Xplore, IEEE Catalog Number:CFP16H47-POD, ISBN Print-On-Demand:978-1-5090-2939-6, ISBN Online:978-1-5090-2938-9, p. 176. URL; https://ieeexplore.ieee.org/document/7786009 consultada el 13/02/2025 a las 10:17 hs.

[17] Angeleri, P., Sorgen, A., Bidone, P., Fava, A., Grasso, W.; Diseño y desarrollo de un framework metodológico e instrumental para asistir a la evaluación de software, 43 JAIIO - JUI 2014 - ISSN: 1851-2518 - Página 10. URL: https://43jaiio.sadio.org.ar/proceedings/JUI/836-2618-1-DR.pdf consultada el 12/02/2026 a las 13:50 am.

[18] Formulario de presentación del Proyecto de Investigación "Metodologías y Framework para la Evaluación de Productos Software: Evaluación comparativa de productos software de propósito similar (MyFEPS2)", Universidad de Belgrano, Buenos Aires, 2016.

[19] Abraham Dávila; Rolando Titiosky; Paula Angeleri; Agile new process for software product evaluation: Agile-MyFEPS | IEEE Conference Publication | IEEE Xplore. URL: https://ieeexplore.ieee.org/abstract/document/9082432, consultada el 18/07/2022, DOI: 10.1109/CIMPS49236.2019.9082432.
[20] Malerba P., Comparación de herramientas para desarrollo de videojuegos FPS, UB, 2018.
[21] Favaloro F., Evaluación comparativa de tecnologías Backup a Disco vs Backup a Cinta, usando el Framework MyFEPS, UB, 2018.
[22] Grupo de Investigación y Desarollo en Ingeniería de Software del Departamento Académico de Ingeniería, PUCP, URL: https://departamento-ingenieria.pucp.edu.pe/grupo_investigacion/gidis-pucp/ consultada el 12/02/2026 a las 19:15 hs.
[23] Grupo de Investigación y Desarrollo en Ingeniería de Sistemas (GIDIS), URL: https://sites.google.com/pucp.pe/gidis/inicio, consultada el 12/02/2026 a las 19:23 hs.
[24] Formulario de presentación del Proyecto de Investigación "Framework MyFEPS para evaluación de la característica básica de Seguridad Informática de un producto software (MyFEPS 3)", Universidad de Belgrano, Buenos Aires, 2016.
[25] Paula M. Angeleri, Jorge L. Ceballos, Gustavo Aldegani; Miguel Ángel Blanco. Framework para evaluar la seguridad informática de un producto software XXI Workshop de Investigadores en Ciencias de la Computación WICC 2024, Libro de Actas - XXVI Workshop de Investigadores en Ciencias de la Computación, URL: https://sedici.unlp.edu.ar/bitstream/handle/10915/54072/Documento_completo.pdf-PDFA.pdf?sequence=1 consultada el 16/02/2026 a las 11:45 hs.
[26] Angeleri, P.M.; Ceballos, J.L.; Aldegani, G.N.; Blanco, M.A.; Menal, C.M.; Comparación de dos modelos de calidad de producto software respecto de la Característica "Seguridad Informática": QSAT y SQuaRE. Congreso Nacional de Ingeniería Informática y Sistemas de Información CONAIISI 2024, San Fernando del valle de Catamarca, 2024.
[27] Protocolo de presentación de proyectos de investigación SIGEVA UNLaM: Protocolo de verificación del uso de buenas prácticas de "programación segura" en un producto software, UNLaM, 2024.
[28] IRAM 17610-1, Ecosistema Digital de Integrabilidad. Parte 1: Requisitos, IRAM, Buenos Aires, 2024.
[29] ISO/IEC 25010 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) - Product quality model, ISO, Ginebra, 2023.
[30] Santi, M., **Medición de la Usabilidad de un Ecommerce empleando el Framework MyFEPS**, tesis de grado de Universidad de Belgrano, 2014. https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=49786, 08/03/2018 14:50 hs.
[31] Santi, M., **Evaluación de la Usabilidad de un sitio de Comercio Electrónico empleando el Framework MyFEPS,** CONAIISI 2014.
[32] OWASP: Open Web Application Security Project ® (Estándar de Verificación de Seguridad en Aplicaciones). URL: https://owasp.org/www-project-top-10-ci-cd-security-risks/ consultada el 14/02/2026 a las 09:56 hs.

Formalización de heurísticas en Ingeniería de Requisitos mediante IA generativa

Gladys Kaplan[1], María Inés Bertín[1], Gerardo Riera[1], Gabriel Blanco[1]

[1]Departamento de Ingeniería e Investigaciones Tecnológicas,
Universidad Nacional de La Matanza. San Justo, Buenos Aires
{gkaplan,mbertin,griera,gblanco}@unlam.edu.ar

RESUMEN

La construcción de artefactos en ingeniería de requisitos es una actividad intensiva en conocimiento experto, sustentada en heurísticas metodológicas que suelen permanecer implícitas y depender de la experiencia del ingeniero de requisitos. Esta situación dificulta la reproducibilidad de los procesos, incrementa el esfuerzo cognitivo y limita su adopción sistemática en contextos educativos e industriales.

En los últimos años, la inteligencia artificial generativa (IAG) ha despertado un creciente interés en su aplicación a la ingeniería de requisitos. Sin embargo, muchos enfoques existentes utilizan estos sistemas como generadores autónomos de contenido, con escaso control metodológico y riesgos asociados a inconsistencias, omisiones y pérdida de trazabilidad.

El proyecto se desarrolla en el marco del *Proceso de Requisitos Basado en Escenarios*, seleccionado como ámbito inicial por la disponibilidad de heurísticas de construcción consolidadas. En este contexto, el trabajo se orienta a la formalización de heurísticas expertas como reglas operativas y a su ejecución automatizada mediante IAG, respetando la forma de construcción previamente definida en el proceso.

Los resultados preliminares sugieren una reducción del esfuerzo cognitivo y del tiempo de ejecución, manteniendo consistencia y reproducibilidad bajo supervisión humana.

Palabras clave: ingeniería de requisitos; inteligencia artificial generativa; conocimiento experto; heurísticas metodológicas; reproducibilidad; colaboración humano-IA.

CONTEXTO

El presente artículo fue desarrollado en el contexto del proyecto "Impacto de la inteligencia artificial en la ingeniería de requisitos" como parte de la línea de investigación "Ingeniería de requisitos: Un enfoque a los negocios" que se realiza en la Universidad Nacional de La Matanza (UNLaM) en el Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT).

1. INTRODUCCIÓN

La ingeniería de requisitos (IR) constituye una de las actividades más críticas del desarrollo de software, dado que los problemas introducidos en esta etapa tienden a propagarse y amplificarse en fases posteriores, impactando negativamente en costos, calidad y satisfacción del cliente [4] [5] [6]. Una característica transversal de la disciplina es su fuerte dependencia del conocimiento experto: gran parte de las decisiones metodológicas se apoya en heurísticas adquiridas a través de la experiencia profesional, muchas veces implícitas y difícilmente transferibles. Esta situación introduce variabilidad en los resultados, incrementa el esfuerzo cognitivo y dificulta la reproducibilidad de los procesos.

En este contexto, el Proceso de Requisitos Basado en Escenarios se ha consolidado como un enfoque eficaz para comprender el dominio del problema y facilitar la comunicación con clientes y usuarios [1]. Este enfoque favorece la validación temprana y continua al integrar acciones, decisiones, restricciones del contexto y objetivos del sistema. En el presente trabajo se adopta dicho proceso como ámbito inicial de aplicación debido a la existencia de un cuerpo consolidado de heurísticas de construcción y

validación desarrolladas en investigaciones previas.

En los últimos años, la inteligencia artificial generativa (IAG), en particular, los modelos de lenguaje, ha despertado un creciente interés en su aplicación a la ingeniería de requisitos [2] [3]. Revisiones recientes evidencian un aumento sostenido de trabajos que exploran su uso para asistir actividades de elicitación, análisis y validación [7] [8]. Sin embargo, muchos de estos enfoques utilizan la IAG como generador autónomo de contenido, con escaso control metodológico, lo que introduce riesgos de inconsistencias semánticas, omisiones y pérdida de trazabilidad [9] [10].

Frente a estas limitaciones, investigaciones recientes proponen avanzar hacia esquemas de colaboración humano–IA, en los cuales la inteligencia artificial actúe como soporte del proceso y no como sustituto del juicio profesional [11]. En esta línea, resulta necesario desarrollar mecanismos que permitan integrar la IAG de manera controlada, preservando el método y haciendo ejecutable el conocimiento experto que guía la construcción de artefactos de requisitos.

En el marco del proceso, el presente proyecto trabaja sobre desgrabaciones de entrevistas con usuarios del dominio, principal fuente de información específica para el software a construir. Estos documentos suelen contener formulaciones incompletas, referencias implícitas y supuestos compartidos propios del lenguaje natural. En este contexto, la IAG aporta una capacidad diferencial respecto de la automatización tradicional, al poder interpretar el contexto del discurso, completar formulaciones elípticas y articular el contenido explícito con conocimiento general pertinente, asistiendo la desambiguación bajo validación del ingeniero de requisitos.

El enfoque propuesto se basa en la formalización de heurísticas expertas previamente definidas para las principales actividades del proceso, incluyendo la construcción del Léxico Extendido del Lenguaje (LEL), los Escenarios Actuales (EA), los Escenarios Futuros (EF) y el Léxico Extendido del Lenguaje de Requisitos (LEL_R) así como su posterior **representación** como reglas y restricciones ejecutables por un modelo de lenguaje. De este modo, la IAG actúa como ejecutor controlado de heurísticas explícitas, sin redefinir el método ni introducir conocimiento arbitrario, bajo supervisión del ingeniero de requisitos.

En este contexto, el presente trabajo presenta un proyecto orientado a explorar el uso de la IAG como ejecutor de heurísticas metodológicas explícitamente formalizadas, con el objetivo de reducir el esfuerzo cognitivo, mejorar la reproducibilidad del proceso y preservar el control metodológico humano. La implementación inicial se desarrolla en el marco del Proceso de Requisitos Basado en Escenarios, con proyección hacia su extensión a otras actividades de la ingeniería de requisitos.

En particular, la formalización de heurísticas no solo busca hacer explícito el conocimiento experto, sino también habilitar su ejecución sistemática por un agente computacional, permitiendo evaluar su impacto en términos de esfuerzo, tiempo, consistencia y reproducibilidad del proceso.

La contribución principal del trabajo radica en transformar heurísticas metodológicas expertas previamente definidas en conocimiento formalmente representado y ejecutable, preservando el control humano del proceso y habilitando su evaluación empírica. Hasta donde alcanza nuestro conocimiento, no se han reportado trabajos que formalicen heurísticas metodológicas del Proceso de Requisitos Basado en Escenarios como reglas ejecutables controladas por IAG.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo se inscribe en una línea de investigación orientada al fortalecimiento metodológico de la IR y este proyecto en particular, al uso controlado de IA, con énfasis en la formalización y reutilización del conocimiento experto que guía la ejecución de los procesos de requisitos. En particular, la investigación se enmarca dentro de los estudios que abordan la IR como una disciplina intensiva en conocimiento, donde gran parte de las decisiones críticas se sustentan en

heurísticas adquiridas a través de la experiencia profesional y no siempre documentadas de manera explícita.

Desde sus orígenes, la IR ha reconocido la necesidad de contar con procesos sistemáticos que permitan reducir la ambigüedad, detectar omisiones tempranas y facilitar la validación con clientes y usuarios. No obstante, numerosos trabajos han señalado que la aplicación efectiva de estos procesos continúa dependiendo fuertemente del ingeniero de requisitos, lo que introduce variabilidad en los resultados y dificulta la reproducibilidad del proceso.

En paralelo, investigaciones previas han destacado el valor de explicitar heurísticas y reglas metodológicas como mecanismo para mejorar la calidad y consistencia de los artefactos de requisitos, así como para apoyar la formación de ingenieros menos experimentados. Por ejemplo:

Heurística original: "Identificar símbolos relevantes del dominio".

Regla formalizada: "Toda entidad nominal recurrente con función operativa explícita debe proponerse como símbolo candidato."

La formalización de heurísticas constituye, además, la condición necesaria para su ejecución automatizada y posterior evaluación empírica, permitiendo analizar el impacto del enfoque en variables operativas del proceso tales como esfuerzo, tiempo de ejecución, consistencia y reproducibilidad.

La contribución de esta línea se orienta a integrar la IAG dentro del *Proceso de Requisitos Basado en Escenarios*, preservando el control humano y el rigor metodológico, y promoviendo procesos más reproducibles, escalables y sostenibles en contextos académicos e industriales. Esta investigación busca, además, generar bases conceptuales y operativas para futuras extensiones hacia otras actividades del proceso de requisitos, manteniendo una postura crítica y responsable frente al uso de tecnologías emergentes.

El objetivo es explorar el uso de la *IAG como ejecutor de heurísticas expertas explícitamente definidas* para asistir actividades del *Proceso de Requisitos Basado en Escenarios*, con el propósito de reducir el esfuerzo y la carga cognitiva del ingeniero de requisitos, preservando el control metodológico y la calidad del proceso. La Tabla 1 resume las heurísticas expertas existentes.

Actividad del proceso	Heurísticas de construcción	Finalidad principal	Tipo de heurística
Léxico Extendido del Lenguaje (LEL)	Identificación y clasificación de símbolos; definición de noción e impacto	Capturar y estructurar el vocabulario del dominio, reduciendo ambigüedad	E
Escenarios Actuales (EA)	Identificación de actores, contexto, episodios y excepciones	Describir el funcionamiento actual del proceso del negocio	S
Escenarios Futuros (EF)	Proyección de acciones, decisiones y restricciones	Representar el proceso proyectado con el sistema en operación	S
LEL de Requisitos (LEL_R)	Ampliación léxica orientada al comportamiento futuro	Incorporar nuevo vocabulario requerido por el sistema	C
Extracción de requisitos desde EF	Identificación de comportamientos esperados, restricciones y excepciones	Derivar requisitos trazables a partir de los Escenarios Futuros	T

E: Estructural; **S**: Semántica; **C**: Consistencia; **T**: Trazabilidad

Tabla 1 – Heurísticas metodológicas formalizadas en el Proceso de Requisitos Basado en Escenarios

Las heurísticas de la Tabla 1 constituyen la base del proceso de formalización y fueron transformadas en reglas operativas ejecutables por IAG.

Se espera:

- Identificar y explicitar el conjunto de heurísticas metodológicas que guían la ejecución de actividades del Proceso de Requisitos Basado en Escenarios.
- Formalizar dichas heurísticas en forma de reglas y restricciones susceptibles de ser ejecutadas por un sistema de IAG.
- Analizar la viabilidad del uso de IAG como ejecutor controlado de heurísticas, evitando su uso como generador autónomo de contenido.

- Evaluar el impacto del enfoque propuesto en términos de esfuerzo requerido, tiempo de ejecución y consistencia de los resultados obtenidos.
- Identificar los límites, riesgos y condiciones de aplicabilidad del enfoque dentro del *Proceso de Requisitos Basado en Escenarios*.

A partir de los objetivos planteados, se formulan las siguientes preguntas de investigación:

PI1: ¿Es posible formalizar heurísticas expertas del Proceso de Requisitos Basado en Escenarios de manera que puedan ser ejecutadas por una IAG?

PI2: ¿Puede una IAG actuar como ejecutor de heurísticas metodológicas sin perder el control del proceso por parte del ingeniero de requisitos?

PI3: ¿En qué medida el uso de IAG reduce el esfuerzo y el tiempo requeridos para la ejecución de actividades del Proceso de Requisitos Basado en Escenarios?

PI4: ¿Qué impacto tiene este enfoque en la consistencia y reproducibilidad de los resultados obtenidos?

PI5: ¿Cuáles son los principales riesgos y limitaciones asociados al uso de IAG en este contexto?

3. RESULTADOS OBTENIDOS/ESPERADOS

Como primera etapa del proyecto, se realizaron pruebas exploratorias para evaluar la viabilidad del enfoque y su coherencia con los objetivos de formalización y ejecución controlada. Las pruebas se concentraron en las heurísticas asociadas a la construcción del Léxico Extendido del Lenguaje (LEL), las cuales fueron formalizadas como reglas y restricciones explícitas y ejecutadas mediante un sistema de IAG, utilizando un modelo de lenguaje de gran escala (ChatGPT) como entorno de ejecución, bajo supervisión del ingeniero de requisitos. Si bien fue posible traducir las heurísticas metodológicas a reglas operativas ejecutables, el proceso resultó altamente demandante debido a la diferencia entre el carácter relativamente flexible de las heurísticas y la precisión requerida por reglas formales. Además, se observó una marcada interdependencia entre reglas, donde modificaciones menores en una pueden afectar el comportamiento de otras. Se construyeron entre 40 y 45 reglas operativas.

Los resultados preliminares indican coherencia con los objetivos planteados, mostrando que la IAG puede ejecutar las heurísticas respetando patrones estructurales y restricciones metodológicas, siempre que estén formuladas con precisión. En particular, se observó una reducción del tiempo de ejecución respecto de la realización manual de las mismas tareas, así como una disminución del esfuerzo cognitivo requerido por parte del ingeniero de requisitos, quien pasa a desempeñar un rol predominantemente de supervisión y validación.

La Tabla 2 muestra la responsabilidad de las diferentes actividades.

Actividad	Ingeniero de Requisitos	IAG
Definición de heurísticas	✓	✗
Interpretación del contexto organizacional	✓	✗
Ejecución de heurísticas	✗	✓
Incorporación de conocimiento general	Define límites y valida	Incorpora solo conocimiento permitido por reglas
Validación de resultados	✓	✗

Tabla 2 – Distribución de responsabilidades humano–IA

Asimismo, las pruebas realizadas permitieron constatar un alto grado de consistencia interna y reproducibilidad de los resultados, especialmente en comparación con ejecuciones humanas realizadas por analistas con distinto nivel de experiencia. Estos resultados iniciales sugieren que la IAG puede actuar eficazmente como ejecutor de heurísticas expertas, sin sustituir el control metodológico humano.

La formalización de heurísticas expertas permite transformar conocimiento

metodológico acumulado en un artefacto ejecutable, evaluable y reproducible, reduciendo la dependencia del juicio individual sin sustituir el control humano.

Límites y riesgos. El enfoque depende críticamente de la precisión con que las heurísticas se formalicen como reglas operativas: formulaciones ambiguas o incompletas pueden generar ejecuciones inconsistentes. Además, la interdependencia entre reglas introduce sensibilidad a cambios locales, requiriendo refinamiento iterativo y control de versiones. Finalmente, aun bajo restricciones, los modelos de lenguaje pueden introducir inferencias no solicitadas o generalizaciones indebidas; por ello, la supervisión del ingeniero de requisitos y la validación sistemática de resultados constituyen condiciones necesarias de aplicabilidad.

Resultados esperados: en etapas posteriores del proyecto se espera ampliar el conjunto de heurísticas formalizadas y extender su aplicación a otras actividades del *Proceso de Requisitos Basado en Escenarios*. En particular, se espera:

- consolidar la reducción del esfuerzo y del tiempo requerido para la ejecución de actividades del proceso;
- mejorar la reproducibilidad y uniformidad de los resultados obtenidos;
- facilitar la transferencia del conocimiento metodológico a ingenieros de requisitos menos experimentados;
- identificar de manera sistemática los límites del enfoque y los casos en los que la intervención humana resulta indispensable.

En conjunto, se espera que estos resultados contribuyan a una integración más controlada, responsable y metodológicamente fundada de la IAG en la IR.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación presentada es parte de la formación inicial en investigación de la Ingeniera María Inés Bertín y del Ingeniero Gerardo Agustín Riera.

5. BIBLIOGRAFÍA

[1] Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., Hadad, G.D.S., Ridao, M.N.: Defining system context using scenarios. In: Cheng, B.H.C. et al. (eds.) *Perspectives on Software Requirements*, pp. 169–199. Kluwer Academic Publishers, Boston (2004)

[2] Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* In: FAccT '21. ACM, 610–623. DOI: 10.1145/3442188.3445922

[3] Ji, Z., et al.: *Survey of Hallucination in Natural Language Generation*. arXiv:2202.03629 (v6).

[4] Bell, T. E., Thayer, T. A., *Software requirements: Are they really a problem?* Proceedings of the Second International Conference on Software Engineering (ICSE), San Francisco, CA, IEEE Computer Society Press, 1976, pp. 61–68.

[5] Pohl, K.: *Requirements Engineering: Fundamentals, Principles, and Techniques*. Springer (2010).

[6] Sommerville, I., *Software Engineering*, 9th ed., Addison-Wesley, Boston, 2011.

[7] H. Cheng, J. H. Husen, Y. Lu, T. Racharak, N. Yoshioka y N. Ubayashi, "Generative AI for Requirements Engineering: A Systematic Literature Review," Software: Practice and Experience, 2024.

[8] Ferrari, A. and Ginde, Gouri., "Handbook on Natural Language Processing for Requirements Engineering" Springer, 2025.

[9] Vogelsang, A., "Prompting the Future: Integrating Generative LLMs and Requirements Engineering," in Proceedings of the NLP4RE Workshop, 2024.

[10] Saleem, S., et al., "Generative language models potential for requirement engineering," Journal of Intelligent & Fuzzy Systems, 2025.

[11] Abbasi, M. A., Ihantola, P., Mikkonen, T., & Mäkitalo, N., "Towards Human-AI Synergy in Requirements Engineering: A Framework and Preliminary Study," arXiv preprint, 2025.

Go como Lenguaje de Diseño: Reducción de Complejidad Accidental en Sistemas Concurrentes

Valentín Torassa Colombero, Santiago Enrique Roatta, María Eugenia Casco

CAETI – Universidad Abierta Interamericana
Av. Montes de Oca 725 – Buenos Aires – Argentina

{Valentin.TorassaColombero, Santiago.Roatta, MariaEugenia.Casco}@uai.edu.ar

Resumen

La complejidad accidental constituye uno de los principales factores de fallo en sistemas de software modernos, especialmente en contextos concurrentes y distribuidos. Este trabajo analiza el lenguaje Go como caso de estudio en diseño de lenguajes orientado a reducir dicha complejidad. Desde un enfoque conceptual y técnico, se examinan decisiones centrales del lenguaje: simplicidad sintáctica, concurrencia estructurada, composición sobre herencia, modelo explícito de errores y tooling integrado; y su impacto en mantenibilidad, razonamiento y seguridad del software. El objetivo es argumentar que Go reduce deliberadamente el espacio de diseño posible para favorecer modelos mentales más simples, consistentes y auditables en sistemas reales.

Palabras clave: diseño de lenguajes, complejidad accidental, Go, concurrencia, ingeniería de software, mantenibilidad.

CONTEXTO

Los proyectos radicados en el CAETI se organizan en líneas vinculadas a Robótica, Ingeniería de Software y Ciberseguridad. Este trabajo se enmarca en Ingeneria de Software mostrnado decisiones de *diseño de lenguaje* y *tooling* que afectan propiedades deseables en software concurrente de larga vida (robustez, auditabilidad, mantenibilidad y reducción de defectos no deterministas). En entornos actuales (microservicios, pipelines de datos, agentes, sistemas de observabilidad), la concurrencia deja de ser opcional y se vuelve un requisito estructural; al mismo tiempo, incrementa el riesgo de fallos de difícil diagnóstico, con consecuencias operativas.

1. INTRODUCCIÓN

El desarrollo de software moderno enfrenta un problema persistente y ampliamente documentado: la brecha entre la complejidad inherente del dominio que se desea modelar y la complejidad introducida por herramientas, lenguajes y abstracciones usadas para implementarlo. En sistemas concurrentes, distribuidos o de larga vida, esta brecha tiende a ampliarse, generando software difícil de razonar, mantener y auditar.

En concurrencia, la complejidad se manifiesta mediante estados compartidos, interacciones no deterministas, dependencias temporales implícitas y flujos de control difíciles de seguir. Estos factores aumentan el costo de desarrollo y habilitan fallos que emergen tarde y con baja reproducibilidad, degradando confiabilidad y seguridad. En la práctica, muchos incidentes operativos se originan en defectos de diseño más que en errores algorítmicos.

Brooks introdujo la distinción entre complejidad esencial y complejidad accidental, argumen-

tando que esta última es consecuencia directa de las herramientas y abstracciones elegidas [4]. En ese marco, el diseño del lenguaje adquiere un rol central como mecanismo capaz de amplificar o mitigar la complejidad accidental. Este trabajo sostiene que Go fue diseñado explícitamente para reducirla en software real y que su valor principal reside en la coherencia de sus decisiones de diseño.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se propone una línea de I/D enfocada en **divulgación técnica y transferencia** de criterios de diseño aplicados a Go, estructurada en ejes:

- **C**omplejidad accidental y criterios de calidad.
- **C**oncurrencia estructurada (goroutines, channels, `select`, cancelación).
- **C**omposición e interfaces implícitas.
- **M**anejo explícito de errores (errores como valores) y control del flujo.
- **T**ooling integrado como extensión del lenguaje (formato, tests, análisis, dependencias).

El objetivo teórico es caracterizar cómo el lenguaje *restringe* el espacio de diseño para favorecer soluciones homogéneas y legibles. El objetivo experimental es producir *artefactos reproducibles* (ejemplos, guías, repositorios, plantillas) que permitan transferir estos principios a contextos académicos y proyectos reales.

3. GO COMO LENGUAJE DE DISEÑO

Go fue creado en Google con el objetivo explícito de mejorar la ingeniería de software a gran escala [7]. A diferencia de lenguajes que priorizan expresividad o generalidad, Go adopta un enfoque pragmático centrado en legibilidad, simplicidad y previsibilidad. Esto se refleja en un conjunto reducido de características: ausencia de herencia clásica, no uso de excepciones, metaprogramación limitada y un sistema de tipos explícito y simple.

La filosofía subyacente enfatiza la simplicidad deliberada: “When in doubt, leave it out” [10]. En Go, esto reduce la cantidad de formas equivalentes de escribir lo mismo, favoreciendo consistencia entre equipos y disminuyendo el costo de lectura y revisión del código. Pike también discute la importancia de priorizar claridad y economía conceptual en diseño [8].

En términos pedagógicos y de ingeniería, Go funciona como lenguaje *opinionated*: las restricciones son una estrategia para limitar soluciones frágiles y evitar complejidad accidental.

4. CONCURRENCIA Y RAZONAMIENTO EN GO

La concurrencia en Go se apoya en goroutines y channels, inspirados en CSP [6]. En lugar de centrar el diseño en memoria compartida y locks, Go se centra en la comunicación. Facilitando representar el sistema como *flujos de datos* y reduciendo estados compartidos.

Listing 1: Concurrencia estructurada con cancelación.

```
func worker(ctx context.Context, in <-
    chan int, out chan<- int) {
  for {
    select {
    case <-ctx.Done():
      return
    case v, ok := <-in:
      if !ok { return }
      out <- v * 2
    }
  }
}
```

El uso de `select` hace explícitas decisiones de recepción, salida y cancelación. Este patrón se alinea con el modelo de memoria de Go, basado en relaciones *happens-before*, que especifica condiciones bajo las cuales los efectos de una goroutine son visibles para otras [2]. Para un tratamiento práctico y sistemático del tema, Cox-Buday desarrolla patrones y anti-patrones de concurrencia en Go [5].

5. COMPOSICIÓN Y MODELO DE OBJETOS

Go evita jerarquías rígidas mediante composición explícita e interfaces implícitas. En lugar de relaciones “es-un”, promueve “tiene-un”, reduciendo acoplamiento estructural y favoreciendo evolución incremental. Las interfaces se definen por comportamiento; esto facilita sustituir imple-

mentaciones sin reestructurar jerarquías.

Listing 2: Composición e interfaces implícitas.

```
type Logger interface {
  Log(msg string)
}

type Service struct {
  Logger
}
```

Bodner resalta que este modelo favorece testabilidad y mantenibilidad en software real, donde los requerimientos cambian y el diseño debe adaptarse sin romper contratos [3].

6. ERRORES COMO VALORES: CONTROL EXPLÍCITO DEL FLUJO

Una decisión central de Go es evitar excepciones como mecanismo general de control de flujo. En su lugar, usa valores de error retornados explícitamente. Pike resume esta idea como "Errors are values" [9]. Desde la perspectiva de complejidad accidental, esto reduce caminos implícitos de ejecución y obliga a visibilizar fallos locales en el lugar donde ocurren.

Listing 3: Manejo explícito de errores y propagación.

```
func LoadConfig(path string) (*Config,
    error) {
  b, err := os.ReadFile(path)
  if err != nil {
    return nil, fmt.Errorf("read config:
        %w", err)
  }
  var c Config
  if err := json.Unmarshal(b, &c); err
      != nil {
    return nil, fmt.Errorf("parse config:
         %w", err)
  }
  return &c, nil
}
```

Este patrón mejora auditabilidad (el flujo de error es rastreable) y promueve contratos claros entre capas. En sistemas concurrentes incentiva a diseñar protocolos de cancelación y finalización explícitos, evitando *panics* difíciles de observar.

7. TOOLING COMO EXTENSIÓN DEL LENGUAJE

Go incorpora tooling como parte esencial del ecosistema: `gofmt` estandariza estilo; `go test` integra pruebas; `go vet` aporta análisis; y el detector de carreras refuerza detección temprana de defectos de concurrencia [1]. Esta integración reduce divergencias entre equipos y baja el costo de adoptar buenas prácticas.

Listing 4: Ejecución de tests con detección de carreras.

```
go test -race ./...
```

Desde una perspectiva sociotécnica, el lenguaje más tooling define un "camino feliz" de desarrollo: menor fricción y análisis implica mayor probabilidad de uso sistemático en equipos grandes.

8. RESULTADOS OBTENIDOS Y ESPERADOS

Resultados obtenidos: se consolidó una caracterización técnica de Go como lenguaje orientado a la reducción de complejidad accidental, identificando decisiones que afectan directamente el razonamiento (concurrencia estructurada, composición, errores explícitos y tooling integrado). Se elaboraron ejemplos didácticos y patrones de comunicación/cancelación, adecuados para su uso en clases, seminarios y formación aplicada.

Resultados esperados:

- Consolidar una línea de investigación aplicada al estudio de lenguajes modernos y su impacto en mantenibilidad.
- Transferir estos conceptos a proyectos reales (microservicios, herramientas internas, automatización).
- Generar material académico/didáctico y repositorios guía para programación concurrente segura en Go.

9. FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación se desarrolla en el CAETI e involucra docentes-investigadores y estudiantes avanzados.

10. DISCUSIÓN Y LÍMITES

Go no elimina la complejidad esencial ni garantiza ausencia de errores. Su enfoque restrictivo puede ser incómodo en dominios donde metaprogramación avanzada o jerarquías profundas son comunes. Sin embargo, este trabajo sostiene que la restricción deliberada funciona como control para reducir complejidad accidental y facilitar el razonamiento humano, especialmente en software de larga vida.

11. CONCLUSIONES

Go puede interpretarse como un caso de estudio exitoso en diseño de lenguajes orientado a reducir complejidad accidental. Sus decisiones coherentes promueven modelos mentales simples y auditables. El aporte principal del lenguaje es su capacidad de guiar a los desarrolladores hacia diseños mantenibles y consistentes en sistemas e implementaciones reales.

Referencias

[1] Introducing the go race detector. `https://go.dev/blog/race-detector`, 2013.

[2] The go memory model. `https://go.dev/ref/mem`, 2022.

[3] Jon Bodner. *Learning Go: An Idiomatic Approach to Real-World Go Programming*. O'Reilly Media, 2021.

[4] Frederick P. Brooks. No silver bullet—essence and accidents of software engineering. *IEEE Computer*, 20(4):10–19, 1987.

[5] Katherine Cox-Buday. *Concurrency in Go: Tools and Techniques for Developers*. O'Reilly Media, 2017.

[6] C. A. R. Hoare. Communicating sequential processes. *Communications of the ACM*, 21(8):666–677, 1978.

[7] Rob Pike. Go at google: Language design in the service of software engineering. `https://talks.golang.org/2012/splash.article`, 2012.

[8] Rob Pike. Less is exponentially more. `https://commandcenter.blogspot.com/2012/06/less-is-exponentially-more.html`, 2012.

[9] Rob Pike. Errors are values. `https://go.dev/blog/errors-are-values`, 2015.

[10] Ken Thompson. Reflections on trusting trust. `https://www.archive.ece.cmu.edu/~ganger/712.fall02/papers/p761-thompson.pdf`, 1984.

Hacia la Gestión y Reuso de Activos en Sistemas de Big Data basado en CBR orientado a Servicios

Andres Flores, Agustina Buccella, Alejandra Cechich, Alan De Renzis
GIISCo Research Group
Departamento de Ingeniería de Sistemas – Facultad de Informática
Universidad Nacional del Comahue
Neuquén, Argentina
Contacto: [andres.flores, agustina.buccella]@fi.uncoma.edu.ar

RESUMEN

El desarrollo de Sistema de Big Data (SBD) involucra grandes desafíos. La diversidad de los datos involucra distintas características, o Variedad: estructural, de las fuentes, de contenido y de procesa-miento. Identificar estas variedades genera desafíos de modelado que permitan una amplia reusabilidad: tanto de los datos como en la gestión del proceso de reuso. El proyecto de investigación actual propone un Proceso de Modelado de Variedad y Reuso de Activos en SBD, para identificar aspectos relevantes de variedad en distintos casos de analítica de datos. Se está desarrollando una plataforma que soporte el proceso de identificación, almacena-miento y reuso de casos de analítica de datos, incorporando capacidades de recomendación, trazabilidad y visualización. En un proyecto de investigación previo, enfocado en Computación Orientada a Servicios, se definió un Proceso de Razonamiento basado en Casos (CBR) para Descubrimiento y Selección de Servicios Web. La combinación de ambas investigaciones implicaría considerar que los activos de dominio de SBD podrían ser encapsulados como servicios, para ser tratados como casos en el proceso de CBR, y posteriormente ante la necesidad de analizar un caso nuevo, realizar descubrimiento de casos, y analizar los activos de dominio almacenados para encontrar oportunidades de reuso, tanto de activos de procesamiento como de conocimiento (y analítica de datos).

***Palabras Clave*:** Big Data – Variedad Reuso – Software Orientado a Servicios.

CONTEXTO

La línea presentada se inserta en el contexto del Proyecto 04/F019 "Modelado de Variedad en Sistemas Big Data". Financiado por UNCo. (2022-2026).

INTRODUCCIÓN

Los Sistemas Big Data (SBD) son sistemas inherentemente distribuidos, cuyas tecnologías y frameworks evolucionan rápidamente, lo cual genera grandes desafíos para arquitectos y diseñadores, particularmente sobre los requerimientos que una arquitectura para SBD debe cumplir; entre ellos, considerar las cinco "Vs" (*Volumen*, *Velocidad*, *Variedad*, *Variabilidad* y *Veracidad*) [DL20]. Un cambio importante con respecto a depósitos de datos tradicionales, es la naturaleza no estructurada de algunos datos que pueden provenir de diversas fuentes: sensores, redes sociales, entorno y la propia empresa cliente. La diversidad de esos datos puede analizarse considerando distintas características, lo cual en SBD se denomina *Variedad*, y es posible categorizarlas de acuerdo a una taxonomía que divide el análisis de variedad en cuatro casos de diversidad: *estructural*, de las *fuentes*, de *contenido* y de *procesamiento* [AJ15]. La diversidad *estructural* denota la variedad en formatos y tipos de datos, clasificándolos como estructurados, semi-estructurados y no estructurados; la diversidad de las *fuentes* se clasifica en tres grupos – datos generados por humanos (resultados de interacciones hombre-computador), generados por máquinas (automáticamente a partir de dispositivos, como sensores), o mediados por procesos (correlación

con datos generados por humanos); la diversidad de *contenido* considera diferentes tipos de soporte (único medio, multimedio o gráfico); y la diversidad de *procesamiento* se enfoca en las distintas necesidades de procesamiento algorítmico (batch, interactivo, streaming o gráfico). Identificar la variedad en SBD abre el desafío a alternativas de modelado, sobre todo analizando posibilidades de reusabilidad futura de los aspectos comunes (no variantes) de un SBD. En ese sentido, han surgido diferentes propuestas que enfocan en aspectos diversos del reuso – tanto de los datos como de la gestión del proceso de reuso. Un reuso efectivo requiere de un soporte que permita almacenar, identificar y eventualmente reusar recursos.

En este proyecto estamos desarrollando un Proceso de Modelado de Variedad y Reuso de Activos en SBD [BCMetal24], para lo cual actualmente se efectúan aplicaciones en el dominio de la agricul-tura inteligente, en colaboración con INTA Alto Valle, Río Negro. El proceso alcanzado permite identificar aspectos relevantes de variedad en distintos casos de analítica de datos, y se espera almacenar diferentes activos que faciliten el reuso en problemas similares. Con lo cual se está desarrollando una plataforma que soporte el proceso de identificación, almacenamiento y reuso de casos de analítica de datos, incorporando capacidades de recomendación, trazabilidad y visualización.

En un proyecto previo [FGCetal21] se enfocó la investigación en el paradigma de Computación Orientada a Servicios (SOC), que promueve fuertemente el reuso de software. La funcionalidad a ser reusada adopta la forma de servicios, o unidades lógicas que presentan entornos heterogéneos de ejecución y pueden ser ensambladas para formar otras unidades lógicas de mayor nivel de abstracción que resuelvan (directamente o en parte) los procesos de negocios para un contexto de aplicación [PTDL07]. El paradigma SOC encontró una plataforma potencial de aprovechamiento mediante la Web, desde donde se desarrolló la tecnología de Servicios Web [Wetal05], con notaciones formales específicas para descripción de las interfaces de servicios. Los Servicios Web pueden adoptar diferentes estilos: SOAP y RESTful [GFMCZ16]. Los servicios SOAP se describen mediante WSDL (Web Service Description Language) para especificar las operaciones e intercambio de mensajes. Existen diferentes versiones del lenguaje WSDL (1.0, 1.1, 2.0), siendo la versión WSDL 2.0 un estándar OMG. Los servicios RESTful se pueden describir mediante diferentes lenguajes, siendo WADL (Web Application Description Language) el primer intento de notación como analogía de WSDL. La industria ha propuesto otros lenguajes: OpenAPI/Swagger, REST API Modeling Language (RAML), o MIT's API Blueprint.

Evaluar y consumir tal variedad de servicios heterogéneos podría reducir el espacio de solución de los servicios a consumir, de acuerdo con la tecnología que se pueda soportar. De este modo, la amplia gama de servicios probablemente reusables, se vuelve seriamente limitada.

Para ello, hemos desarrollado un Metamodelo de Servicios Heterogéneos basado en estándares [GF19] para ampliar la gama de compatibilidad de servicios de diversas tecnologías, que pudieran ser potenciales candidatos para aplicaciones consumidoras. Los principales estándares para la descripción del servicio incluyen: WSDL 2.0, WADL y OpenAPI, considerando así ambos estilos de servicios, basados en SOAP y REST.

En función del Metamodelo de Servicios se ha construido un Proceso de Razonamiento basado en casos (CBR) para el almacenamiento, descubrimiento y selección de servicios [FGCetal21]. CBR resuelve problemas utilizando o adaptando soluciones a problemas antiguos [RS13], gracias a la función de similitud utilizada para cuantificar la semejanza entre un par de casos, basados en servicios, evaluando las descripciones de la interfaz WSDL. El Metamodelo de Servicios permite representar servicios candidatos heterogéneos como casos reutilizables. Los nuevos casos se pueden resolver a partir de varios casos similares anteriores, que se recuperan mediante funciones de similitud basadas en diferentes algoritmos *K-nn* configurables. El proceso CBR junto con el Metamodelo de Servicios proveen

soporte a la selección de servicios heterogéneos, agnósticos a la tecnología de una manera transparente. Esto aprovecha nuestro proceso CBR como un enfoque de gestión del conocimiento, para capturar y enriquecer el conocimiento obtenido de las selecciones sucesivas de servicios.

LÍNEAS DE INVESTIGACIÓN, Y DESARROLLO

Actualmente, abordamos aspectos sobre la efectiva detección y reuso de casos previos en analítica de datos, lo cual depende en gran medida de la representación de la documentación de los activos de dominio, incluyendo la variedad. El proceso de identificación desarrollado, es soportado actualmente por un primer prototipo, denominado COVAMAT, que requiere ser extendido para cumplir con los propósitos planteados [OCBetal23]. Entre esas extensiones se incluyen:

Reuso de casos: la versión actual de COVAMAT incluye una función de similitud para encontrar casos similares, de acuerdo a una consulta. Aunque la función está simplificada a una igualdad del contexto consultado y de los casos almacenados, que implica bajas capacidades de reuso de casos. Esta función debe ser ampliamente extendida en términos de estudios más avanzados sobre similitud, para identificar las variables más relevantes dentro de un dominio, que permiten determinar la similitud de casos; o para reconocer combinación de valores de contexto, que bajo ciertas condiciones, podrían presentar una similitud no inicialmente evidente.

Almacenamiento de recursos de los casos: hasta el momento COVAMAT posee una facilidad de almacenamiento acotada a los metadatos de los casos de dominio, considerando las variedades más relevantes. Por lo tanto, se requiere una gran extensión sobre el almacenamiento de diferentes activos del dominio con artefactos que permiten la definición de variedades para luego ser instanciadas al implementar un caso de dominio específico. De esta forma, para cada caso de dominio, la extensión requerida incluye agregar los mecanismos necesarios para almacenar: (1) los recursos que permitan identificar y recuperar las fuentes originales utilizadas, (2) las fuentes transformadas, es decir el resultado de las transformaciones durante el preprocesamiento, (3) los notebooks (códigos) trabajados, y (4) los resultados de las analíticas aplicadas. De esta forma, no sólo se almacenarían las variedades y las instanciaciones realizadas para casos de dominio, sino también los artefactos de software y recursos creados para su completo funcionamiento.

RESULTADOS OBTENIDOS / ESPERADOS

En [BCMetal25, FBCR25] hemos enumerado una serie de contribuciones previas, enfocadas en la identificación y modelado de variedad en SBD. El proceso desarrollado plantea dos enfoques para identificar variedad en un dominio: (1) identificación desde los datos o bottom-up (B-VIP) y (2) identificación desde los requerimientos o top-down (T-VIP). B-VIP considera la definición de un problema de un dominio e intenta detectar características variantes a partir de los datos existentes en uno o varios repositorios. T-VIP considera el análisis de requerimientos y del dominio para contrastar suposiciones con los datos. Las variedades de dominio son documentadas considerando sus metadatos mediante una hoja de datos (datasheet), que permiten la identificación posterior para un posible reuso de casos, detectando variedades de contexto.

Para analizar el impacto que diversas características de contexto podrían tener en las inferencias que se realizan a partir de los datos, se han efectuado aplicaciones en colaboración con INTA Alto Valle [CBetal23, BCMetal24, BCMetal25, BCGM26]: Identificación de variedad bottom-up para el análisis de calidad de cuerpos de agua superficiales; Caracterización de variables que influencian la turbidez de canales de riego; Caracterización de Variables para el Análisis del Índice de Vegetación; Caracterización de variables meteoro-lógicas que influencian las fluctuaciones del nivel freático; entre otros.

Actualmente, la facilidad en la recolección y almacenamiento de datos, y las tecnologías e infraestructuras para manipularlos, ha motivado

la creación de sistemas intensivos en datos (Big Data). En la práctica, el crecimiento de SBD obliga a plantear un cambio de paradigma de desarrollo: de dirigidos por la funcionalidad, a dirigidos por los datos. Es decir, el valor de los sistemas se asocia especialmente a la correcta identificación y explotación de los datos (y a su correspondiente calidad) [GK15].

En base a la posibilidad de reuso que provee el paradigma SOC, se han propuesto soluciones arquitectónicas para SBD con orientación a servicios, dado que la arquitectura de referencia estándar para SBD [NIST16] propone de manera conceptual una división de componentes funcionales lógicos conectados mediante interfaces, en la forma de servicios, que permiten interoperabilidad, y soportan la gran variedad en entornos de negocios, de una manera desacoplada, facilitando com-prender que Big Data complementa y difiere de los enfoques existentes en analítica, inteligencia de negocios, bases de datos y sistemas. Se han propuesto algunas arquitecturas para SBD [GK15, KB19]: Lambda, Kappa, basada en Microservicios, y Zetta, entre otras. La arquitectura Lambda, se convirtió en un estándar en SBD, y está dividida en tres capas: la capa de procesamiento batch de datos históricos; la capa de interacción con el usuario; y la capa de velocidad para procesar flujos de datos en tiempo real. La arquitectura Kappa, elimina la capa batch, con capacidades avanzadas en la capa de velocidad. La arquitectura de Micro-servicios [WT24], ofrece la ventaja de computación en la nube, con una estructura desacoplada de servicios REST. La arquitectura Zetta, es una variante orientada a Microservicios, pero a un nivel más alto de granularidad, para interconectar arquitecturas de negocio o empresariales.

En el contexto del Proceso de Modelado de Variedad y Reuso de Activos en SBD, y su herramienta de soporte COVAMAT, el almacenamiento de recursos y activos de dominio podrían ser tratados mediante servicios en la nube, como versiones de las arquitecturas Lambda y de Micro-servicios. Tanto los notebooks o códigos para procesamiento de colecciones de datos, como las mismas colecciones de datos, con sus versiones (original, post-procesado, etc.) podrían ser encapsulados en la forma de servicios, para ser almacenados en la nube. En este sentido, el Proceso de CBR [FGCetal21] para razonamiento y selección de servicios heterogéneos, podría ser ajustado de forma pertinente al contexto de SBD, para actuar como un motor de descubrimiento de servicios, con la capacidad de almacenamiento de activos y recursos de los casos de dominio generados, en el proceso B-VIP; y luego con la potencialidad de la recuperación y reuso, para el proceso T-VIP. Por otro lado, se deberán abordar los desafíos relacionados con el reuso de conocimiento de dominio, que plantea el proceso T-VIP en su totalidad. lo cual plantea nuevas oportunidades de investigación y desarrollo en SBD. La visión de esta línea de investigación se resume en [FBCR25]:

"Extender el prototipo COVAMAT como plataforma orientada a servicios basada en CBR, que soporte el proceso de identificación, almacenamiento y reuso de casos de analítica de datos, y gestión de activos de software, con capacidades de recomendación, trazabilidad y visualización".

FORMACIÓN DE RECURSOS HUMANOS

El proyecto completo reúne aproximada-mente a unos 15 investigadores, entre los que se cuentan docentes y estudiantes de UNComa, y colaboradoras expertas del dominio de aplicación, pertenecientes al INTA. Se cuenta con cuatro doctores y un magister. Tres de los docentes-investigadores están realizando carreras de postgrado en la temática del proyecto de investigación.

BIBLIOGRAFÍA

[AJ15] Abawajy, J. (2015). *Comprehensive analysis of big data variety landscape*. International Journal of Parallel Emerging Distributed Systems, 30(1):5–14.

[BCGM26] Buccella, A., Cechich, A., Garrido, W., Montenegro, A. (2026). *Contextual Reuse of Big Data Systems: A Case Study Assessing Groundwater Recharge Influences*. Applied Sciences, 16(3), 1650, 30p.

[BCMetal25] Buccella, A., Cechich, A., Mazalu, R., Luzuriaga, J., Villegas, C., Cruz, C., Corgatelli, F., Martinez, R., De Renzis, A., Arias, M., Torres, F. (2025). *Hacia el Desarrollo de Sistemas Big Data Basados en Reuso Intra e Inter-Dominio*. WICC, Workshop de Investigadores en Ciencias de la Computación, Mendoza.

[BCMetal24] Buccella, A., Cechich, A., Mazalu, R., Luzuriaga, J., Villegas, C., Cruz, C., Corgatelli, F., Martinez, R., Saurin, F., Torres, F. (2024). *Hacia una Plataforma para la Gestión de Variedad en Sistemas Big Data*. WICC, Workshop de Investigadores en Ciencias de la Computación, Santa Cruz.

[CBetal23] Cechich, A., Buccella, A., Villegas, C., Montenegro, A., Muñoz, A., and Rodriguez, A. (2023). *A model of reusable assets in AIE software systems*. Journal of Computer Science & Technology, 23(2):e13.

[DL20] Davoudian, D., Liu, M. (2020). *Big Data Systems: A Software Engineering Perspective*. ACM Computing Surveys, 53(5).

[FBCR25] Flores, A., Buccella, A., Cechich, A., De Renzis, A. (2025). *Hacia la Gestión de Activos en Sistemas de Big Data basado en CBR orientado a Servicios*. WICC, Workshop de Investigadores en Ciencias de la Computación. Mendoza.

[FGCetal21] Flores, A., Garriga, M., Cechich, A., Moyano, M., De Renzis, A., Anabalon, D., Corgatelli, F. (2021). *Reuso de Servicios Heterogéneos basado en CBR*. WICC, Workshop de Investigadores en Ciencias de la Computación. La Rioja.

[GF19] Garriga, M., Flores, A. (2019). *Standards-driven Metamodel to Increase Retrievability of Heterogeneous Services*. SAC, ACM/SIGAPP Symposium on Applied Computing, pp. 2507-2514.

[GFMCZ16] Garriga, M., Flores, A., Mateos, C., Cechich, A., Zunino, A. (2016). *RESTful service composition at a glance: A survey*. Journal of Network and Computer Applications, 60:32-53.

[GK15] Gorton, I., Klein, J. (2015). *Distribution, data, deployment: Software architecture convergence in Big Data systems*. IEEE Software, 32(3):78–85.

[KB19] Kalipe, G., Behera, R. (2019). *Big Data architectures: A detailed and application oriented review*. International Journal of Innovation, Technology and Exploratory Engineering, 8:2182–2190.

[NIST16] NIST (2016). NIST Big Data Interoperability Framework: vol6 Reference Architecture. Technical report, National Institute of Standards and Technology.

[OCBetal23] Osycka, L., Cechich, A., Buccella, A., Montenegro, A., and Muñoz, A. (2023). *COVAMAT: Functionality for variety reuse through a supporting tool*. Cloud Computing, Big Data & Emerging Topics, pp 57–74, Cham. Springer.

[PTDL07] Papazoglou, M.; Traverso, P.; Dustdar, S.; Leymann, F. (2007). *Service-Oriented Computing: State of the Art and Research Challenges*. IEEE Computer, 40(11): 38-45.

[RS13] Riesbeck, C., Schank, R. (2013). *Inside Case-based Reasoning*. Psychology Press.

[Wetal05] Weerawarana, S.; Curbera, F.; Leymann, F.; Storey, T.; Ferguson, D. (2005). *Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging, and More*. Prentice Hall PTR.

[WT24] Werner, S., Tai, S. (2024). *A reference architecture for serverless Big Data processing*. Future Generation Computer Systems, 155:179–192.

Hacia un Observatorio de Tendencias en Ingeniería de Sistemas y de Software en Contextos de Transformación Digital

Silvina Migani[1], María Inés Lund[1], Juan Aranda[1], Erica Minet[2], Juan Cruz Pantano[1], Cristina Vera[1], Oscar Flores[1] , Leonardo de los Rios[1*], Zamira Ale[1*], Germán Alvarez[1*], Maria Eugenia Moreira[1*], Juan Capdevila[1*], Franco Gil[1*], Marcelo Silva[1*], Francisco Oldecop[1*]

[1]Departamento de Informática, FCEFN, Universidad Nacional de San Juan
[1*]Alumno próximo a defender su trabajo final, Departamento de Informática, FCEFN - UNSJ
[2]Facultad de Arquitectura, Urbanismo y Diseño - UNSJ.

{silvina.migani, mlund}@unsj-cuim.edu.ar, jaranda@unsj.edu.ar, {ericaminet, juancruz17620, civerados, oscarflores.ing, leodelosriosfacultad, zamiranahirale, geralvarezn100, mariaeugeniamoreira086, juan.capdevila27, fmgil.44, mdsmarce, oldecop.francisco}@gmail.com

RESUMEN

La Ingeniería de Sistemas y la Ingeniería de Software atraviesan un proceso de transformación impulsado por la convergencia entre metodologías ágiles, automatización avanzada, tecnologías emergentes y la creciente centralidad de los datos como activo estratégico. En este escenario, se vuelve necesario distinguir entre tendencias ya consolidadas, enfoques en adopción temprana y desarrollos que podrían marcar nuevas direcciones en la evolución disciplinar.

El presente trabajo presenta los fundamentos conceptuales y el diseño inicial de un Observatorio de Tendencias orientado a relevar, organizar y analizar metodologías, arquitecturas, procesos y tecnologías emergentes en contextos de transformación digital. El modelo se estructura en ejes de investigación que contemplan la evolución de los procesos de desarrollo, los cambios en las arquitecturas de sistemas y la incorporación de la analítica y la inteligencia de negocios como componentes cada vez más relevantes en la ingeniería contemporánea.

Palabras claves: Ingeniería de Sistemas; Ingeniería de Software; Transformación Digital; Analítica e Inteligencia de Negocios; Metodologías de Desarrollo.

CONTEXTO

Esta presentación se enmarca dentro del proyecto "RADAR TECNOLÓGICO EN INGENIERÍA DE SISTEMAS Y DE SOFTWARE: ANÁLISIS INTEGRADO DE METODOLOGÍAS, PRÁCTICAS Y TECNOLOGÍAS ACTUALES Y EMERGENTES".

El equipo de trabajo está constituido por docentes/investigadores y alumnos de grado y postgrado que trabajan dentro del ámbito de la UNSJ.

1. INTRODUCCIÓN

El campo de la Ingeniería de Sistemas y de Software se encuentra inmerso en un proceso de transformación, impulsado por la evolución de tecnologías digitales, la adopción generalizada de metodologías ágiles, la incorporación sistemática de inteligencia artificial y la consolidación de modelos de desarrollo y operación altamente

automatizados. Las normas internacionales ISO/IEC/IEEE constituyen una base fundamental para la integración conceptual y operativa de los procesos de ingeniería en todo el ciclo de vida de sistemas y software, proporcionando definiciones, actividades y modelos de procesos que habilitan análisis comparativos sistemáticos, libres de ambigüedades.

La Ingeniería de Sistemas, de acuerdo con ISO/IEC/IEEE 15288 [1] y la guía del SEBoK [2], se define como un enfoque interdisciplinario e integrador orientado a transformar necesidades, expectativas y restricciones en soluciones viables y sostenibles durante todo su ciclo de vida. Esta disciplina articula conocimientos provenientes de la ingeniería, la gestión, las ciencias computacionales y las ciencias sociales, con el fin de asegurar que los sistemas diseñados respondan a requisitos operativos, tecnológicos y organizacionales. Su énfasis en la visión sistémica y en la trazabilidad desde las necesidades hasta la implementación y el retiro del sistema permite comprender fenómenos de creciente complejidad, como la integración de servicios distribuidos, la operación en infraestructuras híbridas, el uso de modelos predictivos o generativos y la interacción entre componentes físicos y digitales en arquitecturas híbridas.

La Ingeniería de Software, complementariamente, se concibe como la aplicación de enfoques sistemáticos, disciplinados y cuantificables al desarrollo, operación y mantenimiento del software. La definición establecida en ISO/IEC/IEEE 12207 [3] y recogida en el SwEBOK [4] incorpora tanto la perspectiva metodológica como la epistemológica, entendiendo la disciplina no solo como el ejercicio profesional que estructura el ciclo de vida del software, sino también como el estudio de sus métodos, teorías y prácticas. Esta doble naturaleza resulta esencial en un escenario donde la ingeniería de software generativa, el desarrollo asistido por inteligencia artificial, la automatización de pruebas y la inferencia automática de componentes están redefiniendo las capacidades de los ingenieros, las formas de organización del trabajo y los criterios de calidad del producto.

Dentro de ambas disciplinas pueden distinguirse cuatro dimensiones que permiten caracterizar sus enfoques y prácticas esenciales:

- Arquitectura de Software: Ocupa un lugar central dentro de este marco conceptual. Su definición como organización fundamental del sistema, expresada por sus componentes, sus relaciones y los principios que guían su diseño y evolución, sitúa a la arquitectura como un punto de convergencia entre la ingeniería de sistemas, la ingeniería de software y la ingeniería de requisitos. La norma ISO/IEC/IEEE 42010 [5] establece que toda arquitectura debe ser descripta a partir de modelos, puntos de vista y vistas que permitan capturar las preocupaciones de los distintos interesados y garantizar la consistencia entre la representación conceptual y su implementación física o lógica. De este modo, la arquitectura se convierte no sólo en un conjunto de artefactos, sino en un mecanismo de gobernanza que permite gestionar la complejidad, anticipar riesgos, evaluar calidad y mantener la coherencia del sistema durante su evolución.

- Ingeniería de Requisitos: Regulada por ISO/IEC/IEEE 29148 [6], constituye el puente conceptual que vincula necesidades y soluciones, y es un elemento crítico para la coherencia del ciclo de vida. Permite la especificación verificable y trazable de los requisitos que alimentan tanto el proceso arquitectónico como la planificación de desarrollo. En un entorno de gestión de proyectos mediante metodologías ágiles, enfoques orientados a productos y modelos DevOps o DevSecOps, la ingeniería de requisitos adquiere importancia al proporcionar mecanismos para gestionar incertidumbre, priorizar valor y asegurar alineación entre necesidades estratégicas y capacidades del sistema.
- Metodologías: Constituyen otro componente esencial del marco teórico. Según ISO/IEC/IEEE 15288 [1], una metodología se entiende como un marco sistemático que combina métodos, prácticas, procedimientos, roles, modelos de ciclo de vida y notaciones para guiar el desarrollo de un sistema o software. En la actualidad, la coexistencia de metodologías con enfoques ágiles automatizados o semiautomatizados, plantea el desafío de comprender su aplicabilidad en contextos diversos. Metodologías como Scrum, Kanban, XP, RUP, SAFe y LeSS conviven con marcos más recientes como DevOps, DevSecOps, MLOps y AIOps, los cuales redefinen la relación entre desarrollo, operación y seguridad mediante flujos automatizados, integración continua, despliegue continuo, infraestructura como código y observabilidad avanzada. Estos enfoques se articulan con tecnologías como contenedores, orquestadores, plataformas cloud, servicios gestionados, pipelines de datos y herramientas basadas en aprendizaje automático que transforman radicalmente las prácticas de ingeniería.
- Paradigmas emergentes y analítica avanzada: La expansión de tecnologías emergentes ha incorporado nuevos paradigmas como el desarrollo asistido por IA, el testing generativo, los agentes autónomos de ingeniería, la incorporación de plataformas low-code/no-code y el empleo de ciencia de datos y analítica BI/BA para la toma de decisiones [7].

En este escenario, se propone el diseño de un Observatorio de Tendencias que funcione como marco de referencia para el análisis estructurado de metodologías y tecnologías relevantes en la disciplina.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las tareas se han organizado alrededor de ejes temáticos que reflejan las principales dimensiones tecnológicas y metodológicas de la disciplina en su configuración actual. Esta delimitación inicial integra aspectos estructurales, organizacionales y orientados a datos, considerando su articulación en los procesos contemporáneos de ingeniería.

Los ejes temáticos que describen las líneas de investigación planteadas son:

1. Arquitectura y diseño de sistemas.
2. Procesos de desarrollo en contextos de transformación digital.
3. Ingeniería orientada a datos: bases de datos, analítica e inteligencia de negocios.

4. Tecnologías emergentes y nuevos paradigmas de ingeniería.

3. RESULTADOS

Si bien el proyecto marco de este trabajo se encuentra en su etapa inicial (presentado en la convocatoria de Proyectos de la UNSJ 2026-2027), su formulación se apoya en una trayectoria consolidada del equipo de investigación en temáticas vinculadas a arquitectura de sistemas, procesos de desarrollo e ingeniería orientada a datos. Estas experiencias permitieron advertir la necesidad de un análisis sistemático de las transformaciones en curso en la disciplina, lo que motivó la creación del Observatorio como ámbito de sistematización y reflexión académica.

Entre los resultados del 2025 se mencionan:

- Publicación del Trabajo: "De los datos al conocimiento: Construcción de soluciones de Analítica e Inteligencia Empresarial". XXVII Workshop de Investigadores en Ciencias de la Computación – XXVII WICC 2025 (UNCuyo, Mendoza, 10 y 11 de Abril). Autores: María Inés Lund, Silvina I. Migani, y otros. ISBN 978-987-575-267-2.
- Defensa del trabajo final de la Maestría en Inteligencia de Negocios y Explotación de Datos, de la Universidad Empresarial Siglo XXI (Escuela de Negocios y Posgrado de la Universidad Siglo XXI), desarrollada por el alumno Fernando Andrés Más, titulada: "Implementación de un Selfservice de Datos para la gestión de KPIs en la Administración Tributaria". Fecha de defensa: 21/03/2025.

Culminación de Becas de Investigación y Creación, categoría Estudiantes avanzados, financiadas por la UNSJ, para el año 2025. Res. N° 3182-R-2024:

- "Toma de decisiones apoyada en una solución de Inteligencia de Negocios. Caso de Aplicación: Ministerio de Educación de la provincia de San Juan", alumno Leonardo de los Rios.
- "Integración de Datos en el contexto de la Analítica e Inteligencia Empresarial: Herramientas Computacionales aplicadas en una Empresa de Telecomunicaciones de San Juan", alumno Juan Manuel Capdevila.
- "Apache Spark en un Contexto de Inteligencia de Negocios: Análisis de Datos de Transporte Público de Pasajeros de Corta Distancia de la Provincia de San Juan.", alumno Germán Álvarez Manzini.

Trabajos finales de grado en instancias de defensa:

- "Desarrollo de una solución de inteligencia de negocios aplicada al programa de finalización de estudios secundarios técnicos 2024". Leonardo de los Rios.
- "FARMADIP: Plataforma digital para la geolocalización de farmacias y solicitud de medicamentos mediante recetas electrónicas". Zamira Alé.
- "Apache Spark en un Contexto de Analítica e Inteligencia de Negocios: Análisis de Datos de Transporte Público de Pasajeros de Corta Distancia de la Provincia de San Juan". Germán Álvarez.
- "Sistema de integración de datos en contexto de la analítica e inteligencia empresarial, para contribuir a la mejora del servicio de telecomunicaciones de una empresa en la provincia de San Juan". Juan Capdevila.

- “Integración de Ciencia de Datos y Análisis Geoespacial para la generación de propuestas de Gestión eficiente y distribución inteligente de contenedores. Caso de estudio: Gestión de Residuos del Departamento Rivadavia”. María Eugenia Moreira y Franco Gil.
- “Migración de un Sistema Monolítico a una arquitectura de microservicios en el ámbito de una Fintech”. Francisco Oldecop.
- “Integración de la inteligencia artificial en plataformas de gestión de aprendizaje - LMS”. Marcelo Silva.

Los avances alcanzados hasta el momento corresponden al diseño preliminar del modelo del Observatorio, el cual recupera aprendizajes derivados de experiencias previas en investigación aplicada, particularmente en el análisis e integración de datos. En este sentido, se ha definido:

- una estructura de dominios temáticos;
- un esquema de categorización por niveles de adopción y madurez;
  una estrategia metodológica de carácter mixto, que combinará técnicas cualitativas y cuantitativas;
- un plan de trabajo organizado en fases sucesivas de preparación, relevamiento, análisis y sistematización.

4. FORMACIÓN DE RECURSOS HUMANOS

Hay una marcada propuesta de formación de recursos humanos, articulando las actividades de investigación con instancias de formación de grado y posgrado. En este marco, el Observatorio se proyecta como espacio de referencia para el desarrollo de trabajos finales y tesis vinculadas a metodologías de desarrollo, arquitectura de sistemas, transformación digital e ingeniería orientada a datos.

El equipo de trabajo ha dirigido y acompañado tesis de grado y posgrado en temáticas relacionadas con metodologías de ingeniería, tecnologías emergentes, bases de datos y analítica. Actualmente, el proyecto cuenta con la participación de 18 alumnos: 2 tesistas de posgrado (Maestría) y 16 tesistas de grado, de los cuales 8 se encuentran en instancias de defensa para la obtención del título de Licenciado (marzo 2026).

7. BIBLIOGRAFÍA

[1] *ISO/IEC/IEEE 15288:2023 - Systems and software engineering — System life cycle processes, 2nd ed. ISO/IEC/IEEE, 2023.*

[2] Guide to the Systems Engineering Body of Knowledge v2.10.pdf - SEBoK, vol. 2.10. INCOSE- IEEE SYSTEMS COUNCIL - STEVENS, 2024.

[3] ISO.org, ISO/IEC/IEEE 12207:2017 - Systems and software engineering - Software life cycle processes. 2017.

[4] IEEE Computer Society, Guide to the Software Engineering Body of Knowledge (SWEBOK v4.0). 2024.

[5] ISO/IEC/IEEE 42010:2022 - Software, systems and enterprise — Architecture description, 2nd ed. ISO/IEC/IEEE, 2022.

[6] ISO/IEC/IEEE 29148:2018 - Systems and software engineering — Life cycle processes — Requirements engineering, 2nd ed. ISO/IEC/IEEE, 2018.

[7] J. Reis and M. Housley, *Fundamentals of Data Engineering: Plan and Build Robust Data Systems*. O’Reilly Media, 2022.

HERRAMIENTA TECNOLÓGICA PARA LA GESTIÓN DEL CONOCIMIENTO EN LABORATORIOS DE INFORMÁTICA FORENSE

Carlos Orellana, Cecilia Lara, Liliana Figueroa, María de los Ángeles Valdez

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero ;
orellanacarlos3003@gmail.com, laraceciliacristina@gmail.com; lmvfigueroa@yahoo.com.ar;
mariavaldez_lsi@hotmail.com

RESUMEN

Desde el año 2017 a la fecha, a través de distintos proyectos de investigación los autores vienen consolidando una línea de investigación sobre herramientas / tecnologías que permitan gestionar y apoyar el trabajo de manera eficiente de los Laboratorios de Informática Forense.

Debido a que actualmente no existe una cultura de intercambio de experiencias que se adquieren durante la realización de las pericias ni se fomenta el autoaprendizaje organizacional en estos entornos de trabajo.

A partir de lo enunciado, surge esta propuesta como una continuidad a la línea de investigación que propone una herramienta tecnológica que promueva la colaboración y la gestión del conocimiento de los peritos informáticos y promueve mejorar el desempeño durante los procesos de obtención y análisis de evidencias digitales.

Palabras clave:

Informática forense, peritos informáticos, evidencias digitales, herramientas tecnológica gestión del conocimiento, laboratorios de informática forense.

CONTEXTO

Este artículo documenta el resultado de la línea de investigación Estudio y Desarrollo de herramientas tecnológicas para el apoyo de la Gestión de los Laboratorios de informática forense, enmarcado en el proyecto de investigación "Tecnologías de Apoyo a la Informática Forense" (23/C223-bint-2025) [2], vigente desde el 1 de enero de 2025 hasta el 31 de diciembre de 2026, que se desarrolla en el Instituto de Investigación en Informática y Sistemas de Información (IIISI), perteneciente a la Universidad Nacional de Santiago del Estero (UNSE).

1. INTRODUCCIÓN

En el ámbito de la informática forense, para un perito resulta cada vez más complejo mantenerse actualizado, siendo un reto constante identificar dónde y cómo extraer evidencias digitales en los dispositivos emergentes, debido a la diversidad de esquemas de almacenamiento entre fabricantes y a la continua evolución de los modelos [10]. En este contexto, surge el siguiente interrogante: ¿Y si existiera una herramienta de software que integre conocimientos y experiencias diversas en un repositorio unificado, facilitando así la colaboración y el intercambio de información entre los peritos?

Además de estar al tanto de las tecnologías emergentes y nuevas fuentes de evidencia, los peritos deben evaluar críticamente la precisión y la aplicabilidad de cada técnica o herramienta utilizad [5] A partir de esto surge el interrogante de ¿cómo la experiencia de un perito sirve de ayuda para abordar la tarea solicitada respecto de un dispositivo en particular?

Como respuesta a estas problemáticas, el presente desarrollo propone la creación de una herramienta tecnológica que sirva como un espacio colaborativo orientado al aprendizaje organizacional y a la capitalización de conocimiento en el ámbito forense digital, contribuyendo al fortalecimiento de la labor de los peritos y al avance de la disciplina. Esta solución tecnológica no gestiona evidencia ni investigaciones en curso, sino que se limita a facilitar la gestión y reutilización de experiencias validadas, sirviendo como apoyo complementario a la práctica forense.

El trabajo se alinea con los objetivos del IIISI al promover la integración entre investigación y desarrollo tecnológico, aplicando metodologías de ingeniería de software, bases de datos e ingeniería web [9].

1. INTRODUCCIÓN

La Informática Forense, es una disciplina criminalística, que tiene por objeto, la investigación en sistemas de Tecnologías de Información de hechos con relevancia jurídica o para simple investigación privada. Para conseguir sus objetivos, los expertos se basan en técnicas idóneas para ubicar, reproducir y analizar evidencia digital.

De esta manera, el análisis forense es una disciplina que permite identificar, analizar, preservar y presentar evidencia digital obtenida de infraestructura tecnológica que sea válida en un proceso legal [3]. Los elementos informáticos que se disponen en los celulares, las computadoras o los dispositivos de almacenamiento como imágenes, mensajes de texto, conversaciones de WhatsApp, videos, ubicación, etc. constituyen las evidencias necesarias para resolver diferentes tipos de casos judiciales [6]. Todos estos elementos son trabajados en los laboratorios de informática forense, en donde se realizan las extracciones forenses mediante hardware y software forense de alta precisión. La información resultante es procesada bajo estrictos protocolos de cadena de custodia para transformarla en evidencia digital reproducible y veraz [4].

No obstante, la labor diaria del perito informático enfrenta desafíos dinámicos derivados de la obsolescencia acelerada y la sofisticación de los vectores de ataque. Factores como la diversidad de sistemas operativos, el cifrado de datos y la complejidad intrínseca de la evidencia digital, sumados a las exigencias éticas y legales, demandan una actualización constante.

Ante este escenario, la gestión de experiencias surge como una alternativa estratégica para abordar la complejidad de la disciplina. Esta gestión no solo implica la recopilación de datos, sino la creación de un ecosistema que facilite la transformación del conocimiento tácito (aquel adquirido por la práctica y la intuición del perito) en conocimiento explícito (documentado y estandarizado). Al fomentar un entorno de colaboración y aprendizaje organizacional, se garantiza que el saber acumulado sea accesible, aplicable y escalable frente a nuevos y más complejos contextos de investigación [7].

2. DESCRIPCIÓN DE LA INVESTIGACIÓN Y DESARROLLO

La propuesta de solución consistió en desarrollar un prototipo de aplicación web que permita a los peritos informáticos registrar, organizar y compartir sus experiencias profesionales de manera estructurada y segura [8].

Disponer de esta solución permite a la institución gestionar de manera sistemática las experiencias de los peritos informáticos, centralizando y facilitando el acceso a los conocimientos previamente adquiridos. Esto posibilita su consulta y reutilización de forma ágil y estructurada, favoreciendo su aplicación eficiente en las investigaciones forenses digitales [1]. Asimismo, contribuye al fortalecimiento del aprendizaje organizacional y a la mejora continua de las prácticas dentro de la oficina forense.

El objetivo general de la investigación propuesta es: desarrollar una herramienta tecnológica que permita gestionar y compartir la experiencia de los peritos informáticos.

El objetivo general se alcanzó a partir de la definición de los siguientes Objetivos Específicos:

- Investigar estrategias para capturar y almacenar experiencias.
- Realizar el modelado del prototipo.
- Identificar y especificar los metadatos que son necesarios para caracterizar una experiencia de un perito informático.
- Estudiar herramientas para gestionar experiencias.

Las actividades de investigación se desarrollaron en cuatro fases:

En la Fase 1 se realizó una revisión de la literatura sobre gestión del conocimiento y

experiencias en el ámbito forense digital, complementada con un análisis de necesidades y requerimientos de los peritos informáticos. A partir de estos hallazgos, se formularon criterios clave para la gestión y estructuración de experiencias dentro de un sistema.

La Fase 2 consistió en el modelado del prototipo, donde se definieron los requerimientos funcionales y no funcionales, y se diseñaron los casos de uso, diagramas UML y diseño de interfaces de usuarios que representan la estructura general y las interfaces de usuario del sistema.

Durante la Fase 3, se llevó a cabo un análisis de herramientas existentes, incluyendo el Repositorio Nacional de Inteligencia Forense Digital (NRDFI) y la plataforma IBM QRadar SIEM, con el objetivo de identificar ventajas, desventajas y funcionalidades aplicables al sistema propuesto. Este análisis permitió incorporar mecanismos avanzados de búsqueda por palabras clave, metadatos y búsquedas combinadas a un sistema de gestión de conocimiento, mejorando la accesibilidad y recuperación de información.

Finalmente, en la Fase 4, se trabajó en la definición de los metadatos que describen las experiencias forenses, identificando los atributos esenciales para su caracterización, clasificación y posterior recuperación.

Los criterios orientadores para el diseño y desarrollo de la herramienta tecnológica para la gestión de experiencias de los peritos informáticos se presentan en la tabla 1.

3. RESULTADOS OBTENIDOS

El resultado de esta investigación fue el desarrollo de una herramienta tecnológica para la gestión de experiencias de los peritos informáticos, que integra principios de gestión del conocimiento, ingeniería web y usabilidad. Esta solución representa un aporte innovador al ámbito de la informática forense, al permitir preservar, organizar y reutilizar experiencias profesionales, fortaleciendo la formación de los peritos y la eficiencia en las investigaciones futuras.

Tabla 1: Criterios para el diseño y el desarrollo del prototipo.

Criterio	Descripción	Implicancia práctica
Capitalización del conocimiento	El sistema debe permitir recopilar y documentar las experiencias de los peritos informáticos, transformando el conocimiento tácito en explícito para su reutilización.	Evita la pérdida de aprendizajes valiosos y posibilita que la experiencia acumulada se convierta en un activo organizacional.
Accesibilidad y recuperación	Las experiencias registradas deben almacenarse en un repositorio estructurado que facilite la búsqueda por distintos criterios relacionados a la experiencia.	Permite el acceso rápido a antecedentes relevantes y mejora la eficiencia en futuras investigaciones.
Actualización y aprendizaje continuo	La plataforma debe admitir la incorporación constante de nuevas experiencias.	Promueve la actualización profesional y la adaptación a los avances tecnológicos en informática forense.
Usabilidad y escalabilidad	El sistema debe ser intuitivo, flexible y adaptable a las necesidades cambiantes de los peritos.	Facilita su adopción por los usuarios y su ampliación futura con nuevos módulos o funcionalidades.
Apoyo a la toma de decisiones	La información registrada debe servir como base para el análisis de casos y la mejora de las prácticas periciales.	Reduce errores por falta de experiencia y fortalece la calidad y consistencia de las investigaciones.

A si mismo los resultados obtenidos están alineados con los objetivos de la línea de investigación debido a que se ha podido:

a. Estudiar sistemáticamente herramientas de informática forense.

b. Desarrollar aplicaciones que permitan dar apoyo a las tareas del perito informático.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por un grupo de investigadores que vienen trabajando en la línea de informática forense desde el año 2017 en proyectos relacionados con la temática.

Además, para trabajar en el desarrollo de aplicaciones que mejoren la funcionalidad de las herramientas forenses y el desarrollo de soluciones informáticas para la gestión eficiente de los laboratorios forenses se incorporaron egresados y estudiantes avanzados de la carrera Licenciatura en Sistemas de Información.

La participación de profesionales que desarrollan su actividad profesional en el ámbito judicial y vinculados a las ciencias forenses posibilita concretar soluciones acordes a las necesidades de las instituciones judiciales y además realizar pruebas e implementaciones de los resultados de la investigación.

5. REFERENCIAS

[1] Biros, D. P., Weiser, M., & Witfield, J. (2007). Managing digital forensic knowledge an applied approach.

[2] Figueroa L, Lara C, Viaña G, Méndez A. Tecnologías de apoyo a la informática forense. Workshop de Investigadores en Ciencias de la Computación. Facultad de Ingeniería de la UNCUYO. 10 y 11 de abril de 2025.

[3] Iglesias LR. Herramientas open source para Informática forense. Carrera de Especialización en Seguridad Informática, Facultades de Ciencias Económicas, Ciencias, Exactas y Naturales e Ingeniería, Universidad de Buenos Aires. (2015). http://bibliotecadigital.econ.uba.ar/download/tpos/1502-0187_IglesiasLF.pdf

[4] ISO/IEC 27037:2012. Information technology — Security techniques — Guidelines for identification, collection, acquisition and preservation of digital evidence https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27037:ed-1:v1:en

[5] Kahvedžić, D., & Kechadi, T. (2009). DIALOG: A framework for modeling, analysis and reuse of digital forensic knowledge. digital investigation, 6, S23-S33.

[6] Kebande, V. R., Karie, N. M., Choo, K. K. R., & Alawadi, S. (2021). Digital forensic readiness intelligence crime repository. Security and Privacy, 4(3), e151.

[7] Nonaka, I., & Takeuchi, H.. The Knowledge-Creating Company. Oxford University Press. (1995**).**

[8] Orellana, C., Lara, C., & Figueroa, L. M. (2024). Prototipo web para gestión de experiencias de Informática Forense. En Libro de Actas — 30° Congreso Argentino de Ciencias de la Computación – *CACIC 2024* (pp. 1616–1620). Universidad Nacional de La Plata. Repositorio SEDICI. http://sedici.unlp.edu.ar/handle/10915/17752 (Sedici)

[9] Pressman, R. S. Ingeniería del Software. Un enfoque práctico. McGraw-Hill Interamericana, 9ª edición. México, 2021.

[10] X. Zhang, K. -K. R. Choo and N. L. Beebe, "How Do I Share My IoT Forensic Experience With the Broader Community? An Automated Knowledge Sharing IoT Forensic Platform," in IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6850-6861, Aug. 2019, doi: 10.1109/JIOT.2019.2912118.

IMPACTO DE LA GESTIÓN DEL CONOCIMIENTO SOBRE LA MOTIVACIÓN DE LOS/AS TRABAJADORES/AS

Nocenti, Florencia[1]; Straccia, Luciano[1,2]; Gonzalez, Natalia[3]

[1]Universidad Tecnológica Nacional, Facultad Regional La Plata, Programa de Maestría en Administración de Empresas, La Plata, Argentina
florencianocenti@gmail.com

[2]Universidad Tecnológica Nacional, Facultad Regional Buenos Aires, Grupo GEMIS.BA, Ciudad Autónoma de Buenos Aires, Argentina
lstraccia@frba.utn.edu.ar

[3]Universidad Nacional de General Sarmiento, Instituto de Industria, Provincia de Buenos Aires, Argentina
ngonzale@campus.ungs.edu.ar

RESUMEN

La gestión del conocimiento (GC) aborda la creación y el uso del conocimiento en las organizaciones desde un enfoque multidisciplinar. El Grupo GEMIS, a través de diversos proyectos, se propone analizar la gestión del conocimiento desde sus diversos enfoques y perspectivas integrales. Recientemente se identifica la necesidad de ampliar el análisis asociado a la perspectiva de los trabajadores en relación con su motivación y la influencia de la gestión del conocimiento, profundizando la línea de investigación en Gestión del Conocimiento. Además, este proyecto articula propuestas de diversas Unidades Académicas generando aportes sustantivos para un trabajo interinstitucional en la formación de recursos humanos.

Palabras clave: gestión del conocimiento, motivación.

CONTEXTO

La Universidad Tecnológica Nacional (UTN) posee un Programa de Investigación, Desarrollo e Innovación, de Sistemas de Información e Informática. En la Facultad Regional Buenos Aires (UTN.BA), en el marco del Departamento de Ingeniería en Sistemas de Información, se llevan adelante proyectos de investigación y desarrollo. Allí desarrolla sus actividades, entre otros grupos de investigación, el grupo GEMIS.BA que trabaja en el campo de la Ingeniería en Sistemas de Información, Ingeniería de Software e Ingeniería del Conocimiento.

El Grupo GEMIS.BA realiza, en el marco de las temáticas objeto de este trabajo, diversos proyectos asociados a la línea de Gestión del Conocimiento con proyectos de investigación y desarrollo y actividades de docencia e investigación en diferentes instituciones contando con trabajos finales de especialización, tesis de maestría y un plan de tesis doctoral en proceso.

Diversos integrantes del Grupo forman parte de los equipos docentes de los programas de Maestría de la Universidad Tecnológica Nacional, Facultad Regional La Plata a través del dictado de diversos seminarios y la dirección de tesis de maestría.

1. INTRODUCCIÓN

El conocimiento organizacional es una herramienta fundamental para el desarrollo sostenible y la obtención de ventajas competitivas [1]. Según Nonaka [2], en un contexto de incertidumbre, la única fuente confiable de ventaja competitiva es el conocimiento. La gestión del conocimiento (GC) implica planificar, organizar, dirigir y controlar los recursos de una organización, transformándolos en valor agregado [3]. Patrick Sullivan [4] define el capital intelectual como aquellos conocimientos propiedad de la organización (explícito) o de sus miembros (tácito), que crean o producen valor. Una gestión efectiva del conocimiento convierte el tácito en explícito, fortaleciendo el capital intelectual de la organización y con ello su capacidad de diferenciarse. Por ser el trabajador el principal portador del conocimiento organizacional, resulta esencial reconocer y gestionar de forma adecuada su aporte a través del liderazgo participativo, el trabajo en equipo y una cultura enfocada en el alto rendimiento, ya que constituye un pilar clave para el desarrollo y la competitividad de la organización.

Para que una organización pueda implementar un sistema integral de GC resulta esencial que adopte una visión centrada en las personas. Esto implica poner el foco en la motivación de las personas y en el modo en que ésta influye tanto en el logro de los objetivos individuales como en los organizacionales. Cada colaborador posee necesidades y motivaciones únicas, por ello, las organizaciones deben identificarlas y atenderlas, orientando su potencial hacia un desempeño que impulse la productividad y fomente la innovación. La motivación, entendida como el motor interno de la conducta humana [5], surge de necesidades fisiológicas o psicológicas que generan tensión hasta ser satisfechas. Soriano [6] resalta que las personas motivadas extrínsecamente buscan recompensas externas, mientras que las motivadas intrínsecamente actúan por el placer de realizar la tarea. Lograr un equilibrio entre la motivación extrínseca e intrínseca es clave. Un equipo motivado no solo obtiene beneficios económicos, sino que también potencia el sentido de pertenencia y la generación de nuevo conocimiento. Para promover una cultura de GC que fomente la innovación y la competitividad, a través del talento humano que es lo más valioso que tiene una compañía, éstas deben propender a capacitar, calificar, desarrollar y retener el capital humano, potenciando no sólo las habilidades prácticas, sino también complementar el conocimiento existente [7]. La GC implica dinámicas de interacción continua y depende de la disposición de los colaboradores a compartir su conocimiento [8], especialmente el conocimiento tácito que, por su naturaleza subjetiva e intuitiva, requiere de incentivos para su transferencia [9]. Según este autor, el liderazgo comprometido y la formación continua del talento humano son esenciales para transformar el conocimiento tácito en explícito, fomentar la creatividad, el pensamiento crítico y fortalecer la innovación.

De las principales investigaciones relevadas sobre la temática de estudio, se observa que ninguna de ellas aborda directamente el tema propuesto, sino que aportan elementos clave desde otras perspectivas.

Han y Anantatmula [10] buscaron examinar los factores culturales, tecnológicos, motivacionales y organizacionales que influyen en la difusión de conocimiento dentro de la organización desde la perspectiva de los trabajadores. En la misma línea, Chung Jen y Jing Wen [11] examinaron los efectos del clima y la estructura organizacional sobre la GC desde una perspectiva de interacción social. Straccia et al [12] se enfocaron en que la GC cuenta con muchas aristas que permiten desarrollar iniciativas dentro de las organizaciones para mantener y potenciar el conocimiento de las personas. Por ello proponen identificar cuáles son los aspectos que deben considerarse para

construir un modelo integral de gestión del conocimiento cuestionando cuáles son estas posturas en la literatura académica, cuál es su presencia actual y si hay alguna sin identificar. El trabajo de investigación de Fadaie et al. [13] aporta un enfoque interesante sobre las causas que pueden hacer fallar a un sistema de GC. Los autores investigaron si existe una relación entre GC, rendimiento organizacional y satisfacción laboral entre los empleados de la Red de Salud y Tratamientos de la ciudad de Haji Abad en Irán. El enfoque de los autores Atapattu y Huybers [14] nos permite un mayor acercamiento al tema de estudio. Ellos se plantearon como objetivo investigar el efecto de diferentes prácticas organizacionales de GC sobre el compromiso de los colaboradores con la GC y sobre el desempeño general de la compañía a través del compromiso individual de los colaboradores con la GC. Alineados con este punto de vista, en [15] evaluaron en qué medida la GC influye en el rendimiento de los individuos y de las organizaciones y analizaron el efecto mediador de la motivación y la innovación en la relación entre la GC y el rendimiento individual y organizacional. Meher y Mishra [16] buscaron comprender los efectos de las prácticas de GC (creación, difusión y retención) en empresas IT de India sobre los beneficios percibidos y la satisfacción de los trabajadores. Es importante también considerar el aporte de Pangaribuan y Satrya [17] y el enfoque puesto no sólo en la GC sino también sobre el estilo de liderazgo de las organizaciones. Los autores investigaron la influencia de las prácticas de GC y el liderazgo transformacional sobre el rendimiento de los trabajadores del organismo financiero gubernamental en Indonesia, tomando el compromiso organizacional como mediador. Nuralfya, Priatna y Burhanuddin [18], realizaron aportes sobre la relación que existe entre la motivación de los trabajadores y su consecuente productividad laboral.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el año 2018 se llevó a cabo el primer proyecto de investigación y desarrollo (PID) del Grupo GEMIS.BA asociado a GC en la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional. Este proyecto buscaba indagar cuál era el grado de implementación de GC en las pequeñas y medianas fábricas de software del Área Metropolitana de Buenos Aires (AMBA), cuáles eran los problemas presentes en esta implementación y qué modelos se podrían proponer para que esas organizaciones pudieran implementarlos (PID UTN, 2018-2020). Este proyecto obtuvo resultados construyendo un modelo general pero que se centraba especialmente en las actividades, desconociendo (o minimizando) otros aspectos.

A partir de los resultados obtenidos en el proyecto anterior y atendiendo a la necesidad de indagar en algunos aspectos particulares de la GC, en el año 2021 se iniciaron proyectos asociados a aspectos tecnológicos y a la medición de algunos de esos aspectos. En la UTN.BA se realizó un proyecto que busca analizar modelos de representación del conocimiento y proponer tecnologías y arquitecturas tecnológicas para la implementación de GC en las organizaciones (PID UTN 2021-2024). Además, entre los años 2021 y 2023 uno de los investigadores del Grupo dirigió un proyecto en la Universidad Argentina de la Empresa (UADE) cuyo objetivo fue definir un modelo de medición de GC construyendo indicadores sobre las tecnologías para la medición de GC y proponiendo un modelo de implementación de los indicadores en las organizaciones (PID UADE, 2021-2023). Finalmente, en los últimos años se inició el proyecto de desarrollo de un observatorio tecnológico para relevar y monitorear el estado de aplicación de la gestión del conocimiento en organizaciones de la Ciudad de Buenos Aires (PID UTN, 2023-2025) y la creación de un

modelo integral de gestión del conocimiento para las PyMes argentinas (PID UTN, 2025-2027).

A partir de los resultados obtenidos de los proyectos mencionados se identifica la necesidad de ampliar el análisis asociado a la perspectiva de los trabajadores en relación con su motivación y la influencia de la gestión del conocimiento. Este proyecto se inscribe como una continuidad y profundización de la línea de investigación en Gestión del Conocimiento, enfocándose específicamente en el análisis de estos aspectos.

3. RESULTADOS OBTENIDOS/ESPERADOS

En [12] se exploran diferentes vistas o dimensiones a considerar en la gestión del conocimiento a través de un modelo integral: 1) individuos/personas; 2) aspectos organizacionales; 3) actividades y procesos; 4) tecnología y representación del conocimiento; 5) medición. En [19] se hace foco en los aspectos tecnológicos analizando y categorizando los conceptos asociados a la tecnología para la gestión del conocimiento, posteriormente ampliando en [20] en un enfoque de algunas tecnologías y la innovación.

El objetivo principal de este proyecto consiste en analizar la relación entre un sistema de gestión del conocimiento y la motivación de los/as trabajadores dentro de las organizaciones. Específicamente se busca: 1) analizar los conceptos de gestión del conocimiento y motivación laboral para establecer las bases conceptuales del estudio; 2) diseñar y aplicar una encuesta para medir el grado de conocimiento que poseen los trabajadores de distintas organizaciones sobre la existencia e implementación de Sistemas de Gestión del Conocimiento (SGC); 3) identificar variables relacionadas con la motivación de los trabajadores que surjan como consecuencia de la implementación de prácticas de gestión del conocimiento con el fin de orientar a las organizaciones en su aplicación.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo propuesto se encuentra conformado por investigadores formados y tesistas de diversas Unidades Académicas. De esta forma resulta un proyecto que integra esfuerzos de diferentes Universidades como la Universidad Tecnológica Nacional y la Universidad Nacional de General Sarmiento

Esta línea de trabajo busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para que asciendan dentro de la carrera de investigadores, además de fomentar la aplicación de las temáticas en su propia actividad profesional.

Además el equipo cuenta con investigadores estudiantes que realizan sus primeras participaciones en el ámbito de la investigación y cuya experiencia en el proyecto propuesto contribuirá especialmente al inicio de trayectos de formación académica. Se espera la incorporación de nuevos becarios alumnos y graduados a partir de las diferentes propuestas de becas de investigación que son llevadas a cabo por las Unidades Académicas involucradas.

Este proyecto ya cuenta con el desarrollo de un plan de tesis para la Maestría en Administración de Negocios de la Universidad Tecnológica Nacional, Facultad Regional La Plata de la autora principal de este trabajo.

5. BIBLIOGRAFÍA

[1] Davenport, T. H., & Prusak, L. (1998). Working Knowledge: How Organizations Manage What They Know. Harvard Business School Press.

[2] Nonaka, I. (1991). The knowledge-creating company. Harvard Business Review, 69 (6).

[3] Rivera Martínez, W., Canay Pazos, J. (2019). Ecosistema de emprendimiento e innovación en Cauca, Colombia. Revista Venezolana de Gerencia, 24 (87), 922–937.

[4] Sullivan, P. H. (1999). Profiting from intellectual capital. Journal of Knowledge Management, 3 (2), 132-143.

[5] Carrillo, M., Padilla, J., Rosero, T., Villagómez, M. S. (2009). La motivación y el aprendizaje. Alteridad. Revista de Educación, 4 (2), 20-32.

[6] Soriano, M. M. (2001). La motivación, pilar básico de todo tipo de esfuerzo. Proyecto Social: Revista de Relaciones Laborales (Universidad de Zaragoza), 9, 163-184.

[7] Obando Changuán, M. P. (2020). Capacitación del talento humano y productividad: Una revisión literaria. Revista de investigaciones en Ciencias Administrativas y Economía (Universidad Técnica de Manabí), 11 (2).

[8] García, M. A. (2008). Modelo de acción sobre las dimensiones del clima organizacional: impacto de implementación de gestión de conocimiento en las organizaciones. Tesis de Maestría en Ingeniería y Gestión de Conocimiento. Brasil: Universidad Federal de Santa Catarina.

[9] Andrade León, J. R. (2023). La Gestión del Conocimiento y su relación con el Éxito de las Empresas. Recuperado de https://www.linkedin.com/pulse/la-gestion-del-conocimiento-y-su-relacion-con-el-de-jesus-rodolfo/.

[10] Han B. M., Anantatmula V. S. (2006). Knowledge Management in IT Organizations from Employee's Perspective. Proceedings of the 39th Annual Hawaii International Conference on System Sciences, Kauai, HI, USA.

[11] Chung Jen C., Jing Wen H. (2007). How organizational climate and structure affect knowledge management - The social interaction perspective. International Journal of Information Management, 27, 104–118.

[12] Straccia, L., Pollo Cattaneo, M.F., Maulini, A. (2022). Visões da Gestão do Conhecimento para a construção de um modelo integral. Congresso Brasileiro de Gestão do Conhecimento. KM BRASIL 2022, São Paulo, Brasil, septiembre 19-22.

[13] Fadaie, N., Lakbala, P., Ghanbarnejad, A. (2023). Impact of knowledge management on job satisfaction and organizational performance among healthcare employees: A structural equation modeling approach. Health Science Reports, 6 (9), 1-9.

[14] Atapattu, M. M., Huybers, T. (2022). Motivational antecedents, employee engagement and knowledge management performance. Journal of Knowledge Management, 26 (3), 528-547.

[15] Ştefan, S. C., Popa, I., Olariu, A. A., Popa, Ş. C. and Popa, C. F. (2024). Knowledge management – performance nexus: Mediating effect of motivation and innovation. Business Process Management Journal, 30 (8), 27-48.

[16] Meher, J. R., Mishra, R. K. (2022). Evaluation of perceived benefits and employee satisfaction through knowledge management practices. Global Knowledge, Memory and Communication, 71 (1/2), 86-102.

[17] Pangaribuan, L., Satrya, A. (2024). The Role of Knowledge Management, Transformational Leadership, and Organizational Commitment on Employee Performance: Empirical Study in Public Sector. Journal of Theory and Applied Management (Journal Management Teori Dan Terapan), 17 (3), 355-371.

[18] Nuralfya, A., Priatna, W. B., Burhanuddin, B. (2023). Factors affecting productivity of plantation employees at Pt. Great Giant Pineapple, Central Lampung. Agricultural Socio-Economics Journal, 23 (2), 141-150.

[19] Straccia, L.; Maulini, A.; Bongiorno, M.G.; Giorda, M.; Pollo-Cattaneo, M.F. Knowledge Representation and Technologies in the Latin American Academic Literature. ICAIW 2022: Workshop at ICAI, Arequipa, Perú.

[20] Straccia, L.; Pollo-Cattaneo, M.F. (2024) Knowledge Management: Innovation, Technology and Ontologies. ICAIW2024: Workshop at ICAI, Viña del Mar, Chile.

Incorporación de IA Generativa en la Reestructuración de Escenarios

María C. Elizalde[1], Graciela D.S. Hadad[1,2], Lourdes Sánchez[1], Agustín Catanzaro[1], Abril A. Ruiz Blagine[1], Santiago Kalafa[1]

[1] Instituto de Ingenierías y Nuevas Tecnologías, Universidad Nacional del Oeste
[2] Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano
{melizalde, ghadad}@uno.edu.ar, , misaki.mt36@gmail.com, agustincatanzaro@gmail.com, abrilruizblagine2019@gmail.com, santiagokalafa8@gmail.com

RESUMEN

Las narrativas que describen las actividades de los usuarios con el futuro sistema de software y su contexto facilitan las vinculaciones entre la parte cliente y la parte desarrolladora para acordar lo que el sistema proveerá y bajo qué condiciones o limitaciones lo hará. El modelo de Escenario es una técnica que describe, en lenguaje natural, interacciones en el contexto de aplicación, entre usuarios, el software a construir y otros sistemas o dispositivos, con una estructura simple, pero que representa no solo requisitos funcionales sino también no funcionales, mostrando situaciones normales, situaciones alternativas y excepcionales. Este modelo es de utilidad no solo para acuerdos entre las partes, sino además para diversas actividades durante el desarrollo. Los escenarios son entrada para priorizar el desarrollo de los servicios del software, validarlos con usuarios, definir la arquitectura y las interfaces de usuario y generar pruebas del software. El orden en que se presentan los escenarios puede favorecer o dificultar cada uno de estos usos, por lo que resulta conveniente tener organizaciones diferentes para cada caso. En el presente proyecto se explorarán las nuevas tecnologías basadas en Modelos de Lenguaje de Gran Tamaño como recursos para automatizar estas actividades de reestructuración de conjuntos de escenarios.

Palabras clave: Ingeniería de Requisitos, Escenarios, Inteligencia Artificial, LLMs.

CONTEXTO

La propuesta que se presenta describe los lineamientos del proyecto de investigación acreditado "Evolución de los requisitos desde una visión organizacional hacia una visión técnica" del Instituto de Ingeniería y Nuevas Tecnologías de la Universidad Nacional del Oeste.

1. INTRODUCCIÓN

En la Ingeniería de Requisitos suelen utilizarse tempranamente modelos escritos en lenguaje natural, dada la comunicación intensiva entre usuarios e ingenieros de requisitos, por un lado, para la adquisición de conocimiento y, por otro, para el intercambio de ideas y propuestas sobre el futuro sistema de software (Hadad et al., 2025).

El presente proyecto de investigación se lleva a cabo en el marco de un proceso de requisitos cuyos modelos se focalizan en describir las actividades del contexto tanto en el presente como en el futuro cuando se disponga del nuevo sistema de software. Esto se materializa mediante la construcción de tres modelos que utilizan narrativas en lenguaje natural: los escenarios actuales, los escenarios futuros y el documento de especificación de requisitos (Leite et al., 2004). Los escenarios actuales son necesarios para comprender las características actuales del contexto, mientras que los escenarios futuros preconizan la forma en que llevarán a cabo las actividades en el contexto planificado para el futuro, incluyendo los servicios que prestará el sistema de software y bajo qué condiciones. Finalmente, el tercero, de carácter opcional, es principalmente de características esencialmente contractuales, aunque su construcción suele ser de utilidad durante el desarrollo al ofrecer un ancla para mejorar la completitud.

Cada escenario describe una situación que

muestra el comportamiento de los actores, del sistema de software y de otros sistemas o dispositivos, ya sea con la visión del contexto actual o del futuro.

Cabe destacar que este proceso de requisitos y sus modelos han sido estudiados por distintos autores (Maclen et al., 2024) (Pérez et al., 2025) y puestos en la práctica (Mighetti & Hadad, 2016) (Mira et al., 2018).

Este proceso organiza tanto los escenarios actuales como los escenarios futuros en secuencias producidas por la forma en que son construidos (Leite et al., 2004). En ambos grupos de escenarios se construye una jerarquía, donde los escenarios de alto nivel suelen representar procesos de negocio y los de más bajo nivel actividades y subactividades de esos procesos.

Usualmente esta organización jerárquica no es la más apropiada para utilizar como entrada en algunas actividades del proceso de software. Por ejemplo, una validación de alto nivel de procesos debe permitir visualizar claramente los escenarios integradores (escenarios de alto nivel en la jerarquía), mientras que una validación de detalles en las actividades deberá permitir visualizar claramente los escenarios relacionados con un actor o un grupo de actores.

Por su parte, un arquitecto de software preferirá un orden compatible con la estrategia que utilice. Para facilitar todas estas actividades es necesario realizar reorganizaciones de los escenarios actuales o los escenarios futuros, según sea el caso. Eventualmente puede llegar a ser necesario la construcción de algún artefacto auxiliar que incremente la visibilidad de los aspectos a ser evaluados y contribuya a la exhaustividad de la actividad a ser llevada a cabo.

Esto implicará en algunos casos focalizar la organización y la visualización en las restricciones de los escenarios, en las relaciones entre los actores, en sus responsabilidades, en las situaciones de excepción o en las trasformaciones de los recursos utilizados, entre otros posibles casos.

En este sentido, el presente proyecto se aboca a elaborar procedimientos para la reestructuración de escenarios futuros de acuerdo a su uso para validar requisitos, priorizarlos, definir aspectos arquitectónicos, gestionar iteraciones en función de la implementación de escenarios o elaborar pruebas de aceptación.

Se asume que esta reestructuración de escenarios podrá de alguna manera automatizarse parcialmente utilizando herramientas de la Inteligencia Artificial Generativa.

La Ingeniería de Requisitos se ve influenciada de dos maneras distintas por la difusión y el incremento del uso de los Modelos de Lenguaje de Gran Tamaño (LLMs). Por un lado, la generación de modelos de arquitectura del software a partir de requisitos (Schmid et al., 2025), la generación automática de bloques de código o prototipos a partir de narrativas derivadas de los requisitos (Rashed et al., 2025) o la posible creación semiautomática de documentación de usuario, implican la necesidad de mejorar aún más la calidad de los requisitos, ya que sutiles ambigüedades u omisiones pueden resultar en defectos mayores debido a la disminución de la intervención humana.

Por otro lado, los LLMs están mostrando que pueden influenciar significativamente en el propio proceso de requisitos (Krishna et al., 2024) (Zhu et al., 2025).

Con los avances significativos en Inteligencia Artificial Generativa, el uso de los LLMs se ha instalado en muchas actividades de la Ingeniería del Software (Brown et al., 2020) (Hou et al., 2024), aunque es un área que aún requiere más estudios rigurosos, que eviten alucinaciones (Perković et al., 2024) y grandes consumos de recursos. Uno de los problemas críticos identificados es justamente el fenómeno de las "alucinaciones" y la falta de conocimiento específico del dominio en los LLMs generalistas.

Esto se puede lograr, por un lado, con estrategias de prompting orientadas a actividades específicas (Zhou et al., 2023) y con adaptaciones al LLM, tales como técnicas de fine-tuning y retrieval augmented generation (Lewis et al., 2020) (Di Palma, 2023).

El diseño de los prompts es de vital importancia para obtener resultados de calidad, ya que las inferencias que haga el LLM dependen fuertemente del prompt (Liu et al., 2023).

Por otro lado, dentro de las adaptaciones al LLM, el fine-tuning ha mostrado eficacia para adaptar el estilo y formato de salida, sin embargo, enfoques como retrieval-augmented generation están ganando interés por su capacidad para agregar información de bases de conocimiento propias evitando las alucinaciones (Lewis et al., 2020). En (Zhang et al., 2023) han explorado cómo la integración de bases de conocimiento externas mejora la trazabilidad y la consistencia en tareas de ingeniería de software.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Conciliar las necesidades de los clientes y usuarios, con los recursos técnicos disponibles y las restricciones legales, culturales, temporales y de costos existentes al definir los servicios que prestará un sistema de software, hace que esta no sea una actividad que meramente se dedique a buscar requisitos. Por el contrario, es necesario adquirir conocimiento de diversas fuentes de información, conciliar intereses y lograr una comprensión adecuada del contexto que luego permita elaborar propuestas sobre lo que el software podrá hacer (Hadad et al., 2025). Sin embargo, no debe perderse de vista que estos servicios y las restricciones o condiciones que se imponen al software, es decir, los requisitos deben documentarse de una manera precisa en modelos que puedan cumplir el rol de guiar los restantes pasos del desarrollo. Que estos modelos posean una calidad tal que les sea posible cumplir ese rol requiere que ellos sean precisos y de alta legibilidad. Justamente la organización de los escenarios es un factor clave en esta legibilidad.

Se ha planteado como objetivo general del presente proyecto de investigación: "especificar mecanismos sistematizados de reestructuración de escenarios para facilitar actividades de aseguramiento de calidad y de construcción de software". Ello involucra identificar actividades del proceso de software que se puedan alimentar de los escenarios futuros y qué tipo de información precisan de los escenarios, definir las mejores formas de presentación de los escenarios según la actividad y especificar las los modelos auxiliares de ser necesarios. Adicionalmente, se definió un objetivo específico referido al estudio exploratorio del uso de herramientas de Inteligencia Artificial que permitan automatizar parcialmente posibles formas de reestructuración de escenarios.

Se está asumiendo que la reestructuración de escenarios es una tarea que puede sistematizarse aplicando procedimientos establecidos y también puede automatizarse con Inteligencia Artificial Generativa. Es esperable que esta nueva forma de reestructuración sea simultáneamente más eficiente y más eficaz, más eficiente porque reduzca los esfuerzos involucrados y más eficaz porque los resultados contribuyan a mejorar los resultados obtenidos en las diferentes actividades en las que participan los escenarios.

Por ejemplo, en el caso de la validación de escenarios, es importante que haya una lectura de revisión con distintos puntos de vista para cada escenario. Además, en el caso de escenarios integradores se espera que estos sean validados por usuarios con niveles gerenciales, mientras que escenarios específicos de una actividad sean validados por actores que intervienen en el escenario. Ello involucra generar varios agrupamientos de escenarios por actores con distintos niveles organizacionales o distintos roles; es decir, se requiere entonces establecer criterios específicos de agrupamiento por actores para validar escenarios.

Es altamente probable que estas tareas se puedan automatizar parcialmente utilizando LLMs. Ello precisa de estrategias de prompting adecuadas, ya que es conocido que el desempeño de los LLMs es muy dependiente de los prompts (Liu et al., 2023) (Zhao et al., 2024) (Ronanki et al., 2024). Para ello se explorarán variadas técnicas de Ingeniería de Prompting, tales como one-shot, few-shot y Chain-of-Thought (Wei et al., 2022) (Kojima et al., 2022).

Otra cuestión de relevancia a considerar en los LLMs son las alucinaciones y cómo mitigarlas (Ji et al., 2023) (McKenna et al., 2023). Ello requiere un estudio en profundidad de las técnicas existentes de acuerdo a los recursos disponibles.

Se ha armado un corpus de escenarios, acompañados de un glosario y el documento de objetivos del sistema, que corresponden a

diversos dominios y de diferentes tamaños. Este corpus debe ser normalizado para su uso con el LLM.

Con algunos casos de este corpus normalizado se harán pruebas piloto para estudiar el desempeño de las técnicas de prompting, como también las posibles adaptaciones que serían convenientes para el LLM, en función de las distintas organizaciones de escenarios necesarias para ser entrada en actividades específicas en el desarrollo de software.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el presente proyecto se espera obtener procedimientos para reestructurar escenarios orientados a facilitar la validación de requisitos, la priorización, el diseño arquitectural y la gestión de iteraciones en el desarrollo. Asimismo, se espera elaborar pautas para el uso de LLMs que colaboren en el agrupamiento y reestructuración de los escenarios. En este sentido, se han iniciado un estudio experimental. Un grupo del equipo inició algunas pruebas piloto usando estrategias simples de prompting, con un solo caso de estudio, que contiene 24 escenarios, con 70 requisitos funcionales y 12 requisitos no funcionales.

Con esta entrada de datos y el meta-modelo de Escenario, Copilot permitió clasificar los escenarios en aquellos que describen macro-procesos, los que describen actividades específicas de usuarios, aquellos con tareas operativas que se reúsan en otras actividades y aquellos que detallan situaciones excepcionales. En otras palabras, el LLM reconoció la estructura jerárquica que presentan los escenarios. Pudo armar el flujo lógico de proceso que describen los escenarios

Otro grupo, usando ChatGPT, con el mismo conjunto de escenarios, al darle el objetivo general y los objetivos específicos del sistema, el LLM asoció correctamente los escenarios que cumplían cada objetivo específico. Al contrastar este resultado, con las relaciones previamente establecidas por los ingenieros de requisitos, se comprobó que ellos habían omitido identificar una relación. Con este mismo LLM se probó la identificación de requisitos no funcionales en los escenarios, pudiendo tipificarlos correctamente.

Ambas experiencias solo fueron preliminares, aunque prometedoras por lo que el siguiente paso será la definición del flujo de trabajo basado en LLMs, con prompts refinados y la generación de un corpus con conjuntos de escenarios de diversos dominios, realizando pruebas sistematizadas.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto participan tres investigadores formados, un becario de grado, dos alumnos adscriptos y un alumno colaborador. También participa un alumno de grado que está realizando su Trabajo Final de carrera dentro de los alcances de este proyecto.

5. BIBLIOGRAFÍA

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33:1877-1901.

de Oliveira. O.S., & Silva, C. (2015). Modelando SOA a partir de Modelos Organizacionais. 18th Workshop on Requirements Engineering (WER), Lima, Perú.

Di Palma, D. (2023). Retrieval-augmented recommender system: Enhancing recommender systems with large language models. 17th ACM Conference on Recommender Systems, pp.1369-1373.

Hadad, G.D.S., Doorn, J.H., & Leite, J.C.S.P. (2025). Gestión de Requisitos para la Ingeniería de Software. Alfaomega Grupo Editor. ISBN: 978-987-8983-90-5.

Hou, X., Zhao, Y., Liu, Y., Yang, Z., Wang, K., Li, L., Luo, X., Lo, D., Grundy, J. & Wang, H. (2024). Large Language Models for Software Engineering: A Systematic Literature Review. ACM Transactions on Software Engineering and Methodology, 33(8), 220.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12):1–38.

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners. Advances in neural information processing systems, 35, 199-213.

Krishna, M., Gaur, B., Verma, A. & Jalote, P. (2024). Using LLMs in Software Requirements Specifications: An Empirical Evaluation. 2024 IEEE 32nd International Requirements Engineering Conference (RE), Reykjavik, Iceland, pp. 475-483.

Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., & Hadad, G.D.S., Ridao, M.N. (2004). Defining System Context using Scenarios. En: Perspectives on Software Requirements, Kluwer Academic Publishers. ISBN:

1-4020-7625-8, pp.169-199.
Leshob, A., Blal, R., Mili, H., Hadaya, P., Hussain, O.K. (2020). From BPMN Models to SoaML Models. In: Barolli, L., Hussain, F., Ikeda, M. (eds) Complex, Intelligent, and Software Intensive Systems. CISIS 2019. Advances in Intelligent Systems and Computing, vol 993. Springer, Cham.
Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. 34th International Conference on Neural Information Processing Systems (NIPS'20). Curran Associates Inc., NY, EE.UU., article 793, 9459–9474.
Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM computing surveys, 55(6), 1-35.
Maclen, G., Torres, D., & Antonelli, L. (2024). Una herramienta para componer escenarios a través de procesamiento del lenguaje natural y derivar casos de prueba. 27th Workshop on Requirements Engineering (WER).
McKenna, N., Li, T., Cheng, L., Hosseini, M., Johnson, M., & Steedman, M. (2023). Sources of Hallucination by Large Language Models on Inference Tasks. Conference on Empirical Methods in Natural Language Processing.
Mighetti, J.P., & Hadad, G.D.S. (2016). A Requirements Engineering Process Adapted to Global Software Development. CLEI Electronic Journal, 19(3), 7.
Mira, N.C., Boggio, M.A., Perez, S.B., & Salamon, A.G. (2018). Una propuesta de análisis de problemas utilizando LEL y Escenarios en entrenamientos de equipos que gestionan situaciones de crisis, 6to Congreso Nacional de Ingeniería Informática – Sistemas de Información, pp.1041-1048.
Pérez, G., Mostaccio, C., & Antonelli, L. (2025). Análisis comparativo de arquitecturas de NLP para detectar similitudes entre escenarios en español. 28th Workshop on Requirements Engineering (WER).
Perković, G., Drobnjak, A., & Botički, I. (2024). Hallucinations in LLMs: Understanding and Addressing Challenges. 2024 47th MIPRO ICT and Electronics Convention (MIPRO), Opatija, Croacia, pp. 2084-2088.
Rasheed, Z., Waseem, M., Kemell, K. K., Ahmad, A., Sami, M. A., Rasku, J., Systä, K., & Abrahamsson, P. (2025). Large language models for code generation: The practitioners perspective. arXiv:2501.16998.
Ronanki, K., Cabrero-Daniel, B., Horkoff, J., & Berger, C. (2024). Requirements Engineering using Generative AI: Prompts and Prompting Patterns. En: Nguyen-Duc, A., Abrahamsson, P., Khomh, F. (eds) Generative AI for Effective Software Development. Springer.
Schmid, L., Hey, T., Armbruster, M., Corallo, S., Fuchß, D., Keim, J., Liu, H., & Koziolek, A. (2025). Software Architecture Meets LLMs: A Systematic Literature Review. ArXiv, abs/2505.16697.
Zhang, Q., Fang, C., Xie, Y., Zhang, Y., Yang, Y., Sun, W., Yu, S., & Chen, Z. (2023). A Survey on Large Language Models for Software Engineering. arXiv:2312.15223.
Zhao, Z., Fan, W., Li, J., Liu, Y., Mei, X., Wang, Y., Wen, Z., Wang, F., Zhao, X., Tang, J., & Li, Q. (2024). Recommender Systems in the Era of Large Language Models (LLMs). IEEE Transactions on Knowledge and Data Eng. 36(11), 6889–6907.
Zhou, D., Schärli, N., Hou, L., Wei, J., Scales, N., & Wang, X. (2023). Least-to-most prompting enables complex reasoning in large language models. 11th International Conference on Learning Representations.
Zhu, T., Cordeiro, L.C., & Sun, Y. (2025). ReqInOne: A Large Language Model-Based Agent for Software Requirements Specification Generation. 2025 IEEE 33rd International Requirements Engineering Conference (RE), Valencia, España, pp. 449-457.

INGENIERÍA DE REQUERIMIENTOS ACELERADA CON CÓMPUTO PARALELO APLICADA A IDIOMAS INDÍGENAS

Andrés Rojas Paredes, Brenda Mareco, Christian Chamizo

Instituto de Ciencias (ICI) Universidad Nacional de General Sarmiento – UNGS

arojas@campus.ungs.edu.ar, brendaamareco@gmail.com, christianchamizo@live.com.ar

RESUMEN

Este trabajo se enfoca en el procesamiento de reviews, comentarios en formato texto plano, que los usuarios hacen de una determinada aplicación. La idea, según Guzmán et al., 2017 y Villarroel et al., 2016, es rankear los reviews de tal forma que queden al top del ranking los reviews más importantes o más relevantes para el ciclo de vida de la aplicación.

En este trabajo se pretende aplicar técnicas para la obtención de un mejor ranking. Esta tarea se enfrenta a barreras de complejidad a la hora de encontrar los mejores parámetros para un ranking (ver Rojas et al., 2024). Por esto, un objetivo es utilizar técnicas de procesamiento de alto rendimiento. El trabajo de Rojas et al., 2024 también sugiere que existen atributos como la Entropía que ofrecen mejores resultados en un ranking. En esta línea, se pretende encontrar atributos similares.

Finalmente, en sintonía con el trabajo de Guzmán et al., 2021 donde se hace un estudio cultural de los reviews, se pretende corroborar la efectividad del modelo Hofstede en su dimensión cultural a la hora de aplicarlo a idiomas indígenas y cual es su impacto en el ranking de los reviews escritos en estos idiomas.

Palabras Clave: reviews, ranking, alto rendimiento, Hofstede, idiomas indigenas

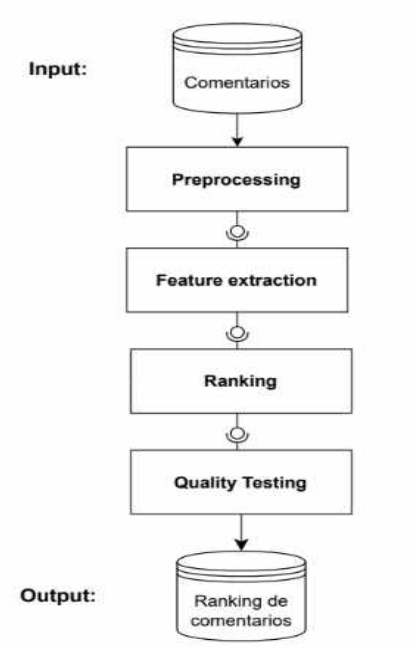
CONTEXTO

Este trabajo se inserta en el proyecto de investigación UNGS 30/1201 Metaheurísticas y optimización aplicada con cómputo eficiente, bajo la dirección del dr. Guillermo Tetzlaff en el Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

1. INTRODUCCIÓN

El presente trabajo tiene como área de estudio el procesamiento de comentarios de usuarios de aplicaciones móviles en el contexto de ingeniería de requerimientos. El principal objetivo de esta área de estudio es procesar los comentarios de los usuarios y extraer información útil para mejorar las versiones existentes de las aplicaciones y así promover la evolución del software. Ejemplo de la información que se puede extraer de los comentarios es la categoría, por ejemplo, según Guzmán et al., 2014 y Khalid et al., 2015 las principales categorías son los reportes de bugs (Bug Reports) y las solicitudes de nuevas funcionalidades (Feature Requests). Extraer información de los comentarios en crudo tiene dificultades inherentes como la naturaleza no estructurada de los comentarios y la gran cantidad de datos existente. Por ejemplo, se estima que más de 500 millones de mensajes cortos (tweets) son escritos por día y que en el rango de 30000 tweets diarios corresponden a aplicaciones populares como Facebook o Snapchat. Las

aplicaciones móviles reciben alrededor de 23 comentarios por día y las aplicaciones populares como Facebook reciben un promedio de 4275 comentarios por día. Estas características plantean la necesidad de herramientas automáticas que se encarguen de procesar los comentarios. Estas herramientas están implementadas típicamente como un pipeline que clasifica los comentarios en útiles o más urgentes y comentarios que se pueden descartar. La figura a continuación muestra las principales etapas que tiene un pipeline de procesamiento automático de comentarios:



El pipeline de procesamiento de comentarios recibe como input los comentarios de los usuarios sobre las aplicaciones móviles y como output retorna un ranking de los comentarios procesados, donde los comentarios rankeados con mayor prioridad serán atendidos con mayor urgencia. El pipeline cuenta con cuatro principales etapas:

1. **Preprocessing** prepara los comentarios crudos (es decir, directamente descargados de la App Store) para ser procesados en las siguientes etapas. Por ejemplo, se convierte los textos en minúscula y se los divide en tokens, entre otros pasos de preprocesamiento.

2. **Feature extraction** extrae atributos a partir de los comentarios preprocesados. Un atributo representa una característica de un comentario (por ejemplo, longitud) que se tiene en cuenta a la hora de realizar el ranking. Por ejemplo, en esta etapa, es estándar extraer el sentimiento y la categoría de los comentarios.

3. **Ranking** clasifica los comentarios según urgencia, exponiendo los comentarios que probablemente sean los más importantes a tener en cuenta para analizarlos y darles seguimiento. Para realizar el ranking de comentarios se emplean principalmente dos técnicas: función de pesos o clasificador automático (ver Guzmán et. al. 2017). Por un lado, la función de pesos depende de parámetros llamados pesos que se asignan a los atributos calculados en la etapa anterior. Los pesos representan la importancia de un atributo con respecto a los otros atributos utilizados en la función de pesos. Por otro lado, los clasificadores automáticos dependen de hiperparámetros para su configuración y de un dataset de entrenamiento. En trabajo previo se ha demostrado que la función de pesos es mejor que los clasificadores automáticos para ciertos valores dados a los parámetros e hiperparámetros (ver Shankar et. al. 2020). Sin embargo, el cálculo de parámetros óptimos es computacionalmente demandante (ver Rojas et. a. 2024).

En este trabajo adoptamos la función de pesos como herramienta para determinar el ranking de reviews. La función de pesos que se utilizó fue la del pipeline de Guzmán et. al. 2017, pero fue adaptado a los siguientes atributos: Category, Sentiment, Score, Revlen y Entropy.

Category indica a qué categoría pertenece un comentario (Reporte de bug, Solicitud de funcionalidad u Otros).
Sentiment implica el sentimiento del usuario que escribió el comentario. Puede ser un valor dentro del rango [-2, 2] de números enteros. 2 se refiere al sentimiento más positivo, mientras que -2 representa el sentimiento más negativo. El valor 0 corresponde a un sentimiento neutro.
Score es la cantidad de estrellas otorgadas en un comentario, representando una calificación que varía en un rango del 1 al 5.
Revlen es la longitud de un comentario.
Entropy implica la cantidad de información útil en un comentario, utilizando la fórmula de entropía de Shannon.
Teniendo estos atributos, la fórmula para calcular el ranking R de un comentario c es:

$$R(c) = \sum_{i=1}^{4} w_i * f_i(c)$$

donde wi representa el peso del atributo i, mientras que fi representa el factor del ranking del atributo i.

4. **Quality Testing** determina la calidad del ranking de los comentarios utilizando varias métricas.

Al pipeline se pueden agregar más etapas, por ejemplo, un módulo crawler para conseguir los datos iniciales es escencial.

En el trabajo de Rojas et al., 2024 se realizó una búsqueda exhaustiva con tal de determinar la combinación de parámetros óptima para lograr el mejor ranking. Este resultado se logró discretizando el espacio de búsqueda a través de la cantidad de dígitos binarios permitidos para representar los valores, en particular se utilizaron dos dígitos. En este trabajo se pretende aplicar técnicas de procesamiento de alto rendimiento para realizar una búsqueda similar pero elevando el número de dígitos de precisión.
Por otro lado, el trabajo de Rojas et al., 2024 indica que existen atributos como la Entropía que ofrecen mejores resultados en un ranking. En particular, se demuestra que Entropía puede reemplazar atributos que son costosos computacionalmente como el sentimiento. En esta línea, se pretende buscar atributos similares a la Entropía en el sentido de ser atributos fáciles de computar y que pueden reemplazar atributos costosos de la función de pesos.

Finalmente, siguiendo la línea de investigación de Guzmán et al., 2021 donde se hace un estudio cultural de los reviews, se pretende corroborar la efectividad del modelo Hofstede. Este modelo ofrece una descripción de los países en términos de variables como el nivel de individualismo. El trabajo de Guzmán et al., 2021 describe que hay correlación entre variables del modelo Hofstede y características del feedback de usuarios. El objetivo de este trabajo es aplicar este modelo a idiomas indígenas y verificar cual sería su impacto en el ranking de los reviews escritos en estos idiomas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se plantean las siguientes líneas de investigacion.

Línea 1 - Cómputo Paralelo: Utilizar técnicas de procesamiento de alto rendimiento y cómputo paralelo para encontrar los mejores parámetros para una función de pesos.

Línea 2 - Eliminación de Atributos: Determinar qué otros atributos standard en el procesamiento automático de reviews se pueden eliminar utilizando la Entropía.

Linea 3 - Idiomas Indígenas: Verificar si existe una influencia de características típicas y comunes de las culturas indígenas en los

atributos que definen el mejor ranking de un conjunto de reviews.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se esperan los siguientes resultados.

Línea 1 - Cómputo Paralelo: Se espera mejorar la precisión los pesos de la función de pesos a tres o cuatro dígitos. Esta mejora permitirá comparar resultados óptimos contra resultados de trabajos previos que están hechos de manera aproximativa y con heurísticas.

Línea 2 - Eliminación de Atributos: Se espera poder eliminar los atributos más costosos computacionalmente, esto permitirá una implementación del pipeline mas liviana. Se podrán eliminar las librerías de aprendizaje automático requeridas para trabajar con atributos como sentimiento y categoría. Con esta simplificación se podra portar el pipeline a infraestructuras como por ejemplo un cluster para realizar experimentos que requieren mayor poder de cómputo.

Línea 3 - Idiomas Indígenas: Si se demuestra que existe una influencia de las características culturas indígenas en los atributos que definen el mejor ranking de un conjunto de reviews, entonces, estaríamos en presencia de una situación de posible bias en detrimento de los reviews escritos en idioma indígena. Esta situación tendría que ser verificada y corregida.

4. FORMACIÓN DE RECURSOS HUMANOS

Un miembro del equipo de la línea de investigación presentada posee formación de postgrado a nivel de doctorado, siendo un especialista en Teoría de la Complejidad, Ingenieria de Requerimientos y Cómputo de Altas Prestaciones. El resto de los investigadores está compuesto por una graduada de la Licenciatura en Sistemas de la UNGS-Universidad Nacional de General Sarmiento que actualmente se desempeña en la industria, y por un estudiante de la Tecnicatura en Informática de la UNGS.

5. BIBLIOGRAFÍA

Shankar A. and Guzmán E.. Algorithm Bias and User Feedback. M.s. thesis, Universiteit van Amsterdam, 2020.

Guzmán E. and W. Maalej. How do users like this feature? A fine grained sentiment analysis of app reviews. In; (RE'14), p. 153–162, (2014).

Guzmán E., M. Ibrahim, and M. Glinz. Prioritizing user feedback from Twitter: A survey report. In International Workshop on Crowd Sourcing in Software Engineering, p. 21–24, (2017).

Guzmán E., Fischer R., Walczuch R: Does Culture Matter? Impact of Individualism and Uncertainty Avoidance on App Reviews. ICSE-SEIS 2021: 67-76

Khalid H., E. Shihab, M. Nagappan, and A. E Hassan. What do mobile app users complain about? In: IEEE Software 32, 3 (2015), p. 70–77, (2015).

Rojas Paredes A., Mareco B.. Shannon Entropy is better Feature than Category and Sentiment in User Feedback Processing. XXI Workshop Ingeniería de Software, XXX Congreso Argentino de Ciencias de la Computación - CACIC 2024. ISBN: 978-950-34-2428-5

Villarroel, G. Bavota, B. Russo, R. Oliveto, and M. Di Penta. Release planning of mobile apps based on user reviews. In ICSE '16. P. 14–24, (2016).

INGENIERÍA DE SOFTWARE PARA EL DESARROLLO DE SISTEMAS INTELIGENTES: OPERACIONALIZACIÓN DE REQUISITOS ÉTICOS

Emanuel Irrazabal; Rubén Alfredo Bernal; Juan Andrés Carruthers; Andrea Lezcano Airaldi; Ríos, Leopoldo José Romera Lourdes, Jesús Zini

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
{eirrazabal, jacarruthers,alezcano}@exa.unne.edu.ar

RESUMEN

Este es el primer año del proyecto 25F005 “Desarrollo y personalización de herramientas para la Inteligencia Artificial Responsable: Ciencia de Datos, ética y aplicaciones para el bien público”, acreditado por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el período 2026–2029. El proyecto constituye una evolución natural de las líneas previas del Grupo de Investigación en Calidad de Software (GICS), orientadas a la evaluación empírica de procesos, productos y herramientas de ingeniería de software, incorporando como nuevo eje estratégico la operacionalización de requisitos éticos en sistemas basados en Inteligencia Artificial.

La línea principal de investigación es el diseño y desarrollo de un modelo de Inteligencia Artificial Responsable acompañado de un sistema de recomendación para asistir a equipos de desarrollo en la identificación, priorización y verificación temprana de requisitos éticos tales como equidad, imparcialidad, explicabilidad, privacidad y seguridad. Se busca integrar estos atributos en el ciclo de vida del software mediante artefactos concretos —modelo conceptual, motor de reglas, catálogos de métricas y protocolos de validación— construidos y evaluados bajo el enfoque de Investigación en Ciencia del Diseño.

Otras líneas secundarias de investigación son el desarrollo de modelos de evaluación para la mejora de visualizaciones, desarrollo de estudios empíricos para la toma de decisiones en grupo y el desarrollo de aplicaciones orientadas entornos científicos.

Palabras clave: inteligencia artificial responsable, ingeniería de software empírica, requisitos no funcionales, ciencia de datos, transferencia tecnológica

CONTEXTO

Las líneas de Investigación y Desarrollo presentadas en este trabajo corresponden al proyecto 25F005 “Desarrollo y personalización de herramientas para la Inteligencia Artificial Responsable: Ciencia de Datos, ética y aplicaciones para el bien público”, acreditado por Resolución Nº 1121/25 C.S. de la Universidad Nacional del Nordeste (UNNE) para el período 2026–2029.

El proyecto se desarrolla en el ámbito del Grupo de Investigación en Calidad de Software (GICS–UNNE), radicado en la Facultad de Ciencias Exactas y Naturales y Agrimensura, y se articula con una beca interna doctoral de CONICET para el período 2026–2030, orientada al diseño y validación de un sistema de recomendación para la gestión de requisitos éticos en sistemas inteligentes.

La propuesta da continuidad a las líneas previas del grupo en evaluación empírica y calidad de software, ampliando su alcance hacia la Inteligencia Artificial Responsable como nuevo eje estratégico de investigación y transferencia regional.

INTRODUCCIÓN

La creciente adopción de sistemas basados en Inteligencia Artificial (IA) en dominios críticos ha intensificado el debate sobre sus implicancias éticas, sociales y regulatorias. Diversos estudios han evidenciado que los algoritmos pueden presentar disparidades sistemáticas en su desempeño según características demográficas o contextuales [1], así como limitaciones teóricas para satisfacer simultáneamente múltiples criterios de equidad [2]. Estas tensiones han impulsado el desarrollo de marcos regulatorios y principios internacionales orientados a promover una IA confiable, tales como los Principios de IA de la OCDE [3], la Recomendación sobre la Ética de la IA de la UNESCO [4] y el reciente Reglamento de Inteligencia Artificial de la Unión Europea (AI Act) [5].

En el plano técnico, se han propuesto herramientas para la evaluación y mitigación de sesgos, así como técnicas de explicabilidad que permiten interpretar modelos complejos [6][7]. Sin embargo, persiste una brecha entre estos desarrollos y su integración sistemática en el ciclo de vida del software. La incorporación temprana y verificable de requisitos éticos —como equidad, explicabilidad, privacidad y seguridad— continúa siendo un desafío metodológico para los equipos de desarrollo.

En el contexto regional, esta problemática adquiere particular relevancia en aplicaciones vinculadas a la salud pública, la criminalística y la gestión pública, donde las decisiones automatizadas pueden tener impacto directo en derechos y políticas públicas. En particular, la situación epidemiológica del dengue en la región de las Américas durante 2024 [8] evidencia la necesidad de soluciones basadas en datos que combinen capacidad predictiva con salvaguardas éticas y mecanismos de validación adecuados.

En este marco, el proyecto propone el diseño y validación de un modelo de Inteligencia Artificial Responsable y un sistema de recomendación para asistir a equipos de desarrollo en la definición y gestión de requisitos éticos, articulando enfoques de Investigación en Ciencia del Diseño [9] con estrategias de validación empírica propias de la ingeniería de software.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación se describe la línea de investigación y desarrollo prioritaria del proyecto 25F005, orientada a la operacionalización de requisitos éticos en sistemas basados en Inteligencia Artificial mediante la construcción y validación de artefactos tecnológicos.

Inteligencia Artificial Responsable y sistema de recomendación para requisitos éticos

La línea principal de investigación se centra en el diseño y desarrollo de un modelo de Inteligencia Artificial Responsable acompañado de un sistema de recomendación que asista a los equipos de desarrollo en la definición, priorización y verificación temprana de requisitos éticos dentro del ciclo de vida del software.

El trabajo se estructura bajo el enfoque de Investigación en Ciencia del Diseño (Design Science Research) [9], lo que implica la construcción iterativa de artefactos —modelo conceptual, motor de reglas, catálogos de métricas y protocolos— y su evaluación en contextos reales. En una primera etapa se prevé la realización de una revisión sistemática de la literatura siguiendo lineamientos

consolidados en ingeniería de software [10], con el objetivo de identificar atributos, métricas y estrategias de integración de principios éticos en proyectos de IA.

A partir de este análisis se definirá la arquitectura conceptual del modelo, contemplando: (i) principios rectores, (ii) procesos de integración en etapas tempranas del desarrollo, (iii) métricas e indicadores verificables, y (iv) roles y responsabilidades. Sobre esta base se desarrollará un sistema de recomendación basado en reglas que sugiera requisitos éticos según el dominio, tipo de datos y nivel de criticidad del proyecto.

La validación de los artefactos se realizará en dos niveles. En primer lugar, mediante evaluaciones controladas y pruebas con conjuntos de datos sintéticos y escenarios académicos. En segundo lugar, a través de pilotos en dominios estratégicos de la región: salud pública, criminalística y gestión pública. Estos casos permitirán evaluar tanto el desempeño técnico como la usabilidad, comprensión y utilidad práctica del sistema en equipos reales de desarrollo.

Esta línea representa una evolución de las investigaciones previas del grupo en calidad de software y evaluación empírica, trasladando los principios de rigurosidad metodológica, validación experimental y construcción de evidencia hacia el ámbito de la Inteligencia Artificial Responsable, con énfasis en su aplicabilidad en contextos regionales.

RESULTADOS OBTENIDOS/ESPERADOS

Durante este primer año de ejecución del proyecto se está trabajando en una revisión rápida de la literatura (rapid review) orientada a identificar enfoques actuales y brechas en la gestión de requisitos de equidad en sistemas inteligentes. La revisión se diseñó siguiendo lineamientos metodológicos para revisiones rápidas en Ingeniería de Software [11], adaptando los principios de revisiones sistemáticas tradicionales [10] a un protocolo acotado en alcance temporal pero manteniendo criterios explícitos de selección, exclusión y extracción de datos.

El proceso fue asistido por herramientas basadas en Inteligencia Artificial para apoyar la búsqueda, filtrado preliminar y organización de estudios primarios, en línea con investigaciones recientes que exploran el uso de técnicas de aprendizaje automático y modelos de lenguaje para apoyar revisiones de literatura en informática [12]. Este enfoque permitió mejorar la trazabilidad del proceso y reducir sesgos en la etapa inicial.

Entre los principales hallazgos de la revisión se identificaron:

- Predominio de enfoques técnicos centrados en métricas de *fairness* a nivel algorítmico, con escasa integración en fases tempranas de ingeniería de requisitos.
- Ausencia de modelos sistemáticos que permitan traducir principios de equidad en requisitos verificables y trazables.
- Limitada articulación entre riesgos reales documentados y mecanismos preventivos formalizados como requisitos no funcionales.
- Falta de estrategias de validación empírica en contextos institucionales reales.

Estos resultados fundamentan la necesidad de avanzar hacia modelos que conecten explícitamente riesgos observados en sistemas de IA con requisitos de equidad formalmente especificados.

En esta línea, se inició un estudio empírico basado en la base de incidentes de IA del MIT, orientado a la construcción de una taxonomía de requisitos de equidad derivada de casos reales. El trabajo, titulado provisionalmente *"From AI Risks to Fairness Requirements: A Data-Driven Taxonomy and Traceability Model"*, propone utilizar dicho corpus como fuente

empírica para identificar patrones de riesgo asociados a inequidad y discriminación.

El estudio se estructura en torno a las siguientes preguntas de investigación:

- PI1: ¿Qué tipos de riesgos documentados están relacionados con inequidad o discriminación?
- PI2: ¿Qué clases de requisitos de equidad podrían haber prevenido dichos incidentes?
- PI3: ¿Cómo establecer trazabilidad entre riesgo, requisito, métrica y mecanismo de validación?

Metodológicamente, se emplea codificación cualitativa de incidentes para identificar categorías emergentes de riesgo, que luego se mapean a tipologías de *fairness* (individual, grupal, contrafactual, procedimental, entre otras). A partir de este análisis se está construyendo un metamodelo de requisitos de equidad inspirado en enfoques clásicos de Ingeniería de Requisitos orientados a metas y atributos no funcionales, incorporando un esquema de trazabilidad riesgo–requisito–métrica–validación.

Como resultados esperados del proyecto se prevé:

- La publicación de la revisión rápida en revista indexada.
- La formalización de una primera taxonomía empírica de requisitos de equidad basada en incidentes reales.
- El desarrollo de un modelo de trazabilidad aplicable en procesos de ingeniería de software.
- La implementación y evaluación experimental del sistema de recomendación en dominios críticos regionales.
- La consolidación de producción científica en congresos y revistas especializadas en Ingeniería de Software y Ética de la IA.

Estos resultados posicionan al proyecto en la intersección entre Ingeniería de Requisitos, Inteligencia Artificial Responsable y validación empírica basada en evidencia.

FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo del Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 3 docentes investigadores, dos becarios internos doctorales de CONICET, uno de ellos trabajando en el desarrollo principal y el segundo becario apoyando en la construcción de la experimentación; y un estudiante de maestría como soporte técnico.

1. REFERENCIAS

[1] D. Pessach and E. Shmueli, “Policy advice and best practices on bias and fairness in AI,” *Ethics and Information Technology*, 2024. doi: 10.1007/s10676-024-09746-w.

[2] J. Kleinberg, S. Mullainathan, and M. Raghavan, “Inherent Trade-Offs in the Fair Determination of Risk Scores,” *arXiv preprint arXiv:1609.05807*, 2016.

[3] OCDE, “Principles on Artificial Intelligence,” OECD Publishing, 2019.

[4] UNESCO, “Recommendation on the Ethics of Artificial Intelligence,” UNESCO, 2021.

[5] Unión Europea, “Reglamento de Inteligencia Artificial (AI Act),” 2024.

[6] R. K. E. Bellamy et al., “AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias,” *arXiv preprint arXiv:1810.01943*, 2018.

[7] M. T. Ribeiro, S. Singh, and C. Guestrin, “‘Why Should I Trust You?’: Explaining the Predictions of Any Classifier,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[8] OPS, “Situación epidemiológica del dengue en las Américas,” Organización Panamericana de la Salud, 2024.

[9] P. Johannesson and E. Perjons, *An Introduction to Design Science*. Springer, 2014.
[10] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Technical Report EBSE 2007-001, Keele University and Durham University Joint Report, 2007.
[11] I. Cartaxo, G. Pinto, and S. Soares, "Rapid Reviews in Software Engineering," in *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2016.
[12] E. Felizardo, R. Ferreira, and S. Nakagawa, "Using Machine Learning and Text Mining to Support Systematic Literature Reviews in Software Engineering: A Systematic Mapping Study," *Information and Software Technology*, 2022.

Ingeniería de Software: Desarrollo de Métodos, Técnicas y Herramientas para la Protección de Bases de Datos

Jorge A. Quevedo, Mario M. Berón
Departamento de Informática / Universidad Nacional de San Luis - U.N.S.L./San Luis/Argentina.
Ejército de los Andes 950, Teléf. +54 (0266) 4520300; int. 2102
qjorge1996@gmail.com, mberon@unsl.edu.ar

25 de febrero de 2026

RESUMEN

La creciente digitalización de la información en organizaciones públicas y privadas ha incrementado la necesidad de compartir datos para su análisis y explotación. Sin embargo, la aplicación de técnicas de procesamiento sobre grandes volúmenes de datos puede derivar en la exposición de información personal, sensible o confidencial, generando desafíos significativos en términos de privacidad y protección de datos.

En este contexto, los lenguajes específicos del dominio (DSL) ofrecen un enfoque declarativo que permite abstraer detalles de implementación y expresar políticas de protección mediante construcciones especializadas. Asimismo, las arquitecturas basadas en plugins facilitan la extensibilidad del software mediante la incorporación de componentes desacoplados que amplían sus capacidades sin modificar el núcleo del sistema.

Este trabajo presenta dos líneas de investigación. La primera propone el estudio y diseño de un DSL orientado a la protección de bases de datos, capaz de definir políticas y técnicas de protección sobre datos sensibles de forma abstracta y expresiva. La segunda aborda el análisis de una arquitectura de plugins que permita integrar nuevas técnicas de protección mediante componentes extensibles, favoreciendo la mantenibilidad y evolución del sistema. En conjunto, ambas líneas buscan establecer las bases para una plataforma extensible que permita la publicación segura de datos preservando la privacidad en escenarios de análisis y explotación de información.

Palabras clave: Bases de datos - Lenguaje Específico del Dominio - Protección de datos - Arquitectura de Plugins

CONTEXTO

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de Software (LaCIS) de la Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto: “Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”, perteneciente a la misma. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional.

INTRODUCCIÓN

Hoy en día, los avances en el procesamiento y almacenamiento de información han permitido

a las organizaciones gestionar grandes volúmenes de datos, incluyendo información personal y sensible, con diversos fines. Este crecimiento impulsó el desarrollo de áreas como la Minería de Datos, cuyo objetivo es extraer conocimiento y patrones relevantes mediante técnicas de inteligencia artificial, aprendizaje automático y análisis estadístico [11] [2].

Sin embargo, la aplicación de estas técnicas puede derivar en la exposición involuntaria de información privada o confidencial, lo que plantea desafíos significativos en términos de seguridad y preservación de la privacidad. En consecuencia, resulta necesario desarrollar enfoques que permitan aprovechar el valor analítico de los datos sin comprometer la protección de la información sensible [2].

En investigaciones anteriores [10] se abordó el desarrollo de una herramienta orientada al enmascaramiento y protección de repositorios de datos mediante técnicas de Privacy-Preserving Data Publishing (PPDP) [4], permitiendo reducir riesgos asociados a la exposición de información confidencial. Particularmente la técnica implementada, *Binary Search* [12], erradica ataques del tipo de Vinculación de Registros [14]. En la misma, el usuario tiene la posibilidad de elegir qué datos considera sensibles y que parámetros considerar en función de cuanta usabilidad se desea mantener de los datos sin descuidar la seguridad de los mismos.

A partir de la experiencia obtenida en dicha implementación, se identificaron limitaciones relacionadas con la extensibilidad del sistema y la dificultad para incorporar nuevas técnicas de protección sin modificar la arquitectura central. Estas restricciones evidencian la necesidad de avanzar hacia enfoques más abstractos que permitan definir políticas de protección de forma declarativa y desacoplada de los mecanismos de ejecución.

En este contexto, los lenguajes específicos del dominio (DSL) [16] emergen como una alternativa adecuada para expresar políticas de protección de datos mediante construcciones especializadas, facilitando la separación entre la definición de reglas y su implementación técnica. Asimismo, las arquitecturas de plugins [1] ofrecen un modelo de extensibilidad que permite integrar nuevas técnicas de protección mediante componentes independientes, favoreciendo la evolución del sistema sin afectar su núcleo.

El presente trabajo propone dos líneas de investigación: una orientada al diseño de un DSL para la protección de bases de datos y otra al estudio, desarrollo e implementación de una arquitectura de plugins que permita incorporar múltiples técnicas de preservación de privacidad. La primera permite generar estrategias de protección, es decir, cada DSL generado es una estrategia de protección que combina distintas técnicas. Mientras que la segunda línea se enfoca en una herramienta que permita colocar nuevas técnicas mediante plugins, que pueda posteriormente ejecutarse. De esta manera se tendrá la posibilidad de tener un menú de estrategias para aplicar.

LÍNEA DE INVESTIGACIÓN y DESARROLLO

Las líneas de investigación propuestas surgen como una evolución del trabajo previo orientado al desarrollo de herramientas de protección de repositorios mediante técnicas PPDP. A partir del análisis de dichas soluciones, se identificó la necesidad de avanzar hacia un enfoque más abstracto que permita desacoplar la definición de políticas de protección de su implementación técnica.

Lenguaje específico del dominio

La primera línea de investigación se centra en el diseño de un DSL orientado a la protección de bases de datos. Este lenguaje permitirá la especificación del tipo de repositorio, el nivel deseado de privacidad y la clasificación de los tipos de datos. Este lenguaje debe ser lo suficientemente expresivo, simple e intuitivo para ser aprendido por cualquier ingeniero de software.

Las especificaciones generadas por el lenguaje permiten configurar todo el proceso de transformación:

- Configurar el tipo de repositorio y su información básica (tipo, nombre, string de conexión, ruta, etc.).
- Clasificar cada tipo de dato en la información presente sobre el repositorio que se está manejando (puede clasificarse en ID, QID, SA, NSA) [17] [13] [18].
- Elegir qué técnica aplicar para proteger el repositorio. Cada una protegerá contra distintos tipos de ataque: Vinculación de Registros, Vinculación de Atributos, Vinculación de Tablas y Ataques Probabilístico [9] [7].

Lo que el usuario deberá proveer a través del DSL está especificado en el listado 1:

- Nombre de base de datos
- Especificación las tablas donde serán aplicadas las técnicas, y para cada una clasificar los atributos en PPDP:
  - '@' : atributo identificador (ID).
  - '&' : atributo cuasi-identificador (QID).
  - '∼' : atributo sensible (SA).
  - ninguno: si no aplica ninguno de los anteriores.
- Técnicas a aplicar, donde cada una ofrece diferentes grados de privacidad.

Listado 1: Gramática del DSL propuesta

```
nexus : repositoryOptions dataDescription
    privacyOptions ;

repositoryOptions
  : 'Relational' relationalRepositoryOptions
  | 'CSV' csvRepositoryOptions
  ;

relationalRepositoryOptions
  : 'Connection String:' STRING_LITERAL
  ;

csvRepositoryOptions
  : 'Path:' STRING_LITERAL
  ;

dataDescription
  : NAME entities
  ;

entities
  : entity (entity*)
  ;

entity
  : NAME '[' attributeList ']'
  ;

attributeList
  : attribute (',' attribute)*
  ;

attribute
  : '@' NAME
  | '&' NAME
  | '~' NAME
  | NAME
  ;

privacyOptions
  : 'Techniques:' '[' techniqueModelList ']'
  ;

techniqueModelList
  : NAME (',' NAME)*
  ;

techniqueModel
  : 'BinarySearch'
  | 'Top-Down Disclosure'
  | 'Mondrian'
  | 'Cross-Training Round Sanitization'
  ;

NAME : [a-zA-Z][-_a-zA-Z0-9]*;
NUM : [1-9][0-9]*;
STRING_LITERAL : '"' (~["\\\r\n] |
    EscapeSequence)* '"';
WS : (' ' | '\t' | '\r'? '\n')+;
```

Esta es una primera versión de la gramática del lenguaje que a futuro va a evolucionar. La misma fue implementada con ANTLR4 [8] usando primitivas básicas las cuales a medida que avance la investigación también van a evolucionar. Se colocaron en esta versión inicial los nombres de técnicas que dispondrá la herramienta y que luego se podrán combinar tal como la antes mencionada Binary Search, Top Down Disclosure [5], Mondrian [6] [15] y Cross-Training Round Sanitazion [3]. A futuro se le agregarán nuevas técnicas a fin de disponibilizar una amplia gama de posibilidades al ingeniero de software.

Arquitectura de plugins

La segunda línea de investigación aborda el estudio de una arquitectura de plugins capaz de soportar la incorporación dinámica de nuevas técni-

cas de protección mediante componentes desacoplados. En este enfoque, el núcleo del sistema actúa como motor de ejecución, mientras que los plugins implementan estrategias específicas de protección adicionales. Este modelo busca mejorar la mantenibilidad y escalabilidad del sistema, permitiendo su evolución sin modificar la estructura central.

Una arquitectura de plugins se compone de tres partes: el Core o Kernel del sistema, el plugin manager, y los plugins. La interrelación entre estas partes se muestran en la Figura 1. El Core del sistema se encarga del encendido, encontrar los plugins instalados, unirlos con sus correspondientes puntos de extensión, construir el registro de plugins global y almacenar los datos de este registro en Caché para la siguiente ocasión. Por otro lado, el Administrador de plugins es una pieza genérica de código que encuentra y carga los plugins para posteriormente inicializarlos, así como también de descargarlos. Los plugins en sí mismos constituyen el protocolo del administrador de plugins que conforman las expectaciones del sistema principal.

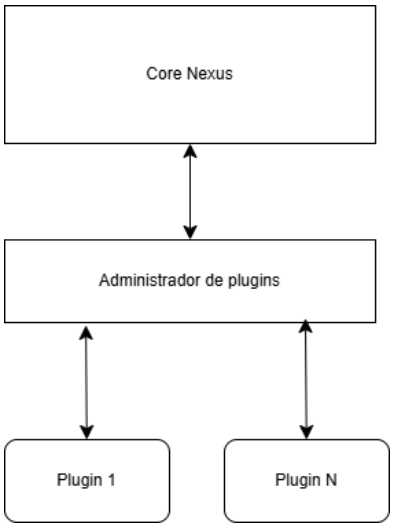


Figura 1: Arquitectura de plugins propuesta

RESULTADOS ESPERADOS/OBTENIDOS

El trabajo llevado a cabo por el equipo de investigación tiene como objetivos principales la obtención de los siguientes resultados:

- Ampliar el estado del arte en cuanto al uso de herramientas y técnicas de protección de bases de datos.
- Profundizar en el conocimiento sobre técnicas existentes para proteger un repositorio.
- Desarrollar herramientas capaces de lograr la protección de un repositorio, basándose en los requisitos del usuario final.

Como parte de los avances obtenidos, se definió una versión preliminar de la gramática formal del DSL propuesto, orientada a describir la configuración de repositorios, atributos sensibles y modelos de ataque. Esta primera aproximación permitió establecer la estructura sintáctica del lenguaje y validar su viabilidad como mecanismo declarativo para la protección de bases de datos.

En relación con la línea de investigación orientada a la arquitectura de plugins se avanzó en su especificación y definición conceptual de un modelo basado en componentes extensibles que permite desacoplar el núcleo del sistema de las técnicas específicas de protección de datos. Este enfoque propone la incorporación de puntos de extensión mediante plug-ins, posibilitando la integración progresiva de nuevos mecanismos sin modificar la lógica central de la herramienta.

FORMACIÓN DE RECURSOS HUMANOS

Los avances logrados en esta área de investigación constituyen la base para el desarrollo de múltiples trabajos académicos en distintos niveles de formación, incluyendo proyectos finales de ingeniería, tesis de maestría y trabajos doctorales en la Universidad Nacional de San Luis. Esta línea

de investigación se enmarca dentro del proyecto institucional mencionado en la sección Contexto, promoviendo la continuidad y evolución del conocimiento en el área.

Referencias

[1] Nishat Anjum and Md Sayed. Plug-in architecture as software extension mechanism: An extensive study of eclipse architecture, 04 2024.

[2] Tânia Carvalho, Nuno Moniz, Pedro Faria, and Luís Antunes. Survey on privacy-preserving techniques for data publishing, 2022.

[3] Shuchi Chawla, Cynthia Dwork, Frank McSherry, Adam Smith, and Hoeteck Wee. Toward privacy in public databases. In *Theory of Cryptography Conference (TCC)*, volume 3378 of *Lecture Notes in Computer Science*, pages 363–385. Springer, 2005.

[4] B.C.M. Fung, K. Wang, A.W.C. Fu, and P.S. Yu. *Introduction to Privacy-Preserving Data Publishing: Concepts and Techniques*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. CRC Press, 2010.

[5] Benjamin C. M. Fung, Ke Wang, and Philip S. Yu. Top-down specialization for information and privacy preservation. In *Proceedings of the International Conference on Data Engineering (ICDE)*, pages 205–216. IEEE, 2005.

[6] Kristen Lefevre, David J. DeWitt, and Raghu Ramakrishnan. Mondrian multidimensional k-anonymity. In *Proceedings of the International Conference on Data Engineering (ICDE)*, pages 25–25. IEEE, 2006.

[7] Mehmet Emre Nergiz, Maurizio Atzori, and Chris Clifton. Hiding the presence of individuals from shared databases. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, SIGMOD '07, pages 665–676. ACM, 2007.

[8] Terence Parr. The definitive antlr 4 reference. 2013.

[9] Jovan Powar and Alastair R Beresford. Sok: Managing risks of linkage attacks on data privacy. *Proceedings on Privacy Enhancing Technologies*, 2023.

[10] Jorge A. Quevedo and Mario M. Berón. Enmascaramiento de bases de datos. In *Workshop de Investigadores en Ciencias de la Computación (WICC)*, San Luis, Argentina, 2024. Departamento de Informática, UNSL.

[11] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Always learning. Pearson, 2016.

[12] Pierangela Samarati. Protecting respondents' identities in microdata release. *Knowledge and Data Engineering, IEEE Transactions on*, 13:1010–1027, 12 2001.

[13] A. Sharma, P. Panday, R. Baghel, and P. Saini. A survey on techniques for privacy preserving data publishing (ppdp). *International Journal of Computer Applications*, 139(11):1–5, 2016.

[14] Anushka Vidanage, Thilina Ranbaduge, Peter Christen, and Rainer Schnell. A taxonomy of attacks on privacy-preserving record linkage. *Journal of Privacy and Confidentiality*, 12(1), Jul. 2022.

[15] Ke Wang, Benjamin C. M. Fung, and Philip S. Yu. Handicapping attacker's confidence: An alternative to l-diversity. *Knowledge and Information Systems*, 11(3):345–368, 2007.

[16] Andrzej Wasowski and Thorsten Berger. *Domain-Specific Languages*. Springer, 2023.

[17] Raymond Chi-Wing Wong and Ada Wai-Chee Fu. *Privacy-Preserving Data Publishing: An Overview*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2010.

[18] Li Xu, Chunxiao Jiang, Jian Wang, Jian Yuan, and Yong Ren. Information security in big data: Privacy and data mining. *IEEE Access*, 2:1149–1176, 2014.

INNOVACIÓN EN DESARROLLO DE SOFTWARE: VALIDACIÓN DE ARQUITECTURA BACKEND-DRIVEN UI EN LA APLICACIÓN "+65 AGENDA SALUD"

Emilio Gustavo Ormeño[1], Gustavo Antonio Sevilla[1], Sergio Gustavo Zapata[1], Cintia Ferrarini Oliver[2] Estela Torres [3], Sergio Alejandro Gaitan[1], Fabricio Carrio[1], Facundo Gallardo[1], Martín Varela[1], Diana Bruno[4], Luciana María Vita[4]

1. Instituto de Informática, 2.Departamento de Informática, 3.Departamento de Matemática y Química Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ. 4. Instituto de Investigaciones en Psicología Básica y Aplicada (IIPBA). Fac. de Filosofía y Humanidades. U. Católica de Cuyo

eormeno@gmail.com, gustavoasevilla@gmail.com, sergio.zapata.diaz@gmail.com, ferrarinicintia@unsj-cuim.edu.ar, etorres271@gmail.com, Alejandrosg197@gmail.com, fabriciocarrio@gmail.com, facundoggallardo@gmail.com

RESUMEN

El presente proyecto tiene por objeto la validación científica y transferencia al sector productivo de una tecnología de desarrollo de software basada en arquitectura Backend-Driven. Esta herramienta, resultado de un Proyecto de Investigación (PIC) previo del Instituto de Informática, propone centralizar la lógica de interfaz en el servidor para acelerar los tiempos de desarrollo de software. La propuesta intentará validar el framework en dos fases de ejecución: primero, se utilizará el framework para construir y desplegar "+65 Agenda Salud", una aplicación móvil de autogestión sanitaria para adultos mayores. Segundo, se ejecutará un experimento controlado en dos empresas de Cuyo: Aconcagua Software Factory y Nodos Hub, en las cuales se comparará la productividad de este enfoque frente al desarrollo tradicional, consolidando la transferencia de la tecnología a la industria de software regional.

CONTEXTO

Este proyecto ha sido presentado en la Convocatoria 2025, que aún se encuentra en proceso de evaluación. Según UNESCO pertenece a la disciplina desagregada: Sistemas de Información y Medicina -Salud Pública, con campo de aplicación en Ciencia y cultura-Ciencia y tecnología y Tecnología sanitaria y curativa-Varios. En la Especialidad: Ingeniería de Software / Optimización de Procesos de Desarrollo / Información y Comunicación (TIC) - Software y desarrollo y análisis de aplicativos.

El grupo responsable cuenta con una sólida trayectoria que le permitirá desarrollar la propuesta, dado que este proyecto es una continuación del Proyecto de Investigación (PIC) denominado “Framework USIM para el desarrollo de Aplicaciones Web” , donde se concibieron y prototiparon las bases del framework de arquitectura Backend-Driven, pero con énfasis en la transferencia en el medio. El equipo de integrantes posee conocimientos profundos de la tecnología generada por ser sus creadores.

También antecedente a esta propuesta el proyecto denominado "Software para Gamificar Terapias de Rehabilitación Ad-hoc" (Código SIGEVA: 80020190200072SJ), financiado por SECITI, donde se demostró capacidad para desarrollar software móvil complejo y gestionar vínculos con adoptantes del sector salud. En dicho proyecto, fueron miembros el Mag. Ormeño y la Mag. Ferrarini. También constituye un antecedente importante el proyecto PIC USIM, en el que fueron respectivamente Director y Co-directora de ese proyecto, el Mag. Lic. Emilio Ormeño y la

Mag. Lic. Cintia Ferrarini.

Por otra parte, existen antecedentes específicos sobre el Caso de Uso (+65), realizado en el marco del proyecto “+65 Agenda Salud”. En el que su equipo pudo adquirir sólidos conocimientos sobre la problemática a resolver. Los integrantes del mismo Mag. Lic. Sergio Zapata, el Lic. Fabricio Carrió y Téc. Facundo Gallardo; son autores del diseño conceptual y prototipo funcional de "+65 Agenda Salud". Este desarrollo cuenta con el respaldo del IIPBA, cuyos investigadores participaron como asesores en la validación de los instrumentos de monitoreo cognitivo. Esta sinergia entre Ingeniería (UNSJ) y Psicología (UCCuyo) garantiza que el caso de uso sobre el cual se probará el framework posee un alto nivel de madurez y relevancia científica.

1.INTRODUCCIÓN

A continuación se presentan algunos conceptos relacionados con el proyecto mencionado

1.1.INNOVACIÓN EN DESARROLLO DE SOFTWARE Y ARQUITECTURAS MODERNAS

La innovación en desarrollo de software se ha orientado hacia arquitecturas que permitan mayor flexibilidad, modularidad y eficiencia en ciclos de desarrollo, superando limitaciones del modelo tradicional de separación fuerte entre frontend y backend. La literatura muestra que las estrategias evolutivas incluyen enfoques como Micro-Frontends, arquitecturas desacopladas y modelos centrados en la lógica del servidor para optimizar aspectos de mantenibilidad y despliegue [1].

1.2. ARQUITECTURA BACKEND-DRIVEN UI: CONCEPTO Y BENEFICIOS

La idea de Backend-Driven UI (también denominada Server-Driven UI) implica que el backend no solo responde con datos, sino que define la estructura, la disposición y el comportamiento de la interfaz de usuario en tiempo de ejecución, generalmente mediante un esquema o descripción serializada (JSON). Esto permite actualizar interfaces sin nuevas versiones cliente y coordinar lógica de UI desde un único punto de control para acelerar iteraciones y uniformizar experiencia[2,3].

1.3. MODELOS TRADICIONALES Y ARQUITECTURAS DESACOPLADAS

En contraste con arquitecturas convencionales (REST o GraphQL puramente cliente-lideradas), los modelos backend-driven buscan reducir la duplicidad lógica y el tiempo de despliegue, mitigando costes de mantenimiento y sincronización de versiones. Estas propuestas se inscriben en una tendencia más amplia de desacoplamiento y modularidad (p. ej., headless o micro-frontends), favoreciendo tanto la productividad del equipo como la escalabilidad técnica [4].

1.4 FRAMEWORKS Y MOTORES DE VIDEOJUEGOS

Desde una perspectiva conceptual, un framework puede entenderse como un puente programático entre conceptos abstractos y sus implementaciones concretas de bajo nivel. Taylor (2018), citado en Politowski et al. [5], define los frameworks como estructuras que permiten mapear estilos arquitectónicos en implementaciones reutilizables, proporcionando una base sólida para el diseño de arquitecturas de software.

En el ámbito del desarrollo de videojuegos, los motores de videojuegos (game engines) constituyen frameworks especializados. Según Lewis y Jacobson (2002), citados en Politowski et al. [6], un motor de videojuegos

es una colección de módulos de código de simulación que no define directamente la lógica del juego ni los datos de nivel, sino que provee los servicios fundamentales para su construcción. Motores como Unity, Unreal Engine y Godot integran funcionalidades esenciales —gráficos, física, audio y scripting— ofreciendo una plataforma robusta para el desarrollo de proyectos interactivos.

1.5. PARADIGMAS DE DISEÑO EN EL DESARROLLO DE SOFTWARE

El desarrollo de software contemporáneo, particularmente en motores de videojuegos y sistemas de simulación en tiempo real, se sustenta en patrones arquitectónicos que priorizan la flexibilidad y la modularidad por sobre estructuras jerárquicas rígidas [7]. En este contexto, resulta central analizar la dicotomía entre herencia y composición como mecanismos de reutilización y extensión del comportamiento de objetos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto, atendiendo a que el modelo tradicional de desarrollo web obliga a duplicar la programación en servidor y cliente, generando altos costos, ineficiencia y dependencia de programadores senior difíciles de contratar en la provincia, debido a la alta rotación del sector señalada por la Cámara de Servicios de Software de Argentina y la migración de talento al exterior. Esto debilita a las PyMEs, encarece proyectos públicos como +65 Agenda Salud y amplía la brecha frente a empresas como Uber y Airbnb, que utilizan tecnologías más eficientes pero inaccesibles. Por estos motivos, se requiere una herramienta más simple y accesible que reduzca costos y fortalezca la competitividad local. en la sección anterior, en este sentido se propone generar los siguientes productos:

- Tecnología "USIM" : Un Framework de desarrollo de software optimizado y documentado, diseñado específicamente para que equipos de desarrollo puedan construir aplicaciones complejas sin duplicar esfuerzos ni costos. El mismo se entregará con documentación técnica en español, guías y tutoriales de implementación adaptadas a PyMEs del sector de software.
- Informe de Validación de Productividad: Un estudio técnico que demuestre, con datos cuantitativos y cualitativos reales, cuánto tiempo y dinero se ahorra al utilizar esta tecnología en comparación con los métodos tradicionales.
- Aplicación "+65 Agenda Salud": se entregará esta aplicación finalizada y en producción, aplicando la potencia del framework. Se trata de una herramienta de mHealth (Salud Móvil) diseñada para empoderar al adulto mayor, permitiéndole la autogestión de su bienestar mediante:
- Registro de indicadores clínicos: (Presión arterial, glucemia, peso).
- Gestión de tratamientos: Alertas y recordatorios de medicación configurables.
- Análisis cognitivo y anímico: Módulos de seguimiento del estado mental y cognitivo. En este punto, el equipo de investigadoras del IIPBA (UCCuyo) cumple un rol fundamental de validación funcional y de usabilidad: ellas certificarán si la interfaz generada automáticamente por el Framework USIM cumple con los estrictos estándares de accesibilidad requeridos para adultos mayores y si los instrumentos digitales son válidos

desde la psicometría. De esta forma, las especialistas en salud validan indirectamente la capacidad técnica del Framework para generar soluciones complejas y útiles.

- Vinculación: Conexión con cuidadores y profesionales de la salud.
- La complejidad de implementar estas reglas de negocio psicológicas para validar la capacidad del framework USIM para manejar lógica experta.
- Paquete de Transferencia: Capacitación y entrega de la herramienta a las empresas Aconcagua Software Factory y Nodos Hub para su uso inmediato.

3. RESULTADOS ESPERADOS

Se espera demostrar que la tecnología propuesta, gestada en la UNSJ, reduce el esfuerzo de codificación entre un 40% y un 60% en entornos reales.

Hito 1: Refinamiento del framework mediante el desarrollo y despliegue de la aplicación "+65 Agenda Salud" (Mes 10).

Hito 2: Transferencia de conocimientos y capacitación al equipo de la empresa adoptante (Mes 14).

Hito 3: Ejecución del experimento controlado de desarrollo paralelo en ambas empresas (Mes 18).

Hito 4: Adopción tecnológica por parte del sector privado y difusión científica de los resultados (Mes 24).

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está conformado por cuatro integrantes en el equipo responsable y siete integrantes del grupo colaborador. Es un equipo interdisciplinario, dado que está compuesto por Licenciados en Ciencias de la computación, Licenciados en Sistemas de Información, Técnicos en programación, Matemáticos, Psicólogos y Educadores. El desarrollo de este proyecto se prevé permita la formación de recursos humanos especializados en las tecnologías requeridas para generar los productos de software mencionados. Esta línea permitirá generar propuestas de trabajos finales de grado y de posgrado en Ciencias de la Computación y Sistemas de Información. Actualmente se han iniciado dos proyectos de tesis de maestría relacionados con el proyecto, correspondiente al Ing. José Pablo Doña y el Lic. Alberto Domínguez.

5. REFERENCIAS

[1] N. Savani, "The Future of Web Development: An In-depth Analysis of Micro-Frontend Approaches", Int. J. Comput. Trends Technol., vol. 71, no. 11, Nov. 2023, doi:10.14445/22312803/IJCTT-V71I11P109.

[2] A. Rota, "Backend Driven UI Architecture at Glovo", Glovo Tech Blog, Dec. 2025. [Online]. Disponible en: tech-blog.glovoapp.com/blog/backend-driven-ui-architecture-at-glovo

[3] A. Goncharov, "Convincing the Team Lead: The Power of Server/Backend-Driven UI", Mish Design, 2026. [Online]. Disponible en: mish.design/en/articles/14-prodayom-timlidu-ideyu-serverbackend-driven-ui

[4] OpenWebinars, "Headless UI: Separando la interfaz de usuario del backend", OpenWebinars Blog, 2024. [Online]. Disponible en: openwebinars.net/blog/headless-ui-separando-la-interfaz-de-usuario-del-backend.

[5] C. Politowski, R. Fontoura, E. Santos, and A. I. Crnkovic, "Are Game Engines Software Frameworks? A Three-Perspective Study," IEEE Access, vol. 9, pp. 113688–113706, 2021, doi: 10.1109/ACCESS.2021.3103281.

[6] R. Nystrom, Game Programming Patterns. Genever Benning, 2014.

Inteligencia artificial como andamiaje cognitivo: eficiencia y calidad estructural en ecosistemas de desarrollo híbridos

Esnaola, Leonardo Martín[1]; Ramón, Hugo Dionisio[1];
Lanzarini, Laura Cristina[2]

[1]Instituto de Investigación y Transferencia en Tecnología (ITT)
Centro asociado a la Comisión de Investigaciones Científicas (CIC)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

[1]Instituto de Investigación en Informática LIDI
Centro asociado a la Comisión de Investigaciones Científicas (CIC)
Universidad Nacional de La Plata (UNLP)

leonardo.esnaola@itt.unnoba.edu.ar; hugo.ramon@itt.unnoba.edu.ar; laural@lidi.info.unlp.edu.ar

Resumen

Como parte del proyecto marco sobre optimización de procesos socio-productivos mediante tecnologías emergentes, el presente trabajo profundiza en una de sus líneas de investigación específicas: el impacto de la Inteligencia Artificial Generativa (IAG) en el desarrollo de software. A través de estudios empíricos cuasi-experimentales [1], evaluamos el potencial de la IA conversacional para asistir a usuarios sin experiencia en programación. Los hallazgos sugieren que, bajo condiciones controladas, estos usuarios logran tasas de éxito funcional comparables a las de desarrolladores tradicionales, frecuentemente en menor tiempo. Asimismo, al auditar los artefactos [2], se observó que la IA actúa como un "andamiaje" que tiende a elevar el piso de la calidad estructural del código. Sin embargo, esto refleja los estándares integrados por defecto en el modelo más que una habilidad de programación superior por parte del usuario. Finalmente, se evidenció que la calidad de la interacción humano-máquina resulta fundamental para el éxito, consolidando a la especificación estructurada (*prompting*) como una competencia indispensable en los nuevos ecosistemas híbridos.

***Palabras clave*:** Inteligencia Artificial Generativa, Ingeniería de Software, Interacción Humano-IA, Optimización de Procesos.

Contexto

Esta línea de investigación forma parte del proyecto marco "Técnicas y algoritmos avanzados para la optimización de procesos: seguridad, sostenibilidad e innovación", presentado y aprobado en el marco de la convocatoria a Subsidios de Investigación Bianuales (SIB 2025) de la Secretaría de Investigación, Desarrollo y Transferencia de la UNNOBA. Su desarrollo se lleva a cabo principalmente en el Instituto de Investigación y Transferencia en Tecnología (ITT), centro asociado a la Comisión de Investigaciones Científicas (CIC) de la Provincia de Buenos Aires. Asimismo, cabe destacar que los resultados experimentales aquí expuestos, y la línea de trabajo abordada, se encuentran estrechamente vinculados a la tesis doctoral en curso de uno de los investigadores del ITT, realizada en el marco del Doctorado en

Ciencias Informáticas de la Universidad Nacional de La Plata (UNLP). Esta articulación no solo nutre la investigación, sino que consolida el trabajo cooperativo y la sinergia interinstitucional entre el ITT (UNNOBA-CIC) y el Instituto de Investigación en Informática LIDI (UNLP-CIC).

Introducción

Nos encontramos inmersos en una acelerada transformación digital (la denominada "Industria 4.0") que exige a las organizaciones adoptar herramientas innovadoras no solo para optimizar sus procesos, sino para sobrevivir y mantener su competitividad en entornos altamente dinámicos [3-4]. En este escenario, la Inteligencia Artificial (IA) y el procesamiento de señales multimodales (imágenes, audio, texto) han dejado de ser promesas a futuro para convertirse en el núcleo de la eficiencia operativa [5-6]. Al combinar y fusionar información de diversas fuentes, es posible obtener una representación mucho más completa, contextual y robusta de los fenómenos analizados, revolucionando sectores que van desde la educación personalizada [7-8] y la manufactura —mediante la creación de gemelos digitales o *digital twins* [9]— hasta la atención médica, la ciberseguridad y la gestión financiera [10-11].

Sin embargo, es dentro de la industria tecnológica y la ingeniería de software donde esta revolución está mostrando una de sus facetas más disruptivas. La evolución de los Grandes Modelos de Lenguaje (LLMs) y los asistentes conversacionales ha abierto un debate fascinante y central tanto en la academia como en la industria: *¿Seguirá siendo estrictamente necesario aprender a programar de la forma tradicional, o la IA permitirá a cualquier usuario desarrollar software?* [12]. Por un lado, diversas corrientes sostienen que la IA, al igual que las plataformas *low-code* o *no-code*, logrará democratizar la creación de aplicaciones derribando las barreras técnicas de entrada [13-14]. Por otro lado, la evidencia sugiere que, más que reemplazar a los programadores, la IA transformará radicalmente sus roles, actuando como un copiloto que acelera rutinas pero que demanda nuevas competencias de supervisión [15-17].

Si bien está comprobado empíricamente que los modelos generativos pueden escribir código rápidamente [18], la ciencia advierte que un software funcional no siempre es un software seguro o mantenible. La simple capacidad de superar pruebas básicas (corrección funcional) no garantiza la calidad arquitectónica a largo plazo. Investigaciones recientes han demostrado que el código generado por IA, aunque pase satisfactoriamente las pruebas de unidad, puede introducir vulnerabilidades de seguridad severas [19] y acarrear una importante deuda técnica expresada en problemas de modularidad, duplicación y falta de robustez [20-21].

Por lo tanto, el foco científico se está desplazando hacia la validación exhaustiva de los artefactos generados y la forma en que los humanos estructuran sus requerimientos. En este nuevo paradigma de desarrollo mediado por el lenguaje natural, la interacción estructurada (ingeniería de *prompts*) emerge como una competencia esencial [22]. Para un usuario, experto o novato, el éxito ya no radica únicamente en conocer la sintaxis de un lenguaje, sino en la capacidad de expresar metas, especificar entradas y salidas, declarar reglas de negocio y proveer casos de prueba de manera inequívoca frente al modelo [23]. En esencia, la interacción con la IA se está consolidando como una forma moderna de ingeniería de requerimientos [24].

Este proyecto explora empíricamente esa frontera: analizamos ecosistemas donde usuarios no expertos se comunican con la máquina mediante lenguaje natural para crear soluciones, estudiando no solo si el código "funciona", sino cuán robusto es y qué habilidades de comunicación iterativa son necesarias para lograrlo de manera exitosa y sostenible [2], [25].

Líneas de Investigación, Desarrollo e Innovación

Para abordar los desafíos de la transformación digital, los ejes de investigación del proyecto se articulan sobre áreas de vanguardia tecnológica y metodológica, abarcando desde la infraestructura de datos hasta la evaluación empírica del software:

- *Procesamiento de señales multimodales e IA aplicada*: esta línea investiga la fusión de diferentes naturalezas de datos (texto, imágenes, audio y video) aplicando algoritmos de aprendizaje automático (Machine Learning) e Internet de las Cosas (IoT). El objetivo es obtener una representación mucho más completa, contextualizada y robusta de fenómenos complejos [6], [26]. Esto permite la optimización de actividades económicas regionales a través de sistemas predictivos, automatización y la creación de gemelos digitales (*digital twins*) para simular y mejorar procesos productivos [9], [10].
- *Evaluación empírica en entornos de desarrollo híbridos*: utilizando marcos rigurosos de experimentación en ingeniería de software [27], se diseñan estudios empíricos para evaluar cómo las interfaces conversacionales de IA moldean la creación de tecnología. Específicamente, se mide y contrasta la eficacia funcional y eficiencia en tiempos de resolución de usuarios sin formación en programación asistidos por IA (NP-AI) frente a desarrolladores operando de manera manual (P-noAI). Esto permite analizar con base científica el verdadero potencial de los Grandes Modelos de Lenguaje (LLMs) para democratizar el desarrollo de software [1], [13-14].
- *Calidad de software e Interacción Humano-IA*: dado que la literatura advierte que superar pruebas básicas no garantiza la mantenibilidad o seguridad a largo plazo [19-20], esta línea aplica profundas auditorías de código (evaluando diseño, modularidad, y separación de responsabilidades) sobre los artefactos generados. Además, se audita el proceso humano: se analiza la dinámica de iteración y la calidad de los mensajes (*prompt engineering*) como una evolución moderna de la ingeniería de especificación de requerimientos, buscando comprender las competencias comunicativas necesarias para el éxito en estos nuevos ecosistemas [22-24].

Resultados y Objetivos

A nivel macro, el proyecto en su conjunto tiene como resultados esperados la identificación de procesos socio-productivos susceptibles de mejora en la región, el establecimiento de Indicadores Clave de Rendimiento (KPIs) y la implementación de tecnologías (IA, IoT, robótica) que incrementen la eficiencia operativa y promuevan la sostenibilidad ambiental [26].

En lo que respecta a los resultados experimentales ya obtenidos dentro de la línea de investigación sobre desarrollo de software, los estudios cuasi-experimentales han arrojado

descubrimientos reveladores sobre cómo la IA reconfigura la creación de tecnología [1-2]:

- *Democratización funcional con alta eficiencia*: al contrastar el desempeño de ambos grupos, se observó que los usuarios no expertos asistidos por IA (NP-AI) lograron resolver los desafíos de programación con tasas de éxito (medidas como la capacidad de superar pruebas funcionales) equiparables a las de los programadores trabajando de forma manual (P-noAI). Sin embargo, la ventaja más notable radicó en la eficiencia: el grupo asistido por IA completó los desafíos en tiempos sustancialmente menores. Esta diferencia temporal a favor de los usuarios no expertos se amplificó de manera estadísticamente significativa frente a los problemas de mayor complejidad [1], [21].
- *La IA como "andamiaje" estructural*: para evaluar si el código generado era sostenible a largo plazo, se implementó una estricta auditoría técnica (evaluando semántica, modularidad y manejo de excepciones). Sorprendentemente, los envíos del grupo NP-AI obtuvieron puntuaciones de calidad interna superiores a las de los programadores manuales. No obstante, este hallazgo no implica que los usuarios sin conocimientos posean habilidades superiores de ingeniería; por el contrario, refleja que la IA actúa como un "andamiaje" cognitivo. El modelo integra por defecto buenas convenciones de diseño y separación de responsabilidades que los programadores humanos, bajo la presión del tiempo, tienden a omitir [2], [20].
- *La interacción estructurada como nueva competencia*: dentro del grupo de usuarios no expertos, el éxito no fue producto del azar ni de la automatización total. Se evidenció una fuerte asociación positiva entre la calidad del artefacto de software obtenido y la "calidad de la interacción" humano-máquina. Quienes lograron estructurar metas claras, declarar reglas de negocio y utilizar casos de prueba en sus mensajes (ingeniería de *prompts*), obtuvieron resultados superiores [2], [25]. Por el contrario, un alto volumen de mensajes (mayor cantidad de interacciones) se correlacionó con mayores tiempos de resolución y fracasos, lo que indica que interactuar en exceso suele ser un síntoma de dificultades iterativas y formulaciones ambiguas [16].

En conjunto, estos resultados sugieren que los Grandes Modelos de Lenguaje no eliminan la necesidad del conocimiento experto, sino que actúan como habilitadores que elevan el piso de la calidad estructural, exigiendo a cambio nuevas y sofisticadas habilidades de especificación de requerimientos [18], [22].

Formación de Recursos Humanos

La estructura del equipo de trabajo concibe a la capacitación continua como el motor del proyecto. Se promueve activamente la inscripción y acompañamiento de los miembros en programas de grado y posgrado, apoyando la formulación de tesis orientadas al análisis de datos y ecosistemas humano-IA. Asimismo, se organizan talleres internos de fortalecimiento técnico y se impulsa la interacción colaborativa con otros grupos de I+D+i de la Universidad, así como redes de contacto a nivel nacional e internacional.

Bibliografía

1. L. M. Esnaola, H. D. Ramón, and L. C. Lanzarini, “Can generative AI bridge the gap? A quasi-experimental study of non-programmers with AI vs. programmers without AI,” Empirical Software Engineering, vol. 31, 2026, doi: 10.1007/s10664-026-10813-7.
2. L. M. Esnaola, H. D. Ramón, and L. C. Lanzarini, “Beyond functional correctness: Code quality and human–AI interaction in a quasi-experimental comparison of AI-assisted non-programmers and programmers without AI,” submitted to Springer Nature Computer Science.
3. J. L. del Val Román, “Industria 4.0: la transformación digital de la industria,” in Conferencia de Directores y Decanos de Ingeniería Informática, Informes CODDII, Valencia, Spain, Mar. 2016
4. L. Trujillo Valdiviezo, L. Rodríguez Alegre, D. Mejía Ayala, and R. D. P. López Padilla, “Transformación digital en América Latina: Una revisión sistemática,” Revista Venezolana de Gerencia, vol. 27, no. 100, pp. 1519–1536, Sep. 2022.
5. Y. Lu, “Artificial intelligence: A survey on evolution, models, applications and future trends,” Journal of Management Analytics, vol. 6, no. 1, pp. 1–29, Jan. 2019.
6. M. S. Javaid and A. Haleem, “The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities,” IEEE Access, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3060863.
7. X. Zhai et al., “A review of artificial intelligence (AI) in education from 2010 to 2020,” Complexity, vol. 2021, Art. no. 8812542, 2021, doi: 10.1155/2021/8812542.
8. W. Holmes y I. Tuomi, “State of the Art and Practice in AI in Education,” European Journal of Education, vol. 57, no. 4, pp. 542-570, Dec. 2022, doi: 10.1111/ejed.12533.
9. F. Emmert-Streib, “What Is the Role of AI for Digital Twins?” AI, vol. 4, no. 3, pp. 721–728, Sep. 2023, doi: 10.3390/ai4030038.
10. A. Mozo, A. Karamchandani, S. Gómez-Canaval, M. Sanz, J. I. Moreno, y A. Pastor, “B5GEMINI: AI-Driven Network Digital Twin,” Sensors, vol. 22, no. 11, p. 4106, 2022, doi: 10.3390/s22114106.
11. A. Alonso-Robisco and J. M. Carbó, “Inteligencia artificial y finanzas: Una alianza estratégica,” Banco de España, Documentos Ocasionales no. 2222, 2022.
12. J. Prather, P. Denny, J. Leinonen, et al., “The robots are here: Navigating the generative AI revolution in computing education,” in Proc. 2023 Working Group Reports on Innovation and Technology in Computer Science Education (ITiCSE-WGR ’23), New York, NY, USA: ACM, 2023, pp. 108–159, doi: 10.1145/3623762.363349.
13. A. J. Ko, R. Abraham, L. Beckwith, et al., “The state of the art in end-user software engineering,” ACM Computing Surveys, vol. 43, no. 3, Art. no. 21, pp. 1–44, 2011, doi: 10.1145/1922649.1922658.
14. M. O. Ajimati, N. Carroll, and M. Maher, “Adoption of low-code and no-code development: A systematic literature review and future research agenda,” Journal of Systems and Software, vol. 222, Art. no. 112300, 2025, doi: 10.1016/j.jss.2024.112300.
15. S. Barke, M. B. James, and N. Polikarpova, “Grounded copilot: How programmers interact with code-generating models,” Proc. ACM Program.

Lang., vol. 7, no. OOPSLA1, Art. no. 78, pp. 1–27, 2023, doi: 10.1145/3586030.

16. H. Hao, K. A. Hasan, H. Qin, et al., "An empirical study on developers' shared conversations with ChatGPT in GitHub pull requests and issues," Empirical Software Engineering, vol. 29, Art. no. 150, 2024, doi: 10.1007/s10664-024-10540-x.
17. A. Ziegler, E. Kalliamvakou, X. A. Li, et al., "Measuring GitHub Copilot's impact on productivity," Communications of the ACM, vol. 67, no. 3, pp. 54–63, 2024, doi: 10.1145/3633453.
18. Z. Zheng, K. Ning, Q. Zhong, et al., "Towards an understanding of large language models in software engineering tasks," Empirical Software Engineering, vol. 30, 2025, doi: 10.1007/s10664-024-10602-0.
19. N. Tihanyi, T. Bisztray, M. A. Ferrag, et al., "How secure is AI-generated code? A large-scale comparison of large language models," Empirical Software Engineering, vol. 30, 2025, doi: 10.1007/s10664-024-10590-1.
20. Y. Liu, T. Le-Cong, R. Widyasari, et al., "Refining ChatGPT-generated code: Characterizing and mitigating code quality issues," ACM Transactions on Software Engineering and Methodology, vol. 33, no. 5, 2024, doi: 10.1145/3643674.
21. M. S. Sheikhaei, Y. Tian, S. Wang, et al., "An empirical study on the effectiveness of large language models for SATD identification and classification," Empirical Software Engineering, vol. 29, 2024, doi: 10.1007/s10664-024-10548-3.
22. X. Amatriain, "Prompt design and engineering: Introduction and advanced methods," arXiv preprint arXiv:2401.14423, 2024, doi: 10.48550/arXiv.2401.14423.
23. R. Feldman and C. Anderson, "Non-expert programmers in the generative AI future," in Proc. 2024 ACM Symp. Human-Computer Interaction for Work (CHIWORK '24), 2024, doi: 10.1145/3663384.3663393.
24. B. Nuseibeh and S. Easterbrook, "Requirements engineering: A roadmap," in Proc. Conf. Future of Software Engineering (FOSE), 2000, pp. 35–46, doi: 10.1145/336512.336523.
25. M. Pickering, H. Williams, A. Gan, et al., "How humans communicate programming tasks in natural language and implications for end-user programming with LLMs," in Proc. 2025 CHI Conf. Human Factors in Computing Systems (CHI '25), 2025, doi: 10.1145/3706598.3713271.
26. C. Surianarayanan, J. J. Lawrence, P. R. Chelliah, E. Prakash, and C. Hewage, "A survey on optimization techniques for edge artificial intelligence (AI)," Sensors, vol. 23, no. 3, p. 1279, Jan. 2023, doi: 10.3390/s23031279.
27. C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, Experimentation in Software Engineering. Berlin, Heidelberg: Springer, 2012, doi: 10.1007/978-3-642-29044-2.

La Ingeniería de Software aplicada al diseño de Interfaces Gráficas de Usuario Accesibles

M. Claudia Albornoz, Mario Berón, Germán Montejano

Departamento de Informática/Universidad Nacional de San Luis-U.N.S.L./San Luis/Argentina
Ejército de los Andes 950, Tel: +54 (0266) 4520300 (int 2127)
{albornoz,mberon,gmonte}@unsl.edu.ar

Resumen

En el mundo actual todas las personas se vean obligados a interactuar con productos de software. Al comenzar a trabajar con cualquier dispositivo (tablet, computadora, cajero, celular, etc.) el usuario comienza a interactuar con la interfaz, ya sea la del sistema operativo, la de un software/aplicación en particular o la de cualquier sitio web. La interfaz no sólo debe ser visualmente atractiva, sino también intuitiva, fácil de usar, adaptable a las necesidades del usuario.

El diseño de interfaces gráficas de usuario es una terea crucial para garantizar que todas las personas, independientemente de sus capacidades, puedan interactuar eficazmente con los sistemas de software. Este artículo aborda técnicas de la ingeniería de software que facilitan el diseño y construcción de interfaces gráficas de usuario (siglas en inglés GUI: Graphical User Intrface) accesibles, destacando la importancia de un enfoque centrado en el usuario. Además, resalta la importancia de contar en el equipo de desarrollo con integrantes especializados en el tema.

El objetivo del presente trabajo es mostrar la importancia del diseño de la GUI utilizando técnicas para su construcción y mostrar la importancia de realizar un diseño centrado en el usuario.

***Palabras Claves**:*

Interfaz Gráfica de Usuario (GUI), Igeniería se Software, Interacción Persona-Computadora (IPC), Diseño Centrado en el Usuario (DCU)

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: "Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube". Dirección: Dr. Daniel Riesco, código P-032020 de la Facultad de Ciencias Físico, Matemáticas y Naturales, de la Universidad Nacional de San Luis. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se han logrado importantes vínculos con diferentes universidades a nivel internacional y nacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. Introducción

La accesibilidad no es sólo una palabra de moda; en cuanto a diseño de software es una necesidad imperativa para la sociedad actual. Las interfaces gráficas de usuario (GUI) deben ser diseñadas para ser accesibles, no sólo por por personas con diversas discapacidades; sino también por una población cada vez más numerosa de adultos mayores. Por esto es imperativo que el diseño de las GUIs considere requisitos ergonómicos cognitivos [15].

Todo producto de software tiene como principal objetivo permitir a los usuarios concretar sus tareas de manera eficiente y efectiva. Para que esto sea posible, no solo se debe poner énfasis en la funcionalidad del producto, sino también en la interfaz.

Es necesario incorporar técnicas que ayuden a captar las necesidades del usuario, y así poder lograr interfaces intuitivas, fáciles de usar y satisfactorias. Algunas de

estas técnicas son el análisis de casos de uso y de tareas [1].

El desarrollo de software a través de la aplicación estructurada de técnicas científicas, se lo considera una ingeniería. Así mismo, el desarrollo y la construcción de interfaces gráfica de usuario se convierte en una ingeniería al utilizar técnicas específicas de ingeniería [2]. La utilización de estas técnicas asegura un producto de software de calidad.

Desde hace ya un tiempo, para el desarrollo de interfaces se aplican metodologías que involucra al usuario en todo el ciclo de vida del diseño, y se las conoce como metodologías centradas en el usuario [10]. Lo cual implica, según la ISO/DIS 13407 (ISO-98) [6]:

- Activa participación del usuario en todas las etapas del desarrollo.
- La iteración de las mejoras de diseño en las que la retroalimentación del usuario es esencial como fuente de información.
- Una perspectiva de diseño multidisciplinar que requiere de gran variedad de disciplinas o especialidades (desarrolladores de software, expertos en usabilidad, accesibilidad, especialistas en ergonomía, etc)

Es necesario tener en claro que los productos de software deben mostrar un fluido diálogo con el usuario; pero el usuario no está interesado en la estructura interna del producto, lo único que le importa es cómo usarlo, cómo comunicarse, cómo concretar las tareas [3].

El diálogo se realiza a través de la interfaz del producto; el usuario se comunica, interactúa por medio de la parte visible del sistema. Es importante considerar que el producto puede ser utilizado por usuarios inexpertos, con discapacidad o adultos mayores.

Lo que antes se tomaba como tarea secundaria y complementaria, hoy cuenta con una metodología y una planificación para lograr la mayor fluidez en la interacción. Actualmente se le ha dado protagonismo al diseño y construcción de la interfaz de usuario. Es por esto que se aplican técnicas de la ingeniería del software para lograr un producto de calidad, las mismas son utilizadas con la participación del usuario.

Además de la fluidez y la satisfacción, es sumamente importante la accesibilidad.

'Accesibilidad' se define como la posibilidad de que un producto de software o servicio web pueda ser usado por el mayor número de usuarios posible, más allá de las limitaciones propias del individuo o de las derivadas del contexto en uso [11].

La accesibilidad hace posible que todo tipo de personas tengan la posibilidad de acceder a distintos productos digitales, dentro de los cuales se ven beneficiadas personas con diferentes discapacidades, incluyendo problemas visuales, auditivos, físicos, cognitivos, neurológicos y del habla; pero también beneficia a personas sin discapacidades, incluyendo personas de edad avanzada que han visto reducidas sus habilidades a consecuencia de la edad [14].

La norma ISO/TC 16027 [12], define 'accesibilidad' como la facilidad de uso en forma eficiente, eficaz y satisfactoria de un producto, servicio, entorno o instrumento por personas con discapacidad. De este modo, accesibilidad tecnológica hace referencia a que productos y servicios tecnológicos puedan ser utilizados por los usuarios con efectividad, eficiencia y satisfacción en un contexto de uso determinado [13].

2. Análisis de técnicas

La técnica de Casos de Uso es muy utilizada en ingeniería del software, dado que permite describir los límites de un producto de software y también describir los requisitos del mismo.

La Ingeniería de Software la define como una técnica que [4]:

- Describe el comportamiento de un sistema desde el punto de vista del usuario en términos de los requisitos funcionales y por tanto se pueden

considerar el punto de partida para poder especificar las necesidades de los usuarios finales.
- Permite definir los límites del sistema.
- El usuario debería poder entender cada caso de uso para poder validarlo.
- Guía todo el proceso de desarrollo del sistema.

Estos cuatro ítems se los puede considerar como los puntos fuertes de la técnica. Sin embargo, es importante mencionar sus debilidades, y es que no capturan requisitos no funcionales, tal como fiabilidad, portabilidad, mantenibilidad, eficiencia y sobre todo usabilidad; y tampoco sirven para capturar las necesidades de los usuarios, en cuanto a la interacción (aspecto esencial a la hora de desarrollar interfaces de usuario).

Es por esto que surge la necesidad de encontrar la manera de revertir esa debilidad. Para esto se utiliza el Análisis de Tareas, técnica que se enfoca en capturar los requisitos de la Interfaz del Usuario. Analizar las tareas permite obtener información respecto a cómo los usuarios realizan su trabajo y así conocer la funcionalidad que el sistema debe mostrar a través de la interfaz.

Los objetivos fundamentales de esta técnica son [5]:
- Identificar usuarios (características principales: experiencia, conocimientos previos, grupo etario, etc).
- Identificar qué debe hacer el usuario (qué quiere hacer el usuario).
- Comprender el dominio del problema.
- Describir el mecanismo de interacción.
- Considerar las diferentes respuestas del usuario.

De manera sintética se puede decir que el análisis de tareas es desarrollar modelos genéricos, abstractos de las tareas del usuario.

Como diferencia entre las dos, se puede mencionar que la técnica de Casos de Uso es utilizada por ingenieros de software que se ocupan del desarrollo del sistema; en tanto que el Análisis de Tareas se utiliza en el área de Interacción Humano-Computadora, donde se hace mayor énfasis en los requisitos para el desarrollo de la interfaz de usuario.

La Interacción Persona Ordenador IPO es una nueva disciplina que surge con el objetivo de lograr el diálogo fluido y satisfactorio; más conocida por sus siglas en inglés HCI (Human Computer Interaction). La IPO se centra en estudiar la interacción entre usuarios y sistemas informáticos; y se ha convertido en un área muy importante dentro de la Ingeniería de Software. A modo de definición, se puede decir que la interacción persona-computadora es el intercambio observable de información, datos y acciones entre un humano (usuario) y la computadora, y viceversa [9].

Ambas técnicas, Casos de Uso y Análisis de Tareas, son utilizadas en las etapas iniciales del diseño y construcción de un sistema de software.

3. Diseño centrado en el usuario

El Diseño Centrado en el Usuario (DCU) es el término que se utiliza para describir el diseño en el que el usuario influye en el resultado final. Se considera que es una filosofía y un proceso. Una filosofía porque sitúa al usuario en el centro con la intención de desarrollar un producto adecuado a sus requerimientos, necesidades y un proceso de diseño porque se centra en los factores cognitivos de las personas y como éstos intervienen en sus interacciones con los productos [7].

Esta metodología consta de diferentes etapas (Fig. 1) [8]. Si bien estas etapas corresponden a todo el proceso de desarrollo, es obvio que el diseño y desarrollo de la Interfaz se ven involucrados en ellas. Como es evidente, el usuario se ve involucrado en cada etapa o fase.

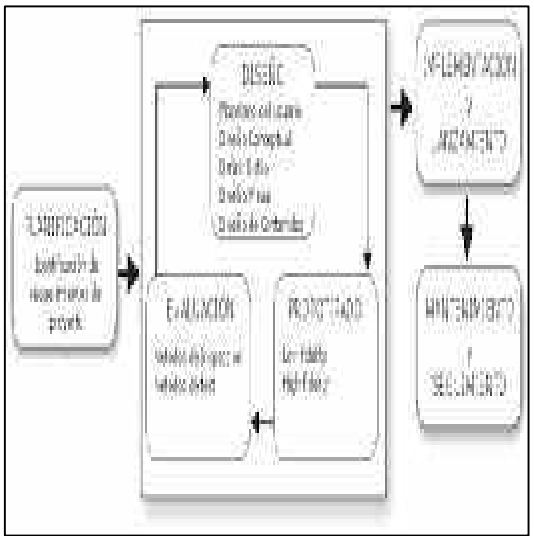


Fig. 1: Etapas del DCU

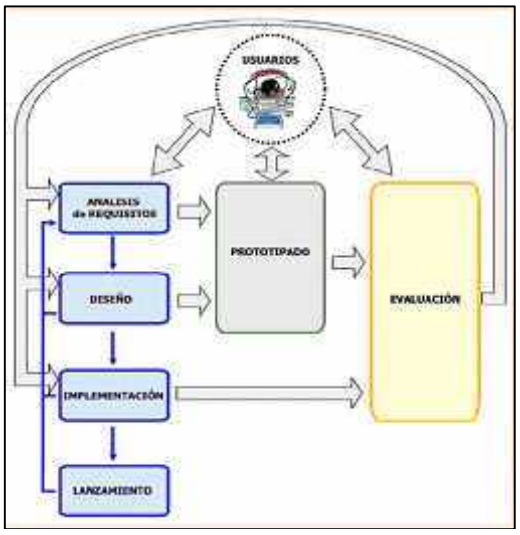


Fig. 2: Ciclo de vida de sistemas interactivos.

El concepto Diseño Centrado en el Usuario nació tras la investigación de Donald A. Norman en la Universidad de California San Diego (UCSD); autor del libro *User centered system design: new perspectives on human-computer interaction,* y posteriormente se consolidó con el libro de Norman *The design of everyday things*. Obra referente en el ámbito del DCU y las disciplinas afines como la interacción persona-computadora, el diseño y la experiencia de usuario.

El usuario final se ve involucrado en cada etapa o fase del proceso de desarrollo, con esto se garantiza que el producto se ajuste a sus necesidades. Ésta forma de diseño le otorga al usuario el rol central del desarrollo. En ocasiones, el usuario es un miembro más del equipo de diseño. El objetivo es obtener productos fáciles de usar, efectivos y eficientes.

La Fig. 2 muestra como característica fundamental del ciclo de vida la presencia constante del usuario, desde la fase de análisis hasta en la evaluación; como también la naturaleza cíclica, dado que hay un a continua retroalimentación a partir de la fase o etapa de evaluación.

Cuanto más se conozcan las características y limitaciones del usuario final, mas cercano estará el equipo de desarrollo a entregar un producto eficiente y accesible.

4. Actividades Llevadas a Cabo en la Línea de Investigación

Hasta el momento, en la línea de investigación descripta previamente, se están llevando a cabo las siguientes tareas:

- Revisión Sistemática de la literatura referente al tema de diseño y construcción de Interfaces Gráficas de Usuario; con el fin de evidenciar la importancia de su diseño y cómo lograr una óptima y satisfactoria interacción con los usuarios. Haciendo una mención especial respecto al rol del usuario en el diseño de la Interfaz.
- Estudio de Métodos de Evaluación Multicriterios. Esta tarea tuvo como objetivo seleccionar un método de evaluación para evaluar interfaces la calidad de las interfaces gráficas de usuario. Hasta el momento se ha encontrado que el método LSP (Logic Scoring of Preference) se adapta muy bien para la evaluación de las GUI.
- Elaboración de Criterios de Evaluación. Esta tarea se lleva a cabo con el objetivo de establecer cuáles son las características que las interfaces gráficas de usuario deben tener para que sean consideradas adecuadas para su utilización por parte del usuario. Estos criterios serán la entrada al

método de evaluación multicriterio mencionado en el ítem precedente.

- Estudio de Reglas, Técnicas y Principios de Diseño de GUIs propuestos por expertos. Esta tarea provee información relevante en el desarrollo de GUIs.
- Análisis de la Influencia Emocional, el objetivo es determinar cómo influye el estado emocional de los usuarios en la interacción con las computadoras y, a su vez, cómo el diseño de la Interfaz influye en el estado emocional del usuario.

5. Conclusiones y Trabajos Futuros

Es esencial reconocer que el diseño de la interfaz no es una tarea secundaria. Además, no sólo involucra un grupo de expertos en programación; sino que es sumamente importante que el equipo involucre a los usuarios. Diversos estudios demuestran que cada vez son más las investigaciones que se realizan con el objetivo de lograr una óptima interacción.

Como resultado, luego de la lectura de diferentes trabajos, esta investigación pretende determinar cómo lograr una Interfaz gráfica de usuario accesible, comprensible, satisfactoria; con el análisis de los diferentes tópicos que se involucran en la Interacción Persona-Computadora. También pretende mostrar la importancia de aplicar técnicas de la Ingeniería de Software en todos los aspectos de un producto de software; no sólo el funcional sino también en la interacción con los usuarios.

La inclusión de personas con discapacidad es fundamental en el mundo tecnológico actual; así también como los adultos mayores, quienes no son nativos tecnológicos pero que se ven afectados en diferentes aspectos de la vida cotidiana.

Futuros trabajos: i) Enriquecer el estudio sistemático de la literatura; ii) Implementar un método de evaluación multicriterio para evaluar GUIs, iii) Definir criterios de evaluación de interfaces gráficas de usuario, iv) Elaborar un manual con las Reglas, Técnicas y Principios de Diseño de GUI, v) Profundizar el estudio sobre la influencia emocional y vi) Profundizar el estudio de técnicas y métodos que garanticen la accesibilidad en el diseño de GUI.

6. Formación de Recursos Humanos

Los progresos obtenidos en esta línea de investigación sirven como base para el desarrollo de tesis de posgrado, ya sea de doctorado, maestrías o especializaciones en Ingeniería de Software; y desarrollo de trabajos integradores finales de las carreras Licenciatura en Ciencias de la Computación, Ingeniería en Informática e Ingeniería en Computación de la Universidad Nacional de San Luis, en el marco de los Proyectos de Investigación.

Referencias

[1] Lozano, María & González, Pascual & Ramos, Isidro & Montero, Francisco & Molina Masso, Jose Pascual. (2002). Desarrollo y generación de interfaces de usuario a partir de técnicas de de Análisis de Tareas y Casos de Uso. Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial. 6. 83-92. 10.4114/ia.v6i16.745.

[2] Losada, Begoña, David López, and Javier Martínez. "Guía de actuación en el desarrollo de interfaces de usuario según la metodología centrada en el usuario INTERGRAM." VIII Congreso Internacional de Interacción Persona-Ordenador Interacción. Vol. 7. 2007.

[3] http://www.lsi.us.es/docencia/get.php?id=4330
Material de estudio del Departamento de Lenguajes y sistemas Informáticos – Universidad de Sevilla. Ultimo acceso: 24/02/2026

[4] https://www.ctr.unican.es/asignaturas/is1/is1-t08-trans.pdf
Material de estudio INGENIERÍA DEL SOFTWARE -Contexto y Requisitos del Sistema (en desarrollo OO) Univ. Cantabria – Fac. de Ciencias. Ultimo acceso: 24/02/2026

[5] Gea, İ. Miguel, and Fco Luis Gutiérrez. "Diseño centrado en el usuario." (2011).

http://lsi.ugr.es/~mgea/docencia/diu/Temario/Diu_Tema3.pdf Ultimo acceso: 24/02/2026

[6] ISO (International Organization for Standardisation), 13407, p.7, Human Centered Design Processes for Interactive Systems, 1997

[7] Domingo, Muriel Garreta, and Enric Mor Pera. Diseño centrado en el usuario, 2010.

[8] Yusef Hassan & Francisco J. Martín Fernández & Ghzala Iazza. Diseño Web Centrado en el Usuario: Usabilidad y Arquitectura de la Información. "Hipertext.net", núm. 2, 2004. <http://www.hipertext.net>

[9] Díaz, F. J., Harari, I., and Amadeo, A. P. (2013). Guía de recomendaciones para diseño de software centrado en el usuario. Editorial de la Universidad Nacional de La Plata (EDULP).

[10] Galeano, Roylan. "Diseño centrado en el usuario." Revista q (2008).

[11] Prefasi Gomar, Salvador, et al. "Tecnologías de la Información y de la Comunicación orientadas a la educación de personas con discapacidad cognitiva". RELATEC: revista latinoamericana de tecnología educativa. 2010.

[12]ISO/TS 16071 International Organization for Standardization/Technical Specification. Ergonomics of Human-System Interaction - Guidance on Accessibility for Human-Computer Interfaces, 2002.

[13] Bustos Torres, Brenda V., Adriana Elba Martín, and Alejandra Cechich. "Extendiendo MVC para diseñar interfaces *de usuario accesibles"*. XIV Congreso Argentino de Ciencias de la Computación. 2008.

[14] CAMILA ANDREA SANTIAGO VALLEJO. Trabajo: "PROPUESTA DE UN DISEÑO CENTRADO EN EL USUARIO DE LA INTERFAZ GRÁFICA A PARTIR DEL MODELO DE PROCESO DE LA INGENIERÍA DE LA USABILIDAD Y ACCESIBILIDAD (MPIU+ a) PARA LA ARQUITECTURA DE INFORMACIÓN DE LAS PLATAFORMAS E-COMMERCE DE LA EDITORIAL." Colombia, (2022).

[15] Maldonado, T. V. P. (2023). Diseño de interfaces de usuario centrado en el usuario: mejores prácticas y tendencias. *Revista Académica y científica VICTEC*, *4*(7).

MANTENIBLIDAD DE SOFTWARE, UN APORTE DE LA ACADEMIA A LA INDUSTRIA

Sonia I. Mariño, Pedro L. Alfonzo, María de los A. Ferraro, Luciana M. Aguirre, Ricardo Ojeda, Yesica D. Vallejos

Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Corrientes, Argentina
soniaimarinio@gmail.com, plalfonzo@hotmail.com

RESUMEN

Esta línea de I+D se establece con el objetivo de sistematizar y profundizar conocimientos teóricos, metodológicos y empíricos en torno al Mantenimiento de Software. Se espera que los resultados contribuyan a la calidad de la educación superior en temas concernientes al Mantenimiento de Software e identificar los aportes entre espacios universitarios y el entorno. Se presentan algunos resultados y las actividades previstas.
Palabras Claves: Ingeniería del Software, Mantenimiento del Software, Guía SWEBOK, Educación Superior

CONTEXTO

En el Grupo de Investigación Sistema de información y TI: modelos, métodos y herramientas acreditado por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste, se desarrollan actualmente, 2 proyectos de investigación: Sistemas informáticos y Gestión del conocimiento: modelos, métodos y herramientas (PI 22F025, Res. Nº 0931/22 CS UNNE) y El Observatorio de Tecnologías Aplicadas a la Educación de la FaCENA-UNNE (21F010 –SGCyT – UNNE

1. INTRODUCCIÓN

La Ingeniería del Software (IS) es una de las disciplinas de ingeniería que comprende los aspectos de la producción de software desde las etapas iniciales de la especificación de un sistema, hasta el mantenimiento de este después de su implementación [1].

En [2] se define la IS como: 1) La aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento de software; es decir, la aplicación de la ingeniería al software. 2) El estudio de enfoques como en 1.

En referencia a la IS, Piattini [3] realiza un recorrido histórico desde la década de 1940 hasta la década de 2010, y parafrasea Boehm, dado que la IS ha seguido un proceso de tesis, antítesis y síntesis.

Una de las actividades fundamentales en el proceso de la IS, es la evolución del software, siendo en esta etapa donde se dónde se modifica el software para reflejar los requerimientos cambiantes del cliente y del mercado. Así, se observa un área de conocimiento relevante en la IS es el mantenimiento.

Por otra parte, la guía SWEBOK v 4.0 a través de 18 capítulos, presenta actualizaciones en todas las áreas de conocimiento (KAS, por sus siglas en inglés) de la Ingeniería del Software, siendo una de ellas el Mantenimiento del Software (MS) [4].

Respecto al MS, el objetivo es modificar el software existente conservando su integridad. El Instituto de Ingeniería Eléctrica y Electrónica define al MS como la modificación de un producto software, después de su entrega para corregir fallos, mejorar el rendimiento u otros atributos, o para adaptar el producto a otro entorno. En este sentido, [1] expresa que el proceso de desarrollo del software continúa durante su tiempo de vida y que inevitablemente debe modificarse para su permanencia y utilidad. En la guía SWEBOK [3] se especifican cuatro aspectos del MS, identificar y difundir las actividades

concernientes a la evolución del software requiere procesos de relevamiento, captura y sistematización de la información, que podría derivar en la definición de una guía de buenas prácticas.

Los beneficiarios de esta línea de I+D son la academia representada a través de los docentes, investigadores y estudiantes de grado y posgrado involucrados, y los destinatarios de estas producciones teóricas y empíricas propuestas como mejoras superadoras.

El resto de este artículo se organiza de la siguiente manera: la sección 2 presenta los objetivos de investigación, diferenciando entre general y específicos, la sección 3 los resultados esperados/obtenidos y, por último, la sección 4 la continuidad en la formación de recursos humanos.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Esta línea de trabajo, en el marco del proyecto y del grupo de investigación, se enmarca en el área de MS de la guía de conocimiento SWEBOK. Se propone indagar, modelar y validar para mejorar la gestión de los procesos desde el MS.

Además, la línea de I+D en torno al MS se plantea para explicitar los estudios y desarrollos que se llevan a cabo con los integrantes del PI, para su sistematización y difusión. Se definen como objetivo general y específicos:

Objetivo general

- Profundizar conocimientos teóricos y metodológicos concernientes al MS y articular actividades de docencia, investigación y extensión, para contribuir al sector de Software y Servicios Informáticos (SSI) en que se insertan los estudiantes y graduados de la carrera.

Objetivos específicos

- Indagar los aspectos concernientes al MS desde bibliografía específica [1] [3] [6] [7] [8] [9].
- Actualizar y relevar prácticas del MS implícitas y explícitas desarrolladas en algunas asignaturas de una carrera de la disciplina Informática como: Taller de Programación II, Ingeniería de Software II, Inteligencia Artificial, Modelos y Simulación y Proyecto Final de Carrera.
- Contribuir a la formación de recursos humanos de grado y posgrado, desde becas, pasantías, proyectos de finalización de carrera, identificando el MS como un aspecto esencial en el ciclo de vida del software, y que promueve articulación entre la industria y otros ámbitos.

Para el logro de las líneas de I+D se realizarán estudios descriptivos y exploratorios:

- Se realizará una revisión de la literatura respecto al proceso de enseñanza aprendizaje respecto al MS desde ámbitos académicos con énfasis en la industria [10] [11] [12].
- Se privilegiará la profundización de la metodología ISBE (Ingeniería del Software Basada en Evidencia) como marco conceptual y empírico.
- Se identificarán y sistematizarán las metodologías de la IS y del MS reportadas [6] [13], así como las adecuaciones publicadas [7] [14] [15] [16] [17].
- Se identificarán estándares que tratan el MS como se menciona en [18], [19].
- Se difundirán los hallazgos con la finalidad de contribuir a la formación de RRHH.

3. RESULTADOS ESPERADOS/OBTENIDOS

Entre los logros del equipo, se mencionan:

- En el año 2024 se defendieron 2 tesis de posgrado en el marco de una maestría interinstitucional de carácter profesional. [24] [25]
- En el año 2025 se inició una tesis de posgrado en el marco de una maestría interinstitucional de carácter profesional. [26]
- Se dirigió y concretó el desarrollo de un Proyecto Final de Carrera que aborda temas de investigación y aplicación de la evolución del software, El MS responde a un sistema

informático diseñado en otra asignatura y de concreta aplicación al sector del emprendedorismo. [27]

- El equipo cuenta con antecedentes de articulación de conocimientos y experiencias en torno al MS y otras áreas de la disciplina Informática. Se propone capitalizar el conocimiento adquirido, y profundizarlo con fines de transformar productos de I+D en recursos educativos.

Se propone aportar a la educación superior, previendo:

- Especificar vinculaciones teóricas, metodológicas y empíricas en torno al MS y otras áreas de conocimiento de la disciplina Informática.
  - Se actualizará la investigación bibliográfica para sistematizar antecedentes.
  - Se indagará en torno a la guía SWEBOK y las distintas posibilidades de definición o redefinición de actividades de docencia, investigación y extensión en el grado y posgrado.
  - Se identificarán actividades en ejecución y los recursos humanos involucrados, que podrían ser sistematizadas desde este enfoque
- Diseñar capacitaciones en torno al Mantenimiento de Software, con la finalidad de lograr aprendizajes significativos en los futuros profesionales del sector.
- Incrementar las habilidades de trabajo en equipo y comunicación efectiva en los recursos humanos involucrados en el proyecto y potenciales pasantes y/o becarios, optando como línea de I+D, el MS.
- Siguiendo a [20] se abordará el estudio de aspectos humanos, sociales, sociotécnicos y organizacionales (por ejemplo, perfiles definidos, características y/o experiencias de las partes interesadas), y aquellos referidos al contexto [21].
-

4 FORMACIÓN DE RECURSOS HUMANOS

En referencia a la formación de recursos humanos, se destaca como valor la transferencia de conocimientos implícita o explícita a la industria o diversos ámbitos laborales a la cual se incorporan quienes participan en este equipo.

- Aportes de vinculación de MS y temas concernientes a la Accesibilidad Web:
  - Desarrollo de un Trabajo Final de Aplicación que abordó temas de investigación y aplicación de MS en relación a la Accesibilidad Web [22],
  - Formación en el marco de una beca de iniciación a la investigación [23].

- Aportes en actividades de posgrado:
  - En el año 2024, se defendieron dos tesis de posgrado, en el marco de una maestría interinstitucional UNNE – UNAM [24], [25].
    En [24], el trabajo consistió en el diseño y validación de un procedimiento para mantener procesos de gestión decisorios en una empresa regional dedicada al transporte. La concreción de este estudio de posgrado permitió explicitar la vinculación universidad – industria.
    En [25], se propuso un proceso de mantenimiento y se validó la propuesta desde una abstracción de un Subsistema de Requerimientos, problemática presente en diversas áreas de la gestión pública. La propuesta se diseñó realizando una adaptación de la guía Mantelasoft.
  - En el año 2025, se aprobó un proyecto de tesis de Maestría en Tecnologías de la Información, Universidad Nacional del Nordeste [26]. El proyecto tiene por objetivo diseñar e implementar un procedimiento para apoyar la Gestión del Conocimiento de los requisitos en una organización de medicina prepaga haciendo hincapié en su documentación.

- Se desarrolla un proceso y su herramienta software que tiene por objetivo diseñar e implementar la Transformación de la Gestión de Calidad en un Sistema de Integrado de Información Financiera Crítico (S.I.I.F.) en una organización del estado adaptando ITIL y aspectos de calidad [28] [29].

Se prevé:

- Capacitar a los integrantes y estudiantes en temas concernientes al MS y su relevancia en ámbitos del SSI.
- Propiciar la elaboración de planes de trabajo de pasantías o proyectos finales para formar recursos humanos de grado y posgrado quienes se insertan en el SSI, y requieren capitalizar en un marco de mejora continua desde el concepto del MS.

5. BIBLIOGRAFIA

[1] I. Sommerville, *Ingeniería de software*. Pearson Educación: Madrid, España. 2011.

[2] IEEE Std 610.12. Institute of Electrical and Electronics Engineers. IEEE Standard Glossary of Software Engineering Terminology. 1990.

[3] M. Piattini, Evolución de la Ingeniería del Software y la formación de profesionales, Revista Institucional de la Facultad de Informática, N 4.

[4] IEEE Computer Society. SWEBOK Guide V4.0 Topics. Software. https://www.computer.org/education/bodies-of-knowledge/software-engineering/topics

[5] IEEE STD 1219. Standard for Software Maintenance. IEEE Computer Society Press. USA. 1993

[6] M. Piattini, J. Villalba, F. Ruiz, T. Bastanchury, M. Polo, M. A. Martínez y C. Nistal, *Mantenimiento del software. Modelos, técnicas y métodos para la gestión del cambio*. 1ª Edición. RA-MA Editorial, Madrid. 2000.

[7] J. Adones Farfán y V. Vega-Zepeda, Mantenibilidad del Software. Consideraciones para su especificación y validación. *Ingeniare. Revista chilena de ingeniería*, *28*(4), 654-667. 2020. https://dx.doi.org/10.4067/S0718-33052020000400654

[8] ISO/IEC/IEEE, Draft International Standard - Software Engineering - Software Life Cycle Processes - Maintenance," in ISO/IEC/IEEE P14764/CD2, 2020.

[9] R. S. Pressman, Ingeniería del software. Un enfoque práctico. Séptima edición. McGraw-Hill. 2010.

[10] K. Gallagher, M. Fioravanti y S. Kozaitis, Teaching Software Maintenance, *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, Cleveland, OH, USA, 2019, pp. 353-362. doi: 10.1109/ICSME.2019.00054.

[11] R. Hardt. A toolset to support a software maintenance process in academic environments. *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, Adelaide, SA, Australia, 2020, pp. 814-817. doi: 10.1109/ICSME46990.2020.00100.

[12] J. P. Gibson y J. Raffy, A "Future-Proof" Postgraduate Software Engineering Programme: Maintainability Issues, *ICSEA 2011: The Sixth International Conference on Software Engineering Advances. Barcelona, Spain.*

[13] M. Almashhadani, A. Mishra, A. Yazici y M. Younas. Challenges in Agile Software Maintenance for Local and Global Development: An Empirical Assessment. *Information*, *14*(5), 261. 2023. https://doi.org/10.3390/info14050261

[14] I. M. Suarez, A. F. Montoya y J. I. Velez, Mantelasoft: una nueva guía para Mipymes desarrolladoras de mantenimiento de software. *Entre Ciencia e Ingeniería, Pereira*, 13(25), 95-99. 2019. https://doi.org/10.31908/19098367.4019

[15] R. A. Aguilar Vera, J. Baeza Pérez, J. C. Díaz Mendoza y A. A. Aguileta Güemez, Reverse engineering for software maintenance: a systematic review of literature. *Revista Ingenierías*

Universidad de Medellín, *23*(44), 1-18. 2024.

[16] E. Juergens. Software maintenance research that is empirically valid and useful in practice. *it – Information Technology*. 58(3). 145–149. 2016. https://doi.org/10.1515/itit-2016-0014

[17] A. A. Rabello de Carvalho y L. E. Galvão Martins. *Maintenance and evolution processes of embedded software: A systematic literature review*. *Journal of Computer Science,* 21(7), 1539–1553. 2025. https://doi.org/10.3844/jcssp.2025.1539.1553

[18] ISO/IEC/IEEE Draft International Standard - Software Engineering - Software Life Cycle Processes - Maintenance, in *ISO/IEC/IEEE P14764/CD2*. pp.1-43. 9 July 2020.

[19] ISO/IEC/IEEE Approved Draft International Standard - Software Engineering - Software Life Cycle Processes – Maintenance. *in ISO/IEC/IEEE P14764/FDIS*. pp.1-46. 13 Dec. 2021.

[20] S. Uchitel, M. Chechik, M. Di Penta, B. Adams, N. Aguirre y G. Bavota, Scoping Software Engineering for AI: The TSE Perspective, in *IEEE Transactions on Software Engineering*, 50(11), 2709-2711, 2024, doi: 10.1109/TSE.2024.3470368.

[21] T. Zhang, Y. Lu, Y. Yu, X. Mao, Y. Zhang y Y. Zhao, How Do Developers Adapt Code Snippets to Their Contexts? An Empirical Study of Context-Based Code Snippet Adaptations, *in IEEE Transactions on Software Engineering*, 50(11), 2712-2731, 2024. doi: 10.1109/TSE.2024.3395519

[22] D. Casaro, P. Alfonzo, S. Mariño y M. Godoy, Mantenimiento Correctivo Aplicado a un Sitio Basado en Joomla. Una Propuesta Centrada en la Accesibilidad. *Revista Latinoamericana de Ingeniería de Software*, 3(2): 101-107.2015.

[23] A. Codutti, S. I. Mariño y P. Alfonzo, Una propuesta integradora de Mantenimiento Correctivo aplicada al Diseño Web Adaptativo y Accesibilidad Web. *Scientia et technica*. Año XXI, 21(1). 2016.

[24] L. M. Aguirre, "Proceso para el mantenimiento del software. Propuesta para una PyME de servicios de la Región NEA". Tesis de Maestría en Tecnologías de la Información, Universidad Nacional del Nordeste, 2024. Directora: Sonia I. Mariño.

[25] R. E. Ojeda, "Diseño de un proceso basado en MANTELASOFT. Caso de estudio: subsistema de requerimientos". Tesis de Maestría en Tecnologías de la Información, Universidad Nacional del Nordeste, 2024. Directora: Sonia I. Mariño.

[26] Y. D. Vallejos Torres, "Procedimiento de Gestión del Conocimiento de los requisitos con enfoque en la Documentación. Estudio de caso en una organización de medicina prepaga. Proyecto de tesis de Maestría en Tecnologías de la Información, Universidad Nacional del Nordeste, 2025. Directora: Sonia I. Mariño. Co-Director: Pedro L. Alfonzo.

[27] D. A. Fernández y N. María, "Solución web para autogestión en fast food". Proyecto final de Carrera, Universidad Nacional del Nordeste, 2025. Profesor orientador. Pedro L. Alfonzo.

[28] M. de los A. Ferraro, S. I. Mariño, L. Antonelli, Innovación en el Estado: Mejora Continua con ITIL y Metodologías Ágiles en el Sistema Integrado de Información Financiera [S.I.I.F.] Simposio de Informática en el Estado (SIE 2025) - JAIIO 54, 2025.

[29] M. de los A. Ferraro, S. I. Mariño. Experiencia en la Industria: Adaptación de Modelos y Métricas de Calidad para la Mejora Continua (10 Años de Evolución). IX Jornada de Calidad de Software y Calidad, 2025.

Modelado y Ejecución de Sistemas Reactivos Distribuidos mediante Redes de Petri No Autónomas y Microservicios

Mauricio Ludemann [1,2], Luis Orlando Ventre [1], Gabriel Valenzuela [1], Orlando Micolini [1]

[1] *Laboratorio de Arquitectura de Computadoras (LAC), Facultad de Ciencias Exactas, Físicas y Naturales (FCEFyN), Universidad Nacional de Córdoba (UNC)*

[2] *Facultad de Informática, Universidad Nacional de La Plata*

mauri.ludemann@unc.edu.ar, luis.ventre@unc.edu.ar, gabriel.valenzuela@mi.unc.edu.ar, orlando.micolini@unc.edu.ar

RESUMEN

Este trabajo presenta una línea de investigación orientada al modelado y ejecución de sistemas reactivos distribuidos mediante la integración de modelos formales y arquitecturas basadas en eventos. En dominios como sistemas industriales, Internet de las Cosas (IoT) y servicios críticos, estos sistemas deben responder a estímulos concurrentes manteniendo propiedades de corrección y confiabilidad, lo que constituye un desafío para los enfoques tradicionales.

La propuesta aborda esta problemática mediante el uso de Redes de Petri No Autónomas como núcleo formal, combinadas con arquitecturas de microservicios orientadas al procesamiento distribuido de eventos. El enfoque incorpora una taxonomía estructurada de eventos y un modelo de ejecución desacoplado que permite trasladar propiedades verificadas en el modelo hacia implementaciones distribuidas reales.

Palabras clave: sistemas reactivos distribuidos, redes de Petri no autónomas, microservicios, procesamiento de eventos, verificación formal de sistemas.

CONTEXTO

El presente trabajo se enmarca en las actividades de investigación que se desarrollan en el Laboratorio de Arquitectura de Computadoras (LAC) de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de Córdoba (UNC). En este ámbito, se lleva adelante una línea de investigación orientada al estudio de modelos formales aplicados al diseño y análisis de sistemas distribuidos, con especial énfasis en sistemas reactivos, arquitecturas basadas en eventos e Internet de las Cosas (IoT) (Coulouris et al., 2011; Gubbi et al., 2013).

Dentro de esta línea, las Redes de Petri y sus extensiones han sido utilizadas como herramientas fundamentales para el modelado de concurrencia, sincronización y comportamiento reactivo en sistemas complejos (David & Alla, 2010; Murata, 1989). Estos modelos permiten razonar formalmente sobre propiedades del sistema desde etapas tempranas del diseño, aspecto especialmente relevante en dominios donde la corrección y la confiabilidad son requisitos críticos.

Paralelamente, el laboratorio aborda problemáticas vinculadas a la implementación de estos modelos en arquitecturas distribuidas modernas, tales como microservicios y sistemas orientados a eventos, ampliamente

utilizados en aplicaciones industriales y plataformas de servicios (Cerny et al., 2018; Etzion & Niblett, 2010). En este contexto, la línea de investigación busca articular el rigor de los modelos formales con las necesidades prácticas de ejecución y escalabilidad, contribuyendo al desarrollo de soluciones que integren especificación, verificación y despliegue en sistemas distribuidos reales.

1. INTRODUCCIÓN

Los sistemas reactivos distribuidos constituyen un componente central de numerosas aplicaciones actuales, donde la capacidad de responder de manera oportuna y coherente a múltiples estímulos concurrentes resulta un requisito fundamental. En arquitecturas modernas basadas en eventos, estos estímulos pueden originarse tanto en el entorno físico como en otros componentes de software, generando comportamientos complejos cuya correcta gestión impacta directamente en la confiabilidad y la mantenibilidad del sistema (Etzion & Niblett, 2010; Harel & Politi, 1999). En este escenario, las arquitecturas de microservicios han demostrado ser efectivas para abordar problemas de escalabilidad, modularidad y resiliencia, permitiendo la evolución independiente de los componentes y favoreciendo la adopción de patrones reactivos y asincrónicos (Bonér, 2016; Cerny et al., 2018). Sin embargo, estas arquitecturas suelen carecer de mecanismos formales que permitan razonar de manera sistemática sobre el comportamiento global del sistema, especialmente en presencia de concurrencia, sincronización y dependencias temporales.

Alternativamente, otros enfoques han abordado la reactividad y la coordinación en sistemas distribuidos desde perspectivas como el Actor Model (Agha, 1986) o los motores de procesamiento complejo de eventos (CEP) (Luckham et al., 2002). Si bien estas propuestas ofrecen mecanismos robustos para la gestión de mensajes y eventos, no necesariamente proporcionan una correspondencia explícita con modelos formales que permitan analizar propiedades estructurales del sistema antes de su implementación.

Por su parte, los modelos formales ofrecen un marco riguroso para la especificación y verificación de sistemas reactivos. En particular, las Redes de Petri y sus extensiones han sido ampliamente utilizadas para modelar concurrencia y causalidad en sistemas basados en eventos, proporcionando herramientas para el análisis de propiedades como ausencia de bloqueos y consistencia del comportamiento (David & Alla, 2010; Murata, 1989). No obstante, la transferencia directa de estos modelos hacia implementaciones distribuidas reales continúa siendo limitada.

A partir de esta problemática, este trabajo presenta una línea de investigación orientada a integrar modelos formales y arquitecturas distribuidas basadas en eventos. El enfoque propone utilizar Redes de Petri No Autónomas como núcleo formal para el modelado del comportamiento reactivo, combinadas con arquitecturas de microservicios para su ejecución distribuida. Esta integración busca reducir la brecha entre especificación y ejecución, facilitando el desarrollo de sistemas reactivos que sean verificables desde el diseño y operables en entornos distribuidos reales. En la sección siguiente se describen en detalle las principales líneas de investigación y desarrollo que conforman esta propuesta.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación desarrollada en el Laboratorio de Arquitectura de Computadoras (LAC) se centra en el estudio y aplicación de modelos formales para el diseño, análisis y

ejecución de sistemas reactivos distribuidos. El objetivo principal es abordar de manera sistemática los desafíos asociados a la concurrencia, la sincronización y la gestión de eventos en sistemas complejos, utilizando formalismos que permitan razonar sobre el comportamiento del sistema desde etapas tempranas del desarrollo.

En este marco, se emplean Redes de Petri y sus extensiones, en particular Redes de Petri No Autónomas, como núcleo formal para modelar la interacción entre estímulos externos y el comportamiento interno del sistema (David & Alla, 2010; Murata, 1989). Estos modelos permiten representar de forma explícita la causalidad y la concurrencia, así como analizar propiedades relevantes tales como ausencia de bloqueos y coherencia del comportamiento reactivo. En particular, se adopta la semántica estándar de Redes de Petri No Autónomas, donde los eventos externos actúan como señales que condicionan la habilitación y el disparo de transiciones, preservando la estructura causal del modelo y permitiendo representar explícitamente la dependencia entre estímulos y evolución del sistema.

De manera complementaria, la línea de investigación explora la integración de estos modelos formales con arquitecturas distribuidas modernas, como microservicios y sistemas orientados a eventos, ampliamente utilizados en aplicaciones industriales, IoT y plataformas de servicios (Bonér, 2016; Cerny et al., 2018; Dragoni et al., 2016; Etzion & Niblett, 2010; Michelson, 2006). El enfoque busca reducir la brecha entre la especificación formal y la ejecución práctica, permitiendo trasladar propiedades verificadas en el modelo hacia implementaciones distribuidas reales.

Finalmente, se investigan mecanismos de ejecución distribuida del modelo formal, incluyendo la definición de taxonomías de eventos y la separación explícita entre eventos, acciones y datos, con el propósito de mejorar la escalabilidad, mantenibilidad y confiabilidad de los sistemas desarrollados. Esta línea articula investigación teórica y desarrollo aplicado, sentando las bases para la construcción de sistemas reactivos verificables y adaptables en distintos dominios. El tratamiento del tiempo contempla la modelización de eventos temporizados y timeouts mediante extensiones temporales del formalismo.

3. RESULTADOS OBTENIDOS

Como resultado del desarrollo de esta línea de investigación, se han obtenido avances relevantes tanto en el plano teórico como en el aplicado. En particular, se ha consolidado un marco formal para el modelado de sistemas reactivos distribuidos mediante Redes de Petri No Autónomas, permitiendo representar de manera precisa la concurrencia, la sincronización y la interacción con estímulos externos en sistemas basados en eventos (Chiola et al., 1993; David & Alla, 2010; Michelson, 2006; Murata, 1989).

Entre los resultados obtenidos se destaca la definición de una taxonomía estructurada de eventos, orientada a clasificar y gestionar de forma sistemática los distintos tipos de estímulos que intervienen en sistemas reactivos distribuidos. Esta taxonomía permite distinguir eventos según su origen, naturaleza temporal y relevancia para el sistema, facilitando el desacoplamiento entre la recepción de eventos y la lógica de control, y contribuyendo a mejorar la escalabilidad y mantenibilidad de las soluciones desarrolladas (Etzion & Niblett, 2010; Michelson, 2006).

Asimismo, se han desarrollado propuestas de implementación basadas en microservicios para la ejecución distribuida de la lógica del sistema modelada mediante Redes de Petri No Autónomas. En este enfoque, el modelo formal

actúa como núcleo de decisión, mientras que los microservicios permiten materializar su ejecución en entornos distribuidos reales, preservando las propiedades verificadas en el diseño y favoreciendo la modularidad y la evolución independiente de los componentes (Bonér, 2016; Cerny et al., 2018). Este núcleo de decisión puede particionarse en subredes distribuidas según criterios funcionales, evitando la centralización del control y reduciendo potenciales cuellos de botella en la ejecución.

Como resultados esperados, se prevé la validación progresiva de este enfoque en distintos dominios de aplicación, tales como sistemas IoT y plataformas de servicios distribuidos, así como la profundización en mecanismos que permitan asegurar la coherencia entre el modelo formal y su ejecución en arquitecturas distribuidas. Estos avances buscan consolidar una metodología que articule rigor formal y viabilidad ingenieril en el desarrollo de sistemas reactivos distribuidos. En este marco, se prevé evaluar el enfoque mediante métricas de latencia, throughput y tolerancia a fallos sobre un prototipo aplicado al dominio de sistemas de pagos, actualmente en fase de implementación.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación presentada se articula estrechamente con actividades de formación de recursos humanos en el ámbito de grado y posgrado, particularmente en el área de ingeniería de software y sistemas distribuidos. En este marco, el trabajo se vincula directamente con la tesis de Maestría en Ingeniería de Software del autor principal, desarrollada en la Facultad de Informática de la Universidad Nacional de la Plata. Dicha tesis, actualmente en la etapa final de su desarrollo, se centra en el modelado y ejecución de sistemas reactivos distribuidos mediante Redes de Petri No Autónomas y arquitecturas basadas en microservicios (Ludemann et al., 2025).

Asimismo, esta línea se integra con la tesis doctoral de Luis Orlando Ventre, actualmente en su etapa final y en proceso de corrección por parte del tribunal. Dicha tesis se orienta al análisis cuantitativo del paralelismo en sistemas modelados mediante Redes de Petri S3PR, proponiendo algoritmos para determinar el grado de concurrencia efectiva y la cantidad adecuada de hilos de ejecución en implementaciones reales (Ventre et al., 2025).

Adicionalmente, se encuentra en desarrollo la tesis de Maestría de Gabriel Valenzuela, vinculada a la implementación y validación de herramientas de soporte para el modelado y ejecución de sistemas basados en Redes de Petri, en coherencia con la línea de investigación presentada.

La articulación entre estas tesis fortalece la coherencia de la línea de investigación, combinando aportes formales orientados al análisis de concurrencia con propuestas arquitectónicas para la ejecución distribuida de la lógica del sistema. Este enfoque integrado contribuye a la formación avanzada de recursos humanos y a la consolidación de un grupo de trabajo especializado en modelos formales aplicados a sistemas reactivos distribuidos.

La línea se articula además con el proyecto de investigación “Diseño e implementación de software y hardware optimizados para sistemas de computación paralelos en ingeniería”, presentado en la convocatoria Proyectos Estimular 2026 y desarrollado en el ámbito del LAC-UNC, el cual contempla el desarrollo de una plataforma modular basada en microservicios para la simulación, análisis y generación automática de código a partir de modelos de Redes de Petri.

5. BIBLIOGRAFÍA

Agha, G. (1986). *Actors- A Model of Concurrent Computation in Distributed Systems*.

Bonér, J. (2016). *Reactive Microservices Architecture*. O'Reilly Media.

Cerny, T., Donahoo, M. J., & Trnka, M. (2018). Contextual understanding of microservice architecture: current and future directions. *ACM SIGAPP Applied Computing Review*, *17*(4), 29–45.

Chiola, G., Marsan, M. A., Balbo, G., & Conte, G. (1993). Generalized stochastic Petri nets: a definition at the net level and its implications. *IEEE Transactions on Software Engineering*, *19*(2), 89–107. https://doi.org/10.1109/32.214828

Coulouris, G. F., Dollimore, J., Kindberg, T., & Blair, G. (2011). *Distributed systems : concepts and design*. Addison-Wesley.

David, R., & Alla, H. (2010). *Discrete, Continuous, and Hybrid Petri Nets* (2nd ed.).

Dragoni, N., Giallorenzo, S., Lafuente, A. L., Mazzara, M., Montesi, F., Mustafin, R., & Safina, L. (2016). *Microservices: yesterday, today, and tomorrow*. http://arxiv.org/abs/1606.04036

Etzion, O., & Niblett, P. (2010). *Event Processing in Action.* Manning Publications.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645–1660. https://doi.org/10.1016/j.future.2013.01.010

Harel, D., & Politi, M. (1999). *Modeling Reactive Systems with Statecharts: The Statemate Approach*.

Luckham, D., Boston, A.-W., San, •, New, F. •, Toronto, Y. •, London, M., Munich, •, Paris, •, Madrid, •, Sydney, C. •, Tokyo, •, Singapore, •, & City, M. (2002). *The Power of Events An Introduction to Complex Event Processing in Distributed Enterprise Systems*.

Ludemann, M., Ventre, L., Valenzuela, G., & Micolini, O. (2025). *Enhanced Modeling of Reactive Systems using Non-Autonomous Petri Nets and Microservices*.

Michelson, B. (2006). *Event-Driven Architecture Overview*. http://www.customers.com/articles/event-driven-architecture-overview

Murata, T. (1989). *Petri Nets: Properties, Analysis and Applications*.

Ventre, L. O., Micolini, O., Ludemann, M., & Valenzuela, G. (2025). *AvPI: Extended Analysis of Average Parallelism in Systems Modeled with S 3 PR Petri Nets Under Deadlock-Prevention Policies*.

MODELOS DE CALIDAD Y MÉTODOS DE EVALUACIÓN BASADOS EN ISO/IEC 25010 PARA SIMULADORES CLÍNICOS EN ENFERMERÍA

María A. Ojeda[1], Ana Garis[2], Carlos Salgado[2], María C. Albornoz[2], Mario Peralta[2]
[1] Facultad de Cs. de la Salud, Universidad Nacional de San Luis
[2] Facultad de Cs. Físico-Matemáticas y Naturales, Universidad Nacional de San Luis
{maojeda, agaris, csalgado, albornoz, mperalta}@unsl.edu.ar

RESUMEN

La simulación clínica constituye una estrategia consolidada en la formación de profesionales de la salud, favoreciendo el desarrollo de competencias en entornos controlados sin comprometer la seguridad del paciente real. En este contexto, el software que opera los simuladores adquiere un rol central como mediador tecnológico del proceso formativo, influyendo en la experiencia de usuario, la fidelidad del escenario y la confiabilidad del entorno de práctica. Sin embargo, la evaluación de estos productos suele centrarse en percepciones subjetivas o criterios comerciales, sin integrar de manera sistemática dimensiones técnicas y pedagógicas.

El presente trabajo presenta avances de investigación orientados al diseño y validación de un modelo integral de evaluación de calidad para software de simulación clínica en Enfermería, basado en ISO/IEC 25010. Se adopta un enfoque metodológico mixto que combina revisión conceptual, definición de métricas e implementación de un estudio de caso. Los resultados preliminares obtenidos evidencian altos niveles de adecuación funcional, usabilidad, fiabilidad y eficiencia de desempeño. Se propone un instrumento estructurado y potencialmente replicable que contribuye a la evaluación objetiva de tecnologías educativas en el ámbito de la formación en salud.

Palabras clave: ISO/IEC 25010, Simulación Clínica, Enfermería, Métricas e Indicadores.

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto Nº P-03-2020 y reconocido por el programa de Incentivos.

Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional.

1. INTRODUCCIÓN

La simulación clínica ha demostrado ser una metodología eficaz para el desarrollo de competencias en Enfermería, mejorando el razonamiento clínico, la toma de decisiones y la seguridad del paciente [1, 2]. Organismos internacionales como INACSL (International Nursing Association for Clinical Simulation and Learning [3]) destacan la importancia de estándares de buenas prácticas en simulación educativa [4].

El componente tecnológico que sostiene estos entornos formativos es el software del simulador, cuya calidad impacta directamente en la experiencia de aprendizaje [4]. Estudios recientes señalan que la usabilidad, estabilidad y desempeño del sistema influyen en la percepción de realismo y en el compromiso del estudiante [4, 5].

Desde la Ingeniería en Software, la calidad de producto se define como el grado en que un sistema satisface requisitos explícitos e implícitos. La norma ISO/IEC 25010 establece ocho características principales de calidad, entre ellas adecuación funcional, usabilidad, fiabilidad y eficiencia de desempeño [6].

Si bien existen investigaciones sobre simulación clínica y sobre calidad de software de forma independiente [7, 8], persiste una

vacancia en la integración estructurada de estándares internacionales con criterios pedagógicos específicos del ámbito sanitario.

Los Modelos de Calidad (MC), son instrumentos o artefactos específicamente diseñados y construidos para soportar evaluación y selección de componentes de software. Permiten la definición estructurada de criterios de evaluación, la especificación de requerimientos, la descripción de componentes en relación a ellos y la identificación de desajustes de manera sistemática facilitando el proceso de evaluación y selección del software [9].

Según Carvallo en [9], las propuestas existentes de modelos de calidad se pueden clasificar según si tienen un enfoque de modelos de calidad fijos, a medida o mixtos. Entre los modelos de calidad fijos se pueden observar los de McCall et al. (1977) [10], Boehm et al. (1978) [11], entre otros. Para los modelos de calidad a medida existen diversas propuestas de métodos para crearlos, como por ejemplo la del estándar IEEE 1061 [12]. Mientras que, Para el caso de los modelos de calidad mixtos se pueden destacar el ADEQUATE Horgan [13], el modelo de Gilb [14] y el modelo propuesto en el estándar ISO/IEC 9126-1 [15], este último es actualizado y reemplazado por el estándar ISO/IEC 25010:2011 [6].

La norma ISO/IEC 25010:11 [6] presenta un modelo de calidad del producto compuesto de ocho características (que luego se subdividen en treinta y dos subcaracterísticas) que se refieren a las propiedades estáticas del software y a las propiedades dinámicas del sistema informático. El modelo es aplicable tanto a sistemas informáticos como a productos software.

Los modelos pueden, por ejemplo, ser utilizados por desarrolladores, adquirentes, personal de aseguramiento, de control de la calidad y evaluadores independientes, particularmente aquellos responsables de especificar y evaluar la calidad del producto software.

Cuando se habla de modelos de calidad se hace referencia a un conjunto de características y subcaracterísticas deseables por parte de algún actor que son esperables que tenga un producto. Identificar y definir estas características permitirán tener los atributos necesarios para poder medir cada producto. En base a estos atributos se definen y/o reutilizan métricas ya sea de la bibliografía o definidas para realizar las mediciones.

Por ejemplo, se detallan a continuación las 8 características que comprende el catálogo del modelo de calidad del producto según la norma ISO/IEC 25010:11 [6] [13].

1-Adaptación funcional
2-Eficiencia del desempeño
3-Compatibilidad
4-Usabilidad
5-Confiabilidad
6-Seguridad
7-Capacidad de mantenimiento
8-Portabilidad

Coincidiendo con lo expresado en [16], la construcción de modelos de calidad viene dificultada por distintas circunstancias relacionadas con: (1) el equipo que realiza la construcción del modelo, en el caso de que este equipo no tenga experiencia en la construcción de modelos de calidad o bien en el contexto del dominio del componente objeto; (2) el dominio para el que se construye el modelo, para el que en muchas ocasiones no existe una terminología común; (3) factores metodológicos, ya que es difícil conocer el nivel de profundidad hasta el que es necesario descomponer los modelos, y por tanto cuándo se puede decir que un modelo de calidad se ha finalizado.

Bajo estas consideraciones, en esta línea de investigación, el interés se centra en la definición y desarrollo de modelos de calidad, métodos de evaluación, métricas e indicadores orientados a la evaluación de software de simulación clínica en Enfermería, sustentados en estándares internacionales de calidad, tales como la norma ISO/IEC 25010 [6], e

integrando dimensiones técnicas y pedagógicas.

2. LÍNEAS DE INVESTIGACIÓN

Los principales ejes de esta línea de I+D están asociados a:

- Estudio de modelos de calidad de productos software.
- Estudio de normas de calidad aplicadas a productos software
- Estudio de estándares y metodologías aplicadas a la construcción de Modelos de Calidad de productos.
- Construcción de modelos de calidad de productos software aplicables a diferentes ámbitos.
- Estudio de simuladores clínicos.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el marco de esta línea de investigación, se avanzó en una primera instancia en la operacionalización del modelo de calidad propuesto, tomando como referencia las características definidas en la norma ISO/IEC 25010 [6] y adaptándolas al contexto específico de la simulación clínica en Enfermería.

Esta etapa preliminar tuvo como finalidad analizar la pertinencia conceptual y la viabilidad de aplicación de un conjunto de dimensiones consideradas prioritarias, entre ellas adecuación funcional, usabilidad, fiabilidad y eficiencia de desempeño. Dichas dimensiones fueron seleccionadas por su relevancia tanto en la evaluación de calidad de producto software [6] como en estudios vinculados al impacto de entornos simulados en la formación en salud [17, 18].

A partir de una aplicación exploratoria del instrumento, se obtuvieron indicadores cuantitativos que permitieron analizar el comportamiento del software en relación con los requerimientos formativos definidos. Los valores observados en adecuación funcional evidenciaron un alto grado de correspondencia entre las funcionalidades implementadas y los escenarios pedagógicos previstos. En cuanto a la usabilidad, los registros iniciales mostraron niveles satisfactorios de completitud de tareas, aspecto señalado en la literatura como un factor clave en la aceptación y efectividad de entornos de simulación virtual [17, 18].

Asimismo, los indicadores vinculados a fiabilidad y eficiencia de desempeño permitieron analizar la estabilidad operativa del sistema y su capacidad de respuesta en contextos de uso simultáneo, en concordancia con los criterios establecidos por el modelo de calidad de producto software [6].

Más allá de los valores obtenidos en esta instancia preliminar, el aporte central radica en la definición de una estructura de evaluación que integra métricas cuantificables con criterios pedagógicos, favoreciendo una aproximación sistemática al análisis de calidad del software de simulación clínica.

Estos resultados evidencian correspondencia entre funcionalidades ofrecidas y requerimientos formativos, coherente con estudios que vinculan desempeño técnico y calidad percibida en entornos simulados [4, 19]. Además de los resultados técnicos del software evaluado, se logró:

- Diseñar una matriz de evaluación con métricas cuantificables.
- Integrar criterios pedagógicos con estándares de Ingeniería de Software.
- Validar la aplicabilidad del modelo en un entorno real de formación universitaria.

Este resultado constituye un aporte metodológico, no solo técnico.

Actualmente se encuentra en desarrollo:

- ✓ La ampliación del modelo incorporando dimensiones pedagógicas más específicas (por ejemplo, impacto en aprendizaje medido mediante desempeño académico).
- ✓ La validación cruzada del instrumento en otros simuladores clínicos.

A futuro se proyecta:

- Integrar técnicas de analítica de aprendizaje para correlacionar métricas técnicas con resultados formativos.
- Extender el modelo a otros campos de la formación en salud.

El estudio demuestra que la norma ISO/IEC 25010 puede ser operacionalizada de manera contextualizada para evaluar software de simulación clínica en Enfermería.

El modelo integra dimensiones técnicas y pedagógicas, permitiendo una evaluación estructurada, objetiva y replicable. Constituye una herramienta estratégica para la toma de decisiones institucionales en adquisición y mejora de tecnologías educativas.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco del proyecto en el cual se inserta esta línea de investigación se desarrollan diversas actividades de formación de recursos humanos en los niveles de grado y posgrado. En particular, la línea contribuye al desarrollo de tesis correspondientes a la Especialización en Ingeniería de Software, la Maestría en Ingeniería de Software, la Maestría en Calidad del Software y el Doctorado en Ingeniería Informática.
Específicamente, esta investigación se articula con el Proyecto de Tesis de la Maestría en Ingeniería de Software (Plan Ord. 017/09-CD) de la Lic. Alejandra Ojeda, de la Facultad de Ciencias Físico-Matemáticas y Naturales de la Universidad Nacional de San Luis, orientado al estudio y desarrollo de modelos y métodos de evaluación de calidad de software.
Asimismo, la línea se vincula con tesinas de grado de la Licenciatura en Computación y trabajos finales de carrera de las Ingenierías en Informática y en Computación, promoviendo la integración interdisciplinaria entre Ingeniería de Software y Ciencias de la Salud.

5. BIBLIOGRAFÍA

[1] P. R. Jeffries. (2012). *Simulation in Nursing Education: From Conceptualization to Evaluation (2da. ed.)*.

[2] S. Görücü, G. Türk, and Z. Karaçam, "The effect of simulation-based learning on nursing students' clinical decision-making skills: Systematic review and meta-analysis " 2024.

[3] INCSL. (2021). *International Nursing Association for Clinical Simulation and Learning*. Available: https://www.inacsl.org/

[4] D. A. Cook, R. Hatala, and R. Brydges, "Technology-enhanced simulation for health professions education: A systematic review and meta-analysis,," 2011.

[5] A. Alharbi, A. Nurfianti, R. F. Mullen, J. D. McClure, and W. H. Miller, "The effectiveness of simulation-based learning (SBL) on students' knowledge and skills in nursing programs: a systematic review," *BMC Med. Educ.,* vol. 24, 2024.

[6] ISO, "ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models," ed, 2011.

[7] J.-H. P. Sujin Shin, Jung-Hee Kim, , "Effectiveness of patient simulation in nursing education: Meta-analysis, ," *Nurse Education Today,* vol. 35, pp. 176-182, 2015.

[8] B. e. a. Kitchenham, "Software quality models," *IEEE Software,* 2017.

[9] J. Bermeo Conto, M. Sánchez, J. J. Maldonado, and J. P. Carvallo, "Modelos de Calidad de Software en la Práctica: Mejorando su Construcción con el Soporte de Modelos Conceptuales," presented at the CEDIA, 2016.

[10] J. A. MCCALL, P. K. RICHARDS, and G. F. WALTERS, "Factors in Software Quality," 1977.

[11] B. W. BOEHM, J. R. BROWN, H. KASPAR, M. LIPOW, G. J. MACLEOD, and M. J. MERRITT, "Characteristics of Software Quality," 1978.

[12] ISO, "IEEE Std 1061-1998 IEEE," ed, 1998.
[13] G. HORGAN, S. KHADDAJ, and P. FORTE, "Anessential Views Modelfor Software Quality Assurance. En: Project Control for Software Quality, Shaker Publishing," 1999.
[14] T. GILB, "Principles of Software Engineering Management," Addison Wesley1988.
[15] ISO, "ISO/IEC 9126-1. Software engineering — Product quality — Part 1: Quality model," 2001.
[16] C. Calero, M. Piattini, and M. Moraga, *Calidad del producto y proceso software*, 2010.
[17] C. L. Foronda, M. Fernandez-Burgos, C. Nadeau, C. N. Kelley, and M. N. Henry, "Virtual Simulation in Nursing Education: A Systematic Review Spanning 1996 to 2018. ," *Simul Healthc,* vol. 15, pp. 46-54., 2020.
[18] J. M. Padilha, P. Costa, P. Sousa, and A. Ferreira, "Clinical virtual simulation: predictors of user acceptance in nursing education," *BMC Med. Educ.,* vol. 24, 2024.
[19] S. Shin, J.-H. Park, and J.-H. Kim, "Effectiveness of patient simulation in nursing education: Meta-analysis, ," *Nurse Education Today,* vol. 35, pp. 176-182, 2015.

PetriNet Studio: Plataforma SaaS para Simulación, Análisis y Generación Automática de Código Concurrente a partir de Redes de Petri

Gabriel Emanuel Valenzuela[1], Orlando Micolini[2], Luis Omar Ventre[2], Mauricio Ludemann[2]

[1]**Maestría en Ingeniería de Software – Facultad de Informática, UNLP**

[2]**Laboratorio de Arquitectura de Computadoras (LAC) – FCEFyN, UNC**

gabriel.valenzuela@unc.edu.ar | orlando.micolini@unc.edu.ar

RESUMEN

Se presenta PetriNet Studio, una plataforma SaaS basada en microservicios para modelado, simulación y análisis de Redes de Petri (RdP) y Redes de Petri No Autónomas (NAPN), con generación automática de código concurrente (monitores e hilos) para modelos S3PR. El trabajo integra formalismos verificados con arquitecturas modernas de software, en el marco de un programa de investigación del Laboratorio de Arquitectura de Computadoras (LAC) de la UNC y la Facultad de Informática de la UNLP. La plataforma elimina las barreras de instalación propias de las herramientas de escritorio existentes, integrando nativamente editor gráfico, simulación interactiva, análisis estructural y generación de código en un único entorno web accesible.

Palabras clave: Redes de Petri, NAPN, SaaS, microservicios, generación automática de código, S3PR, sistemas concurrentes, sistemas embebidos.

CONTEXTO

Este trabajo se enmarca en el proyecto "Diseño e implementación de software y hardware optimizados para sistemas de computación paralelos en ingeniería" (Proyectos Estimular 2026), ejecutado en el LAC de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFyN) de la UNC, en colaboración con la Facultad de Informática de la UNLP. El LAC lleva más de una década de investigación sobre procesadores de Petri implementados en hardware [3], metodologías de diseño de sistemas embebidos y reactivos basados en RdP [4] y algoritmos para la determinación automática de hilos en redes S3PR [2]. PetriNet Studio materializa esa línea de investigación en una plataforma de software moderna, integrando los resultados teóricos y algorítmicos en una herramienta operativa y accesible. El trabajo se articula directamente con la tesis de maestría del primer autor (UNLP), la tesis doctoral de Ventre (UNC) y el trabajo de Ludemann et al. sobre arquitecturas reactivas basadas en NAPN y microservicios.

1. INTRODUCCIÓN

Las Redes de Petri constituyen un formalismo matemático consolidado para modelar, analizar y verificar sistemas concurrentes y distribuidos, con propiedades bien definidas de alcanzabilidad, vivacidad, acotamiento e invariantes [1]. Su extensión a Redes de Petri No Autónomas (NAPN) permite representar estímulos externos mediante una taxonomía estructurada de eventos —simples, compuestos, temporales y de estado—, ampliando el poder expresivo para sistemas reactivos conectados a entornos físicos o de red.

Sin embargo, la adopción práctica de estos formalismos en contextos industriales y educativos continúa siendo limitada. Las revisiones sistemáticas de Kumbhar y Chavan [9] y de Thong y Ameedeen [10] coinciden en el diagnóstico: la mayoría de las herramientas disponibles —PIPE, CPN Tools, Snoopy, SNAKES, PNEditor— son aplicaciones monolíticas de escritorio que requieren instalación local, no soportan redes no autónomas, y ofrecen capacidades limitadas o nulas de generación automática de código. IOPT-Tools representa una excepción parcial al ofrecer una interfaz web con generación de código C/VHDL para controladores embebidos, pero es-

tá restringida a la subclase IOPT y al dominio de automatización industrial [14]. No existe, hasta donde alcanza el relevamiento realizado [11], una plataforma SaaS de propósito general que integre nativamente simulación interactiva, análisis estructural con soporte para NAPN y generación automática de código concurrente a partir de modelos de RdP.

PetriNet Studio aborda esta brecha proponiendo una arquitectura de cinco microservicios independientes: (1) *editor-service*, interfaz gráfica SPA para construcción de modelos RdP y NAPN con asociación de tipos de evento a transiciones; (2) *simulation-engine*, motor en C++ con gestión atómica del marcado y soporte para la regla de habilitación extendida de NAPN; (3) *analysis-service*, módulo en Python para cálculo de invariantes, detección de sifones/trampas y clasificación estructural de la red (incluyendo S3PR y NAPN); (4) *codegen-service*, implementación del algoritmo de Ventre y Micolini [2] para generación automática de monitores e hilos en C++ y Java; y (5) *api-gateway*, punto de entrada único con autenticación y enrutamiento. La comunicación entre servicios emplea REST para operaciones síncronas y mensajería asincrónica para operaciones de larga duración.

La preservación semántica en el entorno distribuido se garantiza centralizando la gestión del marcado en el *simulation-engine*: cada disparo de transición se ejecuta como una transacción atómica protegida por exclusión mutua, reproduciendo fielmente la regla de disparo formal y evitando los problemas de consistencia documentados en simulaciones distribuidas de RdP [12]. Para NAPN, la habilitación de una transición t en el marcado M requiere adicionalmente que el evento $\sigma(t)$ esté presente, condición evaluada por el motor a partir de un flujo de eventos inyectados externamente.

El módulo de generación de código implementa un pipeline de cuatro fases: verificación y descomposición estructural de la red como S3PR válida; determinación de la cantidad de hilos a partir de las instancias de cada componente S2P y del marcado inicial de las plazas ociosas; asignación de responsabilidades de ejecución a cada hilo; y generación del código fuente del monitor y de los hilos. El monitor generado centraliza la sincronización mediante variables de condición, preservando las propiedades verificadas en el modelo formal.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las principales líneas de investigación y desarrollo en curso son las siguientes:

Arquitecturas SaaS para herramientas de modelado formal. Diseño e implementación de plataformas basadas en microservicios que trasladen las capacidades de simulación y análisis de Redes de Petri a entornos web accesibles sin instalación local, con soporte multiusuario y despliegue en nube.

Simulación distribuida de Redes de Petri con preservación semántica. Estudio de los mecanismos de sincronización necesarios para garantizar la atomicidad del disparo de transiciones en entornos de microservicios, extendiendo los resultados de Ferscha [12] al modelo de ejecución basado en contenedores.

Generación automática de código concurrente a partir de modelos S3PR. Implementación y extensión del algoritmo de Ventre y Micolini [2] para la determinación de cantidad y responsabilidad de hilos en redes S3PR, con generación de código en múltiples lenguajes destino (C++, Java, Go) y análisis de corrección del código generado mediante ThreadSanitizer.

Redes de Petri No Autónomas (NAPN) en arquitecturas reactivas. En articulación con el trabajo de Ludemann et al., se investiga el empleo de NAPN como núcleo de decisión para microservicios reactivos, incluyendo la extensión del motor de simulación y el módulo de análisis estructural para soportar la ecuación de estado generalizada [5] y la taxonomía de eventos (simples, compuestos, temporales y de estado).

Validación empírica comparativa. Diseño y ejecución de un plan de evaluación que comprende: comparación de grafos de alcanzabilidad con PIPE y LoLA sobre modelos de referencia del repositorio MCC; verificación de ausencia de condiciones de carrera en el código generado mediante análisis dinámico; y *benchmarks* de rendimiento comparativo frente a PIPE 4 e IOPT-

Tools.

3. RESULTADOS OBTENIDOS/ESPERADOS

El prototipo actual implementa parcialmente la arquitectura propuesta. Los módulos operativos incluyen: el editor gráfico SPA con soporte para plazas, transiciones, arcos con peso, marcado inicial y asociación de tipos de evento para NAPN; el motor de simulación en C++ con disparo atómico, los tres modos de selección de transición (manual, aleatorio y por prioridades) y evaluación de condiciones de evento externo; el módulo de análisis en Python con cálculo de matriz de incidencia extendida para NAPN, P-invariantes, T-invariantes, detección de sifones y trampas, y clasificación estructural de la red; y el servicio de generación de código con descomposición S3PR y generación de monitor en C++ para redes de complejidad pequeña y moderada. Estos módulos han sido verificados funcionalmente mediante casos de prueba manuales comparando resultados con PIPE 4 para los mismos modelos.

Los resultados esperados al completar el trabajo incluyen: (a) la plataforma SaaS completa desplegada en un entorno de nube pública con autenticación multiusuario; (b) la validación empírica formal mediante el plan de evaluación descrito, con resultados cuantitativos de corrección y rendimiento; (c) la extensión del generador de código a redes NAPN con semántica de paso máximo, articulando directamente con el trabajo de Ludemann et al.; (d) la publicación de al menos un artículo en congreso o revista indexada con los resultados de la evaluación.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por investigadores y docentes de la UNC y la UNLP con trayectoria en el área de arquitectura de computadoras, ingeniería de software y sistemas formales.

Tesis de posgrado directamente vinculadas a esta línea:

• Tesis de Maestría en Ingeniería de Software (UNLP) – Gabriel Emanuel Valenzuela: "PetriNet Studio: Diseño e implementación de una plataforma SaaS para la simulación y análisis de Redes de Petri." En curso, Julio 2025. Director: Dr. Orlando Micolini. Codirector: Mg. Pablo Thomas.

• Tesis Doctoral (UNC) – Luis Omar Ventre: "Análisis de paralelismo y determinación de hilos en sistemas embebidos modelados con Redes de Petri S3PR." En curso. Director: Dr. Orlando Micolini.

• Tesis de Maestría (UNLP) – Mauricio Ludemann: "Arquitectura de microservicios reactivos basada en Redes de Petri No Autónomas." En curso. Director: Dr. Orlando Micolini.

Además, los integrantes del equipo participan en actividades de docencia vinculadas a los temas de la línea: Gabriel Valenzuela dicta Sistemas Operativos I y Algoritmos y Estructuras de Datos en la carrera de Ingeniería en Computación (UNC), incorporando el uso de Redes de Petri para la enseñanza de concurrencia.

5. BIBLIOGRAFÍA

REFERENCIAS

[1] T. Murata, "Petri Nets: Properties, Analysis and Applications," *Proc. IEEE*, vol. 77, no. 4, pp. 541–580, 1989.

[2] L. O. Ventre y O. Micolini, "Algoritmos para determinar cantidad y responsabilidad de hilos en sistemas embebidos modelados con Redes de Petri S3PR," en *Memorias del XXVII CACIC*, pp. 490–499, 2021.

[3] O. Micolini, *Arquitectura asimétrica multicore con procesador de Petri*, Tesis doctoral, UNLP, 2015.

[4] O. Micolini, L. O. Ventre y M. Ludemann, "Methodology for design and development of Embedded and Reactive Systems Based on Petri Nets," en *IEEE ARGENCON 2018*.

[5] O. Micolini et al., "Ecuación de estado generalizada para redes de Petri no autónomas y con distintos tipos de arcos," en *Memorias del Congreso*, 2016, pp. 873–882.

[6] M. Ludemann, L. O. Ventre y O. Micolini, "Arquitectura de microservicios reactivos

basada en Redes de Petri No Autónomas," en preparación, LAC-UNC / UNLP, 2025.

[7] G. Decker et al., "RESTful Petri Net Execution," en *WS-FM 2008*, LNCS vol. 5387, pp. 73–87, Springer, 2009.

[8] L. Li y W. Chou, "Design and Describe REST API without Violating REST: A Petri Net Based Approach," en *IEEE ICWS 2011*, pp. 508–515.

[9] V. B. Kumbhar y M. S. Chavan, "A Review of Petri Net Tools and Recommendations," en *Proc. ICAMIDA 2022*, Atlantis Press, pp. 710–721, 2023.

[10] W. J. Thong y M. A. Ameedeen, "A Survey of Petri Net Tools," en *Advanced Computer and Communication Engineering Technology*, Springer LNEE vol. 315, pp. 537–551, 2015.

[11] G. E. Valenzuela, "Análisis de simuladores de Redes de Petri," Informe técnico PPS, UNC, 2020.

[12] A. Ferscha, "Parallel and Distributed Simulation of Petri Nets," Tutorial, ICATPN 1993.

[13] J. Ezpeleta, J. M. Colom y J. Martinez, "A Petri net based deadlock prevention policy for flexible manufacturing systems," *IEEE Trans. Robotics and Automation*, vol. 11, no. 2, pp. 173–184, 1995.

[14] L. Gomes, F. Moutinho y F. Pereira, "IOPT-Tools — Towards cloud design automation of digital controllers with Petri nets," en *Proc. ICMC 2013*, IEEE.

Red bayesiana para elegir un modelo de proceso de software

Mauro Sezella[1], Luján Toledo[1], Diana Palliotto[1], Julio Hurtado-Alegría[2]
María de los Ángeles Menini[1] y Rosanna Costaguta[1]
[1]Universidad Nacional de Santiago del Estero (UNSE)
Instituto de Investigación en Informática y Sistemas de Información (IIISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
rosanna@unse.edu.ar
[2]Universidad del Cauca (UNICAUCA)
Grupo de Investigación y Desarrollo en Ingeniería de Software (IDIS)
Facultad de Ingeniería Electrónica y Telecomunicaciones (FIET)
ahurtado@unicauca.edu.co

RESUMEN

En el ámbito del desarrollo de software, la correcta elección del modelo de proceso es crucial para el éxito de los proyectos. Sin embargo, esta decisión a menudo se realiza de manera intuitiva, lo que puede resultar en la adopción de enfoques inadecuados que afecten la calidad del producto final.

En este artículo se aborda la problemática planteada en el párrafo anterior y para ello se recurre al uso de la Inteligencia Artificial (IA), aprovechando su capacidad para optimizar la toma de decisiones en escenarios complejos. Específicamente, se propone la creación de una Red Bayesiana (RB) capaz de ofrecer una solución precisa y fundamentada para la selección de modelos de proceso en proyectos de software.

La RB se construirá y validará utilizando datos recabados mediante cuestionarios especialmente diseñados, que serán respondidos por un grupo de expertos desarrolladores con experiencia en el uso de las metodologías tradicionales y ágiles.

Con la RB construida se espera fortalecer el proceso de selección del modelo de desarrollo en proyectos de software, incrementando las probabilidades de éxito en términos de esfuerzo, costo y tiempo.

Palabras clave: *Ingeniería del software, Red bayesiana, Inteligencia artificial, Selección de modelos de proceso, Modelo de proceso tradicional, Modelo de proceso ágil.*

CONTEXTO

En este artículo se presenta una de las líneas de investigación del proyecto "*Desarrollo de aplicaciones para colaboración en e-learning*", acreditado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT - UNSE) para el período 2022-2026. El proyecto responde a una de las líneas prioritarias de investigación del IIISI, vinculada con la creación y el uso de tecnologías informáticas destinadas a mejorar los procesos de enseñanza y de aprendizaje. Dicho proyecto, da continuidad a una línea de investigación iniciada en 2010-2011 por el proyecto 23/C089 "*Fundamentos Conceptuales y Soportes Tecnológicos de la Informática Educativa*", continuada en 2012-2016 por el proyecto 23/C097 "*Sistemas de Información Web Basados en Agentes para Promover el Aprendizaje Colaborativo Soportado por Computadora*", y en 2017-2021 por el proyecto 23/C138 "*Mejorando escenarios de aprendizaje colaborativo soportado por computadora*". Todos proyectos acreditados y financiados por SECyT – UNSE.

1. INTRODUCCIÓN

En la industria de software, generalmente, la falta de experiencia y de conocimiento en la toma de decisiones en determinados puntos claves del desarrollo de un producto/servicio, provoca que un considerable porcentaje de los proyectos fracase [1].

Ahimbisibwe et al. [2] afirman que casi

dos tercios de los proyectos de software no cumplen sus objetivos de tiempo y presupuesto y, a menudo, no satisfacen los objetivos de negocio. Aunque hay muchas razones por las cuales los proyectos no tienen éxito, estos investigadores argumentan que los proyectos de software fallan debido a la elección inadecuada del modelo de proceso de desarrollo de software.

Con respecto a la elección de los modelos de proceso, Rajagopalan et al. [3] proponen utilizar la teoría de la contingencia, en la que se estudia el efecto que tienen la elección de los modelos de proceso y el entorno en el resultado del proyecto. Según estos autores, las variables que influyen son: la volatilidad de los cambios en los requerimientos, la criticidad y la complejidad del proyecto.

Por su lado, Hurtado y Bastarrica [4] proponen generalizar la idea de que cada organización defina el contexto y sus variables y que también delimite reglas para adecuar los procesos. Sin embargo, estos autores no establecen cómo se pueden obtener esas reglas y qué prioridad tienen las variables de contexto sobre las decisiones de adaptación que se tomen.

Sin duda, la selección de un modelo de proceso adecuado no es una tarea fácil, dado que el campo de selección es amplio y no existen criterios universales para decidir; en consecuencia, generalmente se realiza una elección informal, intuitiva y poco sistemática del modelo de proceso de software.

En la actualidad, podemos encontrar dos modelos de proceso bien diferenciados que son utilizados en la mayoría de los proyectos de software: modelos de proceso tradicionales y modelos de proceso ágiles. Normalmente se elige uno de estos dos tipos de modelos para gestionar un proyecto [5, 6]. Los primeros se caracterizan por tener un conjunto predefinido de procesos y flujos de trabajo de esos procesos, donde las actividades se planifican previamente siguiendo estándares bien determinados y el progreso de desarrollo se mide sobre la base de esta planificación. En contextos dinámicos, en los que la innovación y la adaptación son clave, este enfoque puede limitar la capacidad de respuesta y, consecuentemente, la satisfacción del cliente. Por el contrario, en los modelos de proceso ágiles la planificación es incremental y continua a medida que se desarrolla el software. Estos métodos son flexibles con respecto a la documentación del proceso, poniendo más énfasis en las personas, la interacción, el desarrollo del software en sí, la colaboración con el cliente y el cambio [7].

Desde la perspectiva de la Ingeniería de software (IS), la selección adecuada del modelo de proceso es esencial para el éxito de los proyectos. Elegir el modelo más adecuado puede determinar la eficiencia de un proyecto, la calidad del producto final y la satisfacción del cliente. La toma de decisiones en este ámbito suele estar influenciada por diversos factores, que a menudo presentan un grado considerable de incertidumbre.

Además, para que una organización pueda seleccionar el modelo de proceso adecuado para un determinado proyecto, es indispensable contar con dos aptitudes: poseer cierta experiencia en el desarrollo de software y haber utilizado en varias ocasiones dichos modelos. Frecuentemente, las organizaciones no cuentan con ambas facultades.

Desarrollar un modelo basado en IA permitiría capitalizar la experiencia y el conocimiento necesarios para resolver este problema, a fin de que sea posible en cualquier momento y lugar, sugerir qué modelo de proceso seleccionar conforme a las necesidades de un proyecto específico.

Considerando que, dentro de la IA, las redes bayesianas son el enfoque adecuado para aplicar en contextos de incertidumbre, se propone crear una enfocada a la problemática descripta.

Una RB permite representar la incertidumbre inherente al dominio mediante probabilidades condicionales. Esto posibilita modelar escenarios donde las relaciones entre variables no son necesariamente de causa y efecto absoluto, asimismo, se podrían realizar inferencias incluso cuando la información es

parcial, incompleta o basada en estimaciones subjetivas. Esta capacidad convierte a la RB en una herramienta robusta para la toma de decisiones en contextos complejos. Además, una RB presenta alta escalabilidad y mantenibilidad. La estructura del modelo puede ampliarse mediante la incorporación de nuevas variables o relaciones sin necesidad de rediseñar la red completa; basta con definir las nuevas dependencias y ajustar las distribuciones probabilísticas condicionales correspondientes. Esta propiedad es especialmente valiosa en entornos como el desarrollo de software, donde las condiciones y prioridades pueden cambiar frecuentemente. Otra ventaja técnica destacable es su tolerancia a la ausencia de datos. Una RB puede realizar inferencias aun cuando algunos nodos no tienen evidencia observada, ya que utiliza la información disponible en el resto de la red para estimar las probabilidades posteriores de las variables de interés.

Al momento de llevar a cabo esta investigación, solo se ha identificado un antecedente vinculado con nuestra línea de trabajo por cuanto crea una RB para la misma problemática [8]. Sin embargo, a diferencia de nuestra propuesta, donde los datos empleados para la creación, entrenamiento y evaluación de la RB serán proporcionados por expertos en el área de la IS, Kokkeler [8] construyó y validó su red con datos obtenidos desde fuentes bibliográficas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO/OBJETIVOS

El objetivo de esta línea de investigación es construir una RB que permita tomar la decisión de adoptar el modelo de proceso (tradicional o ágil) según sea adecuado para las características específicas del proyecto a desarrollar.

En particular, esta línea de trabajo busca dar respuesta a los siguientes interrogantes:

- ¿Qué uso se ha dado a las técnicas de IA en la elección de un modelo de proceso de software tradicional o ágil?
- ¿Qué variables influyen en la selección entre un modelo de proceso de software tradicional y un modelo de proceso de software ágil?
- ¿Cómo están relacionadas estas variables entre sí y con respecto a las variables de salida usadas para la toma de la decisión?
- ¿Es factible obtener datos empíricos de las variables seleccionadas?
- ¿Es posible construir una RB que soporte la relación entre las variables de entrada y de salida definidas?
- ¿La RB construida es capaz de brindar resultados similares a los que propondría un desarrollador con experiencia en la toma de este tipo de decisiones?

Para alcanzar el objetivo y dar respuesta a los interrogantes planteados, se concretarán diferentes actividades. Primero, se realizará un relevamiento bibliográfico vinculado con la aplicación de técnicas de IA en la selección de un modelo de proceso de software. Luego, se identificarán las variables relevantes para la selección de un modelo de proceso, y se analizaron las relaciones entre dichas variables, permitiendo modelar sus interdependencias. Considerando las variables y dependencias establecidas entre ellas, se construirá una RB. Para entrenar esa red, se recabarán los datos necesarios a través de cuestionarios realizados a expertos desarrolladores del área de la Ingeniería del Software (IS). Estos datos serán fundamentales para ajustar los parámetros de la red, asegurando que el modelo refleje con precisión las condiciones del mundo real. Para validar la RB construida se utilizarán datos recolectados a través de un cuestionario especialmente diseñado que será respondido por el grupo de expertos desarrolladores. Este grupo estará formado por al menos diez profesionales que cuenten con 2 años de experiencia, como mínimo, en el desarrollo de software y en el uso de algunos de estos dos modelos de procesos (tradicional o ágil).

De esta manera se espera obtener una RB que modele las incertidumbres y las

influencias mutuas entre los factores clave asociados con la problemática. Además, se desarrollará una interfaz para que los usuarios puedan interactuar con la red, probar diferentes escenarios y extraer conclusiones basadas en las recomendaciones generadas.

Contar con una herramienta que contenga el conocimiento de expertos en IS facilita la creación de una base de conocimiento sólida, la que suele perderse debido a la rotación de personal. Además, esta investigación podría fomentar que otros investigadores generen herramientas basadas en IA para aplicarlas en los problemas de IS que necesiten de conocimientos expertos y datos históricos, sistematizando etapas o decisiones del desarrollo que aún siguen siendo intuitivas y poco claras.

3. RESULTADOS ESPERADOS

Algunos de los resultados a los que se espera arribar al concretar este trabajo son:

- Estado del arte referido a la aplicación de técnica de IA en la elección de un modelo de proceso de desarrollo de software.
- Una RB que tenga utilidad en la resolución del problema planteado y que efectivamente ayude a mejorar la calidad de los procesos de software.
- Resultados empíricos de uso de la RB construida.
- Presentación de resultados en eventos científicos nacionales e internacionales, y publicaciones en revistas indexadas.

A la fecha se han concluido las actividades vinculadas con el relevamiento bibliográfico, a través de un mapeo sistemático de literatura [9], que está permitiendo a los autores de este artículo elaborar el estado del arte de la situación problemática abordada.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación está compuesto por tres docentes investigadores formados y uno en formación. El equipo también cuenta con dos integrantes estudiantes que están desarrollando sus trabajos finales de graduación para obtener el título de Licenciado en Sistemas de Información en el marco de esta investigación.

5. REFERENCIAS

[1] Johnson, J. (2020). *CHAOS 2020: Beyond Infinity,* The Standish Group, https://www.standishgroup.com

[2] Ahimbisibwe, A., Cavana, R., & Daellenbach, U. (2015). A contingency fit model of critical success factors for software development projects: A comparison of agile and traditional plan-based methodologies. *Journal of Enterprise Information Management,* 28(1), 7–33, https://doi.org/10.1108/JEIM-08-2013-0060

[3] Rajagopalan, S., Mathew, S., & Sugumaran, V. (2016). Fit of development methodologies in software projects. In *Proceedings of 22nd Americas Conference on Information Systems AMCIS*, 2, 21–10, https://aisel.aisnet.org/amcis2016/SAND/Presentations/2

[4] Hurtado-Alegría, J., & Bastarrica, M. (2012). Building software process lines with CASPER. In *Proceedings of 2012 International Conference on Software and System Process*, 170–179, https://doi.org/10.1109/ICSSP.2012.6225962

[5] Ramesh, B., Mohan, K. & Cao, L. (2012). Ambidexterity in Agile Distributed Development: An Empirical Investigation, *Information Systems Research*, 23, 323-339, https://www.jstor.org/stable/23274426

]6] Ahimbisibwe, A., Daellenbach, U., & Cavana, R. (2017). Empirical comparison of traditional plan-based and agile methodologies: Critical success factors for outsourced software development projects from vendors' perspective. *Journal of Enterprise Information Managemen*t, 30(3), 400–453, https://doi.org/10.1108/JEIM-06-2015-0056

[7] Beck, K., Beedle, M., Bennekum, A. van, Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R., Mellor, S.,

Schwaber, K., Sutherland, J., & Thomas, D. (2001). *Manifesto for Agile Software Development*, https://agilemanifesto.org

[8] Kokkeler, C. (2021). *Project management method selection using bayesian networks: a novel approac*h. University of Twente, https://essay.utwente.nl/fileshare/file/86859/Kokkeler_BA_EEMCS.pdf

[9] Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literatura review. *Information and Software Technology*, 51(1), 7–15, https://doi.org/10.1016/j.infsof.2008.09.009

Reentrenamiento del modelo de redes neuronales profundas (PhaseNet) con datos sismológicos locales para optimizar la detección de sismos en Argentina.

[1]Marcelo Moreno, [1]María Murazzo, [1]Nelson Rodríguez, [1]Valeria Nicolia, [1]Ricardo Sifon, [2]Ciro Delgado, [2]Juan José Caputo, [2]Carlos Adriano Luzi Mestre, [2]Juan Ignacio Riveros Herrera.
[1] Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
[2]Alumno Avanzado Licenciatura en Sistemas de Información y Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.
Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129
marcelomoreno@unsj-cuim.edu.ar , marite@unsj-cuim.edu.ar, nelson@iinfo.unsj.edu.ar, valenicolia@gmail.com, ricardosifon@gmail.com, cirodelgado7@gmail.com, juanjose27caputo@gmail.com, adriano.luzi16@gmail.com, riverosjuan828@gmail.com.

RESUMEN

La presente línea de investigación tiene como objetivo el desarrollo, adaptación e implementación del modelo de redes neuronales profundas PhaseNet para la detección automática de fases sísmicas, mediante su reentrenamiento utilizando datos provenientes de catálogos sismológicos argentinos. PhaseNet es un modelo basado en redes neuronales convolucionales profundas diseñado para identificar automáticamente los tiempos de arribo de ondas sísmicas a partir de registros continuos, demostrando un alto nivel de precisión en comparación con métodos tradicionales [1], [2].

Dado que el modelo original ha sido entrenado con datos internacionales, su desempeño puede verse afectado por diferencias regionales en condiciones geológicas, instrumentación y niveles de ruido. En este contexto, el reentrenamiento con datos locales constituye una estrategia adecuada para mejorar su adaptación a las características específicas de las señales regionales. El trabajo contempla la preparación de datos, el reentrenamiento del modelo mediante técnicas de transferencia de aprendizaje y su posterior implementación como microservicios dentro de una arquitectura distribuida [3].

Esta línea de investigación tiende a la futura integración de modelos de aprendizaje profundo en sistemas distribuidos para el procesamiento automático de datos sismológicos.

Palabras clave: *deep learning, computational seismology, PhaseNet, microservices, distributed processing, deep neural networks*

CONTEXTO

Esta línea de investigación se desarrolla en el marco del proyecto *"Aplicaciones Distribuidas en Arquitectura de Microservicios. Desafíos técnicos y mejores prácticas"*, el cual aborda el análisis y desarrollo de aplicaciones distribuidas basadas en arquitecturas desacopladas y escalables.

Las arquitecturas distribuidas constituyen un componente fundamental en el desarrollo de sistemas que requieren el procesamiento eficiente de grandes volúmenes de datos y alta disponibilidad operativa [4]. En este contexto, el paradigma de microservicios permite estructurar sistemas complejos mediante la implementación de servicios independientes que pueden desarrollarse, desplegarse y escalarse de manera autónoma [3].

El procesamiento automático de datos sismológicos representa un caso de aplicación relevante de estos sistemas, debido a la naturaleza continua de los registros (formas de onda), su volumen y la necesidad de procesamiento automatizado en tiempo casi real. La integración de modelos de aprendizaje

profundo como PhaseNet dentro de arquitecturas basadas en microservicios permite mejorar la automatización y escalabilidad de estos sistemas, contribuyendo al desarrollo de soluciones aplicadas al análisis de datos científicos.

Esta línea de investigación contribuye tanto al desarrollo de modelos computacionales basados en aprendizaje profundo como al estudio de su integración dentro de arquitecturas distribuidas, en concordancia con los objetivos del proyecto marco.

Además, el proyecto marco de esta línea de investigación cuenta con la participación de investigadores de la Universidad Nacional de San Luis, la Universidad Nacional de Salta y la Universidad Champagnat. Estos grupos poseen experiencia en el desarrollo e implementación de sistemas distribuidos de adquisición y procesamiento de datos provenientes de redes de sensores (en particular en aplicaciones orientadas al monitoreo de incendios forestales). Esta experiencia resulta directamente transferible al procesamiento de datos sismológicos y al entrenamiento de modelos de aprendizaje profundo. Esta colaboración interdisciplinaria fortalece la presente línea de investigación, aportando conocimientos en el diseño de sistemas robustos, escalables y que estén orientados al procesamiento automático de señales provenientes de redes de sensores.

1. INTRODUCCION

El monitoreo de la actividad sísmica requiere el procesamiento continuo de registros provenientes de redes de estaciones sismológicas distribuidas geográficamente. Uno de los procesos fundamentales en el análisis de estos datos es la identificación precisa de los tiempos de arribo de las ondas sísmicas primarias (P) y secundarias (S), lo cual constituye la base para la localización de eventos sísmicos y la generación de catálogos sísmicos confiables.

Tradicionalmente, esta tarea ha sido realizada mediante análisis manual o mediante algoritmos automáticos basados en técnicas clásicas de procesamiento de señales, como el método STA/LTA, los cuales presentan limitaciones en términos de precisión y robustez frente a señales con bajo nivel de señal o alto nivel de ruido.

El avance reciente en el campo del aprendizaje profundo ha permitido el desarrollo de modelos capaces de identificar automáticamente fases sísmicas con mayor precisión y robustez. En este contexto, el modelo PhaseNet, desarrollado por Weiqiang Zhu y Gregory C. Beroza, representa uno de los enfoques más relevantes para la detección automática de fases sísmicas [2].

PhaseNet utiliza una arquitectura de red neuronal convolucional profunda inspirada en el modelo U-Net, que permite estimar, para cada instante temporal, la probabilidad de ocurrencia de una fase sísmica P, S o ruido. Este enfoque permite la identificación precisa de fases sísmicas incluso en condiciones adversas de ruido y ha demostrado un desempeño superior en comparación con métodos tradicionales [2], [5].

Diversos estudios han demostrado que los modelos basados en aprendizaje profundo pueden mejorar significativamente la detección de eventos sísmicos y la precisión en la identificación de fases [5], [6]. En este contexto, han surgido herramientas especializadas que facilitan la aplicación de técnicas de aprendizaje automático en sismología, como SeisBench, una biblioteca diseñada específicamente para el entrenamiento, evaluación e implementación de modelos de aprendizaje profundo aplicados a datos sísmicos [7]. Sin embargo, el desempeño de estos modelos depende en gran medida de las características de los datos utilizados durante su entrenamiento. Factores como las condiciones geológicas regionales, las características instrumentales y las condiciones de ruido influyen en la forma de las señales registradas.

Por este motivo, el reentrenamiento del modelo utilizando datos provenientes de catálogos sismológicos argentinos constituye un paso fundamental para su adaptación al contexto regional.

En el presente trabajo se propone el desarrollo de un entorno computacional que permita el reentrenamiento del modelo PhaseNet utilizando datos locales, así como su implementación dentro de una arquitectura basada en microservicios. El enfoque adoptado se centra en los aspectos computacionales, de ingeniería de software y arquitectura de sistemas distribuidos necesarios para el entrenamiento e integración del modelo, constituyendo una base tecnológica para futuras investigaciones orientadas a su aplicación operativa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las principales líneas de investigación contempladas son:

- Construcción de conjuntos de datos de entrenamiento, que consiste en la extracción, procesamiento y estructuración de formas de onda sísmicas y tiempos de arribo provenientes de catálogos sismológicos argentinos, generando conjuntos de datos adecuados para el entrenamiento del modelo.
- Diseño e implementación de flujos de trabajo automatizados que permitan la preparación de datos, el entrenamiento del modelo y la generación de resultados reproducibles.
- Reentrenamiento del modelo PhaseNet. Es la aplicación de técnicas de transferencia de aprendizaje para adaptar el modelo PhaseNet a las características específicas de los datos sismológicos argentinos.
- Implementación del modelo. Es el desarrollo de los microservicios que sean necesarios para encapsular el modelo reentrenado, permitiendo su integración en sistemas distribuidos.
- Integración en sistemas distribuidos. Es el análisis e implementación de arquitecturas distribuidas que permitan el despliegue eficiente de modelos de aprendizaje profundo para el procesamiento de datos científicos.

3. RESULTADOS ESPERADOS

Se espera obtener los siguientes resultados:

- Desarrollo de un modelo PhaseNet reentrenado utilizando datos provenientes de catálogos sismológicos argentinos.
- Desarrollo de herramientas de software que permitan automatizar el proceso de entrenamiento.
- Implementación del modelo. Desarrollo de microservicios que encapsulen el modelo reentrenado, permitiendo su integración en sistemas distribuidos.
- Desarrollo de una arquitectura de software basada en microservicios para la integración del modelo.
- Generación de documentación técnica completa que permita la reproducibilidad del sistema desarrollado.

El presente trabajo se limita al desarrollo computacional y a la implementación del modelo de microservicios, sin abordar su validación desde el punto de vista sismológico, la cual será objeto de investigaciones futuras y correspondientes a otras áreas, fuera del alcance de las líneas de investigación enunciadas en el proyecto de referencia.

4. FORMACIÓN DE RECURSOS HUMANOS.

Esta línea de investigación contribuirá a la formación de recursos humanos en áreas interdisciplinarias que integran informática, aprendizaje automático y procesamiento de datos científicos.

Se promoverá activamente la incorporación de alumnos avanzados mediante la postulación a becas de investigación, tanto interna como externa, con el objetivo de fomentar su formación temprana en actividades científicas y tecnológicas.

Asimismo, se incentivará la realización de trabajos finales de grado y proyectos integradores vinculados a esta línea de investigación por parte de estudiantes que se encuentren en condiciones de culminar sus estudios. De esta manera, se espera que el proyecto contribuya no solo al desarrollo tecnológico, sino también a la consolidación de trayectorias académicas en sistemas distribuidos, microservicios y aprendizaje profundo.

5. BIBLIOGRAFIA

[1] Zhu, W., & Beroza, G. C. (2018). PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method. arXiv:1803.03211.

[2] Zhu, W., & Beroza, G. C. (2019). PhaseNet: A Deep Neural Network Approach to Seismic Phase Picking. *Journal of Geophysical Research: Solid Earth*, 124(1), 261–273.

[3] Newman, S. (2021). *Building Microservices: Designing Fine-Grained Systems*. O'Reilly Media.

[4] Bahrami, M., & Singhal, M. (2015). The role of cloud computing architecture in big data. In *Information Granularity, Big Data, and Computational Intelligence* (pp. 275–295). Springer.

[5] Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake Transformer—An Attentive Deep-Learning Model for Simultaneous Earthquake Detection and Phase Picking. *Nature Communications*, 11, 3952.

[6] Ross, Z. E., Meier, M. A., Hauksson, E., & Heaton, T. H. (2018). Generalized Seismic Phase Detection with Deep Learning. *Bulletin of the Seismological Society of America*, 108(5A), 2894–2901.

[7] Woollam, J., Rietbrock, A., Bueno, A., et al. (2022). SeisBench—A Toolbox for Machine Learning in Seismology. *Seismological Research Letters*, 93(3), 1695–1709.

Reuso de Sistemas Big Data basado en Conocimiento Contextual

Agustina Buccella, Alejandra Cechich, Rafaela Mazalu,
Juan Luzuriaga, Marcos Cruz, Franco Corgatelli,
Rodolfo Martínez, Alan de Renzis, Maximiliano Arias,
Lara Acuña, Malena Rivera, Francisco Repetto, Nicolás Jacznik
GIISCO Research Group
Departamento de Ingeniería de Sistemas
Universidad Nacional del Comahue
Neuquen, Argentina
agustina.buccella@fi.uncoma.edu.ar

23 de febrero de 2026

1. Resumen

Debido a la gran proliferación de datos, tanto dentro como fuera de las organizaciones, los Sistemas de Big Data (SBDs) han emergido como soluciones eficaces para potenciar la integración y aprovechamiento de los mismos en la toma de decisiones. El desarrollo de este tipo de sistemas posee particularidades específicas, muchas veces no contempladas en el desarrollo de software tradicional, que deben considerarse para lograr explotar la información y dar soporte a los requerimientos específicos de los usuarios finales. Al mismo tiempo, el desarrollo debe también contemplar atributos de calidad incluidos en todos los sistemas software. Por ejemplo es sumamente importante que estos sistemas se construyan sobre bases sólidas que permitan flexibilidad, adaptabilidad, portabilidad, reusabilidad, entre otras.

En nuestro proyecto nos abocamos a la creación y diseño de métodos, técnicas, recursos y herramientas que, además de considerar los requerimientos de todo sistema de big data, considere el reuso como mecanismo principal para lograr sistemas de alta calidad en tiempos de desarrollo mas cortos.

Palabras Clave: Reusabilidad - Variedad - Plataforma de Soporte - Big Data

2. Contexto

La línea presentada se inserta en el contexto del *Proyecto UNComa 04/F019: Modelado de Variedad en Sistemas Big Data*. Directora: Dra. Agustina Buccella, Co-directora: Dra. Alejandra Cechich (2022-2026).

3. Introducción

Detectar y reusar efectivamente casos previos en análisis de datos es un problema que depende en gran medida de la representación de la documentación de los activos de dominio, incluida la variedad. En trabajos anteriores, hemos definido cuatro tipos de variedades [8, 10]: (1) variedad de fuentes, que denota los diferentes tipos de fuentes de datos que pueden ser útiles para el problema del dominio, como estructuradas, semiestructuradas y sin estructura, y los métodos de recopilación, como publicación-suscripción, procesamiento de eventos, etc.; (2) **variedad de contenido**, que identifica variables o factores extraídos de fuentes que son útiles cuando se consideran como una característica de análisis; (3) variedad de proceso o procesamiento, que recopila información sobre técnicas para la limpieza de datos y el análisis de datos; y (4) **varie-**

dad de contexto, que denota la situación o las condiciones bajo las cuales se puede aplicar el dominio del problema. Por ejemplo, podríamos establecer un caso con condiciones climáticas y suelo específicos (variedad de contexto), recopilar datos de conjuntos de datos estructurados (variedad de fuentes), seleccionar características relevantes para el caso como temperatura, precipitaciones, etc. (variedad de contenido) y analizar los datos utilizando diferentes técnicas de correlación (variedad de procesamiento). La variedad de contexto es el principal artefacto que proporciona la base para la reutilización. Define variaciones de dominio que limitarán y afectarán los resultados del proceso.

Previamente, nos hemos enfocado en dos ejes principales. Por un lado, hemos trabajado en la creación de metodologías, técnicas y recursos para el desarrollo de SBDs reusables mediante el modelado y la gestión de la variedad [1, 2, 5, 8, 10]. Por el otro lado, hemos aplicado estas técnicas y metodologías creadas para el desarrollo de casos de aplicación (SBDs) en el dominio de la agricultura, en colaboración con INTA-Alto Valle[1] [3, 4, 13, 14].

3.1. Trabajos Previos

El proceso de identificación de variedad en SBDs puede iniciarse desde la evidencia que presenten los datos en las fuentes (enfoque bottom-up); desde el análisis conceptual del dominio (enfoque top-down); o con una mezcla de ambos casos (izquierda de la Figura 1). En [9] esos enfoques se discuten con cada uno de sus pasos, incluyendo la caracterización de dominio que deriva en la construcción de artefactos mediante los cuales se resguardan variables y relaciones relevantes para describir contextos similares. Esos artefactos son gestionados por una herramienta de soporte (CoVaMaT), que permite almacenarlos con una estructura que facilita su recuperación y reuso.

La primera versión de CoVaMaT, presentada en [10], ofrece tres funcionalidades (derecha de la Figura 1): (S-1) Servicio de documentación de variedad, compuesto por la funcionalidad que permite documentar a cada uno de

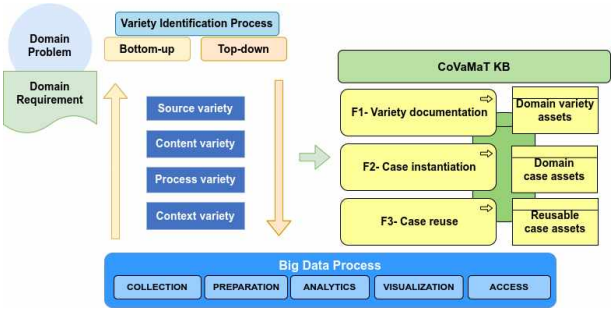


Figura 1: Visión general del proceso y servicios de CoVaMaT (extraída de [10])

los tipos de variedad especificados; (S-2) Servicio de instanciación de caso y (S-3) Servicio de consulta y reuso de activos de dominio. Esos servicios se han modelado a nivel lógico, para luego ser implementados en una arquitectura software con estilo *Cliente-Servidor* usando diversas tecnologías, como *React js* [2], *Apollo* [3], *MongoDB*[4], etc.

La implementación actual permite asistir los tres servicios mencionados, gestionando los diversos casos de variedad [10, 12]. Sin embargo, el modelo actual, basado en una base de datos documental con arreglos de referencias a instancias de estructuras de datos en MongoDB, posee diversas limitaciones en el momento de explorar relaciones. Una estructura con mayores posibilidades de representación (ontologías, grafos de conocimiento, etc.) permitiría inferencias sobre los datos almacenados en un contexto. Por ejemplo, los grafos de conocimiento han sido ampliamente usados en el modelado de diversas aplicaciones de recuperación de información, incluyendo la organización de conocimiento espacio-temporal [6, 15, 16].

En cuanto al uso de ontologías en agricultura, por ejemplo *EUREKA* propone un modelo centrado en el reuso de conceptos y relaciones de ontologías existentes. El modelo más abarcativo se utiliza para describir todas las características relevantes para cada objeto digital del repositorio. Durante el proceso de su diseño, *EUREKA* se apoyó en la metodología para crear ontologías de Noy y McGuiness [7], considerando el reuso de distintas ontologías genéricas

[1] https://www.argentina.gob.ar/inta/cr-patagonia-norte/eea-altovalle

[2] https://es.reactjs.org/

[3] https://www.apollographql.com/

[4] https://www.mongodb.com/es

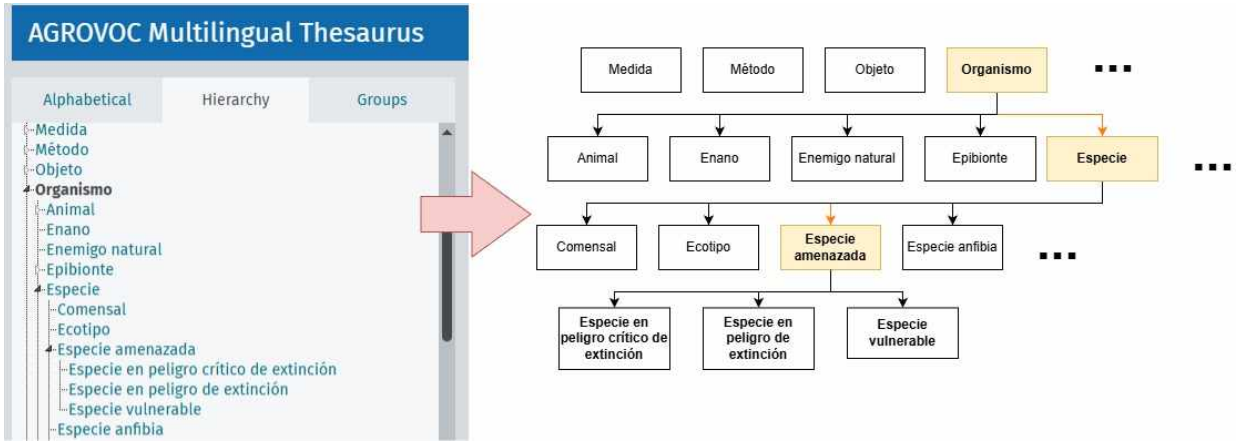


Figura 2: Vista de las jerarquías de *AGROVOC*, expandiendo algunos conceptos de *Organismo*

como *Schema.org*[5], *FRAPO*[6], *FOAF*[7], *DCMI*[8], y ontologías relacionadas a la agricultura como *AGROVOC*, *FoodON*[9] y *AgrO*[10] [11].

4. Líneas de Investigación y Desarrollo

En este instancia de la investigación, se propone extender la herramienta CoVaMaT para incluir información semántica relacionada a dos variedades relacionadas: el contexto de aplicación del caso (o dominio) y el contenido seleccionado a través de las variables relevantes al caso. Para ello, se enfocará en el dominio de la agricultura, tomando como punto inicial ontologías existentes como AGROVOC[11], que es un tesauro relevante publicado como datos abiertos vinculados sobre alimentación y agricultura, disponibles para el uso público. Ofrece una colección estructurada de conceptos, términos, definiciones y relaciones agrícolas que se utilizan para identificar recursos de manera inequívoca, permitiendo procesos de indexación estandarizados y haciendo que las búsquedas sean más eficientes.

Debido a que *AGROVOC* es un tesauro, contiene jerarquías de conceptos. Específicamente, contiene 25 conceptos "superiores"[12], a partir de los cuales surgen especializaciones, como: *Factor*, *Localización*, *Organismo*, *Sustancia*, entre otros. En la Figura 2 se visualiza parte de la jerarquía que constituye el concepto *Organismo*.

Esta ontología será contrastada con requerimientos de casos de analítica de datos provistos por expertos del Instituto Nacional de Tecnología Agropecuaria (INTA) - Estación Experimental Alto Valle[13].

Volviendo a la Figura 1, las contribuciones futuras se centran en la representación del problema y los requerimientos de dominio (extremo superior izquierdo), mediante la identificación de la variedad de contexto/contenido con cualquiera de los procesos de identificación mencionados. Además, actividades de la etapa de recolección de información (análisis de características) se verán favorecidas en la estandarización de conceptos que provengan de las ontologías.

De esta forma, nuestro objetivo actual es **desarrollar una versión de CoVaMaT que permita gestionar variedades de contexto y contenido enriquecidas con información semántica para el dominio agrícola**.

En la Figura 3 podemos observar los pasos para determinar el reuso necesitado. Un analista de datos debe definir sobre qué domino va

[5] https://schema.org/

[6] https://sparontologies.github.io/frapo/current/frapo.html

[7] https://xmlns.com/foaf/spec/

[8] https://www.dublincore.org/specifications/dublin-core/dcmi-terms/

[9] https://foodon.org/

[10] https://bigdata.cgiar.org/resources/agronomy-ontology/

[11] https://www.fao.org/agrovoc/es

[12] La lista completa con sus descripciones se encuentra en: https://www.fao.org/agrovoc/top-concepts-agrovoc

[13] https://www.argentina.gob.ar/inta/cr-patagonia-norte/eea-altovalle

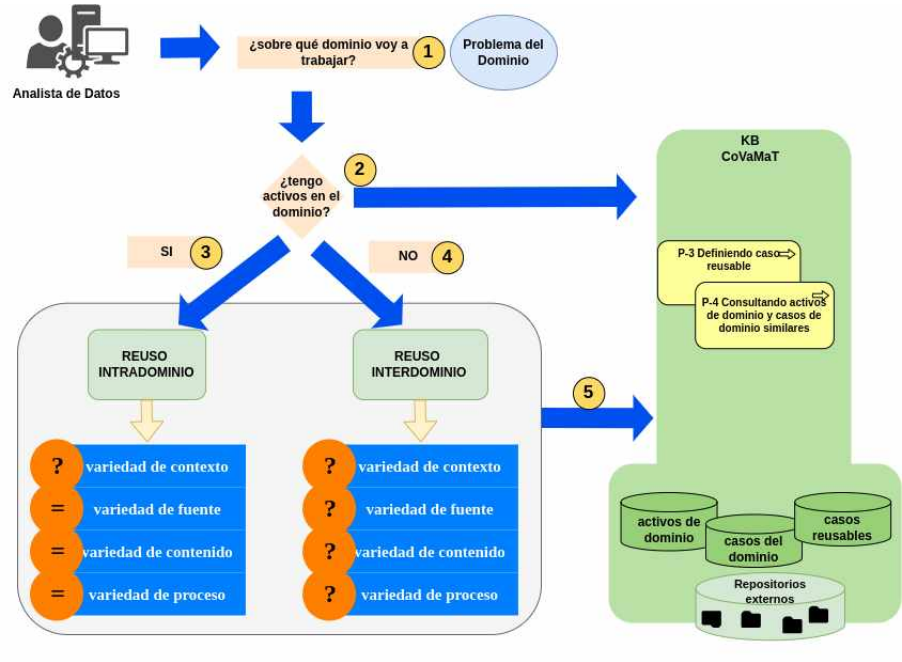


Figura 3: Proceso para reutilizar utilizando CoVaMaT

a construir su SBD definiendo el problema a resolver (1). Luego, se acude a CoVaMaT por medio de los procesos P-3 y P-4 para determinar si hay activos creados en el dominio o no (2). Esa consulta debería extenderse con vocabulario estándar del dominio consultado, en nuestro caso AGROVOC. Si la respuesta es SI (3), entonces se produce un reuso el cual utiliza la variedad de contexto como aspecto diferenciado entre el SBD a construir y los almacenados, dejando las demás variedades iguales.

Cuando la respuesta es NO, es decir, no hay activos creados en el mismo dominio a trabajar, se ingresan nuevos activos basándonos en la estandarización de conceptos de AGROVOC o en la incorporación de un vocabulario compartido propio del sitio de desarrollo (en nuestro caso EE-INTA Alto Valle).

5. Resultados Obtenidos / Esperados

Los resultados obtenidos, que fueron descriptos en los trabajos previos, se refieren a la implementación de metodologías y herramientas para el desarrollo de SBDs reusables. En particular, resaltamos como eje de trabajo principal la construcción de los mecanismos necesarios para lograr un reuso efectivo durante la construcción de los SBDs.

Actualmente, ademas de continuar trabajando en la creación de nuevos SBDs en conjunto con el INTA Alto Valle, nos enfocamos en la forma de mejorar la búsqueda y aprovechamiento de los activos de dominio ya generados en sistemas previos para la utilización de los mismos en los nuevos desarrollos.

6. Formación de Recursos Humanos

El proyecto reúne aproximadamente a unos 15 investigadores, entre los que se cuentan docentes y alumnos de UNComa, y colaboradoras expertas del dominio de aplicación, específicamente pertenecientes al INTA. También el proyecto cuenta con 4 becarios EVC-CIN. Actualmente posee, a su vez, cuatro doctores y un magister. Tres de los docentes-investigadores de GIISCo-UNComa están trabajando para terminar sus carreras de postgrado.

Referencias

[1] A. Buccella and A. Cechich. Un modelo para medir reusabilidad en el desarrollo de sistemas big data basados en contexto. In *XXX Congre-*

so Argentino de Ciencias de la Computación (CACIC), pages 792–801. Red de Universidades con Carreras en Informática, 2024.

[2] A. Buccella, A. Cechich, F. Saurin, A. Montenegro, A. Rodríguez, and A. Mu noz. A context-based perspective on frost analysis in reuse-oriented big data-system developments. *Information*, 15(11), 2024.

[3] Agustina Buccella, Alejandra Cechich, Walter Garrido, and Ayelén Montenegro. Contextual reuse of big data systems: A case study assessing groundwater recharge influences. *Applied Sciences*, 16(3), 2026.

[4] M. Campetella, A. Cechich, A. Buccella, A. Montenegro, A. Muñoz, and A. Rodríguez. Identificación top-down de variedad de contexto: Un caso de estudio en fluctuaciones de la napa freática. In *XXIX Congreso Argentino de Ciencias de la Computación - XX WIS/CACIC*, pages 308–317, 2023.

[5] A. Cechich, A. Buccella, A. Montenegro C. Villegas, and A. Muñoz. A Rodriguez. A model of reusable variation in big data system development. In *XI Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET)*, 2023.

[6] Yu Liu, Jingtao Ding, Yanjie Fu, and Yong Li. Urbankg: An urban knowledge graph system. *ACM Trans. Intell. Syst. Technol.*, 14(4), May 2023.

[7] N. Noy and Deborah Mcguinness. Ontology development 101: A guide to creating your first ontology. *Knowledge Systems Laboratory*, 32, 01 2001.

[8] L. Osycka, A. Buccella, and A. Cechich. Data variety modeling: A case of contextual diversity identification from a bottom-up perspective. pages 124–138, 2022.

[9] L. Osycka, A. Buccella, and N. A. Cechich. Data variety modeling: A case of contextual diversity identification from a bottom-up perspective. In *27th Argentine Congress, CACIC 2021, Salta, Argentina, October 4-8, 2021, Revised Selected Papers. Communications in Computer and Information Science 1584*, pages 124–138. Springer, 2022.

[10] L. Osycka, A. Cechich, A. Buccella, A. Montenegro, and A. Muñoz. Covamat: Functionality for variety reuse through a supporting tool. In *XI Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET)*, 2023.

[11] Hercules Panoutsopoulos, Christopher Brewster, and Spyros Fountas. A semantic data model for a fair digital repository of heterogeneous agricultural digital objects. In *CEUR Workshop Proceedings*, 2021.

[12] Francisco Torres. La herramienta covamat: Componente cliente y extensiones del lado del servidor. Tesis de Licenciatura en Sistemas de Información, 2025.

[13] G. Vidart, A. Cechich, A. Buccella, and A. Montenegro. Análisis de turbidez basado en caracterización de contextos. *JAIIO*, 8(4):170–183, dic. 2022.

[14] C. Villegas, A. Buccella, A. Cechich, , and A. Montenegro. Caracterización de variables para el análisis del índice de vegetación. In *XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)*, pages 978–982, 2022.

[15] Zhu Wang, Fengxia Han, and Shengjie Zhao. A survey on knowledge graph related research in smart city domain. *ACM Trans. Knowl. Discov. Data*, July 2024.

[16] Hao Zhang, Hejie Si, Xinming Ma, Lei Xi, and Xin Xu. Research and application of agriculture knowledge graph. In *Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering*, EITCE '21, page 680–688, New York, NY, USA, 2021. Association for Computing Machinery.

SIXAI: EXPLICABILIDAD INMUTABLE MEDIANTE LÓGICA DIFUSA Y BLOCKCHAIN

Fransoy, Marcelo[1]; Hossian, Alejandro[1]; Merlino, Hernán[1]
[1] Grupo de Estudio en Metodologías de Ingeniería en Software (GEMIS)
Departamento de Sistemas de Información (GEMIS)
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
marcelo.fransoy@gmail.com, alejandrohossian@yahoo.com.ar, hmerlino@frba.utn.edu.ar

RESUMEN

La inteligencia artificial (IA) afectiva enfrenta hoy una encrucijada ética y técnica: mientras los modelos profundos (Deep Learning) logran precisiones inéditas, su opacidad estructural ("caja negra") y la centralización de los datos biométricos impiden su adopción confiable en entornos críticos como la salud mental. Este trabajo presenta **SiXAI (Sovereign Immutable eXplainable AI)**, una evolución arquitectónica que implementa el concepto de **Explicabilidad Inmutable**. A diferencia de enfoques previos centrados únicamente en la precisión métrica, SiXAI prioriza la soberanía del usuario mediante una capa **SSI** (Self-Sovereign Identity) y garantiza la trazabilidad forense de cada decisión mediante lógica difusa anclada en Blockchain. Los resultados validan que, incluso bajo condiciones de estrés con modelos *Zero-Shot*, la arquitectura asegura la integridad del proceso sin degradar la inferencia, estableciendo un nuevo estándar para sistemas auditables.

Palabras clave: Interacción por voz; Lógica Difusa; Blockchain; SSI; Explicabilidad Inmutable.

CONTEXTO

La Universidad Tecnológica Nacional ha definido diversos Programas de Investigación, Desarrollo e Innovación, entre los cuales se encuentra el Programa de Sistemas de Información e Informática que tiene como objetivo "intensificar y focalizar las acciones tendientes a fortalecer y promover el crecimiento de temáticas de investigación en Sistemas de Información e Informática, y promover la interacción con la industria informática en general y de desarrollo de sistemas de información y de software en particular" [1].

En la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional, desarrolla sus actividades, entre otros grupos de investigación, el grupo GEMIS.BA aprobado por la ordenanza N° 1292 de la Universidad Tecnológica Nacional y la resolución N° 3245/2023 del Consejo Directivo de la Facultad Regional Buenos Aires, el cual trabaja en el campo de la Ingeniería en Sistemas de Información, Ingeniería de Software e Ingeniería del Conocimiento.

El Grupo GEMIS.BA realiza, en el marco de las temáticas objeto de este trabajo, diversos proyectos asociados a la línea de la IA con proyectos de investigación y desarrollo y actividades de docencia e investigación en diferentes instituciones contando con trabajos finales de especialización y tesis de maestría.

1. INTRODUCCIÓN

La convergencia entre la Computación Afectiva y la Inteligencia Artificial Generativa ha permitido el desarrollo de asistentes de voz capaces de decodificar estados emocionales con una sutileza sorprendente [2][3]. Sin embargo, esta capacidad técnica ha avanzado más rápido que los mecanismos de control, seguridad y ética necesarios para gestionarla. En dominios sensibles como la psicoterapia asistida o el monitoreo de crisis, la adopción de estas tecnologías se encuentra estancada por tres limitaciones estructurales que la literatura actual no ha resuelto de manera integral: **la opacidad algorítmica, la centralización de la**

identidad y la falta de auditoría forense [4][5].

1.1. La Problemática de la "Caja Negra" y la Soberanía. Los modelos del estado del arte (SOTA), basados en arquitecturas Transformer (como Wav2Vec o BERT), operan como "cajas negras": procesan entradas y generan salidas sin ofrecer una explicación causal inteligible para el usuario final [6][7]. En un contexto clínico, no basta con saber que el sistema detectó "Ira"; es imperativo comprender por qué lo hizo (¿fue por el tono de voz o por el uso de una palabra específica?) para validar el diagnóstico.

Paralelamente, existe un problema crítico de soberanía de datos. Las arquitecturas actuales operan bajo un paradigma de "silo": los datos biométricos de voz y las bitácoras de emociones residen en servidores controlados por el proveedor del servicio [8]. Esto expone al usuario a riesgos de privacidad y manipulación, impidiéndole ejercer una verdadera auditoría sobre su propia información. Si bien el Aprendizaje Federado (Federated Learning) ha surgido como una solución para la privacidad distribuida, presenta una falla crítica en la trazabilidad inmutable: la naturaleza efímera de las actualizaciones de gradientes y la falta de un registro global auditable impiden verificar ex-post si una inferencia fue manipulada en el borde [5]. Aquí es donde la integración de Blockchain se vuelve indispensable, no como almacenamiento masivo, sino como un registro de fe pública que garantiza que la evidencia de la decisión (el hash de la explicación) sea cronológicamente coherente y resistente a cualquier intento de alteración retrospectiva.

1.2. Limitaciones de los Antecedentes Inmediatos. Propuestas recientes han intentado mitigar estos riesgos, pero con enfoques parciales.

BAXDT introduce el registro de decisiones en Blockchain, logrando trazabilidad, pero mantiene un modelo de identidad centralizado donde el usuario no posee las claves criptográficas de sus datos [9].

FIeSER utiliza lógica difusa dinámica para maximizar la precisión, pero sacrifica la interpretabilidad semántica al utilizar pesos de atención abstractos en lugar de reglas lingüísticas legibles por humanos [10].

1.3. La Propuesta SiXAI: Evolución y Aporte Diferencial. Evolucionando la propuesta conceptual presentada por los autores en WICC 2025 [11], este trabajo formaliza y materializa la arquitectura **SiXAI**. El giro epistemológico fundamental de esta investigación es dejar de competir exclusivamente en el eje de la "precisión métrica" para priorizar el eje de la "**confianza arquitectónica**". SiXAI propone un sistema donde:

El Razonamiento es Explícito: Se utiliza un motor de inferencia difusa (Mamdani) que actúa como un "chasis" transparente, generando reglas legibles (ej. SI Arousal es Alto Y Texto es Negativo ENTONCES Ira, basadas en la teoría de conjuntos difusos de Zadeh [12]) que explican la decisión antes de emitirla.

La Identidad es Soberana: Mediante la implementación de una arquitectura SSI, se desacopla la identidad del proveedor [13]. El usuario recibe sus diagnósticos como Credenciales Verificables (VCs) cifradas, **siendo el único custodio de su historial emocional** [14].

La Auditoría es Inmutable: Cada inferencia genera una huella criptográfica (hash) que se ancla en una Blockchain pública, permitiendo auditorías de terceros sin revelar el contenido sensible de la conversación [4].

En este sentido, SiXAI no se presenta sólo como un clasificador de emociones, sino como una infraestructura de cumplimiento (*compliance*) técnica, diseñada para alinearse con normativas emergentes como el EU AI Act, que exigen transparencia y explicabilidad en sistemas de alto riesgo.

En la Figura 1 se visualiza la arquitectura general de SiXAI:

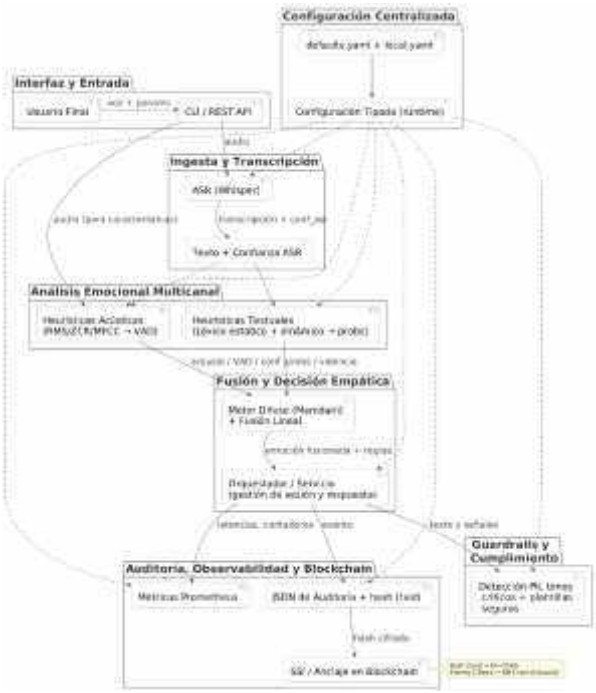


Figura 1. Arquitectura SiXAI

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La Universidad Tecnológica Nacional (UTN) de Argentina ha demostrado un firme compromiso con el avance de la investigación científica a través de la implementación de diversos programas de Investigación, Desarrollo e Innovación. Un ejemplo destacado es el Programa de Sistemas de Información e Informática, establecido en 2016 mediante la Resolución 2508/16. Este programa tiene como objetivo principal fortalecer y expandir las actividades de investigación en áreas clave como los Sistemas de Información y la Informática, fomentando al mismo tiempo una colaboración más estrecha con la industria informática, especialmente en los sectores de desarrollo de software y sistemas de información.

En este contexto, y como parte de una iniciativa más amplia para impulsar la investigación y el desarrollo en Ingeniería en Sistemas de Información e Ingeniería del Software, se iniciaron en 2009 las actividades del Grupo de Estudio e Investigación en Modelos y Sistemas de Información (GEMIS), grupo que queda formalizado y aprobada como GEMIS.BA por la ordenanza N° 1292 de la Universidad Tecnológica Nacional y la resolución N° 3245/2023 del Consejo Directivo de la Facultad Regional Buenos Aires.

Ubicado en el Departamento de Ingeniería en Sistemas de Información de la Facultad Regional Buenos Aires de la UTN (UTN-FRBA), GEMIS.BA reúne a un equipo multidisciplinario de docentes, graduados y estudiantes dedicados a la sistematización y difusión del conocimiento. Su trabajo abarca una amplia gama de aplicaciones y metodologías, lo que refleja el compromiso de la UTN con la investigación aplicada y la colaboración entre la academia y la industria. La creación de GEMIS.BA es un claro ejemplo del enfoque de la universidad hacia la innovación tecnológica y su aplicación en la sociedad.

En el ámbito de GEMIS.BA, se han llevado a cabo diversas investigaciones relacionadas con la IA. Una de las primeras iniciativas fue el Proyecto de Investigación y Desarrollo (PID) titulado "Implementación de sistemas inteligentes para la asistencia a alumnos y docentes de la carrera de ingeniería en sistemas de información", que se desarrolló entre 2016 y 2019. Este proyecto se centró en la aplicación de tecnologías de IA y desarrollos de software inteligente para abordar desafíos en el ámbito educativo.

Posteriormente, entre 2019 y 2022, se implementó otro PID denominado "Prácticas ingenieriles aplicadas para la implementación de Sistemas Inteligentes basados en Machine Learning". En este caso, el proyecto se enfocó en analizar el uso de buenas prácticas de ingeniería para el desarrollo de software inteligente que utiliza Machine Learning.

A partir de 2023, se inició el proyecto "Inteligencia Artificial para el análisis predictivo en salud mental" (SITCBA0008600TC), cuyo objetivo es proponer técnicas, tareas y métodos para la implementación de Modelos Predictivos basados en IA en el campo de la salud mental.

Además, se incorporó el análisis de señales de EEG como complemento a los modelos predictivos desde 2024, aprovechando la información neurofisiológica proporcionada por el EEG.

3. RESULTADOS OBTENIDOS/ESPERADOS

La validación de una arquitectura como SiXAI no puede reducirse exclusivamente a métricas de clasificación tradicionales. Dado que la propuesta introduce un cambio de eje —desde la maximización del rendimiento hacia la garantía de confianza estructural—, los resultados deben analizarse bajo una perspectiva más amplia que contemple gobernanza algorítmica, trazabilidad explicable e integridad criptográfica. En este sentido, la evaluación realizada busca determinar si el sistema no solo funciona, sino si opera conforme a los principios de soberanía, transparencia y auditabilidad que constituyen su fundamento teórico.

3.1 Validación del Marco de Gobernanza (SiXAI). A diferencia de las validaciones tradicionales centradas exclusivamente en la maximización del F1-Score, nuestro experimento se enfocó en verificar la **integridad estructural** de la arquitectura. Utilizando el Spanish MEACorpus 2023, un corpus de alta complejidad acústica en entornos naturales [15], se demostró que el motor difuso actúa eficazmente como una compuerta de validación lógica. En situaciones de conflicto (donde el texto dice "alegría" pero el audio indica "tristeza"), el sistema no generó una alucinación estadística, sino que activó reglas de "incertidumbre explícita". Esto confirma que SiXAI cumple su función como **marco de gobernanza algorítmica**: previene diagnósticos de alta confianza basados en data ruidosa o fallas del ASR (Automatic Speech Recognition) [16][17], garantizando que cada inferencia sea causalmente rastreable.

3.2 Auditoría Forense y Soberanía Efectiva (Módulo SSI). Se logró la materialización técnica de la Explicabilidad Inmutable, resolviendo las vacancias planteadas en la introducción:

Prueba de Auditoría Forense: Se simuló un escenario de disputa legal donde un usuario cuestionaba un diagnóstico. Gracias al anclaje en la testnet Sepolia, el sistema permitió verificar matemáticamente (mediante cotejo de hashes) que el registro local no había sido alterado post-inferencia [11]. Esto constituye una ventaja cualitativa absoluta sobre modelos federados, que no ofrecen inmutabilidad transaccional.

Desacople de Identidad: La emisión de Credenciales Verificables (VCs) firmadas con Ed25519 demostró que es técnicamente viable operar un sistema de salud mental sin que el proveedor custodie la "historia clínica" emocional del paciente [13]. El usuario retuvo el 100% del control de sus claves, validando el modelo de soberanía propuesto.

3.3 Evolución Modular (Trabajo Futuro). Habiendo validado que la arquitectura garantiza transparencia y soberanía, el siguiente paso es la actualización del motor de percepción. Se sustituirán los módulos heurísticos actuales por modelos SOTA (Wav2Vec2 [6], BERT), los cuales se insertarán dentro del entorno de validación SiXAI ya probado. Se espera que esto eleve la precisión diagnóstica al nivel del estado del arte, tal como sugieren los últimos *surveys* en fusión multimodal [18][19], pero con una diferencia crucial: la "caja negra" estará ahora operando bajo una estricta supervisión auditable y segura.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación para este trabajo se encuentra conformado por investigadores formados, investigadores de apoyo y alumnos de grado y posgrado. Esta línea de trabajo busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para su desarrollo en la carrera de investigadores y profesionales, además de fomentar la aplicación de las temáticas en su propia actividad profesional.

En el marco del proyecto de investigación sobre el diseño e implementación de un sistema de interacción por voz en español para asistentes virtuales y chatbots con reconocimiento de emociones, se ha puesto especial énfasis en la formación de recursos humanos, integrando a estudiantes y docentes en diversas etapas de su desarrollo profesional y académico. Este enfoque multidisciplinario permitirá que varios estudiantes de grado y posgrado participen activamente en el proyecto, desarrollando sus tesis y proyectos de investigación en áreas clave como el PLN, la lógica difusa, el reconocimiento de emociones en la voz y la tecnología blockchain.

Se prevé que las investigaciones vinculadas al Proyecto de tesis de Magíster se desarrollen en GEMIS.BA, a cargo de la Dra. María Florencia Pollo Cattaneo, en el marco del Proyecto Inteligencia Artificial para el Análisis Predictivo en Salud Mental, acreditado por la Universidad Tecnológica Nacional, Facultad Regional Buenos Aires.

Asimismo, GEMIS.BA tiene vínculos con otros grupos de investigación que trabajan en áreas vinculadas a las temáticas de tesis. Estas vinculaciones permitirán intercambiar opiniones con tesistas de posgrado e investigadores que trabajen en líneas de investigación afines.

Se espera que esta línea de trabajo permita: la generación de nuevas oportunidades y experiencias para la formación de investigadores; el desarrollo de al menos dos tesis de Maestría en Ingeniería en Sistemas de Información; y la participación de becarios graduados a través de las becas BINID.

5. BIBLIOGRAFÍA

[1] Universidad Tecnológica Nacional (2016). Resolución Nro. 2508/16.

[2] Picard, R. W. (1997). Affective computing. MIT press.

[3] Zhang, J., et al. (2025). Affective computing in the era of large language models: A survey from the NLP perspective. arXiv preprint arXiv:2408.04638.

[4] Awan, K., Zhang, Y., Hussain, S., & Kim, H. (2025). Blockchain and biometrics: A survey. IEEE Access, 13, 42310–42342.

[5] Zichichi, M., et al. (2024). A Survey of Blockchain-Based Privacy Applications: An Analysis of Consent Management and Self-Sovereign Identity Approaches. arXiv preprint arXiv:2411.16404.

[6] Fayek, H., Lech, M., & Cavedon, L. (2022). Evaluating wav2vec 2.0: A study on speech representation learning. IEEE/ACM Trans. Audio, Speech, Lang. Proc., 30, 1348–1359.

[7] Hao, Y., & Liu, D. (2025). Neuro-fuzzy systems for interpretable deep learning: A comprehensive survey. Neural Computing and Applications, 37(5), 8765–8798.

[8] Latif, S., et al. (2023). Privacy-preserving speech processing: A tutorial. arXiv preprint arXiv:2305.05227.

[9] Parlak, İ. E. (2025). Blockchain-Assisted Explainable Decision Traces (BAXDT): An Approach for Transparency. Knowledge-Based Systems, 329.

[10] Nguyen, N. M., et al. (2025). FleSER: Multimodal Emotion Recognition via Dynamic Fuzzy Membership and Attention Fusion. Preprint.

[11] Fransoy, M., Hossian, A., & Merlino, H. (2025). Interacción por voz afectiva y segura: Un enfoque híbrido de PLN, Lógica Difusa y Blockchain. XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025).

[12] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning. Information Sciences, 8(3), 199–249.

[13] Rani, S., Kumar, P., & Singh, R. (2025). The YouGovern secure blockchain-based self-sovereign identity (SSI) management and access control. Applied Sciences, 15(12), 6437.

[14] Wang, Y., Zhang, J., & Zhao, L. (2024). Empowering privacy through peer-supervised self-sovereign identity: Integrating zero-knowledge proofs and blockchain. Sensors, 24(24), 8136.

[15] Pan, R., et al. (2024). Spanish MEACorpus 2023: A multimodal speech–text corpus for emotion analysis in Spanish from natural environments. Expert Systems with Applications, 254, 124317.

[16] Feng, Y., Li, Y., & Wang, L. (2024). ASR-robust speech emotion recognition with multimodal fusion. ICASSP 2024, 1-5.

[17] Sahu, S., et al. (2019). How does ASR performance affect speech emotion recognition? Interspeech 2019, 4365–4369.

[18] Ahmed, N., Aghbari, Z. A., & Girija, S. (2023). A systematic survey on multimodal emotion recognition using learning algorithms. Information Fusion, 91, 521-540.

[19] Dutta, S., & Ganapathy, S. (2025). Multimodal emotion recognition in conversations: A survey. arXiv preprint arXiv:2505.20511.

SYNAPSIS: Historia Clínica Digital y gestión multimodal para el seguimiento de pacientes con Enfermedad de Parkinson

Pedro Salvador Occhipinti[1*], Mónica Giuliano[12], Carlos Ezequiel Gonzalez[1], Ruth Gomez[1], Silvia Noemí Pérez[1], Alan Berduc[13]

[1] Instituto de Ciencias de la Salud, Universidad Nacional del Oeste (UNO), Argentina
[2] Universidad Nacional Arturo Jauretche (UNAJ), Argentina
[3] Hospital Nacional Profesor Dr. Alejandro Posadas (HNP), Argentina
* Autor/a de contacto: pocchipinti@uno.edu.ar

Resumen

SYNAPSIS es una plataforma web desarrollada para el seguimiento de personas con Enfermedad de Parkinson (EP) mediante una Historia Clínica Digital (HCD) integrada con evidencia multimodal (audio, video y documentos). Se enmarca en una línea de I+D orientada al análisis de rasgos vocales, de escritura y de expresión facial para clasificación y monitoreo clínico, en articulación con un hospital público. El sistema se desarrolló colaborativamente entre informáticos y médicos, a medida de necesidades de los profesionales de la salud y con proyección a investigación futura.

La arquitectura implementa un backend en PHP bajo esquema MVC con persistencia en MySQL, organizando pacientes, historia clínica, diagnósticos, estudios y archivos multimedia. La lectura longitudinal se resuelve mediante una línea temporal compacta y expandible que preserva el contexto clínico. Se incorporan mecanismos de trazabilidad con autoría visible y marcas temporales consistentes.

El subsistema de captura contempla tres modalidades: grabación web institucional, captura offline para escenarios sin conectividad y una PWA offline-first con sincronización posterior. El entorno de prueba incluye datos reales de pacientes y estudios y valida flujos completos de registro y asociación clínica. Se encuentran en integración módulos de codificación CIE-10, agenda interoperable y análisis acústico automatizado con retorno de parámetros estructurados.

Palabras **clave**

Historia clínica electrónica; multimodalidad; grabación audiovisual; offline-first; PWA; análisis de voz; Parkinson; trazabilidad; CIE-10; agenda.

1. Contexto de la investigación

El proyecto de I+D se enmarca en el estudio de rasgos vocales, de escritura y de expresión facial orientados a la clasificación y monitoreo en EP, con la articulación entre el Instituto de Ciencias de la Salud de la Universidad Nacional del Oeste (UNO) y el Servicio de Neurocirugía Funcional del Hospital Nacional Profesor Dr. Alejandro Posadas. En ese marco, SYNAPSIS cumple una función instrumental: unificar la gestión clínica y la recolección de evidencia multimodal vinculada a pacientes y estudios, reduciendo fragmentación entre planillas, documentos sueltos y repositorios no conectados (Occhipinti et al., 2024).

El trabajo aborda las críticas a historias clínicas electrónicas que suelen concentrarse en la fricción del flujo de trabajo (tiempo y carga cognitiva) y en la falta de transparencia de la trazabilidad de quién registró qué y cuándo, aspectos especialmente sensibles en entornos hospitalarios (Krause et al., 2023).

2. Objetivos

El objetivo general del proyecto es integrar las HCD con evidencia multimodal derivada del análisis del habla, escritura y videos de pacientes con EP para su utilización en pilotos clínicos en hospitales públicos.

En términos operativos, los objetivos incluyen: (i) gestionar grabaciones y estudios asociados al seguimiento, (ii) construir una base de datos con datos multimodales vinculados a las HCD, y (iii) diseñar la infraestructura para ampliar y sostener una base de datos multimodal a mayor escala.

En esta etapa, el foco estuvo puesto en asegurar que las HCD sea estable y usable y en sostener un diseño de captura multimodal que funcione en condiciones reales, incluyendo escenarios con conectividad limitada.

3. Diseño e implementación de SYNAPSIS

SYNAPSIS se implementa como plataforma web con backend en PHP bajo esquema MVC y persistencia en MySQL. La base de datos incluye 17 tablas que gestionan pacientes, usuarios, historia clínica, diagnósticos, estudios y multimedia.

El módulo de Historia Clínica Digital fue rediseñado para resolver fragmentación previa y mejorar la lectura longitudinal. Se estableció navegación por identificador interno del paciente, autoría visible y registro consistente de fecha y hora, junto con una línea temporal compacta y expandible.

La lectura longitudinal se resolvió con una línea temporal compacta y expandible: cada entrada se presenta con un encabezado resumido y el detalle se despliega bajo demanda. Este patrón reduce fatiga y favorece una visión general primero y detalle después, consistente con propuestas de visualización temporal aplicada a registros clínicos.

La gestión de adjuntos incorpora validaciones básicas de seguridad conforme a OWASP (s. f.) y se orienta a una historia clínica multimedia contextualizada (Plazzotta et al., 2008).

El subsistema de grabación contempla tres modalidades complementarias: captura web institucional mediante APIs estándar (World Wide Web Consortium, 2025), captura offline para trabajo de campo y una PWA offline-first basada en Service Workers (World Wide Web Consortium, 2026) con sincronización posterior.

4. Resultados y estado actual (evidencia operativa)

El entorno de desarrollo cuenta con datos reales de carga: 7 pacientes, 21 entradas de historia clínica, 26 estudios asociados, 20 diagnósticos versionados y 9 archivos multimedia vinculados.

Estos registros permitieron validar flujos completos de selección de paciente, navegación longitudinal, incorporación de nuevas entradas y asociación de estudios y archivos a eventos clínicos.

Se encuentran en integración un módulo de codificación diagnóstica CIE-10 con búsqueda incremental validada por el profesional (World Health Organization, 2019), un módulo de Agenda con integración a Google Calendar (Google, 2025b) y un módulo de análisis acústico basado en microservicios en Python que procesa parámetros con Praat y retorna resultados estructurados (Boersma & Weenink, 2001; Tsanas et al., 2010).

5. Formación de recursos humanos

En el marco del proyecto se incorporaron dos becarios de investigación y un estudiante adscripto. Su participación se orienta al desarrollo incremental del sistema,

documentación técnica de módulos en curso y proyectados para la integración de un módulo de análisis acústico automatizado.

6. Avances y Trabajos futuros

Se incorporó como línea de trabajo un módulo de codificación diagnóstica CIE-10: catálogo estructurado con búsqueda incremental y selección explícitamente validada por el médico, para normalizar diagnósticos y construir una lista longitudinal de problemas por paciente (World Health Organization, 2019).

Asimismo, se avanza en un módulo de Turnos/Agenda con integración de Google Calendar, contemplando agenda personal (calendarId=primary) y agenda general compartida. La integración se implementa bajo una política de privacidad que publica en Google títulos neutros, manteniendo el detalle dentro de SYNAPSIS. La operación de eventos se apoya en documentación oficial (Google, 2025b).

Como proyección inmediata, el equipo de becarios del proyecto desarrolla un módulo de análisis de voz que combina captura web y procesamiento distribuido. El audio capturado se envía a un microservicio en Python expuesto como API, que procesa parámetros acústicos con Praat a través de su interfaz Parselmouth (Jadoul et al., 2018) y devuelve resultados en JSON para integrarlos como parte de la evidencia del estudio. Este diseño modular permite escalar y mantener el procesamiento separado del núcleo web y, a futuro, habilita la incorporación de modelos de aprendizaje automático sobre parámetros como jitter, shimmer, F0, intensidad y HNR (Tsanas et al., 2010; Giuliano et al., 2025).

Como proyección de más largo plazo, se documenta una unidad de expansión orientada a análisis multimodal e inteligencia acústica aplicada, manteniendo el criterio incremental: primero núcleo clínico, luego normalización (CIE-10) y coordinación operativa (Agenda), y finalmente cierre de captura distribuida y análisis avanzado con validación clínica.

7. Conclusiones

Esta etapa consolida un núcleo clínico funcional para SYNAPSIS: Historia Clínica Digital con lectura longitudinal, trazabilidad visible e integración de evidencia multimodal. El trabajo en grabación audiovisual se organiza en tres modalidades complementarias (web, offline y PWA), lo que permite abordar escenarios reales sin perder la asociación clínica de los datos. El foco en consistencia de navegación, reducción de errores expuestos al usuario y organización de la información atiende a aspectos de seguridad y usabilidad en sistemas de salud (Krause et al., 2023). Los próximos hitos se concentran en completar CIE-10, cerrar la integración de Agenda y avanzar en el módulo de análisis de voz con retorno de parámetros acústicos estructurados, sosteniendo criterios de privacidad, auditabilidad y validación clínica.

Referencias

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. Glot International, 5(9/10), 341–345. https://www.fon.hum.uva.nl/paul/papers/speakUnspeakPraat_glot2001.pdf

Giuliano, M., Fernández, L. A., & Legnani, W. E. (2025). Using reduced feature space (2D and 3D) based on entropic measures for detecting Parkinson's disease through voice. *Statistics, Optimization & Information Computing*, *14*(3), 1542-1565, https://www.iapress.org/index.php/soic/article/view/2638

Google. (2025b). Create events | Google Calendar API. Google Developers. Recuperado el 24 de febrero de 2026, de

https://developers.google.com/workspace/calendar/api/guides/create-events

Krause, M., Petracci, M. N., Elicabe, E., Rodríguez Zoya, P. G., Schwarz, P. K. N., & Ridao, M. (2023). Historia clínica electrónica: miradas críticas de médicos del Área Metropolitana de Buenos Aires, Argentina. Interface (Botucatu), 27, e220072. https://doi.org/10.1590/interface.220072

Occhipinti, P. S., Giuliano, M., Pérez, S. N., & Berduc, A. (2024). SYNAPSIS: Plataforma de datos multimodales y gestión clínica para la Enfermedad de Parkinson, Libro de Actas - 30° Congreso Argentino de Ciencias de la Computación (CACIC 2024). Universidad Nacional de La Plata. https://sedici.unlp.edu.ar/handle/10915/176861

OWASP. (s. f.). Input Validation Cheat Sheet (sección: File Upload Validation). OWASP Cheat Sheet Series. Recuperado el 24 de febrero de 2026, de https://cheatsheetseries.owasp.org/cheatsheets/Input_Validation_Cheat_Sheet.html

Plazzotta, F., Luna, D., González Bernaldo de Quirós, F., & Otero, C. (2008). Imágenes más allá del PACS: el Proyecto Historia Clínica Multimedia. INFOLAC 2008 - AAIM. https://www.hospitalitaliano.org.ar/multimedia/archivos/servicios_attachs/4826.pdf

Tsanas, A., Little, M. A., McSharry, P. E., & Ramig, L. O. (2010). Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE Transactions on Biomedical Engineering, 57(4), 884–893. https://doi.org/10.1109/TBME.2009.2036000

World Health Organization. (2019). International Classification of Diseases, Tenth Revision (ICD-10), Sixth Edition, 2019. World Health Organization. Recuperado el 24 de febrero de 2026, de https://icd.who.int/browse10/Content/statichtml/ICD10Volume2_en_2019.pdf

World Wide Web Consortium. (2025). MediaStream Recording (W3C Working Draft, 17 April 2025). Recuperado el 24 de febrero de 2026, de https://www.w3.org/TR/mediastream-recording/

World Wide Web Consortium. (2026). Service Workers. Recuperado el 24 de febrero de 2026, de https://www.w3.org/TR/service-workers/

Jadoul, Y., Thompson, B., & de Boer, B. (2018). *Introducing Parselmouth: A Python interface to Praat*. *Journal of Phonetics, 71*, 1–15. https://doi.org/10.1016/j.wocn.2018.07.00

Tecnologías para el desarrollo de software accesible, inteligente y de calidad

Susana Herrera [1], Maximiliano Budán [1], Sergio Rocabado [2], Federico Rosenzvaig [1], Marina Nasif [1], Adriana Jiménez [1], Aldo Roldan [1], Adrián Belluomini [1], Claudia Avila [1], Luciano Galeano [1]

{sherrera, mbudan, aroldan, cavila}@unse.edu.ar, {sergiorocabado, fedvaig, nasifmarina12, ajfraguas, adriannicolasb, proflgaleano}@gmail.com

(1) Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero.

(2) Facultad de Ciencias Exactas, Universidad Nacional de Salta.

Resumen

En este trabajo se presenta una investigación que da continuidad a los proyectos sobre Computación Móvil, desarrollados en la Universidad Nacional de Santiago del Estero (UNSE) de 2012 a 2024; en los cuales se lograron resultados vinculados a la eficiencia de aplicaciones móviles, el aprendizaje basado en dispositivos móviles (m-learning), realidad aumentada (RA), tecnologías para discapacidad y herramientas y métodos para el desarrollo móvil multiplataforma.

Sobre la base de dichos resultados, esta investigación propone continuar investigando sobre métodos y herramientas para el desarrollo/mantenimiento de aplicaciones móviles accesibles e inteligentes, orientadas a resolver problemáticas en los siguientes dominios: aprendizaje y personas con discapacidad e industria 4.0. Tiene por finalidad la optimización del desarrollo de aplicaciones web y móviles mediante el estudio e integración de tecnologías innovadoras que permitan obtener software accesible, inteligente y de calidad.

La investigación es llevada a cabo por un equipo conformado predominantemente por informáticos, con la colaboración del Ministerio de Salud de la Provincia de Santiago del Estero y el asesoramiento de investigadores de UNLP, UNSa y Universidad Paris 8.

***Palabras clave*:** Aplicaciones móviles multiplataforma, realidad aumentada, inteligencia artificial, accesibilidad, m-learning, aplicaciones móviles para personas con discapacidad.

1 Contexto

El grupo de investigadores de Computación Móvil de la UNSE, integrantes del Instituto de Investigaciones en Informática y Sistemas de Información (IIISI) trabajan sobre métodos y tecnologías móviles que permitan proveer soluciones innovadoras a problemáticas referidas a personas vulnerables, de una manera flexible e interdisciplinaria, relacionando: Computación Móvil, Ingeniería del Software, Inteligencia Artificial, Educación y Salud.

Para realizar esta investigación aplicada, se cuenta con la colaboración del Servicio de Detección Temprana de Hipoacusia de Santiago del Estero, formalizada mediante convenio interinstitucional vigente. Asimismo, cuentan con el asesoramiento del Laboratorio en Investigación en Informática LIDI de la Facultad de Informática de la UNLP (Aguirre et al., 2019; Delia et al. 2018; Delia, 2017) y del Laboratorio en Cognición Humana y Artificial de la Universidad Paris 8, Francia (Hanine et al., 2021), con quienes también existe un acuerdo interinstitucional vigente.

Los antecedentes del grupo se ubican en el dominio de la Ingeniería del Software, especialmente en el área de Computación Móvil (Herrera et al., 2024; Herrera et al., 2023-b; Herrera et al., 2021-b; Diaz et al., 2019; Budan et al., 2018; Fennema et al., 2018; Herrera et al., 2018) y Tecnologías para la Discapacidad (Herrera et al., 2023-a; Zampar & Herrera, 2023; Jimenez et al., 2023; Duran et al., 2023; Zampar & Herrera, 2022; Herrera et al., 2021-a; Herrera, 2020; Rosenzvaig et al., 2020; Barrera Nicholson et al., 2020; Córdoba et al., 20216) y

Aprendizaje Móvil (Macedo et al., 2020; Maldonado et al., 2018, Morales et al., 2018) . Su principal línea de trabajo se centra en el desarrollo de aplicaciones móviles multiplataforma que utilizan técnicas de visualización innovadoras, como la Realidad Aumentada (Saad Elías & Herrera, 2025; Sánchez, 2019), así como herramientas de Inteligencia Artificial.

2 Introducción

Este proyecto de investigación tiene por finalidad la optimización del desarrollo de aplicaciones web y móviles mediante el estudio e integración de tecnologías innovadoras que permitan obtener software accesible, inteligente y de calidad.

Los colectivos impactados corresponden principalmente a personas cuya situación de vulnerabilidad se asocia a algún tipo de discapacidad, principalmente la visual y auditiva. En cuanto a las problemáticas abordadas, se destacan el aprendizaje, el entrenamiento auditivo, la comunicación por reconocimiento automático de voz y la representación de procesos vinculados con la Industria 4.0.

En continuidad con la investigación desarrollada previamente, en Ingeniería del Software, se evoluciona desde arquitecturas híbridas hacia esquemas de compilación cruzada, integrando Inteligencia Artificial Generativa y tecnologías de interacción como Realidad Aumentada y Realidad Virtual.

Actualmente, se pretende optimizar el desarrollo de aplicaciones web y móviles (Londoño et al. 2017), al incluir métodos y tecnologías tales como:

- arquitecturas en la nube,
- arquitecturas móviles,
- desarrollo multiplataforma,
- bibliotecas de RA,
- diseño industrial,
- bibliotecas de IA,
- sistemas de reconocimiento de voz

3 Líneas de investigación y desarrollo

Se trabaja en dos líneas de investigación, las cuales tienen sus objetivos y responsables específicos: 1) Desarrollo de aplicaciones Web y Móvil, RA/RV e Inteligencia Artificial y 2) Desarrollo de Aplicaciones para personas con discapacidad.

En cuando a la línea 1, está integrada por especialistas en Informática y en diseño industrial; se encarga principalmente del análisis y comparación de aplicaciones multiplataforma con arquitectura móvil y web, usando tecnologías Flutter y NET MAUI. Ambas tecnologías son de código abierto, gratuitas, generan aplicaciones móviles y de escritorio compilando en código nativo. Se realizan ensayos para estudiar comparativamente ambas tecnologías, considerando aspectos tales como: curva de aprendizaje, conexión a bases de datos, aplicaciones generadas para diversos sistemas operativos, cantidad de modificaciones que necesita el código para poder ser ejecutado, tiempo de compilación, tamaño de la aplicación, variedad de dispositivos en los que se ejecuta sin error. También se estudia: la eficacia de la migración de Xamarin a MAUI de los desarrollos realizados en el marco de proyectos precedentes, la exploración e incorporación de IA generativa en el desarrollo de las aplicaciones, el procesamiento de lenguaje natural y reconocimiento automático de voz, la incorporación de RA y RV en entornos de diseño industrial y la implementación de una nube de servicios que pueda dar soporte a las aplicaciones desarrolladas en el proyecto.

La línea 2, desde el área HCI, continua con el estudio de tecnologías innovadoras de interacción para el aprendizaje de personas con discapacidad y para la educación superior. Considerando el modelo de enseñanza de inglés para personas con discapacidad visual obtenido en el proyecto precedente, se dedica al diseño de un diccionario bilingüe español–inglés para personas con discapacidad visual y dislexia, que detecte y cambie automáticamente de idioma. Para ello, se trabaja en la integración a proyectos Flutter o MAUI de APIs de IA dedicadas al procesamiento del lenguaje natural y al reconocimiento automático de voz. Se continua

con la investigación vinculada al modelo para la enseñanza de jóvenes hipoacúsicos basado en Logogenia. Por otra parte, se da continuidad al desarrollo de aplicaciones con RA y modelos 3D animados, actualizando las aplicaciones Matemática y Sonidos (entrenamiento auditivo) y Child-Comm (comunicador para niños con parálisis cerebral), migrando su desarrollo a Flutter/.NET MAUI e incorporando componentes inteligentes. Además, se avanzará en el desarrollo de modelos de ambientes empresariales/ industriales/productivos que incorporen interfaces de RA o Realidad Virtual (RV). Además de la RA basada en el uso de dispositivos móviles, se diseñarán modelos que usen cascos o lentes para el despliegue de RA o RV. Se estudiará el impacto del uso de dichos modelos en educación superior y/o en la gestión empresarial.

Como se mencionó en el párrafo anterior, se estudiarán bibliotecas de funciones de IA de procesamiento del lenguaje natural y su uso en Flutter y MAUI. También se estudiarán bibliotecas de funciones de IA generativa (Ej. Open AI). Se pretende diseñar y desplegar una "Nube institucional de Ingeniería del Software & IA" que contenga tecnologías recomendadas para la construcción de software en el ámbito de la UNSE. Esta nube permitirá construir aplicaciones para ámbitos educativos y de desarrollo productivo-empresarial.

4 Resultados esperados

a. Generar nuevo conocimiento tecnológico en Informática, publicados en artículos científicos presentados en congreso, libros y/o revistas; relativos a:

- Ingeniería del Software: métodos y tecnologías blandas para el desarrollo de aplicaciones web y móviles multiplataforma, accesibles e inteligentes, con RA/RV; desarrollo de software con IA; entorno de desarrollo colaborativo y nube de servicios de RA/RV e IA.
- Interacción Hombre-Máquina: tecnologías para mejorar la comunicación, aprendizaje y entrenamiento auditivo de personas con discapacidad.

b. Dar cursos de posgrado sobre Tecnologías para la Discapacidad Visual, como parte de la oferta académica de la FCEyT. Se está programando un curso para 2026.

c. Dar cursos de capacitación en la Escuela de Formación Tecnológica de la FCEyT, sobre recursos accesibles para personas ciegas, entre otros. Se están programando 2 cursos para ser dados en la Escuela de Formación Tecnológica de la FCEyT.

d. Proyectos de extensión a la comunidad santiagueña; en particular con la Escuela para Ciegos de la provincia (CEI 50) que permitan optimizar el uso de las tecnologías informáticas del establecimiento educativo. Se desarrolló uno en 2025.

5 Formación de recursos humanos

Este proyecto contribuye a la formación y capacitación de los investigadores, estudiantes de grado y de posgrado involucrados. Incentivará a los alumnos de la Lic. en Sistemas de Información a iniciarse en las actividades de investigación y favorecerá la realización de trabajos finales de grado y posgrado en las líneas de investigación de este proyecto.

En concreto, en el marco de este proyecto finalizó el desarrollo de una tesis de Magister Tecnologías Educativas, UNdeC (Nasif & Herrera, 2023); y continúa el desarrollo de una tesis de Doctorado en Ciencias Informáticas de la UNLP, una tesis de Maestría en Informática Educativa, dos trabajos finales de grado de la Licenciatura en Sistemas de Información y trabajo final de Ingeniería Industrial, todos los últimos de la UNSE.

Referencias

1. Aguirre, V., L Delía, P Thomas, L Corbalán, G Cáseres (2019) - PWA and TWA: Recent Development Trends. Argentine Congress of Computer Science. Páginas 205-214. Editor Springer.
2. Barrera Nicholson, P., Herrera, S., Macedo, A., Trutner, G. & Diorio, G. (2020). Accessibility in Argentina Initial comparison with the French domain. Conference Handicap 2020. Paris, France.

3. Budán, P., Rosenzvaig, F., Najar P., Saavedra, E, Herrera, S., Morales, M. (2018). MobileRA: Método Ágil para el Desarrollo de Aplicaciones Móviles Multiplataforma. CONAIISI. Actas del Congreso. ISSN 2347-0372. Noviembre 2018. Mar del Plata, Argentina.
4. CÓRDOBA, M., NAJAR, P., BUDAN, P. (2016). Sistema Alternativo de Comunicación para Niños con Parálisis Cerebral Infantil. Argentina. San Salvador de Jujuy. Libro. Artículo Completo. Jornada. XI Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA. CODINOA.
5. Delía, L. (2017). Desarrollo de aplicaciones móviles multiplataforma. Trabajo Final de Especialista en Ingeniería Web. Facultad de Informática, Universidad Nacional de La Plata.
6. Delía, L., Thomas, P., Corbalan, L., Fernandez Sosa, J., Cuitiño, A., Cáseres, G., Pesado, P. (2018). Development Approaches for Mobile Applications: Comparative Analysis of Features. Computing Conference 2018 10-12 July | London, UK.
7. Díaz, D., Herrera, S., Rosenzvaig, F. (2019). Estudio de la capacidad de mantenimiento de las aplicaciones móviles híbridas. XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019). ISBN 978-987-688-377-1. Pag. 662-673. URI: http://sedici.unlp.edu.ar/handle/10915/91084. Rio Cuarto. 2019.
8. Durán, E.B., Herrera, S.I., Fernández, P.M., Alvarez, M.M., Zampar, M. (2023). ICT to Support Inclusive Education in the Republic of Argentina. In: Rioux, M.H., Viera, J., Buettgen, A., Zubrow, E. (eds) Handbook of Disability. Springer, Singapore. https://doi.org/10.1007/978-981-16-1278-7_64-1 .
9. FENNEMA, M., HERRERA, S., PALAVECINO, R., BUDÁN, P., ROSENZVAIG, F., NAJAR RUIZ, P., CARRANZA, A., SAAVEDRA, E. (2018). Aproximaciones para el desarrollo multiplataforma de Aplicaciones Móviles. Workshop en Investigación en Ciencias de la Computación, WICC. ISBN 978-987-3619-27-4. Corrientes.
10. Hanine, M., Herrera, S., Chene, D. & Uzan, G. (2021). Agent virtuel pour l'utilisation effective du web par des aveugles. Colloque « Jeunes Chercheuses, Jeunes Chercheurs » : Handicap, Vieillissement, Autonomie, Insertion, Technologies. IFRATH. ISBN 978-2-9571218-1-6. https://ifrath.fr/jcjc2021/
11. Herrera, S., Rocabado, S., Rosenzvaig, F., Romano Paz, F., Palavecino, R., Jimenez, A., Macedo, A., Zampar, M., Nasif, M., Morales, M., Maldonado, M., Barrera Nicholson, P., Budan, M., Divi, P., Gómez, P. (2024). Tecnologías informáticas accesibles e inclusivas. Investigaciones Científico. La investigación científica en la Universidad Nacional de Santiago del Estero y sus aportes a la Sociedad. Vol II. Pg 541-571. ISBN 978-987-4456-47-2. EDUNSE. Santiago del Estero. Disponible en: https://www.unse.edu.ar/images/secyt/publicaciones/librocicyt2024.pdf
12. Herrera, S., Romano Paz, F., Leyva, J., Tamer, A. (2023-a). Picudo App: aplicación móvil para control del picudo. Congracha. 4to Congreso Internacional del Gran Chaco Americano. ISBN 978-987-8922-58-4.
13. Herrera, S., Rosenzvaig, F., Morales, M., Salvatierra, S., Sanchez, C., Tiberti, A., Budan, P. (2023-b). Matemática y Sonidos: aplicación para entrenamiento auditivo y aprendizaje de matemática. Congreso de Tecnologías para la Accesibilidad. ISBN 978-987-8922-64-5.
14. Herrera, S.I., Manresa-Yee, C. & Sanz, C.V. Mobile learning for hearing-impaired children: Review and analysis. Univ Access Inf Soc (2021-a). https://doi.org/10.1007/s10209-021-00841-z .
15. Herrera, S. I., Budan, P. D., Rosenzvaig, F., Ruiz, P. J., Morales, M. I., Maldonado, M. D., & Sánchez, C. A. (2021-b). Developing Augmented Reality Multi-Platform Mobile Applications. In F. Cipolla-Ficarra (Ed.), Handbook of Research on Software Quality Innovation in Interactive Systems (pp. 371-390). IGI Global.

https://doi.org/10.4018/978-1-7998-7010-4.ch016 .
16. Herrera, S. (2020). Mobile Learning with Augmented Reality to Improve Inclusion of Hearing-Impaired Children. Mobile Learning Week. UNESCO. Paris.
17. HERRERA, S., SANZ, C., MORALES, M., PALAVECINO, R., MALDONADO, M., IRURZUN, I. CARRANZA, A., SUÁREZ, G. (2018). M-learning con Realidad Aumentada basada en Objetos 3D. Workshop de Investigadores en Ciencias de la Computación, WICC. ISBN 978-987-3619-27-4. Corrientes.
18. Jiménez, A., Orieta, L., Herrera, S., Maldonado, M., Nasif, M., Pelaez, M., Barrera Nicholson, P. (2023). Tecnologías para la Evaluación de Accesibilidad Web. Congreso de Tecnologías para la Accesibilidad. ISBN 978-987-8922-64-5.
19. Londoño Rojas, L. F., V. Tabares Morales, M. Rosecler Bez y N. D. Duque Mendez, (2017). "Análisis comparativo de guías para el desarrollo web accesible," Ciencia e Ingeniería Neogranadina, vol. 28, no. 1, pp. 101-115. En: http://dx.doi.org/10.18359/rcin.2683.
20. Macedo, A., Suarez, C., Maldonado, M., Carranza, A. (2020). Noun phrases learning through Imacolab, a tool for m-learning. Argentinian Journal of Applied Linguistics. Trabajo en evaluación.
21. Maldonado, M., Herrera, S., Irurzun, I., Carranza, J., Palavecino, R., Macedo, A., Suárez, C. (2018). M-learning: Aprendizaje en Diversos Niveles Educativos Usando ImaColab. Revista Investigaciones en Facultades de Ingeniería del NOA. ISSN 1853-6662. Nro. 4. Ed. UNSE. Santiago del Estero. 2018.
22. MORALES, M., HERRERA, S., MALDONADO, M., BUDÁN, P., ROSENZVAIG, F. (2018). M-learning con Realidad Aumentada para el aprendizaje significativo en Álgebra Lineal. Revista Tecnologias na Educacao. Ano10-Número/vol24-Edição SITED 2018. ISSN 19844751. Disponible en: http://tecedu.pro.br/ano10-numero-vol24-edicao-tematica-vii/. Brasil.
23. Nasif, M. y Herrera, S. (2023). Diseño tecno-pedagógico para la enseñanza del inglés a personas con ceguera. Congreso de Tecnologías para la Accesibilidad. ISBN 978-987-8922-64-5.
24. Rosenzvaig, F., Córdoba, M., Budan, P., Najar, P. (2020). Child-Comm: An Alternative Communication System adaptive to user's features. Handicap 2020. Paris.
25. Saad Elías, N. & Herrera, S. (2025). Realidad virtual y realidad aumentada en la industria. Una revisión sistemática de literatura. III Congreso de Ciencia y Tecnología del CODINOA. ISBN: 978-987-8922-59-1.
26. Sánchez, C. (2019). MODELOS 3D ANIMADOS PARA APLICACIONES MÓVILES CON REALIDAD AUMENTADA. 8° Jornadas de Estudiantes y Jóvenes Investigadores (JEyJI). FAyA, UNSE. 25 Octubre 2019. Expositor: Carlos Sánchez. Asesora: Marilena Maldonado.
27. Zampar, M. & Herrera, S. (2023). Modelo inclusivo para el desarrollo de competencias lingüísticas de personas sordas. Congreso de Tecnologías para la Accesibilidad. ISBN 978-987-8922-64-5.
28. Zampar, M. & Herrera, S. (2022). Tecnologías de reconocimiento automático de voz en contextos educativos: una revisión sistemática de literatura. XXVIII Congreso Argentino de Ciencias de la Computación - CACIC 2022. ISBN: 978-987-1364-31-2. Pp. 207-216. http://sedici.unlp.edu.ar/handle/10915/149582 . La Rioja.

Transformación de LMS a ALS utilizando una arquitectura de referencia

Alejandro Sartorio, Gustavo Rossi
Centro de Altos Estudios en Tecnología Informática (CAETI)
Universidad Abierta Interamericana (UAI), Argentina
alejandro.sartorio@uai.edu.ar, gustavo.rossi@uai.edu.ar

RESUMEN

Los Sistemas de Gestión de Aprendizaje (LMS) contemporáneos presentan, en su mayoría, arquitecturas monolíticas caracterizadas por una elevada rigidez arquitectónica, lo que limita la implementación de mecanismos de personalización dinámica propios de los Sistemas de Aprendizaje Adaptativo (ALS). Este trabajo propone un marco de modernización incremental basado en el patrón Strangler Fig, integrando lazos de control autonómico MAPE-K (Monitor-Analyze-Plan-Execute-Knowledge) para gestionar la evolución de sistemas legacy hacia entornos pedagógicos adaptativos. La contribución central se articula mediante la Arquitectura de Referencia ARALS, la cual actúa como un facilitador para la sustitución progresiva de componentes mediante el desacoplamiento de la lógica transaccional y la incorporación de componentes que habilitan la implementación de las propiedades PALS (*Propiedades Adaptativas de los Sistemas de Aprendizaje*). Se presenta, así, una estrategia sistemática de transformación arquitectónica para la transición de LMS hacia ALS.

Palabras clave: Strangler Fig Pattern, Sistema de Aprendizajes Adaptativos, Arquitectura de Software.

CONTEXTO

La presente investigación se desarrolla en el marco de la línea de Ingeniería de Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). La propuesta, que integra y amplía resultados difundidos en trabajos previos [16]–[19], forma parte de una tesis doctoral situada en la intersección de dos líneas de investigación del CAETI: Ingeniería de Software y Sociedad del conocimiento y tecnología para la educación. Este encuadre habilita un abordaje interdisciplinario que articula decisiones de arquitectura de software con requerimientos pedagógicos orientados a la personalización educativa.

En función de este encuadre institucional y académico, el estudio se orienta al siguiente objetivo estratégico: proponer un método de transformación arquitectónica que integre el *Strangler Fig Pattern* y ciclos MAPE-K, sustentado en una arquitectura de referencia destinada a operacionalizar acciones de estrangulamiento para el reemplazo controlado de componentes. Asimismo, se propone caracterizar las propiedades PALS (*Propiedades Adaptativas de los Sistemas de Aprendizaje*) como atributos emergentes derivados de la incorporación de componentes autonómicos, consolidando un puente interdisciplinario entre la ingeniería de software y la personalización educativa.

1. INTRODUCCIÓN

Los *Learning Management Systems* (LMS) tradicionales suelen implementar arquitecturas monolíticas fuertemente acopladas, en las que la lógica pedagógica, la gestión transaccional y la integración con herramientas externas quedan entrelazadas, dificultando la incorporación de mecanismos de personalización dinámica característicos de los *Adaptive Learning Systems* (ALS). Esta rigidez limita la evolución incremental y tensiona la estabilidad operativa ante requisitos contemporáneos de adaptabilidad y observabilidad del aprendizaje [1], [2]. Si bien la migración hacia arquitecturas más modulares

puede mitigar estas restricciones, los enfoques *big bang* introducen riesgos significativos de interrupción del servicio y regresión funcional en entornos educativos [3]. En este contexto, el *Strangler Fig Pattern* se presenta como una estrategia de modernización gradual que habilita la coexistencia controlada entre capacidades heredadas y nuevos servicios adaptativos, reduciendo el riesgo de reemplazo abrupto [4], [5].

Bajo este patrón, la modernización progresa mediante la intercepción y el redireccionamiento de responsabilidades desde el monolito hacia componentes externos, lo que resulta especialmente pertinente para arquitecturas como la de la Figura 1, donde el acoplamiento entre módulos y la centralización de la lógica impiden introducir adaptación sin impactar el núcleo del sistema. Para sistematizar dicha transición, se emplea la Arquitectura de Referencia ARALS como estructura de contención e interoperabilidad, que delimita componentes del legado a través de puntos de integración estandarizados. En particular, *LTI* (*Learning Tools Interoperability*) se utiliza para incorporar herramientas externas y *xAPI* (*Experience API*) para capturar y gestionar trazas de aprendizaje [6]. Las acciones de estrangulamiento se operacionalizan mediante primitivas arquitectónicas gobernadas por el ciclo MAPE-K [7], posibilitando la incorporación incremental de propiedades PALS (*Propiedades Adaptativas de los Sistemas de Aprendizaje*) de manera controlada.

A diferencia de dominios de software genéricos, la evolución de sistemas educativos requiere preservar la consistencia del modelo de estudiante y sostener la validez pedagógica durante la transición [8]. Con el fin de estructurar este proceso, la Tabla 1 sintetiza un conjunto de acciones de estrangulamiento y su contribución a propiedades PALS. En particular, la *Intercepción y Control* incorpora sensado no intrusivo para capturar trazas xAPI sin alterar el flujo del monolito; la *Sustitución Funcional* reemplaza gradualmente lógica de negocio por servicios adaptativos, manteniendo la equivalencia funcional; y la *Sincronización de Estado* garantiza que el legado y los nuevos componentes operen sobre un estado del aprendiz (*Learner State*) coherente.

En conjunto, la trayectoria metodológica propuesta busca transformar un *LMS* en un *ALS*. Esto requiere trascender la adición aislada de funcionalidades mediante una hoja de ruta de abstracción que vincule la arquitectura específica del monolito con la Arquitectura de Referencia ARALS, concebida como soporte para ejecutar acciones de estrangulamiento de manera sistemática sobre componentes heredados, preservando la operación del núcleo transaccional durante toda la evolución.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El método de *contención arquitectónica* operacionaliza el estrangulamiento gradual de un LMS monolítico mediante ARALS, concebida como una arquitectura de referencia que estructura la transición en capas y estabiliza la evolución incremental. El enfoque integra tres elementos: (i) una estructura de soporte (ARALS), (ii) *puntos de interoperabilidad* (*interoperability eyelets*) que habilitan instrumentación y ruteo no invasivos, y (iii) acciones de estrangulamiento gobernadas por ciclos MAPE-K que incorporan propiedades PALS de manera controlada. En conjunto, ARALS permite desplazar progresivamente responsabilidades desde el *legacy* hacia servicios adaptativos sin comprometer la operación del núcleo transaccional.

La *Capa de Presentación* concentra el punto de entrada y el control del ruteo durante la coexistencia. En ella, la *Strangler Façade* intercepta solicitudes y decide si cada interacción debe resolverse en el monolito o derivarse a componentes adaptativos, habilitando sustitución gradual con control explícito del flujo. Complementariamente, el *Adaptive UI Engine* aplica ajustes de interfaz derivados de las decisiones del lazo MAPE-K, permitiendo introducir adaptaciones perceptibles para el estudiante sin modificar la *Legacy UI*.

La *Capa de Interoperabilidad* define los *eyelets* como anclajes técnicos de integración y trazabilidad que evitan el acoplamiento directo al código fuente del LMS. Esta capa agrupa (i) un *Gateway/Proxy* para intercepción y redirección, (ii) conectores *LTI 1.3* para inyectar herramientas externas y actuadores, y (iii) un *Telemetry Collector* basado en *xAPI/Caliper* para capturar evidencia de interacción. Asimismo, estándares

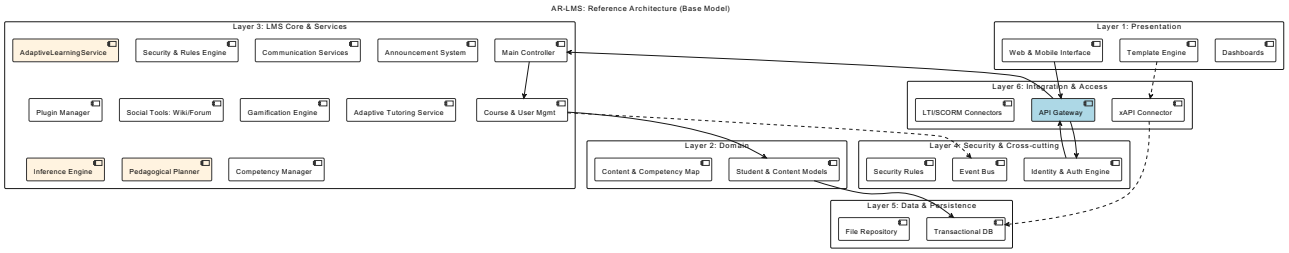


Figura 1. Vista por capas de un LMS típico

Componente	**Acción**	**Propiedad PALS**
Interceptor/Hook	Sniffing (Monitoreo)	Awareness
Proxy/Redirector	Swapping (Sustitución)	Adaptability
Distributed Cache	Shared State (Sincronización)	Memory
Circuit Breaker	Fail-safe (Resiliencia)	Responsiveness

Tabla 1
MAPEO DE ACCIONES DE ESTRANGULAMIENTO Y PROPIEDADES PALS EMERGENTES.

como *QTI/SCORM* proveen puntos de encapsulamiento para contenidos y evaluación. Estos mecanismos sostienen la migración incremental preservando interoperabilidad y auditabilidad de las decisiones.

La *Capa de Adaptación Cognitiva* (*The Laces*) implementa el ciclo MAPE-K como núcleo de control autonómico. El componente *Monitor* consume eventos y métricas recolectadas por la capa de interoperabilidad (Awareness); *Analyze* identifica patrones y brechas de conocimiento (Intelligence) [10]; *Plan* sintetiza decisiones de adaptación (Adaptability) [11]; y *Execute* aplica las acciones mediante conectores LTI y/o ajustes en la interfaz (Responsiveness). Este lazo introduce "presión" de estrangulamiento al desplazar gradualmente lógica pedagógica desde el *legacy* hacia componentes adaptativos.

La *Capa de Servicios* materializa las capacidades adaptativas como microservicios reemplazables, tales como *Personalization Service* y *Assessment Adaptation Service*. Estos servicios encapsulan funcionalidad pedagógica que, en un LMS monolítico, reside típicamente en la lógica de negocio de contenidos y evaluación. Su externalización permite sustitución funcional incremental y evaluación controlada mediante estrategias como *shadowing* y *canary deployment* [9], reduciendo riesgos de regresión.

La *Capa de Persistencia* soporta la continuidad del estado requerido por la adaptación. En particular, un *Learner State Store* mantiene perfiles, progreso y variables del modelo del estudiante, mientras que un repositorio de analítica (*Analytics DB*) consolida métricas históricas para evaluación y mejora continua. Esta capa habilita sincronización de estado entre el monolito y los nuevos servicios mediante *Shared Memory*, *Change Data Capture* y *Event Sourcing*, asegurando consistencia durante la coexistencia [13], [14].

Finalmente, la *Capa de Dominio y Conocimiento* define y mantiene los modelos y políticas que gobiernan la adaptación. Incluye el *Student Model* como representación multidimensional del aprendiz y una *Adaptation Knowledge Base* que almacena reglas, restricciones y preferencias derivadas del proceso de análisis. Para evitar contaminación semántica entre el *legacy* y ARALS, se incorpora una *Anti-Corruption Layer* que traduce modelos y contratos de datos en ambos sentidos, preservando la independencia evolutiva de los servicios adaptativos [9].

La operacionalización del estrangulamiento se realiza mediante *Primitivas de Refactorización Arquitectónica* [5] instanciadas sobre los *eyelets*: (i) *Intercepción y Control* (instrumentación pasiva y ruteo), (ii) *Sustitución Funcional* (swapping/shadowing/canary), (iii) *Composición Estructural* (segmentación por contextos y aislamiento mediante ACL y façade), y (iv) *Sincronización de Estado* (CDC/eventos/almacenamiento compartido). Estas primitivas permiten introducir propiedades PALS sin alterar la semántica

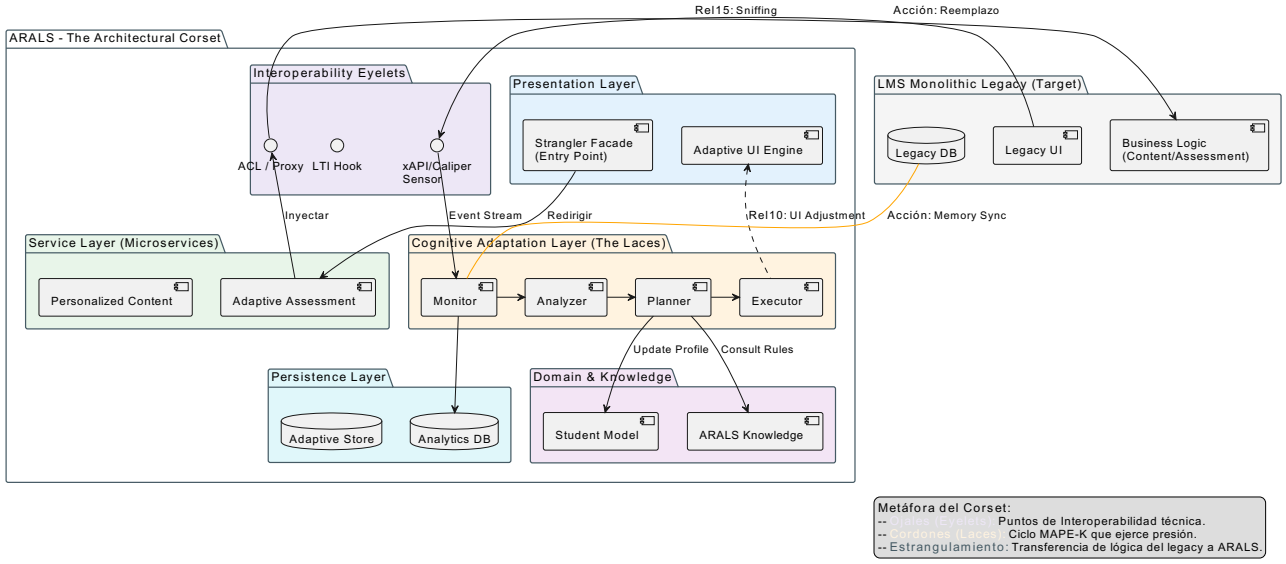


Figura 2. Esquema de Contención Arquitectónica: ARALS integra el sistema *legacy* mediante *eyelets* de interoperabilidad y acciones MAPE-K para estrangulamiento incremental.

base del LMS, guiando la transición hasta reducir el sistema original a un núcleo transaccional mínimo y delegar la inteligencia pedagógica en ARALS, como se ilustra en la Figura 2.

3. RESULTADOS OBTENIDOS Y OBJETIVOS EN CURSO

Se completó la propuesta del método de contención arquitectónica mediante la integración del *Strangler Fig Pattern* con ciclos MAPE-K. En este marco, la Arquitectura de Referencia ARALS quedó especificada a partir de tres elementos estructurales principales: **(i)** una *Strangler Facade* en la capa de presentación para interceptar y enrutar interacciones, **(ii)** una *Anti-Corruption Layer* para traducir modelos y aislar el dominio del *legacy*, y **(iii)** un *Modelo de Dominio y Conocimiento* que define contratos, políticas y vocabulario común para la adaptación.

Como resultado metodológico adicional, se desarrolló y validó una taxonomía de *primitivas de refactorización arquitectónica* organizada en cuatro categorías: (i) *Intercepción y Control* (instrumentación y ruteo no invasivos), (ii) *Sustitución Funcional* (validación y reemplazo gradual de capacidades), (iii) *Composición Estructural* (segmentación por contextos y aislamiento mediante ACL y fachada), y (iv) *Sincronización de Estado* (coherencia del *Learner State* y persistencia). Este marco establece un mapeo explícito entre acciones técnicas y propiedades PALS emergentes, preservando la semántica funcional del LMS durante la transición.

En cuanto a los objetivos en curso, se trabaja en la orquestación de ciclos MAPE-K concurrentes y en su validación en plataformas educativas con requerimientos adaptativos impulsados por la incorporación de inteligencia artificial. Asimismo, se investiga aprendizaje continuo en el componente *Knowledge* y se desarrolla un *framework* de evaluación para vincular decisiones de estrangulamiento con métricas empíricas de desempeño y resultados de aprendizaje.

4. FORMACIÓN DE RECURSOS HUMANOS

Esta investigación se integra en el proceso de formación de posgrado de un doctorando y se desarrolla en el Centro de Altos Estudios en Tecnología Informática (CAETI). El estudio forma parte de proyectos institucionales del CAETI enmarcados en las líneas de investigación sobre Ingeniería de Software y Sociedad del Conocimiento y Tecnología para la Educación.

De los temas que aquí se proponen derivan trabajos de investigación de grado, postgrado y publicaciones en congresos y revistas. Al estar relacionado en el dominio de los tecnología educativos se trabaja de forma interdisciplinar

con proyectos internos en CAETI y externos de CONICET.

Asimismo, se prevé impulsar la formación de recursos humanos mediante la definición de temas de tesis y trabajos finales de maestría alineados con esta línea, así como la incorporación de estudiantes de grado en actividades de investigación y desarrollo vinculadas a arquitecturas adaptativas, en el marco de las iniciativas académicas y de transferencia promovidas por la UAI y el CAETI.

5. CONCLUSIONES

Este trabajo propone una estrategia de transformación incremental para evolucionar plataformas LMS monolíticas hacia Sistemas de Aprendizaje Adaptativo (ALS), integrando el *Strangler Fig Pattern* con ciclos de control autonómico MAPE-K. La contribución se organiza en tres dimensiones: ARALS como estructura de contención y desacoplamiento, primitivas de estrangulamiento gobernadas por MAPE-K para incorporar propiedades PALS, y puntos de interoperabilidad estandarizados que habilitan instrumentación, ruteo y sustitución no invasiva. De esta manera, el enfoque preserva la estabilidad operativa y la integridad del núcleo transaccional del sistema *legacy*, reduciendo riesgos de migraciones disruptivas.

Referencias

[1] S. Meacham, V. Pech y D. Nauck, "AdaptiveVLE: An Integrated Framework for Personalized Online Education Using MPS JetBrains Domain-Specific Modeling Environment," *IEEE Access*, vol. 8, pp. 1-15, 2020.

[2] P. Di Francesco, P. Lago y I. Malavolta, "Migrating Towards Microservice Architectures: An Industrial Survey," en *IEEE Int. Conf. Software Architecture (ICSA)*, 2018, pp. 29-38.

[3] A. Balalaie, A. Heydarnoori y P. Jamshidi, "Microservices Architecture Enables DevOps: Migration to a Cloud-Native Architecture," *IEEE Software*, vol. 33, no. 3, pp. 42-52, 2016.

[4] M. Fowler, "StranglerFigApplication," 2004. [Online]. Disponible: https://martinfowler.com/bliki/StranglerFigApplication.html

[5] S.-P. Ma, C.-Y. Li, W.-T. Lee y S.-J. Lee, "Microservices Migration Using Strangler Fig Pattern and Domain-Driven Design," *J. Information Science and Engineering (JISE)*, vol. 38, no. 6, pp. 1-15, 2022.

[6] IMS Global Learning Consortium, "Learning Tools Interoperability (LTI) v1.3 Core Specification," 2019. [Online]. Disponible: https://www.imsglobal.org/spec/lti/v1p3/

[7] J. O. Kephart y D. M. Chess, "The Vision of Autonomic Computing," *Computer*, vol. 36, no. 1, pp. 41-50, 2003.

[8] J. Reich, *Failure to Disrupt: Why Technology Alone Can't Transform Education*. Harvard University Press, 2020.

[9] S. Newman, *Building Microservices: Designing Fine-Grained Systems*. O'Reilly Media, 2015.

[10] V. Ricci, I. Malavolta, H. Muccini y M. Shahin, "RAMSES: A Microservice Reference Architecture for AI-Enabled Adaptive Cloud Applications," *IEEE Software*, vol. 41, no. 1, pp. 56-65, 2024.

[11] R. F. Kizilcec, J. Reich y M. Yeomans, "Scaling Up Behavioral Science Interventions in Online Education," *Proc. National Academy of Sciences*, vol. 117, no. 26, pp. 14900-14905, 2020.

[12] L. H. Garcia Paucar y N. Bencomo, "The Reassessment of the Knowledge Base in Self-Adaptive Systems," en *IEEE Int. Conf. Autonomic Computing (ICAC)*, 2019, pp. 37-42.

[13] R. B. Halima, K. Guennoun, K. Barkaoui y Z. Guessoum, "A Formal Approach for MAPE-K Loop Based Self-Adaptive Systems," *Comput. Syst. Sci. Eng.*, vol. 42, no. 3, pp. 1009-1028, 2022.

[14] L. C. A. Coelho, "Uma arquitetura de código aberto para a aplicação do Strangler Pattern com Change Data Capture para setores com alto volume de dados," Tesis de Maestría, Universidade Federal do Ceará, 2023.

[15] M. Scozzina, P. A. Toledo, S. E. Roatta, and A. R. Sartorio, "CMS: Cow Monitoring System 2.0," in *Congreso Argentino de Ciencias de la Computación*, Argentina: Universidad Nacional de La Plata, 2023.

[16] S. Ayala, A. R. Sartorio, and A. Hernandez, "Educación superior, plataformas e inteligencia artificial (IA): un acercamiento a los estudios previos," in *Jornadas Argentinas de Informática e Investigación Operativa (JAIIO)*, Argentina: Universidad Nacional de La Plata, 2025.

[17] A. R. Sartorio, "Una arquitectura conceptual para implementar algoritmos de inteligencia artificial en la administración pública," in *Simposio de Informática en el Estado (SIE)*, Argentina: Universidad Nacional de La Plata, 2023.

[18] A. R. Sartorio, S. Ayala, and A. L. Hernandez, "Open Gateway de servicios de Inteligencia Artificial Generativa para la Gestión de Expedientes Digitales en la Administración Pública," in *Jornadas Argentinas de Informática e Investigación Operativa (JAIIO)*, Argentina: Universidad Nacional de La Plata, 2024.

[19] A. R. Sartorio, S. Ayala, and A. Hernandez, "Diseño de aplicaciones e-learning adaptativas: superando limitaciones mediante arquitecturas modulares y escalables," in *Jornadas Argentinas de Informática e Investigación Operativa (JAIIO)*, Argentina: Universidad Nacional de La Plata, 2025.

Transformación Digital. Tecnología en Ciudades Inteligentes Sostenibles. IA en Ingeniería de Software, Calidad de la IA.

Ariel Pasini, Marcos Boracchia, Rocío Muñoz,
Kristian Petkoff Bankoff, Juan Ignacio Torres, Ailen Panigo, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 La Plata Buenos Aires
(apasini, marcosb, rmunoz, kpb, jitorres, apanigo, ppesado) @lidi.info.unlp.edu.ar

RESUMEN

El III-LIDI (Instituto de Investigación en Informática LIDI) posee un subproyecto centrado en cuatro ejes principales
1. Transformación digital y Ciudades inteligentes sostenibles; 2. Gobernanza digital y gobierno abierto; 3. Calidad de I.A.; y 4. Inteligencia Artificial en Ingeniería de Software. Para cada eje se proponen un conjunto de objetivos para analizar durante este año y se describen los resultados obtenidos hasta el momento.

Palabras Claves
Ingeniería de Software – Gobernanza Digital – Ciudades Inteligentes – Transformación Digital – Inteligencia Artificial

CONTEXTO

Esta línea de investigación se enmarca en el proyecto "Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital" (2023-2026), UNLP, Ministerio de Educación de la Nación, en particular el subproyecto "Ciudades inteligentes sostenibles. Gobernanza digital. Buenas prácticas y calidad". Y posee una estrecha interacción con los otros proyectos desarrollados en el Instituto que se relacionan con Inteligencia de Datos y Computación de Alto Desempeño y Distribuida.
Asimismo, la temática se encuentra abordada en diferentes proyectos aprobados por la Facultad de Informática UNLP (2024-2025-2026).
El Instituto posee diversos acuerdos de cooperación con varias Universidades de Argentina y del exterior y con empresas privadas del sector, interesadas en mejorar sus procesos de desarrollo.
Desde esta línea, el III-LIDI participa en proyectos internacionales que abordan los temas de Transformación Digital (TD), Ciudades Inteligentes Sostenibles (CIS) e Inteligencia Artificial (IA):
-Actualmente se trabaja en el proyecto *"UNI-UEAR - Alianza Universitaria Argentina Europea para la Transformación Digital"* cofinanciado por la Unión Europea. El Proyecto UNI UEAR es ejecutado por el Consejo Interuniversitario Nacional (CIN) en colaboración con un consorcio de cinco universidades nacionales UBA, UNLP, UNQ, UNS, UNSaM y cuatro instituciones europeas: European University Institute (Italia), Gdansk University of Technology (Polonia), Technical University of Delft (Países Bajos), University for Continuous Education Krems (Austria). Tiene como objetivo principal contribuir a la promoción de los derechos digitales en Argentina; y en particular, busca fortalecer la creación y difusión de conocimiento aplicado para la transformación digital; y

desarrollar alianzas entre la academia, los sectores público y privado y la sociedad civil para la transformación digital.
-Además, el Instituto posee una importante presencia en el "Consorcio de I+D+I en Cloud Computing-Big Data & Emergent Topics" (CC-BD&ET) con la participación de grupos de investigación de Argentina, España, Chile, entre otros.

1. INTRODUCCION

La línea de investigación se focaliza en 4 ejes. El primero basado en el concepto de la transformación digital (TD) y las ciudades inteligentes sostenibles (CIS), la generación de herramientas para capacitar a la comunidad en el desarrollo de CIS y en el proceso de la TD. El segundo profundiza en el uso de la tecnología en la gobernanza digital. El tercer eje se enfoca en distintos lineamientos referidos a la calidad y buenas prácticas del software, en particular los estándares relacionados con la generación y uso de modelos de IA. Por último, el cuarto eje, avanza en el uso de la inteligencia artificial (IA) en la Ingeniería de Software

Eje 1 – Transformación Digital y Ciudades inteligentes sostenibles

Desde este eje se participa en la generación de contenidos para diferentes cursos del proyecto UNI-UEAR y el dictado de la Maestría en Gestión y Tecnología de Ciudades Inteligentes, que se enfoca en la formación de recursos humanos capacitados para la gestión de ciudades inteligentes y sustentables, con un conocimiento del estado actual de la tecnología y su aplicación, en el contexto de una sociedad digital.

En particular se avanza en el uso de gemelos digitales en la producción agrícola que permiten gestionar la toma de decisiones en el uso de equipos de riego haciendo un uso racional del agua, y el modelo predictivo de residuos electrónicos desde la perspectiva de le ingeniería ambiental en Tesis de Doctorado

Eje 2- Gobernanza Digital y Gobierno Abierto

Desde este eje se analizan diferentes conceptos que buscan optimizar el uso de las TICs para mejorar la forma de gobierno, la entrega de servicios públicos y facilitar las interacciones entre el gobierno y el ciudadano.

En este caso podemos mencionar el trabajo en una Tesis Doctoral sobre el concepto de gobernanza universitaria, en el cual es importante medir el nivel de satisfacción del ciudadano universitario, en particular en el uso de herramientas digitales como SUDOKU, SIU, PILAGA, etc.

Eje 3- Calidad de la I.A.

El avance de la TD en particular el crecimiento en el uso de herramientas de IA requiere analizar desde diferentes perspectivas el nivel de calidad con el que se desarrollan y prestan los servicios que usan ese tipo de tecnología. En particular se propone investigar las normas y modelos de calidad en los sistemas de IA, según las normas ISO/IEC 25058 y 25059, que permiten evaluar la efectividad y transparencia de los sistemas de IA, como parte de una Tesina de Licenciatura.

Eje 4 - IA en Ingeniería de Software

La Inteligencia Artificial (IA) se ha fusionado progresivamente con las prácticas de ingeniería y desarrollo de software actualmente, la IA contribuye al desarrollo de software al automatizar largas tareas rutinarias, como la depuración y la documentación.
Se trabaja en las etapas de elicitación y análisis de requerimientos asistidos por IA en una Tesina de Licenciatura y se continua con el análisis de eficiencia en las etapas de implantación y desarrollo.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Análisis del impacto de la TD en los distintos sectores de la sociedad.
- Desarrollo de prototipos para Voto electrónico presencial y Voto por Internet. Arquitecturas adaptadas a la legislación vigente en distintos organismos.
- Análisis y estudio de normas y modelos orientados a la Calidad relacionados con IA
- Análisis, discusión y estudio de técnicas y herramientas de mejoras de proceso en el desarrollo de software.
- Análisis y discusión del uso de IA en el proceso de desarrollo del software.

3. RESULTADOS OBTENIDOS/ ESPERADOS

- Se inicia la segunda cohorte de la Maestría en Gestión y Tecnología de Ciudades Inteligentes (UNLP-UNS).
- Se avanza en el desarrollo de un modelo de evaluación de satisfacción del ciudadano universitario, con una tesis de doctorado en desarrollo en conjunto con la Facultad de Humanidades de la UNLP.
- Se avanza en el desarrollo de un modelo predictivo de residuos electrónicos desde la perspectiva de la ingeniería ambiental, con una tesis doctoral en curso.
- Se avanza en un modelo basado en gemelos digitales aplicados al sector agrícola, con el inicio de una tesis de doctorado.
- Se realizaron acciones de consultoría y asesoramiento en organismos públicos y privados, en particular se aplicó el Voto por Internet en distintos organismos.
- Se realizaron las elecciones de todos los claustros de la comunidad universitaria: Docentes, Estudiantes, Jefes de Trabajos Prácticos, Auxiliares Diplomados, Graduados y Nodocentes de la Facultad de Informática del 2025 utilizando los equipos de voto electrónico desarrollados recientemente.
- Se mantiene el Sistema de Gestión de la Calidad de la Facultad (SIGCFI).
-En el marco del proyecto UNIUEAR, el Instituto desarrollo y dictara los cursos:
    1. Tecnologías para la TD
    2. Gobernanza de datos
    3. Gestión de la Innovación
- Se analizan las posibilidades de aplicar IA en el proceso de desarrollo de software.

4. FORMACION DE RECURSOS HUMANOS

- Se dirigen tesis de postgrado, tesinas de grado y PPS en temas del área.
- Se dirige la Maestría en Gestión y Tecnología de Ciudades Inteligentes (UNLP-UNS).
- Se participa en el dictado de asignaturas/cursos de grado/postgrado en la Facultad de Informática de la UNLP y en otras universidades del país.

5. BIBLIOGRAFIA

[1] E. Estevez, “Innovación en Servicios Públicos Digitales.” Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance”. 2016

[2] ITU-T, “An overview of smart sustainable cities and the role of information and communication technologies,” 2014.

[3] R. Muñoz, A. Pasini, and P. Pesado, “Innovación en el Sector Público para Ciudades Inteligentes Sostenibles.” III Track de Gobierno Digital y Ciudades Inteligentes (TGDCI). XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020). La Matanza, Buenos Aires, Argentina. ISBN 978-987-4417-90-9. Pág. 797-806. Octubre 2020.

[4] C. Sánchez, Innovación pública: un modelo de aportación de valor. 2014.

[5] S. A. Chun, S. Shulman, R. Sandoval, and E. Hovy, “Government 2.0: Making connections between citizens, data and

government," Inf. Polity, vol. 15, no. 1–2, pp. 1–9, 2010.
[6] A. E. Millat, "Ciudades Resilientes Formando un Mundo Digital," 2020.
[7] S. Esponda, A. Pasini, M. Boracchia, J. Calabrese, R. Muñoz, J. Santiago Preisegger, and P. Pesado, "Fortalecimiento de la calidad en procesos de software y Gobernanza Digital," Sedici.Unlp.Edu.Ar, pp. 574–578.
[8] J. S. Preisegger, A. Greco, A. Pasini, M. Boracchia, and P. Pesado, "Marco de vinculación de datos abiertos aplicado al contexto de datos medioambientales," pp. 684–694.
[9] A. Naser and A. Ramirez, "Plan de gobierno abierto. Una hoja de ruta para los Gobiernos de la Región," CEPAL - Manuales, vol. 81, p. 80, 2017.
[10] A. Pasini, J. S. Preisegger, and P. Pesado, Open Government Assessment Models Applied to Province's Capital Cities in Argentina and Municipalities in the Province of Buenos Aires, vol. 995. 2019.
[11] C. Calderón and S. Lorenzo, Open Government. Gobierno Abierto. 2010.
[12] N. Noy, M. Burgess, and D. Brickley, "Google dataset search: Building a search engine for datasets in an open web ecosystem," in The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019, 2019, pp. 1365–1375.
[13] J. Calabrese, R. Muñoz, A. Pasini, S. Esponda, and M. Boracchia, "Asistente para la evaluación de características de calidad de producto de software propuestas por ISO / IEC 25010 basado en métricas definidas usando el enfoque GQM," pp. 660–671, 2017.
[14] A. Pasini, E. Estévez, P. Pesado, and M. Boracchia, "Una metodología para evaluar la madurez de servicios universitarios," Proc. Congr. XXII Congr. Argentino Ciencias la Comput. Congr. Argentino Ciencias la Comput., pp. 636–646, 2016.
[15] A. Pasini, J. S. Preisegger, and P. Pesado, "Modelos de evaluación de gobiernos abiertos, aplicado a los municipios de la provincia de Buenos Aires," XXIV Congr. Argentino Ciencias la Comput., vol. XXIV, pp. 0–10, 2018.
[16] Barenkamp, M., Rebstadt, J., & Thomas, O. (2020). Applications of AI in classical software engineering. AI Perspectives, 2, 1-15. https://doi.org/10.1186/s42467-020-00005-4.
[17] Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural engineering. Engineering Structures. https://doi.org/10.1016/J.ENGSTRUCT.2018.05.084.
[18] Necula, S. (2023). Artificial Intelligence Impact On The Labour Force - Searching For The Analytical Skills Of The Future Software Engineers. ArXiv, abs/2302.13229. https://doi.org/10.48550/arXiv.2302.13229

UN ENFOQUE UNIFICADO DE VERIFICACIÓN Y VALIDACIÓN DE SOFTWARE BASADO EN EXPRESIONES REGULARES

Martín L. Larrea[1,3,4], Dana K. Urribarri[1,3,4], M. Luján Ganuza[1,3,4], Matías N. Selzer[1,3,4], Antonella S. Antonini[1,3,4] , M. Laura Cobo[1]

{mll, dku, mlg, matias.selzer, antonella.antonini, mlc}@cs.uns.edu.ar

[1]Departamento de Ciencias e Ingeniería de la Computación

[3]Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (UNS-CIC)

[4]ICIC, Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

Universidad Nacional del Sur

RESUMEN

Este trabajo presenta una línea de investigación orientada a la Verificación y Validación de Software mediante el uso de Expresiones Regulares como mecanismo formal para la especificación y evaluación del comportamiento de sistemas informáticos. La propuesta aborda tanto actividades de verificación, vinculadas al testing de implementaciones de software, como actividades de validación, centradas en el análisis del comportamiento observable desde la perspectiva del usuario. La línea de investigación se apoya en una serie de trabajos publicados entre 2017 y 2024, en los que se exploran técnicas de testing de caja negra y caja blanca aplicadas a visualizaciones de información, sistemas orientados a objetos y software industrial. A lo largo de esta evolución, las Expresiones Regulares se utilizan para modelar secuencias de interacciones, intercambio de mensajes y ejecución de comportamientos esperados, permitiendo un enfoque unificado y flexible para el testing. Como resultado de esta línea de trabajo se desarrolló el framework de testing TAPIR [6], basado en la especificación de comportamiento mediante Expresiones Regulares y el uso de Programación Orientada a Aspectos. Actualmente, la investigación se encuentra en una nueva etapa orientada a la extensión del framework para soportar testing a nivel de clase en el paradigma orientado a objetos, mediante una estrategia modal implementada en Java. Esta línea de investigación contribuye al desarrollo de metodologías y herramientas que permiten evaluar la calidad del software en contextos complejos, promoviendo enfoques reutilizables, expresivos y alineados con las necesidades actuales de la Ingeniería de Software.

1. INTRODUCCIÓN

La Verificación y Validación de Software [1,2] constituyen actividades fundamentales dentro de la Ingeniería de Software, ya que permiten evaluar la calidad de los sistemas tanto desde el punto de vista de su correcta implementación como desde su comportamiento observable y su adecuación a las expectativas del usuario. A medida que los sistemas informáticos incrementan su complejidad, incorporan mayores niveles de interacción y asumen un rol central en la toma de decisiones, surge la necesidad de contar con técnicas de testing más expresivas, flexibles y reutilizables.

En este contexto, uno de los desafíos recurrentes es la especificación precisa del comportamiento esperado de un sistema. Tanto en actividades de verificación (orientadas al código y a su estructura interna) como de validación (centradas en el comportamiento externo del sistema), resulta necesario describir secuencias de eventos, interacciones o mensajes que deben ocurrir, o evitarse, durante la ejecución del software. Esta necesidad se manifiesta de manera particular en sistemas interactivos, como las visualizaciones de información, y en aplicaciones desarrolladas bajo el paradigma orientado a objetos.

Las Expresiones Regulares ofrecen un mecanismo formal y compacto para describir patrones y secuencias, y han sido

ampliamente utilizadas en distintos dominios de la informática. Sin embargo, su aplicación sistemática como lenguaje de especificación de comportamiento en actividades de Verificación y Validación de Software ha sido explorada de manera limitada. En particular, existe un potencial significativo en su uso para modelar tanto interacciones de usuario como secuencias de mensajes y ejecuciones de código, permitiendo un enfoque unificado que atraviesa distintas etapas y niveles del proceso de testing.

Este trabajo presenta una línea de investigación que aborda la Verificación y Validación de Software mediante el uso de Expresiones Regulares como herramienta central para la especificación y evaluación del comportamiento de los sistemas. La línea se desarrolla a partir de una serie de trabajos publicados entre 2017 y 2024, que proponen técnicas de testing de caja negra y caja blanca aplicadas a visualizaciones de información, sistemas orientados a objetos y software industrial. A lo largo de esta evolución, las Expresiones Regulares se emplean tanto para la validación de comportamientos observables desde la perspectiva del usuario como para la verificación de implementaciones a nivel de código.

Como resultado de esta línea de trabajo se desarrolló el framework de testing TAPIR [6], que utiliza Expresiones Regulares y Programación Orientada a Aspectos para instrumentar y evaluar el comportamiento de sistemas de software. Actualmente, la investigación se encuentra en una etapa de expansión, orientada al soporte de testing a nivel de clase en el paradigma orientado a objetos, mediante una estrategia modal implementada en Java.

El objetivo de este trabajo es resumir y contextualizar esta línea de investigación, destacando su evolución, sus aportes principales y las direcciones actuales y futuras de desarrollo, en el marco de las actividades de investigación en Verificación y Validación de Software.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación presentada se centra en el uso de Expresiones Regulares como mecanismo formal para la especificación y evaluación del comportamiento de sistemas de software en actividades de Verificación y Validación. El eje central de la propuesta consiste en modelar secuencias esperadas de eventos, interacciones o mensajes, y contrastarlas con el comportamiento observado durante la ejecución del sistema, permitiendo detectar desviaciones respecto de lo especificado.

Desde sus inicios, la línea ha abordado problemáticas vinculadas al testing de sistemas interactivos, en particular visualizaciones de información, donde el comportamiento esperado no puede describirse únicamente en términos de entradas y salidas tradicionales. En este contexto, se propusieron técnicas de testing de caja negra que utilizan Expresiones Regulares para modelar secuencias de interacciones de bajo nivel realizadas por los usuarios, permitiendo verificar restricciones de secuenciación y patrones de uso sobre interfaces gráficas. Este enfoque permitió avanzar en la validación del comportamiento observable del sistema desde la perspectiva del usuario.

A partir de estos primeros resultados, la investigación se extendió hacia el testing de caja blanca en sistemas desarrollados bajo el paradigma orientado a objetos. En esta etapa, las Expresiones Regulares comenzaron a utilizarse para especificar secuencias de mensajes intercambiados entre objetos, permitiendo evaluar el cumplimiento de comportamientos esperados a nivel de implementación. Para ello, se definieron mecanismos de instrumentación que posibilitan capturar eventos de ejecución sin modificar el código fuente, integrando conceptos de especificación de secuencias de mensajes con técnicas de Programación Orientada a Aspectos.

La consolidación de estos enfoques dio lugar al desarrollo del framework de testing TAPIR [6], que integra la especificación de comportamiento mediante Expresiones Regulares con mecanismos automáticos de monitoreo y evaluación de la ejecución. TAPIR permite abordar actividades de Verificación y Validación de manera unificada, soportando tanto el análisis de interacciones de usuario como el testing de implementaciones orientadas a objetos. La utilización de aspectos facilita la separación de las preocupaciones de testing respecto del código funcional, promoviendo soluciones reutilizables y extensibles.

En una etapa posterior, la línea de investigación amplió su alcance hacia contextos industriales y sistemas complejos, en particular aquellos asociados a la Industria 4.0. En este marco, se exploró la aplicabilidad de las técnicas propuestas a escenarios donde el software cumple un rol central en procesos productivos y de toma de decisiones. Asimismo, se avanzó en el uso de Expresiones Regulares para actividades de *runtime verification*, incluyendo implementaciones en lenguajes dinámicos como JavaScript, reforzando la flexibilidad del enfoque propuesto.

Actualmente, la línea se encuentra en una etapa de expansión orientada al testing a nivel de clase dentro del paradigma orientado a objetos. En particular, se está trabajando en una nueva versión del framework TAPIR que incorpora una estrategia modal para el Class Level Testing, implementada en Java. Esta evolución busca fortalecer la capacidad de verificación de factores internos de calidad, manteniendo la expresividad y reutilización que caracterizan al uso de Expresiones Regulares como lenguaje de especificación de comportamiento.

En conjunto, la evolución de esta línea de investigación refleja un enfoque progresivo y consistente, que parte del análisis de interacciones de usuario, avanza hacia la verificación de implementaciones y se proyecta hacia escenarios industriales y de mayor complejidad, manteniendo a las Expresiones Regulares como elemento unificador para la Verificación y Validación de Software.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como resultado principal de la línea de investigación presentada se desarrolló el framework de testing TAPIR[6], una herramienta orientada a la Verificación y Validación de Software basada en la especificación de comportamiento mediante Expresiones Regulares. TAPIR permite modelar secuencias esperadas de eventos, interacciones o mensajes, y contrastarlas con el comportamiento observado durante la ejecución del sistema, proporcionando un mecanismo flexible y expresivo para la detección de desvíos respecto de lo especificado.

Entre los resultados obtenidos se destaca la validación del enfoque en distintos contextos y niveles de testing. En particular, TAPIR ha sido aplicado tanto a escenarios de validación de software, como el testing de visualizaciones de información a partir de interacciones de usuario, como a escenarios de verificación de software, mediante el análisis de secuencias de mensajes y ejecuciones en sistemas orientados a objetos. Esta dualidad constituye uno de los aportes centrales de la línea, ya que demuestra la aplicabilidad de las Expresiones Regulares como un lenguaje unificador para actividades de testing de caja negra y caja blanca.

Otro resultado relevante es la integración del framework con técnicas de Programación Orientada a Aspectos, lo que permite instrumentar y monitorear la ejecución del software sin introducir modificaciones en el código funcional. Esta característica favorece la reutilización de especificaciones, la separación de preocupaciones y la adopción del enfoque en distintos tipos de proyectos y arquitecturas de software.

Asimismo, los trabajos realizados evidencian la extensibilidad del enfoque propuesto, permitiendo su adaptación a contextos industriales y a actividades de *runtime verification*, incluyendo implementaciones en distintos lenguajes y entornos de ejecución. Estos resultados refuerzan la viabilidad del uso de Expresiones Regulares como mecanismo de especificación de comportamiento en sistemas de mayor complejidad y con requerimientos de monitoreo en tiempo de ejecución.

Como trabajo en curso y principal resultado esperado, se encuentra el desarrollo de una nueva versión del framework TAPIR orientada al Class Level Testing dentro del paradigma orientado a objetos. Esta versión incorpora una estrategia modal, implementada en Java, que permite especificar y verificar comportamientos esperados a nivel de clase, fortaleciendo la evaluación de factores internos de calidad del software. Se espera que esta extensión amplíe el alcance del framework, facilitando la aplicación sistemática de técnicas de verificación en etapas tempranas del desarrollo.

Finalmente, se espera que los avances de esta línea de investigación continúen contribuyendo al desarrollo de metodologías y herramientas de Verificación y Validación de Software basadas en especificaciones formales de comportamiento, promoviendo enfoques más expresivos, reutilizables y alineados con las necesidades actuales de la Ingeniería de Software.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos en el transcurso del 2025 se desarrollaron dos tesis de grado de la Ingeniería en Sistemas de Información, así también como un proyecto de grupo de investigación “Verificación y Validación de Software en la Industria 5.0” (PGI 2024-2028 24/N059).

5. BIBLIOGRAFÍA

[1] Jena, Divya, et al. "Systematic Literature Review on Object Oriented Software Testing Techniques." 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA). IEEE, 2023.

[2] Binder, Robert V. "Testing object‐oriented software: a survey." Software Testing, Verification and Reliability 6.3‐4. 1996. 125-252.

[3] Larrea, Martín Leonardo. "Black-box testing technique for information visualization. Sequencing constraints with low-level interactions." Journal of Computer Science & Technology 17 (2017).

[4] Rodríguez Silva, Juan Ignacio, and Martín Leonardo Larrea. "White-Box Testing Framework for Object-Oriented Programming based on Message Sequence Specification." XXIV Congreso Argentino de Ciencias de la Computación (La Plata, 2018). 2018.

[5]. Urribarri, Dana K. "White-Box Testing Framework for Object-Oriented Programming. An Approach Based on Message Sequence Specification and Aspect Oriented Programming." CACIC 2018: 143.

[6] Larrea, Martín Leonardo, and Dana K. Urribarri. "TAPIR: An Object-Oriented Programming Testing Framework based on Message Sequence Specification with Aspect-Oriented Programming." XXVI Congreso Argentino de Ciencias de la Computacion (CACIC)(Modalidad virtual, 5 al 9 de octubre de 2020). 2020.

[7] Larrea, Martín Leonardo, and Dana K. Urribarri. "Expanding the scope of a testing framework for Industry 4.0." XXVII Congreso Argentino de Ciencias de la Computación (CACIC)(Modalidad virtual, 4 al 8 de octubre de 2021). 2021.

[8] Schiaffino, Martín, et al. "A testing tool for information visualizations based on user

interactions." Journal of Computer Science & Technology 22 (2022).

[9] Larrea, Martín L., and Dana K. Urribarri. "New verification and validation tools for industry 4.0 software." Designing Smart Manufacturing Systems. Academic Press, 2023. 61-88.

[10] Larrea, Martín Leonardo, Dana Karina Urribarri, and María Luján Ganuza. "New Testing Techniques to Evaluate the Quality of Information Visualization Implementations." LACCEI 1.8 (2023).

[11]. Pacione, Luciano, et al. "Runtime Verification of JavaScript Implementations Using Aspect Oriented Programming and Regular Expressions." 2024 L Latin American Computer Conference (CLEI). IEEE, 2024.

Un Estudio Comparativo de Especificaciones de Requisitos de Calidad de Software generadas con Copilotos de IA

Carrizo Claudio*, Javier Saldarini*, Juan Pablo Mansilla*, Nicolás Fenoglio, Carlos Salgado+, Alberto Sánchez+, Mario Peralta+
* Facultad Regional San Francisco – Universidad Tecnológica Nacional
Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
{cjcarrizo77, saldarinijavier, juampimansilla17, nicolasfenoglio2004}@gmail.com
+ Departamento de Informática - Facultad de Ciencias Físico-Matemáticas y Naturales Universidad Nacional de San Luis
Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
{csalgado, alfanego, mperalta}@unsl.edu.ar

RESUMEN

El presente trabajo se enmarca en la línea de investigación orientada a integrar IA Generativa en procesos ágiles de desarrollo, con el objetivo de dar soporte a la especificación de requisitos de calidad mediante un enfoque sistemático de ingeniería de prompt alineado con la familia de normas ISO/IEC 25000.

En el presente trabajo se presentan los avances sobre el estudio comparativo de especificaciones de requisitos de calidad del software, generadas mediante cuatro copilotos de Inteligencia Artificial Generativa (ChatGPT, Gemini, Copilot y Deepseek), aplicando una guía estructurada, basada en el proceso de Análisis y Definición de Requisitos de Calidad de ISO/IEC 25030.

Palabras clave: inteligencia artificial generativa, requisitos de calidad del software, ISO/IEC 25000, metodologías ágiles.

CONTEXTO

La presente línea de investigación se enmarca dentro del Proyecto de Investigación y Desarrollo UTN que lleva por título: *"Aplicación de IA Generativa para la Especificación de Requisitos de Calidad en el Desarrollo Ágil de Software conforme a ISO/IEC 25000"*.

El mencionado proyecto fue elaborado y está siendo ejecutado de manera conjunta entre los integrantes del Laboratorio de Calidad e Ingeniería de Software (LaCIS) – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis, y los integrantes del Grupo de I+D denominado "Calidad del Software", perteneciente a la Universidad Tecnológica Nacional, Facultad Regional San Francisco.

El proyecto se encuentra homologado y financiado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, bajo el código SFSIECI876 y según la Disp. SCTyP Nº 45/2025

1. INTRODUCCIÓN

La Inteligencia Artificial Generativa se está consolidando como una herramienta de apoyo en el ámbito del desarrollo de software. No obstante, su implementación plantea múltiples desafíos e interrogantes, especialmente en lo referido a la generación de datos e información que sean precisos y confiables dentro de dominios de aplicación específicos (Tsoftglobal, 2024).

La IA Generativa (GenAI) es capaz de producir contenido en lenguaje natural y

se posiciona como una herramienta prometedora en el soporte a tareas complejas del ciclo de vida del software. Entre sus aplicaciones más destacadas, los Modelos de Lenguaje Extenso (LLM), como, por ejemplo, ChatGPT, GitHub Copilot, entre otros, permiten generar código, diseñar interfaces y asistir en la redacción y validación de requisitos, especialmente mediante el uso de prompts, los cuales deben ser adecuados para cada contexto de aplicación.

Un prompt es un texto que describe un conjunto de instrucciones que personalizan, refinan o mejoran las capacidades de un modelo GPT (Generative Pre-Trained Transformer) (Liu, 2023).

Según (Lo, 2023) los prompts efectivos se caracterizan por los siguientes principios fundamentales: claridad y precisión, información contextual, formato deseado y control de verbosidad. Sin embargo, escribir prompts efectivos puede resultar complicado para usuarios sin conocimientos técnicos, ya que requiere creatividad, intuición y refinamiento iterativo. En consecuencia, las respuestas pueden ser vagas, imprecisas, incorrectas en cuanto a los hechos o inapropiadas para el contexto cuando se utiliza una indicación inadecuada (Jha, 2023).

En este contexto, la ingeniería de prompt surge como una disciplina para diseñar indicaciones mediante las cuales los usuarios pueden programar un Modelo de Lenguaje Grande, como ChatGPT, para producir respuestas precisas (Giray, 2023).

Por otra parte, la calidad en el desarrollo de software continúa siendo un campo en constante evolución. En particular, la especificación cuantitativa y objetiva de los requisitos de calidad sigue siendo una práctica poco frecuente en los proyectos de software (Sommerville, 2011), situación que se evidencia aún más en aquellos que son desarrolados bajo metodologías ágiles (Woubshet, y otros, 2020).

En este contexto, surge la necesidad de desarrollar una propuesta que brinde soporte al proceso de especificación de requisitos de calidad del software, mediante el uso de un copiloto basado en Inteligencia Artificial Generativa (Tsoftglobal, 2024), guiado por prompts efectivos (Lo, 2023) (Jha, 2023) y (Giray, 2023), específicamente diseñados para entornos de desarrollo conducidos por el marco Scrum (Scrum Alliance, s.f.), y tomando como referencia el marco normativo propuesto por la familia de normas ISO/IEC 25000 (ISO/IEC 25000:2014, 2014).

Cabe mencionar que el presente trabajo es la continuación la línea de investigación presentada en WICC 2025 (Saldarini, Carrizo, Mansilla, Salgado, & Sánchez, 2025).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La presente línea de investigación se compone de un conjunto de ejes de trabajo, los cuales se detallan a continuación:

1- **Elaboración del Estado del arte:** elaborar un estado del arte sobre la aplicabilidad de IA Generativa en los

procesos de especificación de requisitos de calidad del software, en entornos de desarrollos ágiles, y utilizando estándares de calidad aplicados al software.

2- **Desarrollo de una Guía para redactar Prompt:** desarrollar una guía basada en el proceso de Análisis y Definición de RC del software propuesto por ISO/IEC 25030 (ISO/IEC 25030:2007, 2007), la cual esté compuesta por un conjunto de pasos, donde cada uno de ellos contendrá un guion para ser utilizado como Prompt, el cual permitirá interactuar con un agente de IA Generativa. Para desarrollar la mencionada Guía, se utilizará: un modelo para historias de usuario y criterios de aceptación; el modelo de calidad de producto definido por ISO/IEC 25010 (ISO/IEC 25010:2011, 2011); y las métricas de calidad, las cuales se obtendrán a través de la norma ISO/IEC 25023 (ISO/IEC 25023:2016, 2016 (E))
3- **Selección de un Agente de IA:** en función de los resultados del Eje 1, se deberán seleccionar al menos 3 agentes de IA Generativa, los cuales puedan actuar como Copiloto, para la aplicación de la guía desarrollada en el Eje 2. La precisión y coherencia que arrojen los resultados obtenidos por cada Copiloto de IA en cada paso de la guía, se compararán en función de trazabilidad con el estándar ISO/IEC 25000 (ISO/IEC 25000:2014, 2014), la historia de usuario de partida y sus criterios de aceptación.
4- **Ajustes de la Guía:** con la experiencia obtenida en el Eje 3, y habiendo analizado los resultados, se realizará un ajuste sobre la Guía desarrollada en el Eje 2.
5- **Prueba de Concepto:** se pretende instanciar la Guía desarrollada en un caso real, a través del uso del Copiloto de IA Generativa seleccionado como resultado de las actividades llevadas a cabo en el Eje 3. Para ello, trabajaremos en conjunto con un Analista QA, sobre un proyecto real basado en el desarrollo de un sistema de gestión de emergencias y urgencias médicas. Como este proyecto se encuentra en desarrollo, se partirá de un conjunto de historias usuarios que ya hayan sido desarrolladas e implementadas, esto nos dará los puntos de comparación sobre las especificaciones de RC que se documentaron originalmente, versus las arrojadas con la implementación de la Guía, haciendo uso de un Copiloto de IA generativa.
6- **Análisis, Discusión y Publicación de los resultados obtenidos:** finalmente se llevará a cabo el análisis y discusión de los resultados; se pretende también publicar al menos un trabajo en una revista o congreso de la especialidad.

3. RESULTADOS

Resultados Obtenidos

En el trabajo presentado en WICC 2025 (Saldarini, Carrizo, Mansilla, Salgado, & Sánchez, 2025) se detallan los primeros resultados de la presente línea de investigación, entre ellos, se destaca el desarrollo de una guía basada en el proceso de Análisis y Definición de RC del software propuesto por ISO/IEC 25030 (ISO/IEC 25030:2007, 2007), la cual está compuesta por un conjunto de pasos, donde cada uno de ellos contendrá un guión para ser utilizado como Prompt, el cual permitirá interactuar con un agente de IA Generativa, y cuya aplicación tendrá como resultado una especificación de requisitos de calidad del software para una historia de usuario.

Específicamente en el presente trabajo se muestra como principal resultado la aplicación de la mencionada Guía en 4 Agentes de IA Generativa (Chat GPT, Gemini, Deepseek y Copilot). La precisión y

coherencia que arrojen los resultados obtenidos por cada Copiloto de IA en cada paso de la guía, se compararán en función de trazabilidad con el estándar ISO/IEC 25000 (ISO/IEC 25000:2014, 2014), la historia de usuario de partida y sus criterios de aceptación.

La aplicación de la guía permitió instanciar los seis pasos para cada Copiloto de IA a través de los Prompt propuestos y para cada paso se fueron registraron los resultados devueltos por cada Copiloto, llegando al final del proceso con cuatro especificaciones de requisitos de calidad del software para la Historia de usuario seleccionada, una por cada Copiloto utilizado.

Si bien cada uno de los Copilotos llegó a la obtención de una especificación de requisitos de calidad con el formato Característica/Sub-característica, los resultados a nivel de cantidad identificada por cada Copiloto no fueron iguales (Ver Tabla 1).

Tabla1-Comparación Copilotos IA

Copiloto IA	Característica	Subcaracterística
Chat GPT	5	7
Gemini	3	6
Deepseek	5	8
Copilot	5	6

Como próximos pasos se plantea llevar a cabo la evaluación con respecto a la precisión y coherencia de los resultados obtenidos por cada Copiloto de IA en cada paso de la guía, en función de trazabilidad con el estándar ISO/IEC 25000 (ISO/IEC 25000:2014, 2014), también se pretende ampliar las pruebas con más historias de usuarios, y llevar a cabo el ajuste a la Guía respecto de la experiencia llevada a cabo.

FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación se trabaja desde el Grupo I+D "Calidad de Software", en colaboración con docentes investigadores formados de la Universidad Nacional de San Luis. El equipo de trabajo se encuentra conformado por 5 docentes investigadores formados, 2 investigadores en formación, 1 graduado y 2 estudiantes de grado.

En el contexto de la ejecución del proyecto, en donde se inserta la presente línea de investigación, se está desarrollando un trabajo final de carrera de un estudiante de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

Extreme Programming. (s.f.). Obtenido de https://www.agilealliance.org/glossary/xp

Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. Ann Biomed Eng, 2629–2633. doi:https://doi.org/10.1007/s10439-023-03272-4

github. (s.f.). github. Obtenido de https://github.com/features/copilot

ISO/IEC 25000:2014. (2014). Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE).

ISO/IEC 25010:2011. (2011). Systems and software engineering-Systems and software Quality Requirements and

Evaluation (SQuaRE)-System and software quality models. ISO.

ISO/IEC 25023:2016. (2016 (E)). Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of system and software product quality.

ISO/IEC 25030:2007. (2007). Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Quality requirements. ISO.

Jha, S. J. (2023). Dehallucinating large language models using formal methods guided iterative prompting. IEEE International Conference on Assured Autonomy (ICAA), 149–152. doi:https://doi.org/10.1109/ICAA58325.2023.00029

Liu, P. Y. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM Computing Surveys, 1-35. doi:https://doi.org/10.1145/3560815

Lo, L. S. (2023). El arte y la ciencia de la ingeniería rápida: una nueva alfabetización en la era de la información. Internet Ref. Serv. Q.

Manifesto for Agile Software Development. (s.f.). Obtenido de http://agilemanifesto.org/

Ocaña, L., Naranjo, F., & Fausto, A. (2023). Estudio neutrosófico sobre la afectación de la. Neutrosophic Computing and Machine Learning, 296-304.

openai. (s.f.). openai. Obtenido de https://openai.com/chatgpt/

Pressman, R. (2005). Ingeniería de Software. (6 ed.). Mcgraw-Hill.

Saldarini, J., Carrizo, C., Mansilla, J., Salgado, C., & Sánchez, A. (2025). Especificación de requisitos de calidad de software en historias de usuario: una propuesta aplicando IA generativa como soporte. WICC 2025, (págs. 316-320).

Scrum Alliance. (s.f.). Obtenido de https://resources.scrumalliance.org/Article/quick-guide-things-scrum

Sommerville, I. (2011). Ingeniería de Software (9° ed.). México: PEARSON EDUCACIÓN.

Tsoftglobal. (09 de 2024). Tsoftglobal. Obtenido de https://lc.cx/MxaAFU

Woubshet, B., Pertti, K., Lidia, L., Xavier, B., Silverio, M.-F., Vollmer, A. M., . . . Oivo, M. (2020). Management of quality requirements in agile and rapid software develop,ment: A systematic mapping study. Information and Software Technology, 123, 106-225.

Uso de técnicas de Ludificación en la gestión de proyectos de sistemas de información

Patricia R. Cristaldo, Lucía I. Thea, Luciana G. Valiente
Grupo de Investigación en Bases de Datos, Departamento Ingeniería en Sistemas de Información,
Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional
Entre Ríos, Argentina
{cristaldop, theal, valientel}@frcu.utn.edu.ar

Daniela López de Luise
CI2S Labs- Computational Intelligence & Information Systems Lab, Buenos Aires
daniela_ldl@ieee.org

Resumen

Numerosos avances en el área de Ingeniería de Software, desde hace años, han conducido a la producción de nuevas tecnologías, modelos, métodos y técnicas que apuntan a mejorar los resultados en la gestión de proyectos. A pesar de los avances logrados hasta el momento el desafío aun perdura. En consecuencia, la motivación y la disciplina se han convertido en elementos aún más cruciales para un buen desarrollo de un proyecto de software. Una forma de introducir y mantener estos dos elementos clave en las actividades consideradas (a veces tediosas y no desafiantes) es el uso de la ludificación. Si bien el término "ludificación" fue creado hace un poco más de dos décadas, tuvo interés generalizado recién a partir de la década del 2010, donde los juegos tuvieron aplicaciones en contextos donde los usuarios los utilizaban para obtener comportamientos deseados. Ha surgido como un fenómeno y está cada vez más presente en la vida cotidiana de las personas, haciendo uso de elementos y mecánicas lúdicas para estimular el comportamiento, mejorar la motivación de las personas y el compromiso en sus tareas en diferentes contextos. Por ello. hoy en día son reconocidos los denominados *serious games*, extensiones ludificadas con objetivos que exceden la recreación. El presente proyecto de investigación propone definir, diseñar e implementar un entorno ludificado general basado en un marco de medición que integra las métricas propuestas en el proyecto que le antecede y que ha generado varias publicaciones en la comunidad. Esto permitirá la realización de las tareas relacionadas con la gestión de un proyecto en forma más amena, ya que logra evaluar numéricamente el estado de los documentos antes de iniciar las actividades de gestión del proyecto. A partir de allí, poder determinar qué metodología de gestión de proyectos es aplicable de acuerdo a la eficacia y la eficiencia del contexto particular. Otra de las ventajas radica en, que el proceso de captura automática y sistemática ludificada es pasible de complementarse con módulos de ciencia de datos que permitan ajustar el propio funcionamiento y establecer un observatorio de la temática a partir de estadísticas y modelos de datos.

En este artículo se presentan los tópicos de interés del proyecto *Uso de técnicas de Ludificación en la gestión de proyectos de sistemas de información.*

Palabras clave: Gestión de proyectos de software, métricas de gestión, ludificación, procesamiento de lenguaje natural.

Contexto

El presente trabajo se desarrolla en el ámbito del proyecto *Uso de técnicas de Ludificación en la gestión de proyectos de sistemas de información (LUDGEPI) (CUSIEC542)* del Grupo de Investigación en Bases de Datos, perteneciente al Departamento Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Concepción del Uruguay.

1. Introducción

La gestión de proyectos de software involucra conocimientos científicos y habilidades gerenciales, combinando aptitudes técnicas y ejecutivas. Esto implica desempeñar diversas funciones, como liderazgo, planificación, seguimiento y control de actividades, administración de requisitos del pr0oyecto y consideraciones empresariales. En este sentido, la dirección de proyectos debe abarcar tanto las actividades de gestión como las funciones tecnológicas. Existen numerosas tecnologías y guías disponibles en la comunidad que proporcionan orientación en este campo, incluyendo PMBOK [1], PRINCE2 [2][3], APM [4], ISO 21500 [5], SCRUM [6][7], KANBAN [8], CRISP-DM [9] [10]. El objetivo primordial de una gestión de proyectos efectiva es lograr un producto exitoso dentro de los plazos y costos previstos [11]. Según el informe CHAOS de 2018, solo el 29% de los proyectos logró cumplir con los plazos, el presupuesto y los requisitos de funcionalidad. Aproximadamente el 37% no cumplió con al menos una de estas dimensiones clave. El informe señala que casi el 52% de todos los proyectos sufrió retrasos, sobrecostos o no cumplió con todos los requisitos [12]. Estas estadísticas representan un aumento del 10% con respecto a las cifras de 2010. Además, el 19% de los proyectos se completaron sin entregar un producto funcional. Los proyectos de software requieren una gestión precisa tanto de los aspectos técnicos como de los gerenciales para tener éxito. A pesar de la disponibilidad de diversas metodologías, aún persisten desafíos en el camino hacia la ejecución efectiva de proyectos dentro de los límites de tiempo y costos establecidos, así como las expectativas de los usuarios. Entre las razones se incluyen una planificación deficiente, falta de adaptación de los cambios del negocio, falta de utilización de metodologías de gestión de proyectos [13][14][15], definición inadecuada de requisitos [16][17][18], falta de habilidades, problemas de disciplina en la gestión y liderazgo insuficiente [19].

El término “ludificación” fue creado en 2002 por el diseñador de juegos Nick Pelling [20] [21] pero obtuvo un interés generalizado recién en la década del 2010, donde los juegos tuvieron aplicaciones en contextos donde los usuarios los utilizaban para obtener comportamientos deseados [22]. Las contribuciones a la literatura sobre ludificación se han centrado particularmente en la educación y el aprendizaje [23] [24] [25] la salud [26] [27] el marketing [28] [29] y en comportamientos sociales [30]. Asimismo, a pesar de la creciente evolución, la ludificación es todavía un concepto creciente con orígenes alrededor de 2006 [31] [32], cobrando impulso en la industria hacia 2015 [33] [34] y estableciéndose hace unos pocos años. Es más que un proceso técnico de aplicar elementos y herramientas del juego que proporcionen diversión y disfrute [35]. Genera diversas dinámicas de juego, como recompensas, competencia, altruismo y autoexpresión, apoyando la satisfacción de las necesidades psicológicas de las personas, tales como, autonomía, competencia y relaciones interpersonales [36] [37]. La selección de los elementos del juego, como las recompensas y

el nivel de competencia, aporta a que las personas se mantengan motivadas y comprometidas [38].

En cuanto a la gestión de equipos, la ludificación juega un papel clave en la coordinación de los miembros del equipo y los altos directivos. Varios estudios sugieren el uso de enfoques de juegos para apoyar el proceso de innovación y la colaboración entre equipos. Puede aplicarse como una práctica de innovación colaborativa que genera oportunidades para apoyar la creatividad colectiva y ayudar a las partes en conflicto a alcanzar un consenso [38]. Por otro lado, permite una autogestión en un equipo con una distribución de poder más equilibrada que brinde igualdad de oportunidades a todos sus miembros.

Asimismo, otros autores [39], ilustran las ventajas de implementar componentes de juegos en los proyectos de gestión del conocimiento, permitiendo entre otras cosas, garantizar un adecuado manejo de este recurso, atendiendo aspectos centrales como la motivación humana y la voluntad individual de compartir el conocimiento. La ludificación ha demostrado ser un enfoque factible para aumentar la motivación de los empleados.

Por otro lado, un reciente estudio [40] establece que la visión lúdica ofrece una de las soluciones más prometedoras para que los gerentes de proyectos de tecnología de la información innoven sus rendimientos. Sin embargo, la evidencia sobre sus ventajas es aún muy limitada. Por lo que este autor, muestra los beneficios de la concepción de videojuegos para el desempeño general del trabajo y los mecanismos de intermediación a través de los cuales ejerce sus efectos ventajosos. La participación de las personas en este tipo de iniciativas altera favorablemente sus cogniciones, transfiriendo su efecto sobre las actitudes y comportamientos en el lugar de trabajo.

Este proyecto se alinea con el desarrollo de métricas orientadas a la evaluación del proceso de aplicación de modelos de gestión de proyectos. Se busca generar una métrica integral para la gestión de proyectos, se proponen nuevas métricas de carácter original, basadas en técnicas de procesamiento de lenguaje natural, que permiten traducir y analizar los requerimientos textuales necesarios para la gestión de los proyectos. [41] [42] [43].

2. Líneas de Investigación, Desarrollo e Innovación

La línea de trabajo principal de nuestro proyecto de investigación es definir, diseñar, implementar y validar un entorno ludificado basado en un marco de medición que integra un conjunto de métricas para la determinación de aplicabilidad y evaluación transversal de metodologías de gestión de proyectos, a partir de expresiones textuales de la descripción de un proyecto, a partir de un trabajo previo [44]. Esto implica:

- Estudiar el comportamiento de diversas técnicas de ludificación y que resulten significativas en la gestión de proyectos.
- Definir y diseñar una plataforma ludificada asentada a un marco de medición que contiene un conjunto de métricas e indicadores aplicados a la gestión completa de un proyecto: alcance, tiempo, costos, calidad, recursos, interesados, comunicaciones, riesgos y adquisiciones.
- Validar la plataforma ludificada en diversos proyectos de contextos diferentes a fin de ampliar la muestra de estudio para generalizar los

hallazgos. Es decir, el propósito de aplicar gamificación en el campo de la gestión de proyectos está directamente relacionado con mejorar la participación de los interesados en el proyecto.

- Establecer un trabajo colaborativo para reotroalimentar el aprendizaje, a partir de la implementación de un entorno lúdico, aprovechando el alto potencial para mejorar los procesos vinculados con la gestión de proyectos.
- Generar modelos de comportamiento predictivo y perfiles de funcionamiento.
- Generar y socializar mejores prácticas en gestión de proyectos de sistemas de información, aportando a los factores de éxito en la gestión de proyectos.
- Capacitar e integrar un grupo de docentes y alumnos en el *know how* específico para trabajar e investigar en la temática.

3. Resultados obtenidos y esperados

Con este proyecto se pretende desarrollar e implementar un entorno ludificado a un marco de medición que permita determinar la aplicabilidad y realizar una evaluación transversal de metodologías de gestión de proyectos, a partir del análisis de expresiones textuales presentes en la descripción de los proyectos. En este contexto, el proyecto ha avanzado y presentado algunos resultados preliminares en [45], donde se muestran la aplicación de técnicas de ludificación que ayudan a la descripción del alcance de un proyecto, y donde es posible generar un documento técnicamente de calidad. Estos primeros resultados admiten confirmar que la ludificación facilita la adquisición de suficiente información, motiva a los interesados, y permite completar un documento formal a través de un juego.

4. Formación de Recursos Humanos

Este proyecto dio origen a una nueva línea de investigación dentro del Grupo de Investigación en Bases de Datos de la Facultad Regional Concepción del Uruguay de la Universidad Tecnológica Nacional. En el marco del mismo, una investigadora ha desarrollado su tesis doctoral y dos investigadoras se encuentran desarrollando sus trabajos de especialización.

Asimismo, el proyecto cuenta con la colaboración de dos becarios graduados con beca de iniciación a la investigación, quienes tienen previsto continuar su formación mediante estudios de posgrado en el área temática abordada. Además, participan cuatro becarios alumnos de la carrera Ingeniería en Sistemas de Información, que inician su formación en actividades de investigación.
Además, se incorporan estudiantes de grado para el desarrollo de trabajos de práctica supervisada, fortaleciendo así la formación académica y la integración temprana a la investigación científica.

5. Referencias

[1] Pmbok Guide (2021) A Guide to the Project Management Body of Knowledge. 7 Ed. Project Management Institute. ISBN: 978-1628256796.

[2] Office of Government Commerce (2009) An introduction to PRINCE2: managing and directing successful projects. Stationery Office, 123 p. ISBN-10: 0113311885.

[3] A. Böhm (2009) Application of PRINCE2 and the Impact on Project Management. ISBN 978-3-640-42634-8.

[4] J. Highsmith (2010) Agile project management: creating innovative products. 2 Ed. Boston, MA: Addison-Wesley. 432 p.
[5] ISO (2012) ISO 21500:2012 Guidance on Project Management, Geneva: ISO.
[6] J. Sutherland (2014) Scrum: The art of doing twice the work in half the time. New York: Crown Business. 256 p. ISBN-10: 038534645X, ISBN-13: 978-0385346450.
[7] R. Van Solingen, R. Van Lanen (2014) Scrum voor Managers. Den Haag: Academic Service. EAN: 9789012585903.
[8] H. Lei, F. Ganjeizadeh, P. Jayachandran, P. Ozcan (2015) Statistical analysis of the effects of Scrum and Kanban on software projects. Mz. onRobotics & Computer-Integrated Manufacturing.
[9] C. Shearer (2000) The CRISP-DM model: The new blueprint for data mining. Journal of Data Warehousing, no. 4, pp. 13-22.
[10] U. Shafique., H. Qaiser (2014) A Comparative Study of Process Models Data Mining). International Journal of Innovation and Scientific Research, pp.217-22.
[11] J. Varajão, C. Dominguez, P. Ribeiro, A. Paiva (2014) Critical success aspects in Project management. Tech. Gazette 21 (3) 583–589.
[12] TSG (2018) The CHAOS Report. The Standish Group, 2018. Disponible en https://secure.standishgroup.com/reports/flyers/CM2018- TOC.pdf.
[13] A. La Paz, R. López (2023). Recommendation method for customized IT project management. https://doi.org/10.1016/j.procs.2023.01.493.
[14] M. Haleem, F. Farooqui, M. Faisal. (2019). A Critical Analysis of Software Product Failure: An Indian & Global Perspective. International Journal of Engineering and Advanced Technology. https://doi.org/10.35940/ijeat.C5941.088619.
[15] W. Sardjono, A. Retnowardhani, (2019). Analysis of Failure Factors in Information Systems Project for Software Implemetation at the organization. Proceedings of 2019 International Conference on Information Management and Technology, pp.141-145, https://doi.org/10.1109/ICIMTech.2019.8843725.
[16] X. Tao, P. Robson, C. Wang. (2023). To learn or not to learn from new product development project failure: The roles of failure experience and error orientation. https://doi.org/10.1016/j.technovation.2023.102830.
[17] T. Chow, D. Chao (2008) A survey of critical success factors in agile software projects. Journal of systems and software, Jun. Available: Science Direct, pp.961–971.
[18] H. Elkadi (2013) Success and failure factors for e-government projects: A case from Egypt. Egyptian Informatics Journal, pp. 165–173.
[19] J. Westenberger, K. Schuler, D. Schlegel. (2022). Failure of AI projects: understanding the critical factors. https://doi.org/10.1016/j.procs.2021.11.074.
[20] G. Galetta. (2013). The gamification: applications and developments for creativity and education. https://doi.org/10.13140/RG.2.2.24817.68965.
[21] R. Patricio, A. Carrizo Moreira (2022). Zurlo F. Gamification in innovation teams. https://doi.org/10.1016/j.ijis.2022.05.003.
[22] S. Deterding, D. Dixon, R. Khaled, L. Nacke. (2011). From game design elements to gamefulness: defining

"Gamification". https://doi.org/10.1145/2181037.2181040

[23] Y. Attali, M. Arieli-Attali. (2015). Gamification in assessment: do points affect test performance? https://doi.org/10.1016/j. compedu.2014.12.012.

[24] D. Gibson, N. Ostashewski, K. Flintoff, S. Grant, E. Knight. (2015). Digital badges in education. https://doi.org/10.1007/s10639-013-9291-7.

[25] A. Toda, P. Valle, S. Isotani (2017). The dark side of gamification: An overview of negative effects of gamification in education, in: HEFA 2017, Higher Education for All. https://doi.org/10.1007/978-3-319-97934-2_9.

[26] S. McKeown, C. Krause, M. Shergill, A. Siu, D. Sweet (2016). Gamification as a strategy to engage and motivate clinicians to improve care. Healthc. https://doi.org/10.1177/0840470415626528.

[27] M. Regalado Chamorro, A. Medina Gamero, M. Hinojosa Núñez, A. Rodríguez Pérez. (2021). Gamification in health: A change in health education. https://doi.org/10.1016/j.appr.2021.100102.

[28] L. Gucassen, S. Jansen. (2014). Gamification in consumer marketing - future or fallacy? https://doi.org/10.1016/j. sbspro.2014.07.034, 2011.

[29] F. Noorbehbahani, F. Salehi, R. Jafar Zadeh. (2019). A systematic mapping study on gamification applied to e-marketing. https://doi.org/10.1108/JRIM-08-2018-0103.

[30] D. Schoech, J. Boyas, B. Black, N. Elias-Lambert, F. Boyas, M. Black (2013). Gamification for behavior change: lessons from developing a social, multiuser, web-tablet based prevention game for youths. https://doi.org/10.1080/15228835.2013.812512.

[31] J. Rice. (2007). "Assessing higher order thinking in video games" (PDF). Journal of Technology and Teacher Education. 15 (1): 87.

[32] D. Djaouti, Damien, J. Alvarez, J. Jessel, O. Rampnoux. (2011). "Origins of serious games". Serious Games and Edutainment Applications. Springer: 25–43. doi:10.1007/978-1-4471-2161-9_3. ISBN 978-1-4471-2160-2.

[33] "Serious Games". (2015). cs.gmu.edu. Retrieved 26 June 2015.

[34] A. Lugmayr, J. Suhonen, H. Hlavacs, C. Montero, E. Suutinen, C. Sedano (2016). "Serious storytelling - a first definition and review". Multimedia Tools and Applications. 76 (14): 15707–15733. doi:10.1007/s11042-016-3865-5. S2CID 207219982.

[35] A. Suh, C. Wagner, L. Liu. (2018). Enhancing user engagement through gamification. https://doi.org/10.1080/08874417.2016.1229143.

[36] R. Mitchell, L. Schuster, H. Jin. (2020). Gamification and the impact of extrinsic motivation on needs satisfaction: making work fun?. https://doi.org/10.1016/j.jbusres.2018.11.022.

[37] M. Suárez-López, A. Blanco-Marigorta, A. Gutiérrez-Trashorras. (2023). Gamification in thermal engineering: Does it encourage motivación and learning? https://doi.org/10.1016/j.ece.2023.07.006

[38] D. de Paula Porto, G. Martins de Jesus, F. Cutigi Ferrari, S. Camargo Pinto Ferraz Fabbri. (2020). Initiatives and challenges of using gamification in software engineering: A Systematic

Mapping. https://doi.org/10.1016/j.jss.2020.110870.

[39] J. Friedrich, M. Becker, F. Kramer, M. Wirth, M. Schneider (2019). Incentive design and gamification for knowledge management. https://doi.org/10.1016/j.jbusres.2019.02.009.

[40] L. Bizzi (2023). Why to gamify performance management? Consequences of user engagement in gamification. https://doi.org/10.1016/j.im.2023.103762.

[41] P. Cristaldo, D. López De Luise, L. La Pietra, A. De Battista, J. Hemanth. (2021). Metrics for the Systematic Evaluation of Project Management Methodologies**.** International Journal of Service Science, Management, Engineering, and Technology (IJSSMET).

[42] P. Cristaldo, D. López De Luise, L. La Pietra, A. De Battista. (2021). Metrics for validation and traceability of Project Management Requirements. Global Research and Development Journal for Engineering (GRDJE) ISSN (online): 2455-5703.

[43] P. Cristaldo, D. López De Luise, L. La Pietra, A. De Battista. (2021). Adjustment Level Metrics for Requirements in Project Management. Proceedings of 2021 8th Intl. Conference on Soft Computing & Machine Intelligence (ISCMI 2021), IEEE ISSN: 2640-0146, ISBN: 978-1-7281-8682-5, IEEE Catalog Number: CFP2144Z-USB, Cairo, Egypt.

[44] P. Cristaldo (2024). Formulación de métricas para la evaluación integral de metodologías de gestión de proyectos. https://doi.org/10.35537/10915/167563.

[45] P. Cristaldo, D. López De Luise, L. Thea , L. Valiente. Ludificación en la Captura de Métricas para la Gestión de Proyectos. 13° Congreso Nacional de Ingeniería en Sistemas de Información (CONAIISI 2025) Córdoba, Argentina.

USO DEL ANÁLISIS DE TEXTO ESTRUCTURADO Y NO ESTRUCTURADO PARA LA GENERACIÓN DE MATERIAL DE ENTRENAMIENTO PARA MODELOS DE IA

Nicolás Hernández, Julio Castillo, Marina Cardenas, Sandra Olariaga, Matias Negrelli, Sofia Britos, Tito Santino
Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/Facultad Regional Córdoba/Universidad Tecnológica Nacional
{damiannicolas05, dr.jotacastillo, ing.marinacardenas, sandramolariaga, matutecba19, soffibritos01, zsantinoz}@gmail.com

RESUMEN

El crecimiento sostenido de la inteligencia artificial y, en particular, de los grandes modelos de lenguaje (LLMs), ha incrementado la demanda de material de entrenamiento de alta calidad, así como la necesidad de mecanismos confiables para la detección de similitudes y redundancias en textos y código fuente. La construcción manual de corpus lingüísticos y conjuntos de datos estructurados constituye un proceso complejo, intensivo y propenso a errores, lo que motiva el desarrollo de herramientas que optimicen estas tareas.

En este trabajo se presenta un conjunto de herramientas desarrollado en la Universidad Tecnológica Nacional – Facultad Regional Córdoba (UTN-FRC), orientado a la generación, análisis y validación de material de entrenamiento, así como a la detección de similitudes en proyectos de software. El conjunto incluye el Programa de Mapeo de Datos (PMD), el Banco de Pruebas de Algoritmos de Semejanza (BPAS), el Sistema de Detección de Similitudes (SDS) y el Sistema de Recuperación de Datos Web (SRDW).

El enfoque integra técnicas de minería de datos, análisis textual y enriquecimiento contextual mediante recuperación de información, facilitando la construcción de corpus y el análisis comparativo de código de manera más eficiente y controlada.

Todas estas herramientas tienen como objetivo la generación de material de entrenamiento de calidad para su uso en la creación de modelos de inteligencia artificial.

CONTEXTO

El presente trabajo se llevó a cabo dentro del grupo de investigación denominado Grupo de Aprendizaje Automático, Lenguajes y Autómatas (GA^2LA) en donde se desarrollan proyectos relacionados con autómatas y lenguajes formales, procesamiento del lenguaje natural, y aprendizaje automático, y en especial proyectos orientados a la aplicación de la inteligencia artificial. Este grupo realiza sus actividades en el Laboratorio de Ingeniería de Software (LIS) del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN - FRC).

El mismo se desarrolló dentro de los proyectos homologados por la SCyT de la UTN denominados “Modelo para el procesamiento de textos estructurados Fase 2” (cód. SIECACO0008518) y “Desarrollo de aplicaciones de análisis de texto para Ingeniería de Software” (cod. COSITCI569).

1. INTRODUCCIÓN

Contar con material de entrenamiento adecuado es fundamental para el desarrollo de modelos de inteligencia artificial.

Originalmente, la generación de corpus lingüísticos representaba un desafío importante; actualmente, este problema se ha intensificado debido a la proliferación de contenido generado por grandes modelos de lenguaje (LLMs) [1]. La IA ha adquirido un rol relevante en múltiples sectores (tecnológico, salud, empresarial y social) donde su precisión es crítica dependiendo de la tarea encomendada, desde mejorar la experiencia de usuario [2] o realizar campañas de marketing [3], hasta apoyar investigaciones científicas [4].

La creación de material de entrenamiento de alta calidad requiere procesos intensivos y artesanales, con un alto esfuerzo humano y riesgo de errores [5][6]. Para abordar esta dificultad, se ha desarrollado un conjunto de herramientas con enfoque semiautomático que reduce la carga de trabajo y mejora la consistencia de los datos. Este conjunto incluye: el Programa de Mapeo de Datos (PMD), que centraliza y analiza información de diversas fuentes; el Banco de Pruebas de Algoritmos de Semejanza (BPAS), que compara algoritmos de similitud de textos estructurados mediante el método de jerarquía analítica (AHP); el Sistema de Detección de Similitudes (SDS), que unifica resultados de diversos programas de detección de similitudes para generar porcentajes comparativos; el Sistema de Recuperación de Datos Web (SRDW), que obtiene información de páginas web mediante análisis HTML, capturas de pantalla y lectura de archivos locales; y finalmente, The Oracle, que combina métricas clásicas de similitud (TF-Coseno, Jaccard y Levenshtein) con LLMs para detectar cambios y redundancias en código fuente.

Estas herramientas proporcionan soporte en tareas como extracción de información, minería de datos [7] e implicación textual [8][9], facilitando la generación de material de entrenamiento de calidad y el análisis de similitudes en proyectos de software.

2. LÍNEAS DE INVESTIGACIÓN

Las líneas de investigación de este trabajo se centran en el desarrollo y aplicación de técnicas de lingüística computacional y aprendizaje automático, tanto supervisado como no supervisado, orientadas a la construcción y mejora de material de entrenamiento para modelos de inteligencia artificial. Se aborda la escasez de datos de calidad mediante la extracción de información de diversas fuentes, incluyendo textos estructurados, no estructurados y proyectos de código fuente.

Otra línea clave se enfoca en la detección de plagio y redundancia en archivos de software, explorando la comparación de código empleando métricas clásicas de similitud y el uso de grandes modelos de lenguaje (LLMs), incorporando estrategias de fine-tuning y generación aumentada por recuperación (RAG) para mejorar la identificación de modificaciones semánticas y estructurales en proyectos de software.

Asimismo, se investigan métodos de integración entre herramientas semiautomáticas de análisis de datos, minería de información y implicación textual, con el objetivo de optimizar la generación de corpus de alta calidad, evaluar algoritmos de similitud y enriquecer el contexto de LLMs para obtener resultados más precisos en tareas críticas.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1. Programa de Mapeo de Datos (PMD)

Este programa tiene por objetivo el procesamiento de diferentes fuentes y orígenes de datos para su normalización y almacenamiento en una estructura estándar que aprovecha las capacidades que proporciona SQL Server.

Este cuenta con dos versiones una de escritorio desarrollada en C# y una versión web. Este enfoque dual de administración permite un equilibrio entre el control local para trabajar de manera individual y el trabajo en grupos de investigadores en entornos de investigación académica.

Su funcionamiento se basa en realizar análisis exploratorio y análisis sobre textos estructurados.

En la figura 1 se puede observar la interfaz de la aplicación de escritorio.



Figura 1. Interfaz de PMD versión escritorio.

3.2. Banco de Pruebas de Algoritmos de Semejanza (BPAS)

El Banco de Pruebas de Algoritmos de Semejanza (BPAS) es una herramienta diseñada para brindar un análisis comparativo de diversos algoritmos de detección de similitudes, con el objetivo de evaluar su capacidad para identificar similitudes en textos y código fuente. La herramienta emplea el método de jerarquía analítica (AHP) para establecer niveles de comparación y generar una jerarquía de opciones basada en criterios previamente definidos.

BPAS analiza principalmente textos estructurados, como archivos de código fuente escritos en lenguajes de programación como Java, C, C#, Perl y Python, aunque también es aplicable a textos no estructurados. La aplicación permite determinar cuál de los algoritmos evaluados presenta mejor desempeño según los criterios establecidos, proporcionando información cuantitativa y comparativa sobre la eficacia de cada algoritmo en la detección de similitudes.

De esta manera, BPAS constituye un componente clave dentro del ecosistema de herramientas desarrollado, facilitando la selección de algoritmos óptimos para tareas de análisis de código y generación de material de entrenamiento, contribuyendo a la eficiencia y precisión en proyectos de software y estudios de similitud textual. En la figura 2 se muestran los criterios utilizados y los pesos asignados a cada uno.

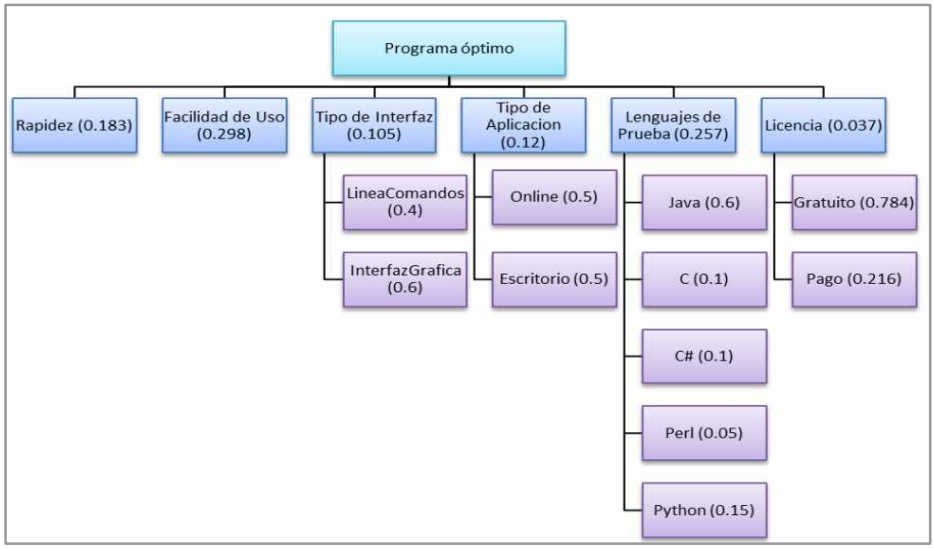


Figura 2. Criterios de Selección.

3.3. Sistema de Detección de Similitudes (SDS)

Se confeccionó una herramienta que permite realizar comparaciones entre archivos de código fuente con los lenguajes de programación antes mencionados, utilizando la salida de los diferentes programas de detección de similitudes seleccionados con el BPAS, tales como Simian, SimTest, Sherlock y Jplag. Este sistema utiliza las salidas de los programas para brindar los resultados en una tabla de porcentajes de similitud por programa comparando de a pares de archivos.

La misma fue desarrollada con C# y ASP.NET para brindar una interfaz web que permite la gestión de las comparaciones, y brinda funciones adicionales como impresión y exportación a archivos.

En la figura 3 se puede observar la interfaz web.

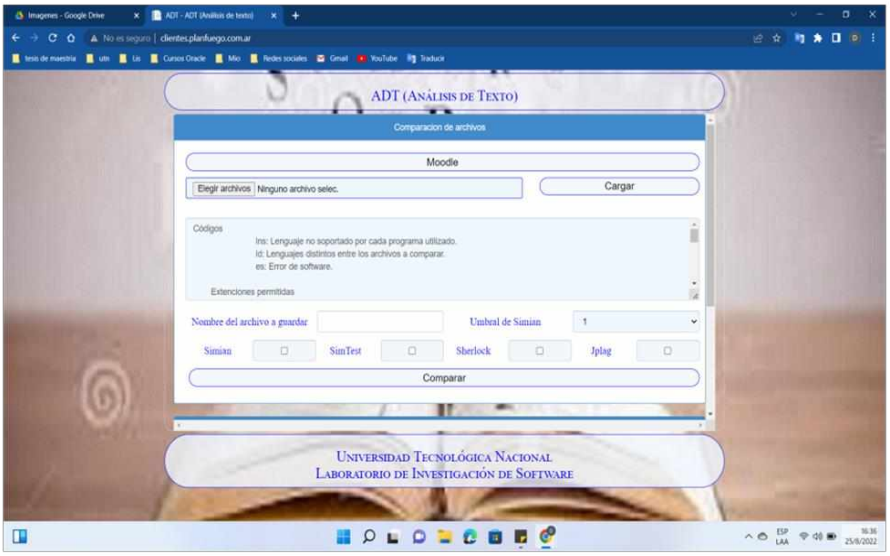


Figura 3. Interfaz del Sistema de Similitudes.

3.4. Sistema de Recuperación de Datos Web (SRDW)

Este sistema permite obtener información actualizada a partir de páginas web con el fin de enriquecer el contexto de los LLMs para que esto brinde resultados más precisos a las consultas realizadas sobre eventos recientes valiéndose del Protocolo de Modelo de Contexto o Model Context Protocol (MCP) que permite dotar de herramientas a los LLM para extender sus capacidades.

Este dispone de tres herramientas:

FetchInfoHTML: permite recuperar el contenido de los sitios web en formato html.

CaptureWeb: permite capturar la información mediante capturas de pantalla.

ReadFile: permite leer el contenido de archivos locales.

Estas herramientas permiten ampliar el contexto de LLMs como Claude Sonnet con su cliente de escritorio Claude Desktop[1] o cualquier modelo que soporte el uso de funciones MCP.

3.5. The Oracle

Esta aplicación permite detectar la similitud entre archivos de código fuente mediante métricas de similitud empleando tres algoritmos clásicos como son Term Frequency [10] con similitud coseno, distancia de Levenshtein [11] e índice de Jaccard [12], también brinda la opción mediante un prompt bien estructurado con el enfoque few-shot para que el mismo detecte las modificaciones de los archivos como por ejemplo cambios en el nombre de las variables, espacios, indentación y comentarios, cambios en bloques de código y sustitución de ciclos manteniendo siempre la funcionalidad original de los archivos.

Además como funcionalidad extra brinda una función para la generación de archivos de código fuente modificados con las alternativas descritas para la generación de material de calidad de manera automática y con la supervisión humana. En el futuro se planea ampliar los lenguajes soportados y extender su funcionalidad a proyectos enteros de software.

4. FORMACIÓN DE RECURSOS HUMANOS

- Una doctoranda en ingeniería con mención en sistemas de información en la UTN-FRC.
- Un doctor en ciencias de la computación.
- Un maestrando en Ingeniería en Sistemas de Información de la UTN-FRC.
- Docentes investigadores ingenieros del departamento de Ingeniería en Sistemas de Información de la UTN FRC.
- Alumnos becarios.
- Alumnos que realizan sus prácticas supervisadas.
- Adicionalmente, se realizan charlas de difusión y jornadas de capacitación.

5. BIBLIOGRAFÍA

1] X. Hou, Y. Zhao, Y. Liu, Z. Yang, K. Wang, L. Li, X. Luo, D. Lo, J. Grundy, y H. Wang, “Large language models for software engineering: A systematic literature review,” arXiv preprint arXiv:2308.10620, 2023.

[2] N. S. Reyes, “Uso de la inteligencia artificial en la personalización de la experiencia del usuario en plataformas digitales,” Polo del Conocimiento, vol. 8, no. 6, pp. 1190–1206, 2023.

[3] F. Zúñiga, D. A. M. Poveda, y D. P. M. Mora, “La importancia de la inteligencia artificial en las comunicaciones en los procesos marketing,” Vivat Academia, pp. 19–39, 2023, doi: 10.15178/va.2023.156.e1474.

[1] https://claude.ai/download

[4] L. B. Díaz Subieta, “El uso de la inteligencia artificial en la investigación científica,” Journal History of Latin American Education/Revista Historia de la Educación Latinoamericana, vol. 26, no. 43, 2024.

[5] J. Castillo, M. Cardenas y N. Hernández, “Preparación de corpus y evaluación de algoritmos de similitud: herramientas de apoyo para el análisis de textos,” CONAIISI 2025, Córdoba, 2025.

[6] T. Stefan y A. Stefanowitsch, Corporain in Cognitive Linguistics, Corpus-Based Approaches to Syntax and Lexis, Berlin: Mouton, 2006, pp. 1–17.

[7] M. D. C. Justicia de la Torre, Nuevas técnicas de minería de textos: Aplicaciones (Tesis doctoral), Universidad de Granada, 2017.

[8] J. Castillo y M. Cardenas, “Using sentence semantic similarity based on WordNet in recognizing textual entailment,” en Iberamia 2010, LNCS, vol. 6433, pp. 366–375, 2010.

[9] J. Castillo, “Using machine translation systems to expand a corpus in textual entailment,” en Icetal 2010, LNCS, vol. 6233, pp. 97–102, 2010.

[10] E. Moyotl-Hernández y H. Jiménez-Salazar, “An analysis on frequency of terms for text categorization,” Procesamiento del Lenguaje Natural, vol. 33, 2004.

[11] D. Jurafsky y J. H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 3ra ed. Stanford, CA, USA: Pearson, 2023.

[12] S. M. Abdullah, S. M. Ali, y M. A. Makttof, “Modifying Jaccard coefficient for text similarity,” Revista de Ciencias Humanas y Sociales, no. 19, pp. 28, 2019

VALIDACIÓN Y TRANSFORMACIÓN DE CONTRATOS INTELIGENTES A SOLIDITY

Anabella Pessini, Valentina Rio, Lorena Baigorria, Ana Garis

Departamento de Informática, Facultad de Ciencias Físico Matemáticas y Naturales

Universidad Nacional de San Luis, Ejército de los Andes 950, San Luis

{pessinianabella@gmail.com,valentina.rio.srz@gmail.com,flbaigor@unsl.edu.ar, agaris@unsl.edu.ar}

RESUMEN

Blockchain permite realizar transacciones seguras en redes descentralizadas sin intermediarios. Frecuentemente, estas transacciones requieren lógica automatizada mediante Contratos Inteligentes. Un Contrato Inteligente (CI) es esencialmente código almacenado en la Blockchain que se ejecuta automáticamente cuando se cumplen ciertas condiciones. Los errores en los CI pueden tener consecuencias graves, especialmente en finanzas descentralizadas. Este trabajo propone una línea de investigación que utiliza UML y OCL para la especificación de CI, su posterior transformación automática al lenguaje formal Alloy para Verificación y Validación (V&V), y finalmente la generación automática de código en Solidity. Esta última transformación incorpora un enfoque asistido por inteligencia artificial basado en Retrieval-Augmented Generation (RAG), el cual utiliza conocimiento del dominio y documentación técnica para apoyar la traducción automática desde Alloy hacia Solidity, con el objetivo de preservar las restricciones y el comportamiento definidos a nivel formal y reducir errores durante la implementación.

Palabras clave: Contratos Inteligentes, UML, OCL, Alloy, Modelos, RAG

1. INTRODUCCIÓN

La inmutabilidad de Blockchain exige que los CI [1] funcionen correctamente desde su despliegue. No obstante, la ambigüedad del lenguaje natural y los errores de programación suelen derivar en comportamientos no deseados. Para mitigar estos riesgos, se propone un modelo basado en la Ingeniería Dirigida por Modelos (MDE)[2].

El uso de UML/OCL, ampliamente utilizados en la ingeniería de requerimientos, permite a las partes interesadas analizar acuerdos. Estas restricciones son modeladas utilizando el lenguaje formal Alloy permitiendo así la exploración exhaustiva de escenarios para detectar inconsistencias antes del despliegue.

Las especificaciones formales verificadas y validadas en Alloy se utilizan como entrada para un proceso de generación automática de contratos inteligentes en Solidity, mediante un esquema de Retrieval-Augmented Generation (RAG). Este enfoque incorpora conocimiento del dominio y documentación técnica relevante para guiar la transformación desde el modelo formal a código ejecutable.

De este modo, se contribuye a reducir la brecha entre la especificación formal y la implementación final, preservando las

restricciones y el comportamiento definidos a nivel de modelo, y disminuyendo la probabilidad de introducir errores durante la etapa de codificación.

2. CONTEXTO

La Ingeniería Dirigida por Modelos (MDE, por sus siglas en inglés) propone el uso de modelos abstractos que se refinan automáticamente hasta obtener el código final [2]. En este contexto, se propone la integración de diferentes técnicas y herramientas para lograr la generación automática de CI. Entre éstas se pueden destacar:

UML/OCL: lenguajes de modelado ampliamente utilizados en el ciclo de vida del desarrollo de software para la representación estructurada de requisitos establecidos por las partes involucradas en el CI. Sin embargo, carecen de semántica formal completa para verificación automática [3,4].

Alloy: un lenguaje formal basado en álgebra relacional [5], utiliza Bounded Model Checking y un SAT Solver para verificar y validar modelos. Permite realizar una exhaustiva revisión de los requisitos establecidos en los diferentes escenarios posibles.

Retrieval-Augmented Generation (RAG): es un enfoque que integra la recuperación de información desde una base de conocimiento especializada con modelos de lenguaje generativos, permitiendo generar salidas más precisas y coherentes [6].

En este trabajo, RAG se utiliza para asistir la transformación de modelos formales en Alloy a contratos inteligentes en Solidity [7], incorporando reglas de traducción, patrones y restricciones del dominio. El uso de RAG permite reducir ambigüedades, mejorar la trazabilidad entre la especificación formal y el código generado, y disminuir errores típicos de la generación puramente automática, al fundamentar el proceso de generación en conocimiento explícito y verificable del dominio.

Solidity: es un lenguaje de programación orientado a contratos, estáticamente tipado y de alto nivel, diseñado para desarrollar contratos inteligentes que se ejecutan en la Ethereum apuntando a la máquina virtual Ethereum (EVM). Permite definir estructuras de datos, funciones, modificadores y eventos, incorporando mecanismos de control de acceso y seguridad propios del entorno Blockchain [7].

En la actualidad existen diversas aproximaciones para asegurar la calidad de los CI [8,9]. La propuesta planteada en esta línea de investigación incluye el uso de un enfoque asistido por inteligencia artificial complementado con la técnica RAG que permita transformaciones sucesivas iniciando en las reglas pautadas por las partes interesadas hasta la generación de los CI en el lenguaje Solidity.

Este enfoque busca que el modelo produzca CI validados y verificados [8], basadas no solo en el conocimiento implícito aprendido durante el entrenamiento, sino también en contexto explícito provisto dinámicamente.

3. PROPUESTA

La propuesta integra herramientas de ingeniería de software para automatizar el ciclo de vida correspondiente a la construcción de un CI:

1. Modelado: definición del CI mediante diagramas de clases/estados (UML) y restricciones (OCL).
2. Transformación a Alloy: conversión semántica de los modelos para su análisis formal de forma automática [10].
3. Verificación y Validación Formal: exploración de escenarios utilizando *Alloy Analyzer* para detectar cláusulas incumplidas.
4. Generación de Código: implementación de RAG para la traducción a de las restricciones validadas de Alloy a Solidity.
5. Testing: Pruebas unitarias e integración del código generado.

En la figura 1 se muestran los pasos correspondientes al modelo propuesto y en la figura 2 la etapa 4 de dicho modelo.



Figura 1: Esquema general del modelo



Figura 2: Esquema correspondiente del paso 4 del modelo

4. RESULTADOS ESPERADOS

La integración de UML/OCL con Alloy busca reducir la brecha entre el modelado tradicional y la verificación formal, buscando promover la automatización de la verificación de propiedades sin intervención manual exhaustiva, la reducción de costos debido a la detección temprana de inconsistencias, evitando fallos costosos en etapas posteriores al despliegue y principalmente la accesibilidad ya que posibilita que perfiles no especializados incorporen técnicas de verificación formal en sus desarrollos. Asimismo, se espera demostrar la viabilidad técnica de la generación automática de contratos inteligentes en Solidity a partir de especificaciones formales en Alloy, mediante un enfoque asistido por RAG.

Además de evidenciar su factibilidad, se prevé evaluar la solución en términos de calidad del código generado, considerando claridad estructural y adecuación al dominio, correspondencia semántica entre la especificación formal y el contrato producido, validación por un experto en el dominio, quien analizará la coherencia y adecuación funcional del contrato generado o validación formal complementaria, ya sea mediante la búsqueda de contraejemplos utilizando modelos formales o a través de un proceso de traducción inversa que permita contrastar el contrato generado con la especificación original.

5. FORMACION DE RECURSOS HUMANOS

Dentro del proyecto, donde se encuentra inserta esta línea de investigación, se están llevando a cabo diferentes tesis de grado y posgrado. Las mismas corresponden a la Licenciatura en Ciencias de la Computación y al Doctorado en Ingeniería Informática. En particular, la línea de investigación

forma parte de una tesis del Doctorado en Ingeniería Informática.

BIBLIOGRAFÍA

[1] L. Luu et al., "Making Smart Contracts Smarter," in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16), 2016, doi: 10.1145/2976749.29783098.

[2] D. C. Schmidt, "Model-driven engineering," IEEE Computer, vol. 39, no. 2, pp. 25–31, 2006, doi: 10.1109/MC.2006.5810. https://www.researchgate.net/publication/3794013613.

[3] C. Bartolini, C. Palomba, and F. Lomuscio, "Smart Contract Modeling using UML and OCL for Formal Verification," in *2021 IEEE International Conference on Software Architecture (ICSA)*, 2021, pp. 139–150.

[4] Y. Hsain, N. Laaz, and S. Mbarki, "Ethereum's Smart Contracts Construction and Development using MDE Technologies: A Review," Procedia Computer Science, vol. 181, pp. 250–257, 20216.

[5] A. Garis and A. Sanchez, "Verification and Validation of Domain Specification Languages using Alloy," in Proceedings of the XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015), 2015, pp. 589-5984.

[6] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., … Riedel, S. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. arXiv preprint arXiv:2005.11401.

[7] Solidity 0.8.24 documentation. (s. f.). https://docs.soliditylang.org/en/v0.8.24/

[8]I. Grishchenko, M. Maffei, and C. Schneidewind, "A Semantic Framework for the Security Analysis of Ethereum Smart Contracts," POPL, 20185.

[9] D. Marmsoler, A. Ahmed, and A. D. Brucker, "Secure Smart Contracts with Isabelle/Solidity," in Software Engineering and Formal Methods (SEFM 2024), A. Madeira and A. Knapp, Eds. Springer, 20259.

[10] A. Cunha, A. Garis, and D. Riesco, "Translating between Alloy specifications and UML class diagrams annotated with OCL," Software Syst Model, vol. 14, pp. 5–25, 2015, doi: 10.1007/s10270-013-0353-52.

VERIFICACIÓN Y VALIDACIÓN DE SOFTWARE EN LA INDUSTRIA 5.0

Martín L. Larrea[1,3,4], Dana K. Urribarri[1,3,4], M. Luján Ganuza[1,3,4], Matías N. Selzer[1,3,4], M. Laura Cobo[1], Antonella S. Antonini[1,3,4], Leila Vecslir[2]

{mll, dku, mlg, matias.selzer, mlc}@cs.uns.edu.ar, antoo.antonini@gmail.com, leila.vecslir@uns.edu.ar

[1]Departamento de Ciencias e Ingeniería de la Computación

[2]Departamento de Economía

[3]Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (UNS-CIC)

[4]ICIC, Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

Universidad Nacional del Sur

RESUMEN

La Industria 5.0 amplía el enfoque de la Industria 4.0 al reconocer el potencial de la industria para alcanzar objetivos sociales que van más allá del empleo y el crecimiento económico, colocando el bienestar del trabajador en el centro del proceso productivo. A partir del estado del arte en Verificación y Validación de Software, así como de los principios de la Industria 5.0 en la interacción entre usuarios y software, es posible diseñar, implementar y adaptar metodologías y herramientas que faciliten la detección de errores y problemas en los factores de calidad, tanto internos como externos, del software. El proyecto se estructura en dos ejes complementarios: el desarrollo de herramientas y metodologías de Verificación y Validación de Software alineadas con la Industria 5.0, y la formación de perfiles no informáticos para su participación activa en procesos de validación de software. De esta manera, se promueve la mejora en la capacitación de los recursos humanos y un mayor control sobre la calidad del software.

1. INTRODUCCIÓN

En el año 2021 la Comisión Europea identificó formalmente a la Quinta Revolución Industrial (denominada Industria 5.0), tras debates entre participantes de organizaciones de investigación y tecnología, y otras agencias de toda Europa. Esto quedó reflejado en el reporte producido por la comisión *"Industry 5.0 - Towards a sustainable, human-centric and resilient European industry"* [1].

La Industria 5.0 complementa el paradigma existente de la "Industria 4.0" al reconocer el poder de la industria para lograr objetivos sociales más allá del empleo y el crecimiento, para convertirse en un proveedor resiliente de prosperidad, haciendo que la producción respete los límites de nuestro planeta y colocando el bienestar del trabajador de la industria en el centro del proceso de producción. La investigación y la innovación se centran en impulsar la transición hacia una industria sostenible, centrada en las personas. La Quinta Revolución Industrial mantiene al software como elemento transversal y fundacional, en especial a través de la reconciliación que pretende entre humanos y máquinas [2]. Las metas que se proponen, específicamente aquellas vinculadas a la relación entre los usuarios y el software, redefinen lo que es nuestro entendimiento de la calidad de los sistemas informáticos. En tal sentido, y en sintonía con los objetivos planteados, la experiencia del usuario y su satisfacción se han transformado en las principales metas de calidad [3]. Los equipos interdisciplinarios son una forma cada vez más común de desarrollar software, por lo que las herramientas que garantizan la calidad de los productos de software deben ser utilizables por todos los miembros del equipo, y no sólo por aquellos con un perfil informático [4].

Esta nueva era requiere nuevas metodologías, técnicas y herramientas para verificar y validar el software. La ingeniería busca la producción de software de calidad, entendida como una combinación de varios factores [5]. Por un lado, consideramos cualidades como la facilidad de uso, la adaptabilidad y la accesibilidad cuya presencia o ausencia en un producto de software puede ser detectada por sus usuarios. Estas propiedades se denominan

factores de calidad externos. Otras cualidades aplicables a un producto de software, como ser modular o extensible, son factores internos, perceptibles sólo para los profesionales de la informática que tienen acceso al código del software.

La Industria 5.0 pone la atención en los factores externos. Sin embargo, la clave para lograr estos factores externos está fuertemente vinculada con los internos: para que los usuarios disfruten de las cualidades visibles, los diseñadores y desarrolladores deben haber aplicado técnicas internas que aseguren las cualidades ocultas. Para poder potenciar ambas caras de la calidad, debemos introducir nuevas formas de guiar el desarrollo de software y de también evaluar lo ya desarrollado.

La Verificación y la Validación son dos formas de evaluar la calidad de un producto de software. La Verificación se ocupa de comprobar que el software cumple los requisitos funcionales y no funcionales de su especificación, los factores internos; mientras que la Validación tiene a su cargo comprobar que el software cumple las expectativas que el cliente espera, los factores externos.

En este contexto industrial, para avanzar en software con un enfoque centrado en las personas es necesario repensar y rediseñar las formas de desarrollar y evaluar los sistemas. Los usuarios deben participar en cada paso de estos procesos dando retroalimentación. Nuestra realidad actual nos muestra que cada día más decisiones son tomadas en base a información brindada por un sistema informático, como es el caso del análisis visual mediante el uso de representaciones visuales de grandes conjuntos de datos, Realidad Virtual o Aumentada. No sólo hay un aumento cuantitativo sino también cualitativo de estas decisiones, las mismas son más críticas y con mayor impacto en la sociedad, el medio ambiente y los individuos.

En el nivel de implementación, debemos asegurar la calidad de los factores internos mediante la verificación del software. Los nuevos desafíos que se enfrentan conllevan nuevas aplicaciones de los sistemas informáticos y nuevos paradigmas de desarrollo, en consecuencia, se requieren adaptar las formas de encontrar errores o generar nuevas técnicas de testeo. Desde la perspectiva del usuario, debemos asegurar la calidad de los factores externos mediante la validación del software. La meta que establece la Industria 5.0 de ubicar el factor humano en el centro de la producción nos obliga a trasladar esto al software involucrado en dicha producción. El factor humano se presenta en la Ingeniería de Software a través de las áreas de Usabilidad e Interacción Humano-Computadora; áreas que existen desde hace décadas y tienen mucho para aportar en esta nueva revolución industrial.

Esta línea tiene como finalidad desarrollar nuevas técnicas de verificación y validación de software en sintonía con la Industria 5.0. Además, se busca fomentar e instruir a aquellos perfiles de trabajo no vinculados con los sistemas informáticos en aspectos relacionados con el testeo de software y el diseño centrado en el usuario.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Basándonos en el estado del arte de la verificación y validación de software, y en los objetivos establecidos para la Industria 5.0 en cuanto a la relación usuario/software, es posible definir, implementar y adaptar metodologías y herramientas que permitan detectar errores y/o problemas, tanto en los factores de calidad externos y/o internos del software.

La generación de espacios de formación orientados a perfiles de trabajadores no informáticos, permite su capacitación para poder realizar la validación de software en aspectos funcionales y no funcionales. Esto logra mejorar la capacitación de los recursos humanos y un mayor control de la calidad del software.



Figura 1. Entrega de certificados por la capacitación en Testing Funcional Manual.

El objetivo general consiste en el estudio, desarrollo y difusión de metodologías para la validación y verificación de software en el contexto de Industria 5.0. El objetivo específico de este trabajo de investigación consiste en el diseño y desarrollo de herramientas que asistan en el desarrollo software para asegurar la calidad del producto final, particularmente en aquellos aspectos vinculados con la satisfacción del usuario al usar el software. También es un objetivo específico el de generar capacitaciones a perfiles no técnicos sobre cómo usar estas herramientas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Esta línea de investigación se apoya en resultados previos obtenidos en trabajos publicados, cuyos aportes principales se resumen en [6], donde se presentan herramientas y enfoques iniciales para la verificación y validación de software en contextos industriales. En particular, dichos trabajos constituyen un antecedente relevante para la adaptación de técnicas de testing a escenarios donde el software cumple un rol central en la toma de decisiones y en la interacción humano–máquina.

Actualmente, el equipo se encuentra trabajando en la extensión de estas propuestas mediante la incorporación de estrategias de testing orientado a objetos a nivel de clase, basadas en una metodología modal [8]. Esta línea de trabajo busca fortalecer la verificación de factores de calidad internos, adaptando técnicas clásicas de testing a los nuevos paradigmas de desarrollo y a sistemas de mayor complejidad.

Como resultados esperados, se prevé avanzar en la definición de metodologías de verificación y validación que integren de manera explícita los factores de calidad externos, particularmente aquellos vinculados con la experiencia y satisfacción del usuario. En este sentido, se espera expandir las metodologías de testing hacia niveles que contemplen la funcionalidad percibida y la experiencia de uso, en concordancia con los principios de la Industria 5.0 y su enfoque centrado en las personas.

Asimismo, se espera que los resultados de esta línea de investigación contribuyan al diseño de herramientas y estrategias de validación accesibles a perfiles no técnicos, facilitando la participación activa de los usuarios en los procesos de evaluación de software y

promoviendo un mayor control sobre la calidad de los productos desarrollados.

4. FORMACIÓN DE RECURSOS HUMANOS

Durante el año 2025 se desarrolló una experiencia de formación orientada a la comunidad, enfocada en el desarrollo de habilidades digitales vinculadas a la validación de software. En particular, se implementó una capacitación de ocho clases sobre Testing Funcional Manual, destinada a personas sin formación previa en informática.

Esta capacitación fue realizada de manera conjunta entre el programa de formación profesional del Consorcio de Gestión del Puerto de Bahía Blanca y la empresa de recursos humanos Cronnos Consulting, y contó con la participación de 19 personas (Figura 1). La experiencia permitió explorar estrategias de enseñanza de conceptos de validación de software desde una perspectiva práctica y centrada en el usuario, y constituye un antecedente relevante para el diseño de futuras instancias de formación e investigación orientadas a la inclusión de perfiles no técnicos en procesos de evaluación de software.

Complementariamente se mantiene en ejecución el Proyecto de Grupo de Investigación “Verificación y Validación de Software en la Industria 5.0” (PGI 2024-2028 24/N059).

5. BIBLIOGRAFÍA

[1] European Commission, Directorate-General for Research and Innovation, Breque, M., De Nul, L., Petridis, A., Industry 5.0 – Towards a sustainable, human-centric and resilient European industry, Publications Office of the European Union, 2021, https://data.europa.eu/doi/10.2777/308407

[2] Akundi, Aditya, et al. "State of Industry 5.0—Analysis and identification of current research trends." Applied System Innovation 5.1 (2022): 27.

[3] Riasat, H. A. F. S. A., et al. "Enhancing Software Quality Through Usability Experience And HCI Design Principles." vol 42 (2023): 46-75.

[4] Ustundag, Alp, and Emre Cevikcan. Industry 4.0: managing the digital transformation. Springer, 2017.

[5] Jena, Divya, et al. "Systematic Literature Review on Object Oriented Software Testing Techniques." 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA). IEEE, 2023.

[6] Larrea, Martín L., and Dana K. Urribarri. "New verification and validation tools for Industry 4.0 software." Designing Smart Manufacturing Systems. Academic Press, 2023. 61-88.

[7] Larrea, Martín Leonardo, et al. "Verificación y validación de Visualizaciones." XXV Workshop de Investigadores en Ciencias de la Computación (WICC 2023, Junin, Buenos Aires). 2023.

[8] Binder, Robert V. "Testing object-oriented software: a survey." Software Testing, Verification and Reliability 6.3-4. 1996. 125-252.

Innovación en Educación en Informática

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

ADOPCIÓN DE RISC-V EN LA CARRERA DE INGENIERÍA EN INFORMÁTICA

Gho Edgardo[1,2] , Ferreyra Biron Martín [1] , Hnatiuk Jair[1] , Maidana Carlos [1]

[1] Grupo de investigación en lógica programable, Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza, Buenos Aires, Argentina.
{egho, mferreyra, jhnatiuk, cmaidana} @unlam.edu.ar

[2] Universidad Abierta Interamericana. Facultad de Tecnología informática. Centro de Altos Estudios en Tecnología Informática. Buenos Aires, Argentina.

RESUMEN

En el año 2023 la Universidad Nacional de La Matanza (UNLaM) implementó un nuevo plan de estudio para la carrera de Ingeniería en Informática. Como parte de este nuevo plan se actualizó el contenido de la materia "Arquitectura de Computadoras" incorporando RISC-V como arquitectura de referencia. Este cambio fue planificado años antes cuando el Grupo de Investigación en Lógica Programable (GILP) del Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT) adoptó como línea de investigación dicha arquitectura y desarrolló dos procesadores que la implementan. El plan de estudios anterior utilizaba como objeto de estudio una arquitectura antigua de 8 bits diseñada por Motorola (68HC11) en la década del 80. El objetivo principal fue modernizar los contenidos, trabajando con una arquitectura moderna, abierta y de 32 bits tipo RISC que sirva para el estudio de técnicas de paralelismo a nivel de instrucción, como pipelining y ejecución fuera de orden. En este trabajo se describen los cambios realizados y se analizan los resultados transcurridos tres años desde la adopción de RISC-V a la materia.

Palabras Clave: Educación, Arquitectura de Computadoras, Unidad Central de Proceso, RISC-V

CONTEXTO

Los resultados presentados surgen del proyecto I/D del DIIT C2-ING-082 [1] "Implementación de la expansión de un procesador RISC-Vp en un entorno de desarrollo de lógica programable" y del proyecto antecesor del programa PROINCE C-219 "Desarrollo e implementación de una arquitectura basado en el conjunto de instrucciones RISC-V"[2]. Actualmente el GILP se encuentra en el marco del proyecto I/D del DIIT C2-ING-153 "Implementación del External Debug Support para RISC-V".

1. INTRODUCCIÓN

La enseñanza de las arquitecturas de computadora es un pilar fundamental de las carreras de Ingeniería en Informática o Sistemas[3]. Típicamente cubre temas como diseño de CPU, conjunto de instrucciones (ISA), lenguajes de bajo nivel e interfaces de entrada/salida. Según [4], la arquitectura de computadoras vive en la intersección de la ingeniería en electrónica, el diseño digital, el diseño de compiladores, la teoría de los lenguajes de programación y la alta eficiencia. Queda claro que el año de carrera en el cual esta materia se dicte tendrá alumnos con distintos modelos mentales del funcionamiento de una computadora.

La elección de la arquitectura a enseñar tendrá una suma importancia en función de cómo se explica el diseño de la CPU. Teniendo

en cuenta el soporte a distintos conjuntos de instrucciones complejas (CISC) o reducidas (RISC) [5] , vemos que en ámbitos académicos las arquitecturas RISC son las más elegidas[6]. Realizando un breve mapeo sistemático de la literatura en IEEE y ACM, encontramos que las arquitecturas MIPS, ARM y RISC-V se encuentran presente en más del 94% de los trabajos publicados sobre la enseñanza de arquitectura de computadoras. La arquitectura más utilizada es MIPS[7], probablemente por el hecho de ser la más antigua de las tres. La arquitectura RISC-V[8], siendo la más reciente de todas las arquitecturas mencionadas posee una cantidad de referencias cercana a MIPS.

2. LÍNEAS DE INVESTIGACIÓN / DESARROLLO

Desde el año 2019 el GILP está trabajando en la línea de investigación de arquitectura de computadoras estudiando la arquitectura RISC-V. Fruto de este trabajo se publicó un diseño de procesador RISC-V2p[9] que permitió conocer en detalle el funcionamiento de la arquitectura RISC-V, haciendo especial hincapié en la eficiencia de programas que acceden a datos de forma alineada[10]. También, fruto de esta investigación se cambió el lenguaje descriptivo de hardware utilizado en el GILP por Verilog con la consecuente adopción de este lenguaje en los cursos de diseño sincrónico de circuitos en Ingeniería en Electrónica[11].

En el año 2022 la UNLaM comenzó el proceso de acreditación por la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU) de un nuevo plan de estudios a ser implementado en 2023 para la carrera de Ingeniería en Informática.En el plan anterior (2009) la materia Arquitectura de Computadoras tenía una carga semanal de 6 horas, y la arquitectura de referencia utilizada era Motorola 68HC11, una arquitectura de 8 bits de la década del 80. Con el cambio de plan esta materia se transforma en dos materias correlativas con una carga horaria de 4 horas semanales cada una. En la primera materia, Fundamentos de Sistemas Embebidos (FSE), se dictan los contenidos de electricidad, circuitos integrados, lógica combinacional, lógica secuencial y máquinas de estado, con foco en Internet de las Cosas (IOT). La segunda materia, Arquitectura de Computadoras (AC), si bien comparte el mismo nombre que su antecesora del plan 2009, ahora comienza con diseño de unidad central de proceso (CPU) . El desdoblamiento de la materia “Arquitectura de computadoras” (2009) permitió ampliar los contenidos mínimos para incluir temas de eficiencia como por ejemplo pipelining y ejecución fuera de orden. El resto de los contenidos entre los que se encuentran memorias y unidad de entrada/salida, se renovaron para adaptarse a la arquitectura RISC-V, en particular al uso del procesador Espressif ESP32C3.

La materia AC se divide en 3 grandes unidades con sus respectivos porcentajes de carga horaria:

- CPU (50%)
- Memorias (20%)
- Entrada/Salida (30%)

Debido a la adopción de RISC-V como arquitectura de referencia, se actualizaron los contenidos de la materia. Temas como direccionamiento, alineamiento y ordenamiento se adelantaron en el programa, dando mayor énfasis en la implementación de los mismos hechos por RISC-V. La evaluación de ciertos temas como diseño de CPU y assembler se realizan en laboratorio provisto de computadoras donde corren simuladores de RISC-V para assembler y simuladores de circuitos lógicos para el diseño de CPU. El foco en la evaluación está en que el alumno demuestre competencias en estos temas de forma práctica, sin evaluar la memorización de

conceptos sino la demostración de conocimiento mediante resolución de problemas. Se agregaron circuitos de referencia en simuladores que permiten al alumno verificar sus suposiciones sobre el funcionamiento de los mismos junto a guías de ejercicios donde se obtiene práctica en el uso de estos. La guía de ejercicios de programación destaca las particularidades de RISC-V en cuanto al acceso a memoria y el manejo de registros.

Acorde a la resolución 54/2011 "Régimen de cursada presencial", los alumnos obtienen dos notas parciales durante la cursada, pudiendo recuperar sólo una. Actualmente la primera unidad en la cual se explica CPU se evalúa en dos instancias. En la primera instancia, la evaluación está enfocada en el diseño de CPU y Byte Alignment/Ordering. En la segunda instancia se evalúa programación en assembler de RISC-V. En esta instancia se utiliza un procesador RISC-V simulado con un entorno de desarrollo que permite la programación en Assembler y en C[12]. Luego la segunda y tercera unidad compuesta por Memoria y Entrada/Salida , son evaluadas con un examen integrador de estos temas. Durante la última parte de la cursada se realizan prácticas de laboratorio sobre microcontroladores reales de la familia Espressif con la biblioteca ESP-IDF[13] donde se ejecutan programas escritos en lenguaje de alto nivel paso a paso utilizando compiladores y depuradores (GCC y GDB). Los alumnos que obtengan nota mayor o igual a siete en ambos parciales logran la condición de promoción, evitando tener que dar un examen final. Aquellos que obtienen notas parciales mayores o iguales 4 obtienen la condición de cursada, debiendo presentarse a rendir examen final. Quienes agotan toda posibilidad de examen y no obtienen ambas notas con al menos 4 quedan reprobados. Quienes no utilizan todas las instancias evaluatorias o abandonan la cursada de la materia quedan en condición de ausente. Actualmente la cátedra posee un profesor asociado, dos profesores adjuntos, un jefe de trabajos prácticos, tres ayudantes y dos ayudantes alumnos ad-honorem. Se dictan 4 cursos en modalidad presencial y uno sincrónico presencial mediado por la plataforma Teams (con exámenes presenciales en el laboratorio).

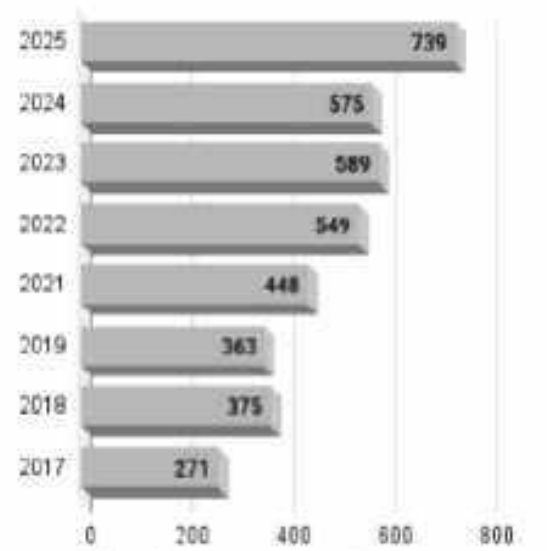


Figura 1 - Alumnos vs Año

3. RESULTADOS OBTENIDOS

Debido a que casi en su totalidad el mismo plantel docente estuvo a cargo durante el plan 2009 y ahora el plan 2023, podemos comparar los resultados luego de la implementación de este último y la incorporación de RISC-V. En la Figura 1 vemos la cantidad de alumnos anotados desde 2017, excluyendo 2020 por haber dictado y evaluado de forma distinta. Se ve como en los últimos nueve años la cantidad de alumnos en la lista aumentó a más del doble.



Figura 2 - Resultados

Si bien la carga horaria en el plan 2023 aumentó un 25%, los contenidos mínimos originales se encuentran cubiertos en el nuevo plan con el agregado de nuevos contenidos. En base a esto realizamos la comparación, considerando que la adopción de RISC-V provocó cambios y luego de tres años presentamos los resultados obtenidos. En la Figura 2 se representan los resultados en cada año. Nuevamente se excluye el año 2020 debido a los cambios en la forma de evaluar a consecuencia de la pandemia Covid-19. El porcentaje de alumnos regularizados, es decir que promocionan o cursan la materia, sufre una caída significativa luego de la pandemia, teniendo su nivel más bajo en 2021. El nuevo plan 2023 comienza con alumnos originalmente inscritos en el plan anterior. Al ser una materia de segundo año recién en 2024 hubo alumnos que comenzaron con el nuevo plan. A partir de este año hay un aumento en la tasa de alumnos regularizados, destacando el incremento significativo de alumnos promocionados, los cuales obtienen notas superiores a siete. En el año 2025 ante el incremento en la demanda de cursos se ofrece un curso sincrónico mediado en cada cuatrimestre, explicando esto el incremento



Figura 3 - Porcentaje de regularizados

significativo de alumnos en 2025. La Figura 3



Figura 4 - Resultados según ingreso

compara el porcentaje de alumnos regularizados en modalidad presencial y modalidad sincrónica mediada. Queda claro que el rendimiento de los alumnos es superior en cursos presenciales y no tan bueno en cursos mediados por la plataforma Teams. Esto explica una caída de 0,27% en alumnos regularizados en 2025.

En cuanto a los alumnos que cursan la materia, durante 2024 un 31,4% de alumnos son provenientes del plan 2023. Luego en 2025 este porcentaje aumentó a 69,25%. En la Figura 4 se compara el desempeño de cada grupo. Se debe tener en cuenta que los alumnos que ingresaron con planes previos a 2023 suelen ser mayores en su edad y en su mayoría tienen trabajos de tiempo completo. Algunos de estos abandonaron la carrera años atrás y se reincorporan a la misma donde su plan original no existe y poseen un sistema de equivalencias para seguir en el plan 2023. A fin de 2025 el 32,9% de alumnos son recursantes, es decir, ya cursaron previamente la materia al menos una vez. El 60,3% de estos alumnos recursantes ingresaron con planes previos. Este dato se explica analizando en la Figura 4 la disparidad de alumnos ausentes entre el nuevo plan y los planes anteriores.

4. TRABAJOS FUTUROS

Dentro del marco de la investigación C2-ING-153 se busca incorporar el funcionamiento de los registros de depuración y la conexión entre estos con las herramientas de desarrollo de GNU, como GCC y GDB.

Actualmente estos contenidos se explican en la cátedra utilizando la biblioteca ESP-IDF de Espressif aunque se aspira integrar estas herramientas con el procesador desarrollado por el GILP. Esto traerá como beneficio una renovación y ampliación de la unidad de entrada/salida cerrando la brecha que poseen los alumnos sobre el funcionamiento de una computadora que aprenden en materias iniciales de programación y el funcionamiento real de una computadora aprendido en Arquitectura de Computadoras. El uso de la arquitectura RISC-V es condición necesaria para esto ya que otras arquitecturas propietarias no comparten su microarquitectura y por ende se hace complicado su estudio y posterior enseñanza.

5. FORMACIÓN DE RECURSOS HUMANOS

Dentro del plantel docente los dos adjuntos se encuentran realizando estudios de posgrado (doctorado y especialización), uno de ellos trabajando en una tesis doctoral sobre la enseñanza de arquitectura de computadoras a nivel de grado y posgrado.

Los dos ayudantes alumnos se encuentran avanzados en la carrera de Ingeniería en Informática y actualmente cumplen tareas de revisión y mejoras del material de clase utilizado.

6. REFERENCIAS

[1] C. E. Maidana, E. A. Gho, J. E. Hnatiuk, C. A. Rodríguez, y J. A. Fiter, "Implementación de la expansión de un procesador RISC-Vp en un entorno de desarrollo de lógica programable". Universidad Nacional de La Matanza. Departamento de Ingeniería e Investigaciones Tecnológicas, el 28 de junio de 2024. Consultado: el 24 de noviembre de 2025. [En línea]. Disponible en: http://repositoriocyt.unlam.edu.ar/handle/123456789/2019

[2] F. I. Szklanny, C. E. Maidana, E. A. Gho, y M. Ferreyra Birón, "Desarrollo e implementación de una arquitectura basada en el conjunto de instrucciones RISC-V". Universidad Nacional de La Matanza. Departamento de Ingeniería e Investigaciones Tecnológicas, el 26 de junio de 2024. Consultado: el 24 de noviembre de 2025. [En línea]. Disponible en: http://repositoriocyt.unlam.edu.ar/handle/123456789/1998

[3] K. Zhang, Y. Chang, M. Chen, Y. Bao, y Z. Xu, "Computer Organization and Design Course with FPGA Cloud", en *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, en SIGCSE '19. New York, NY, USA: Association for Computing Machinery, feb. 2019, pp. 927–933. doi: 10.1145/3287324.3287475.

[4] R. Agrawal, S. Bandara, A. Ehret, M. Isakov, M. Mark, y M. A. Kinsy, "The BRISC-V Platform: A Practical Teaching Approach for Computer Architecture", en *Proceedings of the Workshop on Computer Architecture Education*, en WCAE'19. New York, NY, USA: Association for Computing Machinery, jun. 2019, pp. 1–8. doi: 10.1145/3338698.3338891.

[5] T. Newhall, S. J. Matthews, y K. C. Webb, "A Free Online Textbook Introducing Computer Architecture Topics", en *Proceedings of the Workshop on Computer Architecture Education*, en WCAE '23. New York, NY, USA: Association for Computing Machinery, jun. 2024, pp. 1–8. doi: 10.1145/3605507.3610627.

[6] J. Balkind *et al.*, "BYOC: A 'Bring Your Own Core' Framework for Heterogeneous-ISA Research", en *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, en ASPLOS '20. New York, NY, USA: Association for

Computing Machinery, mar. 2020, pp. 699–714. doi: 10.1145/3373376.3378479.
[7] J. Hennessy *et al.*, "MIPS: A microprocessor architecture", *SIGMICRO Newsl.*, vol. 13, núm. 4, pp. 17–22, oct. 1982, doi: 10.1145/1014194.800930.
[8] K. Asanović y D. A. Patterson, "Instruction Sets Should Be Free: The Case For RISC-V".
[9] E. Gho, *edgardogho/RISC-V2p*. (el 14 de marzo de 2023). VHDL. Consultado: el 24 de noviembre de 2025. [En línea]. Disponible en: https://github.com/edgardogho/RISC-V2p
[10] E. Gho, "Penalidades en lecturas no alineadas dentro de Microcontroladores RISC-V", *Revista Digital del Departamento de Ingeniería e Investigaciones Tecnológicas*, vol. 7, núm. 1, ago. 2022, doi: 10.54789/reddi.7.1.2.
[11] E. Gho, C. Maidana, y J. Hnatiuk, "Adopción de Verilog en materias de diseño digital", en *Libro de Actas*, Junín, Argentina: CEDi Centro de Edición y Diseño. UNNOBA, 2023, pp. 263–267. Consultado: el 15 de octubre de 2025. [En línea]. Disponible en: http://sedici.unlp.edu.ar/handle/10915/163463
[12] M. B. Petersen, "Ripes: A Visual Computer Architecture Simulator", en *2021 ACM/IEEE Workshop on Computer Architecture Education (WCAE)*, jun. 2021, pp. 1–8. doi: 10.1109/WCAE53984.2021.9707149.
[13] *espressif/esp-idf*. (el 5 de diciembre de 2025). C. Espressif Systems. Consultado: el 5 de diciembre de 2025. [En línea]. Disponible en: https://github.com/espressif/esp-idf

DESARROLLO DE UN SIMULADOR EDUCATIVO PARA LA ENSEÑANZA DE SISTEMAS DISTRIBUIDOS EN UNCAUS

Stella Maris Gerzel, Federico Agostini, Carlos Leandro Miguel Latyn

Universidad Nacional del Chaco Austral. Chaco. Argentina
licgerzelstella@gmail.com, agostinifede@hotmail.com, leandrolatyn@uncaus.edu.ar

RESUMEN

La enseñanza de conceptos como sincronización, exclusión mutua, migración de procesos y balanceo de carga presenta dificultades asociadas a su carácter abstracto y dinámico. En este contexto, la simulación computacional constituye una estrategia pedagógica eficaz para el aprendizaje experimental en ingeniería, al permitir la exploración de escenarios controlados sin requerir infraestructura física compleja.

El presente trabajo describe un proyecto de investigación en ejecución en la Universidad Nacional del Chaco Austral (UNCAUS) orientado al diseño e implementación de un simulador educativo para el modelado de sistemas distribuidos. El simulador propuesto integra un motor de modelado configurable, métricas de desempeño cuantitativas y visualización de resultados, posibilitando el análisis comparativo de políticas de asignación y migración. Desde el enfoque educativo, la herramienta promueve aprendizaje activo y basado en experimentación, estrategia respaldada empíricamente en educación en ingeniería. El proyecto busca fortalecer la comprensión conceptual de arquitecturas distribuidas, mejorar el desempeño académico y consolidar la línea de investigación en modelado y simulación del Grupo Investigación en Manejo de Información en UNCAUS.

Palabras clave: simulación educativa; sistemas distribuidos; innovación pedagógica; arquitectura informática; aprendizaje experimental.

CONTEXTO

El proyecto se desarrolla en el ámbito de un Grupo de Investigación de la carrera de Ingeniería en Sistemas de la Universidad Nacional del Chaco Austral y se vincula con las asignaturas de Sistemas Operativos y Redes de Información. Las dificultades detectadas incluyen:

- Baja comprensión de mecanismos de sincronización.
- Dificultad para visualizar políticas de asignación de recursos.
- Limitada experimentación práctica con balanceo de carga.
- Aprendizaje fragmentado de conceptos abstractos.

El simulador surge como respuesta a estas limitaciones, proporcionando un entorno controlado donde los estudiantes puedan modificar parámetros y observar efectos en tiempo real.

1. INTRODUCCIÓN

La enseñanza de sistemas distribuidos constituye uno de los desafíos formativos más complejos en las carreras de ingeniería informática debido a la naturaleza concurrente, asincrónica y parcialmente observable de los sistemas modernos. Conceptos como

exclusión mutua, sincronización distribuida, migración de procesos, consistencia y balanceo de carga requieren un alto nivel de abstracción matemática y comprensión formal [1] [2]. Los fundamentos de coordinación distribuida se sustentan en trabajos clásicos como el algoritmo de relojes lógicos de Lamport [3] y el algoritmo de exclusión mutua distribuida de Ricart y Agrawala [4], los cuales introducen modelos formales para garantizar consistencia sin un reloj global.

Asimismo, el balanceo de carga en entornos distribuidos ha sido ampliamente estudiado como mecanismo para optimizar el uso de recursos y minimizar tiempos de respuesta, particularmente en sistemas heterogéneos [5] [6]. La evaluación de estas políticas requiere análisis cuantitativo del desempeño, tradicionalmente abordado mediante técnicas de simulación discreta y modelado computacional [7] [8].

Desde el punto de vist a pedagógico, múltiples investigaciones en educación en ingeniería demuestran que el aprendizaje activo y basado en experimentación mejora significativamente la comprensión conceptual frente a enfoques exclusivamente teóricos [9]. En este marco, la simulación educativa se configura como una herramienta estratégica para transformar modelos formales en experiencias observables y manipulables, permitiendo a los estudiantes explorar escenarios dinámicos y analizar métricas de desempeño en tiempo real.

En este contexto, la UNCAUS, a través del Grupo de Investigación en Manejo de Información, desarrolla un proyecto de investigación en ejecución orientado al diseño e implementación de un simulador educativo para el modelado de sistemas distribuidos. La propuesta integra fundamentos clásicos de coordinación distribuida, políticas de balanceo de carga y análisis experimental de desempeño en un entorno configurable que permite comparar estrategias de asignación y migración bajo condiciones homogéneas y heterogéneas.

2. INVESTIGACIÓN Y DESARROLLO

Metodología

El desarrollo se organiza en fases de modelado conceptual, implementación modular, diseño de escenarios y validación experimental. Se analizan configuraciones homogéneas y heterogéneas variando parámetros de CPU, memoria y ancho de banda, conforme a los principios de diseño experimental en simulación.

La comparación de políticas de asignación se realiza mediante análisis cuantitativo de eficiencia y estabilidad, considerando métricas de desempeño típicamente empleadas en evaluación de sistemas distribuidos.

La evaluación pedagógica contempla medición del desempeño académico previo y posterior a la incorporación del simulador, en línea con estudios sobre aprendizaje activo en ingeniería.

Arquitectura del simulador

La arquitectura del simulador se estructura en módulos de definición de nodos (CPU, memoria, procesos, recursos, solicitudes), gestión de procesos y recursos, motor de simulación y módulo de métricas. El diseño modular responde a principios de modelado de sistemas distribuidos descritos en la literatura especializada.

El motor permite modelar migración de procesos, umbrales de sobrecarga y políticas conservadoras y dinámicas de asignación. La evaluación comparativa se realiza mediante indicadores cuantitativos tales como carga computacional por nodo, índice de balanceo (desviación estándar de carga) y costo de transferencia, siguiendo los enfoques de análisis de desempeño en simulación discreta.

Asimismo, el proyecto se articula con trabajos previos del Grupo en modelado de políticas de asignación y análisis experimental en entornos distribuidos [10].

Innovación educativa

La propuesta transforma conceptos abstractos en experiencias experimentales medibles. Al permitir modificar parámetros del sistema y observar efectos en tiempo real, el simulador

contribuye al aprendizaje activo y exploratorio, estrategia respaldada empíricamente en educación en ingeniería.

Desde el punto de vista tecnológico, integra modelado computacional, análisis cuantitativo y visualización de métricas en un entorno accesible sin requerir infraestructura física compleja, siguiendo enfoques contemporáneos de simulación aplicada.

3. RESULTADOS ESPERADOS

En el plano científico y técnico, se espera diseñar e implementar un simulador modular funcional capaz de modelar arquitecturas distribuidas con nodos heterogéneos, procesos concurrentes y recursos compartidos, permitiendo configurar distintos escenarios de asignación y migración. El sistema deberá registrar métricas cuantitativas tales como carga computacional por nodo, tiempos de respuesta, costo de transferencia y grado de balanceo, posibilitando la comparación objetiva de políticas conservadoras y dinámicas de gestión de recursos. Se prevé validar experimentalmente el comportamiento del motor de simulación mediante la ejecución de escenarios controlados, verificando consistencia, estabilidad y reproducibilidad de resultados.

En el plano metodológico y experimental, se espera obtener evidencia cuantitativa que permita analizar el impacto de distintos umbrales de sobrecarga y estrategias de migración sobre el equilibrio del sistema. La comparación de configuraciones homogéneas y heterogéneas permitirá identificar patrones de comportamiento y condiciones de mayor eficiencia operativa, contribuyendo al análisis formal de políticas de asignación en entornos distribuidos.

En el plano pedagógico, se proyecta que la incorporación del simulador en asignaturas vinculadas a Sistemas Operativos y Sistemas Distribuidos mejore la comprensión conceptual de fenómenos como sincronización, exclusión mutua, migración de procesos y balanceo de carga. Se espera observar una reducción en errores conceptuales frecuentes y un incremento en el desempeño académico en evaluaciones vinculadas a modelado y análisis de sistemas distribuidos.

Finalmente, en el plano institucional y académico, el proyecto busca consolidar la línea de investigación en modelado y simulación del Grupo, generar publicaciones científicas derivadas de los resultados experimentales y fortalecer la articulación entre investigación y docencia en la Universidad Nacional del Chaco Austral.

4. RESULTADOS OBTENIDOS

En el marco del desarrollo del simulador educativo para modelado de sistemas distribuidos, se logró definir e implementar formalmente el operador de agregación correspondiente al Escenario 1, orientado a modelar el acceso de procesos a recursos compartidos bajo modalidad de exclusión mutua en un entorno homogéneo. El operador integra la migración de procesos como estrategia de balanceo dinámico de la carga de trabajo entre nodos distribuidos. El módulo correspondiente fue sometido a validación experimental mediante la ejecución de escenarios controlados, verificándose consistencia en la asignación y estabilidad en la distribución de carga.

De manera complementaria, se definió el modelo de decisión asociado al Escenario 1, estableciendo criterios formales para la selección de nodos destino en función de indicadores de carga computacional y disponibilidad de recursos. El modelo fue implementado como componente independiente del operador de agregación, permitiendo evaluar su desempeño en condiciones homogéneas. La validación experimental confirmó la coherencia entre el modelo teórico y los resultados observados en la simulación.

En una segunda etapa, se definió e implementó el operador de agregación correspondiente al Escenario 2, extendiendo el modelo anterior a un entorno heterogéneo. En este escenario, los nodos presentan capacidades diferenciadas de procesamiento y almacenamiento, lo que introduce mayor complejidad en la asignación y migración de procesos. El operador mantiene la modalidad

de exclusión mutua, incorporando además criterios adaptativos para la redistribución de carga en función de la heterogeneidad del sistema. El módulo fue desarrollado y validado mediante pruebas experimentales comparativas.

Finalmente, se definió el modelo de decisión correspondiente al Escenario 2. Su implementación permitió analizar el comportamiento del sistema bajo configuraciones heterogéneas, verificándose mejoras en la estabilidad del balanceo respecto a estrategias no adaptativas. La validación experimental confirmó la viabilidad operativa del modelo.

En conjunto, los resultados obtenidos evidencian la funcionalidad del motor de simulación, la correcta implementación de los operadores de agregación y la coherencia entre los modelos de decisión definidos teóricamente y su comportamiento en ejecución experimental.

Conclusiones

El proyecto posiciona a UNCAUS como referente regional en innovación pedagógica aplicada a sistemas distribuidos, integrando modelado computacional y experimentación controlada. La articulación entre investigación en arquitectura distribuida y estrategias de aprendizaje activo, permite consolidar un enfoque formativo basado en evidencia y análisis cuantitativo del desempeño.

5. FORMACIÓN DE RECURSOS HUMANOS

El proyecto contribuye significativamente a la formación de recursos humanos en el ámbito de la Ingeniería en Sistemas de Información de la Universidad Nacional del Chaco Austral, fortaleciendo la articulación entre investigación, docencia y desarrollo tecnológico.

En el plano académico, el desarrollo del simulador involucra a docentes investigadores y a egresados en tareas de modelado conceptual, diseño de arquitectura modular, implementación de algoritmos de asignación y validación experimental de escenarios distribuidos. Esta participación activa permite la formación en áreas estratégicas como sistemas distribuidos, modelado computacional, simulación discreta y análisis de desempeño, adquiriendo competencias técnicas vinculadas al diseño y evaluación de políticas de gestión de recursos.

Asimismo, el proyecto promueve la formación en metodologías de investigación aplicada, incluyendo diseño experimental, validación de modelos, análisis comparativo de resultados y elaboración de publicaciones científicas. La participación en instancias de divulgación académica, como congresos y workshops especializados, favorece el desarrollo de capacidades de comunicación científica y producción académica.

6. BIBLIOGRAFÍA

[1] Tanenbaum, A. S., & Van Steen, M. (2017). Distributed systems: Principles and paradigms (3rd ed.). Pearson.

[2] Coulouris, G., Dollimore, J., Kindberg, T., & Blair, G. (2011). Distributed systems: Concepts and design (5th ed.). Addison-Wesley.

[3] Lamport, L. (1978). Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM*, 21(7), 558–565. https://doi.org/10.1145/359545.359563

[4] Ricart, G., & Agrawala, A. K. (1981). An optimal algorithm for mutual exclusion in computer networks. *Communications of the ACM*, 24(1), 9–17. https://doi.org/10.1145/358527.358537

[5] Casavant, T. L., & Kuhl, J. G. (1988). A taxonomy of scheduling in general-purpose distributed computing systems. *IEEE Transactions on Software Engineering*, 14(2), 141–154. https://doi.org/10.1109/32.4634

[6] Shivaratri, N. G., Krueger, P., & Singhal, M. (1992). Load distributing for locally distributed systems. *Computer*, 25(12), 33–44. https://doi.org/10.1109/2.179516

[7] Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2010). Discrete-event system simulation (5th ed.). Pearson.

[8] Law, A. M. (2015). Simulation modeling and analysis (5th ed.). McGraw-Hill Education.

[9] Prince, M. (2004). Does active learning work? *A review of the research. Journal of Engineering Education*, 93(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.

[10] La Red Martínez, D. L., Gerzel, S. M., Aguirre, C. E., Cáceres, R. A., Berecoechea Galarza, J. A., & Vallejo, S. (2022). Modelado y simulación de políticas de asignación en entornos distribuidos. *En XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022)* (pp. 1–5).

DIDÁCTICA DE LAS CIENCIAS DE LA COMPUTACIÓN: EXPERIENCIAS EN DISTINTOS NIVELES EDUCATIVOS

Ana Casali[1,2] y Natalia Colussi[1]
[1] Facultad de Ciencias Exactas, Ingeniería y Agrimensura,
Universidad Nacional de Rosario, Argentina (UNR)
{acasali, colussi}@fceia.unr.edu.ar
[2] Centro Internacional Franco Argentino de Ciencias de la Información y de Sistemas
(CIFASIS: CONICET-UNR)

RESUMEN

En esta línea de investigación se desarrollan diferentes estrategias en didáctica de las Ciencias de la Computación (CC) para los distintos niveles educativos. Sosteniendo que la educación en CC es fundamental desde el nivel inicial para formar ciudadanos/as digitales con pensamiento crítico, la capacitación de formadores en el nivel primario ha sido muy relevante para el equipo. Las estrategias utilizadas combinan enfoque desconectado (unplugged) con actividades que implican la programación de computadoras y otros dispositivos digitales. Frente al alto desarrollo de las aplicaciones de Inteligencia Artificial (IA) en estos últimos años, se considera que la introducción a esta tecnología y sus fundamentos, debe incluirse en la formación en CC, de toda persona. También, se analiza el impacto del uso de herramientas de la IA como los Chatbots para la enseñanza del pensamiento computacional y la programación en la universidad. Dentro de este marco, por un lado, se continúa con el trabajo de formación de docentes del nivel primario realizado en la Regional VI de Santa Fe (Argentina), analizando los logros y desafíos que enfrentan los/as docentes al realizar experiencias áulicas en procesos de enseñanza de CC. En relación con el nivel universitario, se presenta un proyecto sobre la enseñanza y aprendizaje de las primeras materias de programación a estudiantes de Ciencias Exactas en la Facultad de Ciencias Exactas, Ingeniería y Agrimensura (FCEIA-UNR), donde se trabaja con Aprendizaje Basado en Proyectos y Problemas (ABP). Además, dentro de la enseñanza de programación en carreras STEM se plantea un proyecto innovador que incorpore Chatbot basados en grandes modelos de lenguajes (LLM) con el fin de poder evaluar cómo y cuándo usar estos asistentes. Todas estas experiencias se focalizan en mejorar la didáctica de la enseñanza y aprendizaje de las CC en los distintos niveles educativos.

Palabras clave: Programación, Pensamiento Computacional, Aplicaciones de IA, Formación Docente nivel Primario, Carreras STEM.

CONTEXTO

Esta línea de I+D se está llevando a cabo a través de proyectos que comparten algunos de sus miembros docentes e investigadores, conformando una red de colaboración:

- PID-UNR 80020220700134UR “Didáctica de las Ciencias de la Computación en la Educación Primaria”. (2023-2026)

  Directora A. Casali.

- Proyecto de Tesis del Ing Alex Morales Acosta, Doctorado en Ciencias de la Educación, UNR (2025-2027).

  Directora A. Casali

- PPCT-UNR 80120230300052UR “Aprendizaje Basado en Proyectos y Problemas para Cátedras de Programación en el Ciclo Básico Universitario”. (2024-2027)

  Directora N. Colussi.

1. INTRODUCCIÓN

En un mundo atravesado por computadoras, a fin de formar a ciudadanos/as capaces de comprender la tecnología que nos rodea y de ser usuarios/as críticos/as, crece la relevancia de introducir la enseñanza y aprendizaje de las Ciencias de la Computación (CC) en todos los niveles de la escolaridad. Es importante que los/as estudiantes en formación comprendan desde temprana edad, cómo funcionan las computadoras, cómo procesan la información y que puedan trabajar en la formulación de problemas de forma que un computador pueda ejecutarla. Este modo de pensar se denomina Pensamiento Computacional (PC) y se plantea que su aprendizaje beneficia no sólo a los futuros informáticos, sino a toda la sociedad. Hoy, ante el avance de la Inteligencia Artificial (IA) cuyas aplicaciones impactan cada vez más en las distintas actividades que realizamos, la necesidad de formación en CC y en comprender los fundamentos de esta tecnología, se profundiza.

En el contexto latinoamericano hay iniciativas puntuales y distinto nivel de avance en diferentes países, siendo necesario implementar políticas públicas al respecto [3]. En nuestro país, el Consejo Federal de Educación determinó en 2015 que el aprendizaje de la programación es de importancia estratégica para el Sistema Educativo Nacional durante la escolaridad obligatoria (Res. CFE Nº 263/15, 2015) y en 2018 se aprobaron los "Núcleos de Aprendizajes Prioritarios de Educación Digital, Programación y Robótica" (Res. CFE Nº 343/18, 2018). A partir de dichas resoluciones, cada provincia es responsable de su implementación y es posible señalar que desde el marco institucional existen dos aspectos clave para introducir la enseñanza de las CC en el sistema educativo argentino. Por un lado, la necesidad de realizar una modificación curricular y por otro, la implementación de la formación docente en el área. Teniendo en cuenta esta realidad, desde la UNR en los últimos años, se han llevado a cabo distintos proyectos de formación de docentes del nivel primario en CC atendiendo a las particularidades del contexto.

Por otra parte, se presenta el problema del acceso y retención de estudiantes a las carreras vinculadas a las carreras informáticas [6]. Por esto, en las universidades debemos buscar diferentes estrategias para que el alumnado pueda tener mayor motivación y mejores herramientas para continuar el programa de formación al cual pertenece.

En esta dirección, a nivel universitario se trabaja en un cambio de paradigma de aprendizaje, enseñanza y evaluación a través de la aplicación del Aprendizaje Basado en Proyectos y Problemas (ABP) en el aula del redictado de las cátedras de Programación I y Programación II para las carreras universitarias de ciencias exactas.

También, ante los avances de la IA, el uso de los Chatbot avanzados para los procesos de enseñanza está siendo debatido en toda la educación [11]. A partir de un trabajo doctoral se plantea estudiar este impacto en la enseñanza del PC y la programación, desarrollando una estrategia pedagógica que implemente estos Chatbot como recurso educativo en los primeros trayectos formativos de la programación y algoritmia de carreras de ingeniería informática.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Formación Docente de Nivel Primario en Didáctica de las CC

El equipo, viene trabajando desde 2016 en el desarrollo de distintos trayectos de formación docente del nivel primario en CC, en el marco de distintos convenios de colaboración con la Fundación Sadosky. En primer lugar, se ha trabajado en el desarrollo de la *"Especialización Docente de Nivel Superior en Didáctica de las Ciencias de la Computación: Aprendizaje y Enseñanza del Pensamiento Computacional y la Programación en el Nivel Primario"* (400 hs) la cual fue aprobada por el

Ministerio de Educación de la Provincia de Santa Fe y se dictó una primera cohorte (2017-2019) con resultados muy positivos [2,4]. A través de convenios posteriores se dictaron los siguientes cursos de formación: "La programación y su Didáctica Nivel I" (2021 y 2023), "La programación y su Didáctica Nivel II" (2022) y "La ciudadanía en un mundo atravesado por computadoras" (2023). Este último curso aborda contenidos de organización de computadoras, redes, internet, datos e inteligencia artificial, con el fin de que los/as docentes puedan generar y liderar procesos de reflexión sobre el impacto que estas tecnologías ejercen en nuestras vidas. Destacamos el alto interés de los/as docentes en comprender las tecnologías subyacentes a las aplicaciones de la IA más utilizadas.

Recientemente, en diciembre de 2025, hemos realizado un relevamiento a todos los cursantes, a fines de recabar información sobre sus experiencias áulicas posteriores a sus formaciones en CC. De las respuestas recibidas (#36) destacamos que un 94% sigue trabajando en distintos grados de la docencia primaria, la mayoría de gestión pública y un 85% en las áreas de Tecnología Informática, Computación y afines. Pese a que muchos de ellos refieren problemas de equipamiento informático en sus escuelas, un 53% pudo llevar actividades de aprendizaje de CC al aula, donde la mayoría ha realizado un trabajo mixto (conectado y desconectado). Para socializar estas experiencias, hemos creado tableros que permiten de forma sintética compartir las experiencias áulicas, así como también los desafíos encontrados. La totalidad de los encuestados manifestó que le interesa realizar nuevas formaciones en CC.

Se destaca el gran interés que han tenido los distintos cursos y la necesidad de seguir trabajando en este nivel educativo. Por un lado, hay que continuar revisando la didáctica con que se presentan los distintos contenidos para que se ajuste adecuadamente al nivel primario. Por otro lado, surge la necesidad de contar con políticas y financiamiento para ampliar la capacitación en CC a todos los docentes de la provincia de Santa Fe, en especial a los del área de educación tecnológica.

Estrategias Didácticas para la Enseñanza de la Programación y el PC en Nivel Universitario en Carreras STEM

i) Experiencia en el Ciclo Inicial en carreras Ciencias Exactas

Esta línea de investigación aplica estrategias didácticas activas, centradas en el/a estudiante que permitan un abordaje efectivo a las problemáticas de la enseñanza, aprendizaje y evaluación de los/as estudiantes de primer año de las cátedras introductorias de programación en las carrera de Licenciaturas en Ciencias de la Computación y Matemática. La presentación y evaluación de los contenidos propios de las materias Programación I y II se rediseñaron para aplicar la estrategia didáctica de Aprendizaje Basado en Proyectos y Problemas (ABP), realizando proyectos grupales de programación que invitaron a los estudiantes a cambiar su rol dentro del aula, permitiéndoles un aprendizaje significativo, de mayor autonomía con indagación guiada por los docentes, controles semanales de los avances, despertando curiosidad, colaboración entre pares y creatividad en todo el proceso.

Para la evaluación se utilizaron exposiciones grupales ante los/as compañeros/as, participación en eventos de la carrera LCC, como así también videos expositivos de los distintos grupos de trabajo, recopilados en el sitio llamado Vidriera de Exposición de Proyectos[1], el cual funciona actualmente como repositorio y registro de toda la experiencia [5].

Para medir la receptividad de los/as estudiantes respecto de la forma de aprendizaje se utilizaron cuestionarios grupales e individuales que eran entregados junto a cada proyecto, donde se registraron todas las opiniones para realizar parte de la evaluación didáctica de todo implementado en las cátedras [7, 8, 10].

[1] sites.google.com/view/vidriera-proyectos-fceia

ii) Estrategias pedagógicas implementando Chatbot avanzados para la enseñanza de la programación

En este proyecto se quiere desarrollar y evaluar una estrategia pedagógica que implemente el uso de Chatbots como recurso educativo para el aprendizaje del PC y la programación en la formación de profesionales de carreras de informática [9]. Se tomará como caso de estudio carreras de Ingeniería Informática o similares de instituciones de educación superior en Sincelejo, Sucre, Colombia.

Se han realizado dos tareas fundamentales. Por un lado, se ha realizado una búsqueda sistemática en la bibliografía sobre el uso de herramientas de la IA en el desarrollo del PC y la programación, registrando las experiencias, sus ventajas y también los desafíos que presentan estas tecnologías. Por otro lado, utilizando el test de Marcos Román [1] se ha realizado una evaluación del desarrollo del PC en estudiantes de distintos cursos de carreras vinculadas a Ingeniería Informática en donde se realizará la experimentación. Se analizan los aspectos claves de la estrategia a proponer para el desarrollo del PC, a partir de las dificultades de los estudiantes y de los resultados que se pueden obtener de las posibles interacciones con este tipo de Chatbot [9]. El diseño de la estrategia pedagógica para los diferentes contextos educativos, se construye en un trabajo colaborativo con los/as profesores/as que la implementarán. Se evaluarán los resultados aplicando distintos instrumentos que permitan medir aspectos tales como: el desarrollo del PC, percepción y motivación de los/as estudiantes en los cursos seleccionados.

3. RESULTADOS

Respecto a la formación de docentes del nivel primario, se ha diseñado e implementado una formación superior para docentes del nivel primario en CC, la primera en la provincia de Santa Fe. Se han dictado distintas ediciones de los cursos “La Programación y su Didáctica I y II” y “La ciudadanía en un mundo atravesado por computadoras” en convenio con la Fundación Sadosky donde más de 300 docentes han transitado por los distintos cursos. Las opiniones de los cursantes en cuanto a la didáctica y contenidos de distintos módulos han sido altamente positivas [2, 4]. Dado la falta de financiamiento a través de la Fundación Sadosky para continuar con el dictado de cursos, nos encontramos gestionando otras fuentes de financiamiento enfocadas en la formación en IA. También, sostenemos el contacto con los/as docentes que cursaron alguna formación a través de distintos medios, para compartir tanto sus experiencias áulicas de enseñanza de CC, como los desafíos que enfrentan en su contexto.

Destacamos además, que en este proceso se ha conformado un equipo docente universitario e interdisciplinario, competente en la formación docente en CC para el nivel primario que trabaja en la adecuación de los contenidos y su didáctica para responder a los requerimientos del nivel educativo. Miembros de este equipo han participado en el desarrollo de la Propuesta de Sadosky para el Ministerio de Educación de la Provincia de Santa Fe, de una reforma curricular de la educación primaria y secundaria donde se incluyan contenidos de CC (NAP de Educación Digital, Programación y Robótica).

En cuanto al ciclo inicial universitario en ciencias exactas, es posible señalar que la implementación del ABP en las cátedras de Programación I y II ha sido una experiencia altamente enriquecedora tanto para estudiantes como para docentes. Los resultados de las encuestas muestran que las propuestas fueron percibidas como innovadoras, interesantes y significativas, favoreciendo un aprendizaje genuino. La integración de metodologías activas de aprendizaje centradas en la colaboración entre los/as estudiantes, prepara a los/as alumnos/as para los desafíos reales del desarrollo de software; donde la interacción y el consenso son esenciales. No sólo se desarrollan habilidades técnicas, sino que también se enriquece la experiencia individual y colectiva, fomentando la colaboración, el aprendizaje mutuo y la creación de soluciones conjuntas.

Respecto a la inclusión del uso de Chatbots en los primeros cursos de programación, a partir de relevamientos iniciales, se elabora junto a un equipo docente de la Universidad de Sincelejo, Colombia, la estrategia a utilizar.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por las Dra Ana Casali, la Lic. Natalia Colussi, la Dra Pamela Viale, el Lic. Hernán Galardi y la Prof. Delfina Martín de la FCEIA-UNR. Dentro del marco de esta línea de I+D se desarrolla una tesis de doctorado en Ciencias de la Educación (UNR) del Msc. Alex Morales Acosta quien desarrolla su tesis "Estrategia pedagógica implementando Chatbot avanzados como recurso educativo para el desarrollo del pensamiento computacional en la formación de profesionales en ingeniería de sistemas y afines" dirigido por Ana Casali.

Cabe destacar además, la formación que se lleva a cabo de docentes del Departamento de Ciencias de la Computación (FCEIA-UNR) en la enseñanza de la disciplina a docentes de nivel primario y secundario, es un trabajo interdisciplinario que incluye a la Dra. Natalia Monjelat como especialista en ciencias de la educación.

5. BIBLIOGRAFÍA

[1] Román-González M., Perez-González J. C., Jiménez-Fernández C., Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. Computers in Human Behavior, 72: 678-691, 2017

[2] Casali A., Zanarini D., Monjelat N & San Martín P., Primary Level Teachers Training in Computer Science: Experience in the Argentine Context, In: Pesado P., Arroyo M. (eds) Computer Science–CACIC 2019. CCIS 1184, pp. 389-404, Springer 2019

[3] Brackmann C., Couto Barone D., Casali A., Román-Gonzalez M., Panorama Global Da Adoção Do Pensamento Computacional (Cap.3) Computação na educação básica: fundamentos e experiências (Ed. Raabe,Zorzo e Blikstein), pp. 30-47, 2020.

[4] Casali A., San Martín P., Monjelat N., Viale P., Experiencias y aprendizajes del trayecto proyectual en una Especialización Docente en Didáctica de las Ciencias de la Computación, Revista Teyet no. 27, pp. 9-19, 2020.

[5] Colussi N., Viale P., Monjelat N., Vidriera de proyectos: una modalidad de evaluación posible en tiempos de virtualidad. En II Workshop de Innovación y Transformación Educativa. Transformación Digital. Desafíos de la Educación Superior. 2021.

[6] Marino V., Sustas S. E., Quartulli D., Curcio J. ¿Por qué estudiamos informática? Indagación sobre trayectorias universitarias: instituciones, estudiantes, género y trabajo. Fundación Sadosky, 2021

[7] Colussi N., Monjelat N., Viale P., El ABP en el redictado de materias de programación: una experiencia en la educación superior. En Proceedings JADICC, 2022.

[8] Colussi N., Monjelat N., Proyectos de programación grupal en Python: ABP en el redictado del ciclo inicial universitario. (JAIIO), 2023.

[9] Morales Acosta A. & Casali A.. Developing Computational Thinking in Computer Science Careers: a Strategy Based on AI Chatbots. Libro de Artículos, Congreso Expotecnología. Facultad de Ciencias de la Ingeniería de UNITECNA, Colombia, 2024

[10] Colussi N., Monjelat N., ABP+DataSet: proyectos grupales de programación en primer año de la universidad. (JAIIO), 2025.

[11] Giannini S., AI and the future of education: disruptions, dilemmas and directions, UNESCO, 2025 https://www.unesco.org/en/articles/ai-and-future-education-disruptions-dilemmas-and-directions

Enseñanza de la disciplina de Ingeniería de Requerimientos en Carrera de Grado en la era de la Inteligencia Artificial Generativa

Marcela Samela[1], Constanza Zaninetti[1], Gabriela Caliva[2], Gabriela Figallo[2], Maximiliano Bonaccorsi[1], Alejandro Sartorio[1], Daniel E. Alonso[2], Nicolás Battaglia[1], Leandro Antonelli[13]

[1] CAETI, Facultad de Tecnología Informática, Universidad Abierta Interamericana, CABA, Argentina.
[2] CAETI, Facultad de Tecnología Informática, Universidad Abierta Interamericana, Rosario, Argentina.
[3]Lifia, Facultad. de Informática, Universidad Nacional de La Plata, Calle 50 y 120 - La Plata - Bs. As. – Argentina.

{marcela.samela, constanzasoledad.zaninetti, gabrielalorena.caliva, gabriella.figallo, maximilianoleonardo.bonaccorsi, alejandro.sartorio, danieleduardo.alonso, nicolas.battaglia, Leandro.antonelli}@uai.edu.ar

RESUMEN

Este artículo describe el proyecto de investigación llevado a cabo por el cuerpo docente del curso de Ingeniería de Requerimientos de la carrera Ingeniería en Sistemas Informáticos en la Universidad Abierta Interamericana con el fin de actualizar el curso conforme a los desafíos y necesidades tanto por la forma de motivar a los alumnos, como por la necesidad de actualizar los contenidos conforme a la revolución que ocasiona la Inteligencia Artificial. Este proyecto está integrado por un grupo numeroso de docentes, los cuales poseen una muy variada experiencia en la industria, en docencia y en investigación. Es por eso que el proyecto se basa en 3 elementos: el cuerpo de conocimiento tradicional, la experiencia acumulada de los docentes y pruebas de conceptos sobre nuevas ideas. Como resultado de este proyecto, se elaborará un nuevo diseño de curso para los alumnos del segundo semestre de 2026 el cual se implementará en la plataforma de e-learning UAI Online ultra.

Palabras clave: Ingeniería de requerimientos, curso de grado, diseño de curso, plataforma de e-learning.

CONTEXTO

El proyecto de I/D que se presenta en este artículo lo realiza el cuerpo docente de la cátedra de Ingeniería de Requerimientos de la carrera Ingeniería en Sistemas Informáticos en la Universidad Abierta Interamericana en sus sedes de Buenos Aires y Rosario.

1. INTRODUCCIÓN

La Ingeniería de requerimientos es la etapa en donde se capturan las necesidades, deseos y expectativas de los clientes, para especificar la funcionalidad que se desea desarrollar [1]. En esta etapa también se captura el conocimiento que debe ser incorporado en el sistema. Este conocimiento es crucial para que el sistema a desarrollar pueda cumplir con la funcionalidad correcta.

La Ingeniería de requerimientos es una de las etapas más importantes en el desarrollo del software. Los requerimientos son el insumo necesario para realizar las pruebas del software y por consiguiente, asegurar la calidad del mismo. El modelo de desarrollo en V [2], coloca a los requerimientos como punto de partida y los utiliza para realizar las pruebas de aceptación del usuario. Por otro lado, la ingeniería de requerimientos es clave, dado que los errores en los requerimientos requieren un esfuerzo de retrabajo muy grande para corregirlos. Cuando se comete un error en los requerimientos y se detecta una vez que el software ha sido entregado, se estima que el costo relativo de repararlo es de entre 100 a 200 veces [3].

Luego de la crisis del software, con la definición del método en cascada, se reconoce

a los requerimientos como una etapa en el ciclo de vida, es decir, se lo reconoce como un paso (uno de los primeros) para poder desarrollar software de una forma metódica. La disciplina de Requerimientos se formalizó en los años 90 cuando los autores Loucopoulos y Karakostas editaron el libro *System Requirements Engineering* convirtiéndose en un hito al sentar las bases de la disciplina [1]. Con la difusión de los métodos ágiles, la ingeniería de requerimientos toma otro rol y estilo. En algunos tipos de desarrollo, en donde el costo de oportunidad y el *time to market* es más importante que el consolidar una especificación de requerimientos completa y precisa para planificar y organizar un desarrollo de software, se adoptan métodos agiles [4] como Scrum por ejemplo que priorizan un rol que se ocupa de obtener el conocimiento y comunicarlo por sobre escribir una especificación.

Actualmente, estamos transitando lo que se conoce como la industria 4.0, es decir, la cuarta revolución industrial, haciendo referencia a la primera revolución industrial que se desarrolló en Inglaterra entre mitad del 1700 y del 1800 con la automatización a partir de la máquina a vapor. La cuarta revolución industrial (o industria 4.0) que estamos transitando actualmente se basa en pilares como: *Internet of Things* (IoT), el uso de grandes volúmenes de datos (Big Data), e Inteligencia Artificial (IA) entre otras herramientas. Todas estas componentes vuelven a dar relevancia a la necesidad de realizar una especificación de requerimientos tradicional, completa y rigurosa. Respecto de IoT, el desarrollo de sistemas ciber-físicos, los requerimientos son más críticos en el mundo físico que en mundo del software, dado que hacer o corregir algo en el mundo tangible (físico) tiene más costo (materiales, traslados, dependencias) que en mundo intangible. BigData abre un mundo para la disciplina de requerimientos a partir de la subdisciplina de *Data Driven Requirements Engineering (DDRE)* [5] que trae una visión muy diferente del enfoque tradicional como el orientado a la funcionalidad. En el enfoque tradicional, los requerimientos funcionales son la estrella al pensar los sistemas una función que transforma datos de entradas en datos de salida. Al sistema se lo ve como una automatización de esta transformación. Por el contrario, en el mundo de los datos, las organización están abrumadas por grandes volúmenes de datos, que es necesario entender y organizar para poder sacarles provecho y poder predecir o inferir nuevo datos y conocimiento a partir de los mismos. Por su parte, la IA genera dos desafíos. Por un lado, la gestión de requerimientos para sistemas que usan IA. Esto está muy relacionado con DDRE, dado que se necesitan datos de calidad para entrenar modelos. Y por otro lado, se discute el uso de la IA como herramientas para gestionar requerimientos, es decir, para descubrir, analizar, modelar y evaluar requerimientos.

La enseñanza en la disciplina de requerimientos fue acompañando su evolución histórica y el contexto tecnológico en general. En los inicios fue una unidad en los cursos de Ingeniería de Software. Luego se convirtió en cursos de postgrado. Y en la actualidad se dicta como curso de grado. Si bien es un gran avance para la disciplina, se presenta el desafío que los alumnos necesitan conocer la complejidad de los proyectos de desarrollo para poder entender la necesidad y la pertinencia de la disciplina. Durante una carrera de grado, el alumno no realiza proyectos complejos como parte de su formación. Y si tuviera experiencia en la industria, seguramente se desempeñará en roles iniciales (junior) por lo cual, tampoco enfrentará la complejidad que justifica *hacer* requerimientos.

Es así que los desafíos actuales de la enseñanza de la Ingeniería de Requerimientos son varios. Por un lado cuesta transmitir la importancia de la disciplina si el alumno no participó de un gran proyecto de desarrollo. Si esto no pasa, suelen considerar a los requerimientos como trabajo burocrático que frena al desarrollo. El alumno de carreras de Informática y Tecnología ingresa por e interés en los frameworks y lenguajes de programación, no lo hace para desarrollar

capacidades comunicativas y plasmarlas a través de un archivo de texto. Por otro lado, es muy difícil llevar adelante un curso en donde los alumnos deben realizar experiencia en requerimientos de forma realista, dada la limitación de tiempo para hacer el proyecto durante un curso de carrera de grado. Y por el desafía que implica que el proyecto sea realista considerando cantidad de participantes y dinámica de trabajo. Es muy difícil llevar una actividad de formación, a la vez que se involucra un gran número de participantes y se gestiona un proyecto de software. Finalmente, la disciplina de requerimientos implica comunicar de forma precisa mucha información. En la actualidad la comunicación es cada vez más breve y efímera, no se suele organizar y brindar un gran volumen, sino que se recurre a conversaciones en donde en un periodo muy breve se intercambia información, la cual, tiene una estabilidad de poco tiempo. Además, actualmente se está haciendo un uso masivo de IA Generativa. Por lo cual, las habilidades de escritura (además de lectura y análisis) se están viendo afectadas muy negativamente. Todas esto contexto va en detrimento de las capacidades para desarrollar requerimientos.

Es por eso, que este proyecto busca ajustar y actualizar el curso de Ingeniería de Requerimientos de la carrera Ingeniería en Sistemas Informáticos de la Universidad Abierta Interamericana en sus sedes de Buenos Aires y Rosario a la vez de implementarlo en una plataforma de e-learning UAI Online ultra.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La estrategia general para llevar a cabo este proyecto es la siguiente. En primer lugar se plantea analizar el curso actual, tanto el diseño como su implementación en la plataforma de e-learning. Luego, se propone realizar workshops de trabajo a través de sesiones de *brainstorming* como así también *focus* groups involucrando a todo el cuerpo docente para a partir del estado actual, identificar mejoras y adecuaciones, las cuales se piensan implementar a través de píldoras educativas. Las píldoras estarán categorizadas en 3 tipos: motivación, actividades prácticas y fundamentos teóricos. Esta tipología busca hacer frente a las debilidades identificadas. Por un lado para motivar la necesidad de la disciplina. Por otro lado, realizar actividades prácticas para brindar actividades realistas. Y finalmente, brindar los fundamentos teóricos necesarios para enriquecer el abanico de métodos, técnicas y herramientas que los alumnos deben incorporar, y a partir del criterio que deben desarrollar como resultado del curso, pueden determinar que utilizar en función del contexto en que se encuentren. Finalmente, se busca explorar la eficacia de cada píldora a través de experiencias concretas con alumnos.

3. RESULTADOS ESPERADOS

Se espera tener resultados a tres niveles: pedagógico, académico y formación de talento humano. A nivel pedagógico se espera obtener el rediseño del curso como así también actualizar su implementación en la plataforma de e-learning con la definición de las nuevas píldoras educativas. A nivel académico se espera difundir la experiencia a través de publicaciones científicas. Finalmente, respecto de formación del talento humano, se espera unificar el conocimiento de todo el cuerpo docente a través de los *workshops*.

4. FORMACIÓN DE RECURSOS HUMANOS

Se espera formar el cuerpo docente del curso de Ingeniería de Requerimientos de la carrera Ingeniería en Sistemas Informáticos en la Universidad Abierta Interamericana, el cual está compuesto por 8 docentes.

Además, el punto más importante, es el de desarrollar las habilidades de los estudiantes. El curso de requerimientos busca enseñar métodos, técnicas y herramientas, pero en forma tácita se desarrollan las capacidades de comunicación, escucha activa, análisis y pensamiento crítico. Estas habilidades son fundamentales en una realidad en donde se

confía ciegamente a las respuestas de que brindar un chatbot de IA.

5. REFERENCIAS

[1] Loucopoulos, P., Karakostas, V.: System Requirements Engineering, McGraw-Hill, London (1995).

[2] Forsberg, K., Mooz, H.: The relationship of systems engineering to the project cycle. Proceedings of the First Annual Symposium of the National Council on Systems Engineering (NCOSE) (1991).

[3] Boehm, B. W.: Software Engineering Economics. Englewood Cliffs, NJ: Prentice-Hall (1981).

[4] Manifiesto por el Desarrollo Ágil de Software https://agilemanifesto.org/iso/es/manifesto.html, último acceso 2026/02/16.

[5] Lim, S., Henriksson, A., Zdravkovic, J.: Data-Driven Requirements Elicitation: A Systematic Literature Review. SN Computer Science, 2, 16. https://doi.org/10.1007/s42979-020-00416-4 (2021).

ENSEÑANZA DE REQUERIMIENTOS: EXPERIENCIA CON CLIENTES REALES.

Moyano Ezequiel, Ivan D'Uva, Daniel Aguil Mallea
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: +54-2901-432403
{emoyano, iduva, daguil}@untdf.edu.ar

Resumen

Los procesos de enseñanza-aprendizaje se caracterizan por el papel determinante del docente en identificar, planificar e instrumentar estrategias en la formación de los estudiantes. La didáctica exige la necesidad de implementar estrategias centradas en el sujeto que aprende, no solo en la adquisición de conocimientos, sino el desarrollo de habilidades.

Enseñar ingeniería de requerimientos es fundamental debido a que es la primera, y más importante, etapa de todo paradigma de ingeniería de software. Su enseñanza se desarrolla sobre la base de un modelo conceptual de representaciones, supuestos y restricciones, cuyo producto o resultado sirva para especificar las necesidades, deseos y expectativas de los stakeholders. Su mayor dificultad radica en el proceso de transferencia de información de varias fuentes no siempre consistentes; ésta se incrementa cuando se considera el problema de la enseñanza en las técnicas de elicitación de requerimientos.

Las estrategias tradicionales (sin "interacción con un usuario") no son suficientes, resulta fundamental incorporar estrategias didácticas adecuadas, que permitan mejorar las prácticas y los aprendizajes en los procesos que intervienen en la ingeniería de requerimientos.

Se sintetizan los resultados obtenidos de la información recolectada y analizada luego de implementar estrategias utilizando clientes Reales.

Palabras clave: Estrategias Didácticas, Educación, Ingeniería de Software, Ingeniería de Requerimientos.

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la Universidad Nacional de Tierra del Fuego, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática vinculada a la innovación de educación en informática como en tecnología informática aplicada a la educación.

Actualmente esta línea se desarrolla en el marco de un proyecto de investigación (Oct/2023-Sept/2025) El mismo ha obtenido aval UNTDF PIDUNTDF-RESOL (REC-R/2) No:584/2023.

El proyecto tiene como caso de estudio la cátedra de Ingeniería de software I, donde en sus contenidos mínimos se especifica la enseñanza de ingeniería de requerimientos.

1 Introducción

El proceso de enseñanza-aprendizaje se caracteriza, entre otros aspectos, por reconocer el papel determinante del docente en la identificación, planificación e instrumentación de estrategias adecuadas en la formación de los estudiantes.

Las estrategias de enseñanza y de aprendizaje se encuentran involucradas de forma que el aprendizaje no sea solo la adquisición de conocimientos, sino también la búsqueda de "*...los medios que conducen a la solución de los problemas...*"[1] El término estrategia se utiliza cada vez con mayor frecuencia y presupone la planificación de acciones a corto, mediano y largo plazo; susceptibles al cambio, modificación y adecuación de sus alcances, permitiendo lograr la racionalidad de tiempo, recursos y esfuerzos. Son indiscutibles las ventajas que su adecuada

utilización puede ofrecer en los procesos educativos[2].

La didáctica exige la utilización de estrategias y métodos adecuados, actualmente centrada en el sujeto que aprende, lo cual exige enfocar la enseñanza como un proceso de orientación del aprendizaje, para que los estudiantes no sólo se apropien de los conocimientos, sino que desarrollen habilidades, formen valores y adquieran estrategias que le permitan actuar de forma independiente. La expresión *estrategia didáctica* presupone "...enfocar cómo enseña el docente y cómo aprende el alumno a través de un proceso..."[3].

Entre los métodos de enseñanza-aprendizaje que en los últimos tiempos están siendo reconocidos en la didáctica y, que los docentes deben considerar, se encuentran: el aprendizaje basado en problemas, el aprendizaje basado en proyectos, el método de casos, las simulaciones, etc., los cuales están íntimamente relacionados con la problemática a tratar. Los profesores expertos deben reflexionar continuamente acerca de cómo pueden enseñar aún mejor.

En la enseñanza específica de la Ingeniería de Requerimientos (IR), los futuros profesionales deben adquirir habilidades y aptitudes determinadas (abstracción, trabajo cooperativo y colaborativo, negociación, comunicación, toma de decisiones, comprensión del problema, entre otras) que poco se orientan en la enseñanza tradicional[4].

En la actualidad la enseñanza no se basa sólo en los conceptos impartidos por el docente, sino que se inclina hacia el estudiante como eje central de la clase[5].

Enseñar IR es fundamental debido a que es la primera, y más importante, etapa de todo paradigma de ingeniería de software, según Brooks *"La parte más difícil de construir un sistema de software es decidir qué construir"*. A través de este proceso se obtiene una Especificación de Requerimientos (ERS) que se desarrolla sobre la base de un modelo conceptual y va a representa abstracciones, supuestos y restricciones sobre el dominio de aplicación[6].

Para esto se necesitan técnicas que se incorporen en el proceso de enseñanza en el área de IR, con el propósito de crear estrategias de enseñanza que permitan abordar la complejidad inherente, la cual es una actividad de transferencia dificultosa cuando la fuente del conocimiento es un ser humano, y se requieren habilidades especiales del analista porque son muy sensibles a múltiples factores y restricciones (tales como el tiempo del usuario o dificultades psicológicas de la interacción)[7].

La necesidad mejorar las prácticas de IR en los jóvenes profesionales ha sido mencionado en la literatura y se ha desarrollado como tema de estudio en las carreras de grado en los últimos años.

Existen varios problemas para una buena enseñanza en las aulas de estas técnicas. El primero de ellos es la dificultad de ejecutar una práctica real: las técnicas suelen ser meramente descriptas y, en el mejor de los casos, sobre en un caso simplificado.

Otro problema importante es la subestimación por parte de personas de formación tecnológica de las técnicas "blandas" (metodologías que tratan con las interacciones humanas y los procesos sociales) que requiere la elicitación de requerimientos. Esta subestimación en muchos casos no se limita a los alumnos: se piensa que una sólida formación técnica es suficiente para asegurar el éxito de un proyecto de software.

Para la enseñanza de estos conceptos se percibe que las estrategias tradicionales (simples narrativas descriptivas de un contexto sin "interacción directa con un usuarios") no son suficientes. Resulta fundamental incorporar estrategias didácticas adecuadas a las particularidades y dificultades planteadas como casos pilotos que permitan mejorar las prácticas y los aprendizajes de los estudiantes en los procesos que intervienen en la IR[8] [9].

Las estrategias didácticas permiten en la enseñanza de las técnicas de elicitación y

especificación de requerimientos, minimizar los problemas planteados en los párrafos anteriores.

2 Línea de Investigación y Desarrollo

La línea de investigación general se enfoca en el estudio y desarrollo en estrategias didácticas para la enseñanza de IR, dentro de las temáticas vinculadas a la innovación en educación en informática.

El presente trabajo refleja la experiencia dada en el marco de los procesos de enseñanza y aprendizaje de IR en el espacio curricular Ingeniería de Software I de la Licenciatura en Sistemas; la cual constituye el primer contacto de los estudiantes con los conceptos vinculados al proceso de desarrollo de software, donde los tópicos de IR forman parte de sus contenidos mínimos.

La cátedra se dicta en un cuatrimestre con una importante carga horaria (10h semanales) destinando un mes al dictado de la unidad.

2.1 Problemática

Varios son los problemas que se plantean en la enseñanza de IR en las aulas, como se muestra en Oliveros et al.“...dificultad de ejecutar una práctica real: las técnicas son meramente descriptivas y, en el mejor de los casos, practicadas en un caso simplificado...”.

Hasta hace poco tiempo las prácticas se basaban en dos aspectos puntuales a desarrollar: la Elicitación y la Especificación de Requerimientos; en el primer caso se dividía la clase en 2 grandes grupos (“analistas” y “clientes) dada una situación/problemática especifica, el objetivo era que los “analistas” pudieran obtener los requerimientos a través de consultas a los “clientes”. Para el caso de la especificación, los/as estudiantes se dividían en grupos más pequeños con una narrativa (situación particular) donde debían ser capacs de especificar los requerimientos funcionales y no funcionales del futuro sistema.

Las estrategias, categorizadas como “técnicas tradicionales”, son muy habituales; sin embargo presentan serias dificultades: en cuanto a la necesidad de capacitar a los futuros profesionales en habilidades para diseñar y realizar entrevistas; como en la comprensión por parte del estudiante con respecto a los problemas que presentan las interacciones con usuarios reales al trabajar sobre casos abstractos o ficticios sin “interacción directa con un usuario”.

2.2 Recolección de evidencias

A los efectos de poder determinar la/s estrategia/s adecuadas, se procedió a indagar y relevar cómo se realizan las actuales prácticas en diversas universidades y/o Institutos donde se imparte la enseñanza de IR.

Se procedió a contactar a un importante grupo de universidades a través de sus representantes, a través de la REDUnci, logrando contactar 9 Universidades Argentinas (responsables de cátedras de Ingeniería de Requerimientos), de las cuales con 6 de ellas se realizaron encuentros online y conocer como desarrollaban estas prácticas.

Se analizaron las experiencias recolectadas a los efectos de poder determinar si alguna de ellas podría implementarse en nuestro contexto. En la mayoría de las universidades el tópico de ingeniería de requerimientos es una asignatura, y en algunos casos dos, a diferencia de la UNTDF donde es un contenido mínimo. Por otro lado, los requerimientos se enseñan después del tercer año donde los/as estudiantes tienen un recorrido importante y/o con experiencias laborales. En nuestro caso se dicta en el segundo año donde los estudiantes no tienen experiencia laboral.

Se realizó una comparación con fortalezas y debilidades de las estrategias relevadas

El siguiente paso fue la creación / elección de la mejor estrategia que permita a los estudiantes una mejor comprensión del trabajo de elicitar y especificar requerimientos de la manera más cercana a la realidad, a través de interacciones con usuarios.

Considerando el contexto actual de nuestra asignatura (alumnos, tiempos, recursos, etc),

2.3 Estrategias que mejor se adecúan.

Respecto a las estrategias relevadas y se definió implementar en forma simultánea dos de ellas, reflejadas en la tabla 1.

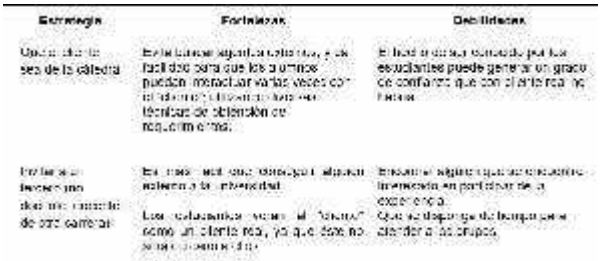


Tabla 1: Fortalezas y Debilidades de las Estrategias

Las experiencias se realizaron durante un mes e involucraron 8 clases prácticas (años 2024 y 2025) con un promedio de 32 estudiantes divididos en 8 grupos (aleatoriamente).

Para la experiencia se utilizaron clientes reales; en el año 2024 con un docente de otra carrera (Ing. Industrial) y un auxiliar de la cátedra; mientras que en el año 2025 otro docente de Ing. Industrial y una No Docente encargada del lab. de Medios Audiovisuales.

Cuatro grupos trabajaron con una estrategia / cliente (caso I) y otros 4 con la otra, diferentes ademas en cada año. Para ambos casos el objetivo era poder realizar una especificación de requerimientos completa.

2.4 Planificación del tiempo de clases

Su utilizaron 8 clases prácticas (5 hs cada una) distribuidas de las siguiente manera:

La primera clase consistió que cada grupo pueda conocer y tener el primer contacto con su "cliente", y lograr un conocimiento general del proyecto. La segunda clase se destinó a trabajar aspectos necesarios al proceso de elicitación de requerimientos, se hizo foco en diferentes estrategias para realizar entrevistas, sobre todo, entrevistas abierta. Posteriormente puntualizando, para su caso, diagramar preguntas y repreguntas posibles.

La tercera clase estuvo centrada en aspectos específicos respecto cómo realizar las entrevistas y que rol iba a desarrollar cada integrante y qué estrategias/herramientas utilizarían (grabar, apuntar, etc.)

Las siguientes clases (4, 5 y 6), los/as estudiantes realizaron el trabajo de campo (cada grupo por separado sin acceso a las entrevistas de los otros grupos). Para lo cual se planificó cada clase de forma que mientras los grupos del caso I realizaban las entrevistas a su "cliente", los grupos del caso II se dedicaron refinar su planificación; a media jornada se invertían roles.

Las últimas dos clases (7 y 8) se destinaron al desarrollo y elaboración del documento de especificación de requerimientos.

Finalizadas las experiencias, todos los grupos lograron cumplir los objetivos propuestos, sin embargo, pasaron por diferentes situaciones encontrando varias diferencias entre ellos (sobre todo en los grupos que compartían clientes). La principal diferencia se detectó en la elaboración de las preguntas y la forma de realizar las entrevistas. Mientras algunos tuvieron resultados satisfactorios rápidamente, otros debieron realizar más encuentros y, sobre todo, reformular las consultas de forma de poder obtener el conocimiento necesario.

3 Resultados Obtenidos

La implementación de las estrategias en la enseñanza de IR, se obtuvieron varios resultados. El cambio más importante fue que los/as estudiantes pudieron experimentar con un cliente real durante el proceso de elicitación de requerimientos.

Para evitar cualquier sesgo en nuestras conclusiones se realizaron encuestas anónimas al finalizar cada uno de los ciclos lectivos. El objetivo de las encuestas fue reunir la opinión general de los estudiantes y utilizarlas como punto de partida para la reflexión y la mejora de futuras experiencias.

El primer análisis fue determinar cómo consideraron la experiencia en forma general. Si bien hubo muy poco porcentaje de alumnos recursantes, el 100% (en ambos años) respondió que la experiencia fue valiosa.

A partir de ese dato y conocer un detalle más profundo se consultó respecto a aspectos generales que consideraron de la experiencia

en general, la figura 1 muestra los resultados obtenidos. En este caso se unificaron ambas experiencias.

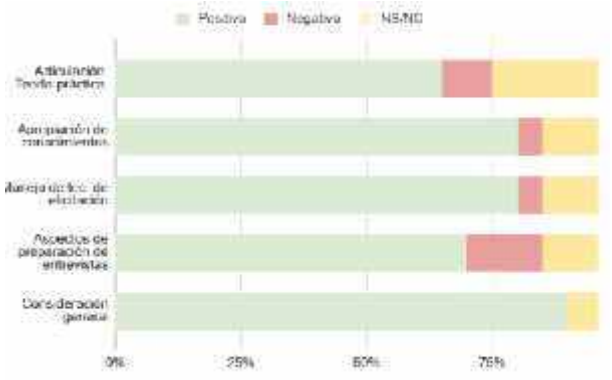


Figura 1: Resultados encuesta de estudiantes

Como principales se pueden sacar importantes conclusiones:

• La elección de las estrategias resultó una experiencia más realista, logrando indagar y obtener más información del dominio.

• Casos concretos y genuinos.

• Los estudiantes pudieron interactuar con un cliente real.

• Implementaron técnicas de elicitación que mejor se adecuaban al dominio.

• Se reflejó que la forma de entrevistar y los aspectos a relevar que cada grupo llevó adelante coincidió en gran mayoría en todos los grupos del mismo dominio.

• Los/as estudiantes fueran capaces de reflexionar sobre aspectos positivos y negativos de los encuentros con su "cliente".

• Alta motivación durante las clases.

• Todos, con mayor o menor medida, aplicaron técnicas adecuadas de elicitación y una especificación de requerimientos acorde.

• Favoreció una mejor comprensión del grado de avances y percepción de los/as estudiantes durante el proceso de enseñanza.

• Benefició el trabajo colaborativo.

La experiencia ha sido sumamente positiva y, con aspectos mejorables, se continuará implementando en los años venideros.

4 Formación de Recursos Humanos

El Equipo de Trabajo está conformado por docentes investigadores de la UNTDF y estudiantes de la Carrera. El Proyecto incluye docentes en etapa de formación de postgrado, dos docentes están realizando su tesis de Magister en Ingeniería de Software en la UNLP sobre IR.

Colaboraron un grupo de estudiantes en formación de grado, con el objetivo de formarlos en los aspectos generales de la investigación y que le sirva para el desarrollo de su futura tesis de grado.

5 Referencias

1. Burón, J. *Aprender a aprender: Introducción a la metacognición*. Bilbao: Ed. Mensajero; , p.94, (1994).
2. Montes de Oca, N.: Estrategias docentes y métodos de enseñanza-aprendizaje en la Educación Superior. Revista Hum Med [online] vol.11, pp.475-488. ISSN 1727-8120, (2011)
3. Santana, T., Kudo, T., Viana, D., Bulcão-Neto F.: Perspective on a Pedagogical Architecture to Requirements Engineering Education: A Qualitative Study. Journal of Software Engineering Research and Development, 13(1), 13:185 – 13:206, (2025)
4. Zapata-Jaramillo, C.M., Awad-Aubad G.: Requirements Game: Teaching Software Project Management, CLEI Electronic Journal, vol.10 (1), (2007)
5. Guitart, M., Rodríguez, E.: Cabot J.: Elección del modelo de evaluación: caso práctico para asignaturas de ingeniería del software: Jornadas Jenui 2006 pp.191–198, (2006)
6. Loucopoulos, P., Karakostas V.: Systems Requirements Engineering, (1995)
7. Oliveros, A.: Aprendizaje de la elicitación y especificación de requerimientos. CACIC (2014).
8. Montoya Suarez L. M., Echeverri Arias J. A.: Estrategias didácticas para enseñar Técnicas de Elicitación de Requisitos. Universidad Católica, Colombia, (2013)
9. M o

Estrategias pedagógicas y desarrollo de recursos tecnológicos para la enseñanza de Ciencias de la Computación: avances del segundo año de un proyecto

Gustavo ASTUDILLO[1], Silvia BAST[1], Paula DIESER[1], Martín LOBOS[1], Yamila MINETTI[1,2] , Lucas DÍAZ[1]
[1] Departamento de Matemática/FCEyN/UNLPam, [2]FCH/UNLPam
astudillo@exactas.unlpam.edu.ar, silviabast@exactas.unlpam.edu.ar

RESUMEN

El presente trabajo expone los avances de un proyecto de investigación orientado al desarrollo de estrategias pedagógicas y recursos didácticos para la enseñanza de las Ciencias de la Computación, en el marco de su segundo año de ejecución. Durante 2025 se profundizaron líneas vinculadas con los estilos de aprendizaje, la evaluación formativa, el desarrollo de un EDI para robótica educativa, la integración de inteligencia artificial generativa y el diseño de materiales educativos digitales. La investigación se centró en la revisión conceptual y psicométrica de CHAEA, así como analizar y ajustar estrategias de evaluación en asignaturas iniciales de programación, incluyendo su implementación en modalidad a distancia. Asimismo, se avanzó con el prototipo funcional *GrIDIE Blocks* y se diseñaron *chatbots* educativos para acompañar a estudiantes. Estos avances consolidan un enfoque integral que articula investigación educativa, innovación didáctica y desarrollo tecnológico.

Palabras clave: Estilos de aprendizaje, Evaluación de los aprendizajes, Robótica educativa, Inteligencia artificial generativa, Enseñanza de las Ciencias de la Computación.

CONTEXTO

El presente trabajo se enmarca en las actividades del Grupo de Investigación y Desarrollo en Innovación Educativa (GrIDIE) y forma parte del proyecto acreditado "Estrategias pedagógicas y recursos didácticos para la enseñanza de las Ciencias de la Computación", en ejecución durante el período 2024–2027 (Res. 16/24 CD FCEyN – UNLPam). Este proyecto da continuidad a líneas de investigación previas vinculadas con la incorporación de estrategias innovadoras en los procesos de enseñanza y aprendizaje de informática, consolidando un recorrido sostenido de trabajo en el área.

El proyecto articula investigación educativa, innovación didáctica y desarrollo tecnológico, con impacto tanto en la formación universitaria como en actividades de extensión y vinculación con la comunidad.

1. INTRODUCCIÓN

En los últimos años, la expansión sostenida de las tecnologías digitales ha modificado las formas de producir conocimiento, comunicarse y participar en la vida social. Este escenario impacta en los sistemas educativos y, en particular, en el área de las Ciencias de la Computación (CC) (Echeveste et al., 2025; Llambi et al., 2023).

En este contexto, resulta necesario repensar las estrategias pedagógicas, los enfoques de evaluación y los recursos didácticos que median los procesos de enseñanza y de aprendizaje. La complejidad de estas prácticas exige miradas integrales que contemplen las características del estudiantado, la diversidad de entornos tecnológicos y los nuevos usos de las tecnologías digitales.

En respuesta a estos desafíos, se vuelve necesario generar propuestas que articulen investigación educativa, innovación didáctica y desarrollo tecnológico. Desde esta perspectiva, el artículo presenta los avances del proyecto de investigación durante 2025.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La hipótesis de investigación sostiene que el diseño de materiales educativos digitales, el desarrollo de entornos tecnológicos específicos para la enseñanza de la programación y la incorporación crítica de herramientas pueden contribuir a fortalecer la motivación, favorecer la comprensión conceptual y promover aprendizajes más profundos. Asimismo, la diversificación de las estrategias de evaluación y la consideración de los estilos de aprendizaje como insumo para el diseño didáctico habilitan procesos de personalización y retroalimentación ajustados al contexto.

A continuación, se desarrollan en los tres ejes que organizan la investigación: (*i*) análisis y utilización de los estilos de aprendizaje, (*ii*) diseño e implementación de estrategias de evaluación formativa integradas al proceso de enseñanza, y (*iii*) desarrollo e integración de recursos tecnológicos en la enseñanza de la programación.

Estilos de aprendizaje

Cada individuo presenta formas particulares de aprender, manifestadas en la elección de entornos, la preferencia por ciertos tipos de tareas, el empleo de estrategias específicas para abordar materiales y la adopción de diferentes niveles de estructura y planificación. A pesar de la diversidad de enfoques, existe consenso en que los EA describen cómo la mente procesa la información y cómo este procesamiento es influido por las percepciones del sujeto (Villarreal-Fernández, 2023).

Diversos instrumentos permiten identificar los EA (Cárdenas-Palomino et al., 2022). Entre ellos, se seleccionó y aplicó CHAEA, pero los resultados llevaron a cuestionar la validez, planteando nuevas preguntas de investigación sobre su confiabilidad en el ámbito universitario hispanohablante y posibles adaptaciones. Para responderlas, se realizó una revisión sistemática, un análisis psicométrico y se utilizó Inteligencias Artificiales Generativas (IAG) como jueces expertos en la validación de contenido del CHAEA, todo con el objetivo de desarrollar una herramienta ajustada al contexto.

Evaluación

Asumir una posición sobre la evaluación supone entenderla no como un acto aislado de medición, sino como un proceso dialógico de comprensión y valoración de las prácticas de enseñanza y de aprendizaje, situado en un contexto sociohistórico y orientado a la toma de decisiones para la mejora de la práctica (Steiman, 2019). Desde este posicionamiento, la evaluación se inscribe en el enfoque del evaluar para el aprendizaje (Anijovich & Cappelletti, 2022), que promueve la publicación de objetivos y criterios, la coherencia entre enseñanza y evaluación, la diversidad de instrumentos, el diálogo, la metacognición y la retroalimentación significativa. Se trata de una evaluación formativa, auténtica y continua, que trasciende la acreditación puntual y articula tareas relevantes para el desarrollo de habilidades cognitivas y sociales complejas.

En entornos digitales, esta concepción adquiere nuevas implicancias. Bongiovanni (2020) señala que la pregunta central es cómo aprovechar el potencial de las mediaciones tecnológicas para ofrecer retroalimentación oportuna, generar evidencias significativas y acompañar las trayectorias estudiantiles.

En esta línea, se profundizó el análisis y la adecuación de la estrategia evaluativa hacia un formato plenamente virtual, en consonancia con los objetivos generales del proyecto de investigación centrados en la enseñanza de las Ciencias de la Computación en entornos mediados por tecnologías digitales.

Recursos y materiales didácticos en la enseñanza de las CC

Esta línea de trabajo se centra en la adaptación e implementación de estrategias de enseñanza en Ciencias de la Computación, abordando tres ejes principales: la Robótica Educativa (RE), la integración de Inteligencia Artificial Generativa (IAG) en el aula y en el diseño de Materiales Educativos Digitales (MED).

En relación con la RE, desde un enfoque centrado en el estudiante, se promueve la interacción con objetos con significado personal (Vaillant, 2013), favoreciendo el

aprendizaje activo y la resolución de problemas. La RE requiere un Entorno de Desarrollo Integrado (EDI) que permita programar y ejecutar instrucciones en los dispositivos. Existen diversos EDIs basados en bloques (Lobos et al., 2021), pero su estructura y funcionalidades no siempre se ajustan a las necesidades específicas de una propuesta didáctica concreta.

Con base en la identificación de un conjunto de requisitos para su adecuación pedagógica (Lobos et al., 2022) se avanzó en el diseño y desarrollo de un IDE propio, alineado con la secuencia didáctica elaborada por el equipo. Actualmente, el prototipo se encuentra en etapa de evaluación y será implementado en un Taller virtual destinado a estudiantes ingresantes de la Facultad, con el objetivo de analizar su usabilidad, adecuación didáctica y contribución al aprendizaje de nociones básicas de programación.

Por otra parte, la IAG se está integrando en el proyecto mediante el análisis, diseño y evaluación de chatbots orientados a acompañar a las/los estudiantes durante el desarrollo de la cursada de distintas asignaturas. Su incorporación plantea interrogantes vinculados con su integración pedagógica (UNESCO, 2024), la gestión de sesgos algorítmicos (Ferrante, 2021; Zou & Schiebinger, 2018) y el fenómeno de las "alucinaciones" en modelos generativos (Cárdenas Herrera, 2023).

El tercer eje se centra en el diseño de MED, concebidos desde un enfoque constructivista. Desde esta perspectiva, se asume que "puesto que el conocimiento no puede transmitirse, la enseñanza debería consistir en experiencias que faciliten la elaboración del conocimiento" (Jonassen, 2000, p. 227). En consecuencia, los MED se conciben como dispositivos que promueven la actividad cognitiva del estudiante. Su desarrollo está pensado sobre tres presupuestos básicos propuestos por Gagné (Basil & Coll, 1990): (*i*) partir de objetivos claramente definidos, (*ii*) establecer una secuencia ordenada en la enseñanza y (*iii*) considerar las características del destinatario, sus saberes previos y los productos o resultados concretos esperados. Asimismo, estos materiales son componentes de un ecosistema de aprendizaje que integra contexto, entorno y objetivos formativos (Schwartzman & Odetti, 2014). En su diseño se incorpora además el modelo SOI (Selección, Organización e Integración) del enfoque constructivista de Mayer (2000), favoreciendo que las/los estudiantes seleccionen la información relevante, la organicen coherentemente y la integren con sus conocimientos previos, promoviendo así aprendizajes significativos y transferibles.

3. RESULTADOS OBTENIDOS/ESPERADOS

Durante 2025 se profundizaron y ampliaron las líneas de trabajo iniciadas en 2024, consolidando resultados previos y abriendo nuevas líneas de análisis e intervención. A continuación, se presentan los avances organizados por eje.

Estilos de aprendizaje

Se continuó la revisión crítica del cuestionario CHAEA. Tras la construcción en 2024 de una versión reducida y el análisis preliminar de su comportamiento psicométrico, en 2025 se amplió el estudio con una muestra mayor de estudiantes universitarios.

Se aplicaron técnicas de Teoría Clásica de los Tests y Teoría de Respuesta al Ítem, en colaboración con una investigadora de la UNTDF. Los resultados evidenciaron debilidades en la relación entre constructos e indicadores y niveles de consistencia interna inferiores a los de la versión original y adaptaciones previas. Esto reforzó la necesidad de revisar la estructura teórica del modelo, más allá de ajustes lingüísticos o reducciones de ítems.

En consecuencia, se inició una nueva etapa junto a un equipo de la Universidad de Zaragoza, orientada a la reevaluación de los constructos mediante análisis factorial exploratorio y juicio de expertos humanos, complementando el uso previo de IAG en la validación de contenido. Este proceso implica un rediseño conceptual del instrumento, cuyos resultados se proyectan para la siguiente etapa.

Evaluación

Se profundizó el análisis y ajuste de la estrategia de evaluación implementada en Introducción a la Computación, consolidando un enfoque formativo y extendiendo el estudio a la modalidad a distancia.

Se completó el análisis de los instrumentos aplicados en 2024 (listas de cotejo, guías, parciales, encuestas y registros de plataforma), identificando fortalezas en la claridad de criterios y coherencia evaluativa, así como aspectos a mejorar en comunicación, gestión de tiempos y acompañamiento.

En 2025 se mantuvo la estructura evaluativa para garantizar comparabilidad entre cohortes, incorporando ajustes en consignas, comunicación y explicitación progresiva de criterios, lo que fortaleció la previsibilidad y comprensión del proceso.

Asimismo, se analizó por primera vez su funcionamiento en modalidad a distancia (Tecnicatura en Informática de Gestión). Los datos de Moodle mostraron participación heterogénea en autoevaluaciones y una brecha significativa entre desempeño conceptual y habilidades de programación, lo que dio lugar al diseño de una nueva línea orientada a optimizar estrategias evaluativas en cursos masivos a distancia.

Integración de TIC en la enseñanza de las Ciencias de la Computación

Durante 2025 se consolidaron y ampliaron acciones vinculadas con la integración de tecnologías digitales como mediadoras del aprendizaje.

EDI para Robótica Educativa

Se continuó el desarrollo de *GrIDIE Blocks*, EDI orientado a la enseñanza de programación con robótica educativa y didáctica por indagación. A la arquitectura y enfoque pedagógico definidos previamente se sumaron la priorización de historias de usuario, la organización bajo metodología ágil (Scrum) y el desarrollo del primer prototipo funcional. Esto posibilitó actividades de transferencia y divulgación, incluyendo talleres de programación con *GrIDIE Blocks* abiertos a la comunidad, que permitieron recoger primeras observaciones en contextos no formales

Diseño de materiales educativos digitales

Se inició una línea orientada al diseño sistemático de MED, con el objetivo de articular objetivos, actividades, recursos y evaluación. En esta etapa se avanzó en la revisión de marcos teóricos sobre diseño instruccional y en el análisis del material existente, proyectando el desarrollo de una plantilla reutilizable que integre consignas, recursos interactivos, autoevaluación y criterios de evaluación. Aunque en fase de conceptualización, esta línea se proyecta como eje articulador entre decisiones pedagógicas, evaluación formativa y tecnologías educativas del proyecto.

Uso de Inteligencia Artificial Generativa en educación

En 2025 se profundizó la exploración del uso de IAG en enseñanza e investigación, analizando sus potencialidades didácticas y desafíos (sesgos, análisis crítico de respuestas y desarrollo de competencias digitales).

Metodológicamente, la IAG se utilizó como apoyo en la validación de instrumentos educativos. En este marco se desarrolló un estudio sobre la validación de contenido del cuestionario CHAEA mediante juicio de expertos asistido por IAG (Astudillo et al., 2025), articulado con la línea de estilos de aprendizaje.

Estas acciones se complementaron con instancias de formación y extensión, incluyendo un taller de posgrado y actividades de capacitación docente orientadas al uso crítico de IAG en educación.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está conformado por investigadores en distintas etapas de formación: dos doctorandos, una investigadora que finalizó su maestría e inició el doctorado, y tres integrantes que continúan estudios de posgrado. Además, se incorporaron dos estudiantes de grado como asistentes de investigación, en el marco de una política activa de formación y renovación del equipo.

5. BIBLIOGRAFÍA

Anijovich, R., & Cappelletti, G. (2022). Evaluaciones: 29 preguntas y respuestas. Editorial El Ateneo.

Astudillo, G. J., Dieser, P., & Bast, S. (2025). Validación de contenido de CHAEA mediante juicio de expertos con el uso de Inteligencia Artificial Generativa. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, (42), 20-29. https://doi.org/10.24215/18509959.42.e2

Basil, C., & Coll, C. (1990). La construcción de un modelo prescriptivo de la instrucción: La teoría del aprendizaje acumulativo. Desarrollo psicológico y educación, 55-68.

Bongiovanni, P. (2020). Evaluar con tecnología, en contextos inesperados. Educación y Tecnología.

Cárdenas Herrera, J. (2023). Inteligencia artificial, investigación y revisión por pares: Escenarios futuros y estrategias de acción. RES, 32(4), 199.

Cárdenas-Palomino, F., Beltran, J. A., Cámaco, D. D. la C., & Rojas-Zuñiga, L. (2022). Identificación de los Estilos de aprendizaje de los Estudiantes Universitarios: Revisión de Literatura. Sinergias Educativas. https://doi.org/10.37954/se.vi.255

Echeveste, M. E., Acosta, A., Rojo, C., Wayar, C., Basel, V., & Zalazar, N. (2025). Reflexiones desde las prácticas educativas para una ciencias de la computación con perspectiva crítica. JAIIO, 11(8), 235-246.

Ferrante, E. (2021). Inteligencia artificial y sesgos algorítmicos ¿Por qué deberían importarnos? Nueva Sociedad, (294).

Jonassen, D. H. (2000). El diseño de entornos constructivistas de aprendizaje. Diseño de la instrucción: teorías y modelos: un nuevo paradigma de la teoría de la instrucción, 225-250.

Llambi, C., Borchardt, M., Klinkovich, V., Locca, N., Martínez, C., & Scasso, M. (2023). Aprendizajes y desafíos para la enseñanza de las Ciencias de la Computación en las escuelas. CAF-Banco de Desarrollo de América Latina y el Caribe y Fundación Sadosky.

Lobos, M., Bast, S. G., & Astudillo, G. J. (2021). Selección de un entorno de desarrollo integrado de robótica educativa para la aplicación de una secuencia didáctica propuesta para el aprendizaje de conceptos de programación. XXIII WICC 2021, Chilecito, La Rioja. http://sedici.unlp.edu.ar/handle/10915/120934

Lobos, M., Bast, S. G., & Astudillo, G. J. (2022). Evaluación de calidad de Entornos de Desarrollo Integrado de Robótica Educativa. En E. A. Irrazábal, G. N. Dapozo, & P. Martínez López (Eds.), Libro de actas: II Jornadas argentinas de didáctica de las Ciencias de la Computación 2022 (pp. 156-167). Fundación Sadosky-Universidad Nacional del Nordeste. http://repositorio.unne.edu.ar/xmlui/handle/123456789/50765

Mayer, R. E. (2000). Diseño educativo para un aprendizaje constructivista. Diseño de la instrucción: teorías y modelos: un nuevo paradigma de la teoría de la instrucción, 1, 153-172.

Schwartzman, G., & Odetti, V. (2014). Experimentación en el trabajo con materiales didácticos. En Schwartzman, G.; Tarasow, F. y Trech, M.(comp.) De la educación a distancia a la educación en línea. Aportes para un campo en construcción. Homo Sapiens.

Steiman, J. (2019). Más didáctica (en la educación superior) (Vol. 3). Miño y Dávila.

UNESCO. (2024). Guía para el uso de IA generativa en educación e investigación—UNESCO Biblioteca Digital. https://unesdoc.unesco.org/ark:/48223/pf0000389227

Vaillant, D. (2013). Integración de TIC en los sistemas de formación docente inicial y continua para la Educación Básica en América Latina. UNICEF Argentina.

Villarreal-Fernández, J. E. (2023). Cuestionario Honey-Alonso de estilos de aprendizaje (CHAEA). Propiedades psicométricas en estudiantes universitarios colombianos. Psicogente, 26(50), 46-69. https://doi.org/10.17081/psico.26.50.6231

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—It's time to make it fair. Nature, 559(7714), 324-326. https://doi.org/10.1038/d41586-018-05707-8

Herramientas de software utilizadas en la enseñanza de la programación desde una perspectiva de la Ingeniería de Software.

Yanina Medina[1], Gladys Dapozo[1], Alejandra Lliteras[2]
(1) Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste. Corrientes. Argentina.
(2) Facultad de Informática. Universidad Nacional de la Plata. Buenos Aires. Argentina. Centro LIFIA y CICPBA
{yanina, gndapozo }@exa.unne.edu.ar; lliteras@lifia.info.unlp.edu.ar

RESUMEN

La programación se ha convertido en una competencia esencial en todos los niveles educativos como respuesta a la necesidad de desarrollar el pensamiento computacional en las nuevas generaciones. La proliferación de herramientas y lenguajes de programación ha generado la necesidad de clasificarlos sistemáticamente. En educación superior, se observa una transición hacia modelos basados en competencias y metodologías activas. La inteligencia artificial también se integra para personalizar la enseñanza. La calidad del software para la enseñanza de la programación es crucial para el aprendizaje, y su evaluación rigurosa es necesaria para asegurar su efectividad. Esta línea de investigación plantea el estudio de las herramientas de software utilizadas en la enseñanza de la programación desde una perspectiva de la Ingeniería de Software, con el propósito de lograr un marco de referencia conceptual y descriptivo que sirva como base para futuras investigaciones orientadas a evaluar su calidad en uso.

Palabras clave: Calidad de Software, Enseñanza de la Programación, Ambientes de Desarrollo, Pensamiento Computacional.

CONTEXTO

Las líneas de I/D corresponden al PI 25F007 “Pensamiento computacional, pensamiento crítico y programación en la universidad: Metodologías, herramientas y recursos para su enseñanza”, aprobado por la SGCyT-UNNE para el periodo 2026-2029.

1. INTRODUCCIÓN

Muchas organizaciones requieren software adaptado a sus necesidades de negocio para alcanzar los objetivos. La programación se ha consolidado como una competencia esencial en los planes de estudio de todos los niveles educativos, desde la educación primaria hasta la superior, como respuesta a la necesidad de desarrollar el pensamiento computacional en las nuevas generaciones [1]. Esta tendencia ha impulsado la proliferación de herramientas, lenguajes y entornos de programación, lo que ha derivado en una necesidad urgente de clasificarlas mediante taxonomías sistemáticas que faciliten su análisis, selección y aplicación educativa [2].

La enseñanza de la programación presenta numerosos desafíos, especialmente en los primeros años de la formación de carreras universitarias vinculadas a la informática, donde los estudiantes deben desarrollar habilidades de pensamiento computacional y resolución de problemas.

En este contexto, el software desempeña un papel clave, facilitando la comprensión de conceptos abstractos y promoviendo el aprendizaje activo. Por ejemplo, el uso de juegos digitales en contextos educativos ha demostrado ser una estrategia prometedora para fomentar la motivación, la adquisición de contenidos y el desarrollo de habilidades cognitivas y sociales, especialmente en el aprendizaje de conceptos de programación en niveles escolares [1]. En el ámbito de la educación superior, las revisiones sistemáticas han identificado una transición paulatina del enfoque tradicional hacia modelos educativos basados en

competencias, que integran metodologías activas como la programación visual, entornos interactivos, e-learning y enfoques ágiles de desarrollo [3]. Asimismo, se ha comenzado a integrar la inteligencia artificial como recurso para personalizar la enseñanza, automatizar evaluaciones y optimizar la retroalimentación, abriendo nuevas posibilidades para la educación programática adaptativa [4].

Desde la perspectiva de la Ingeniería de Software, el software en general y el software para la enseñanza en particular, son productos que deben satisfacer los requerimientos de calidad del cliente-usuario. Para garantizar la calidad de software se debe implementar algún modelo o estándar de calidad que permita la gestión de atributos en el proceso de su construcción teniendo en cuenta que la concordancia de los requisitos con su construcción son la base de las medidas de calidad establecidas [5].

La norma ISO/IEC 25000 SQuaRE [6] define "un modelo de calidad para productos software" y constituye la primera propuesta que incluye de manera específica la calidad en uso como dimensión diferenciada de la calidad tradicional. Además, es una de las normas más actuales y ampliamente adoptadas para este propósito. Según esta norma, una "característica de calidad" refiere a una categoría de atributos que contribuye a la calidad del software.

En consonancia con [7] el "software para el aprendizaje" es un recurso de apoyo al aprendizaje, por lo que la ausencia de calidad afecta directa o indirectamente el aprendizaje alcanzado por los estudiantes. La calidad en uso de herramientas de software para la enseñanza de la programación puede variar, y muchas veces su efectividad no ha sido evaluada de manera rigurosa, lo que puede afectar la experiencia de enseñanza y aprendizaje [8].

En resumen, en todos los niveles educativos los docentes deben seleccionar software para la enseñanza de la programación accesible, intuitivo y alineado con estrategias pedagógicas efectivas. Sin una evaluación clara de la calidad del software, los docentes enfrentan dificultades para elegir las más adecuadas y adaptarlas a sus necesidades.

Esta línea de trabajo propone, en primer lugar, realizar un mapeo sistemático de las herramientas de software usadas en la enseñanza de la programación desde una perspectiva de la Ingeniería de Software, con el propósito de identificar y caracterizar dichas herramientas para luego usar su resultado como base para futuros estudios sobre su calidad, conforme a estándares internacionales como la ISO 25000. El objetivo final es la definición de un modelo de evaluación de calidad en uso para ser aplicado a las herramientas de software utilizadas en la enseñanza de la programación.

La programación y la calidad en uso en el marco de la norma ISO/IEC 25000.

El aprendizaje de la programación demanda la comprensión de conceptos abstractos que no siempre resultan fáciles de asimilar, debido a la falta de referentes concretos que permitan al estudiante experimentar. Para responder a esta dificultad, se han desarrollado múltiples herramientas y entornos de programación con fines educativos específicos, cuyo propósito es favorecer la comprensión de los conceptos básicos de la disciplina, al mismo tiempo que promueven habilidades cognitivas relacionadas con la resolución de problemas y el pensamiento computacional [9]. El pensamiento computacional no se limita a la programación ni a un conjunto de habilidades técnicas, sino que constituye un conjunto de prácticas y modos de razonamiento vinculados al diseño de procesos computacionales y a la interpretación del mundo como sistemas de información [10].

El software en general, y en particular el usado para la enseñanza, debe cumplir con estándares de calidad que garanticen su utilidad y confiabilidad para el usuario final. En este sentido, su calidad es clave, ya que una deficiencia en su diseño o funcionamiento impacta de manera directa o indirecta en los aprendizajes alcanzados por los estudiantes [10].

La norma ISO/IEC 25000 SQuaRE [11] define un modelo de calidad. Este modelo define categorías de atributos que permiten evaluar esta calidad desde distintas perspectivas: proceso, producto y calidad en uso. Esta norma es la primera en incluir un modelo específico de la calidad en uso de forma separada a la especificación habitual de la calidad, y además se trata de una de las normas más actuales y extendidas para este propósito.

La calidad en uso se define como la visión de calidad de los usuarios en un contexto que integra el software, y es medida sobre los resultados de usarlo en ese contexto, antes que sobre sus propiedades en sí mismo [12]. La calidad en uso se centra en si el usuario puede alcanzar sus objetivos y cumplir sus metas utilizando el producto, y no solo en el desempeño técnico intrínseco. Características clave de la calidad en uso son: 1. *Eficacia*: La capacidad de lograr los objetivos especificados. 2. *Eficiencia*: La relación entre el rendimiento del usuario y los recursos que utiliza (tiempo, esfuerzo). 3. *Satisfacción*: El grado en que las necesidades y expectativas del usuario son cumplidas durante el uso. 4. *Libertad de riesgo*: La mitigación de posibles consecuencias negativas (económicas, de salud, de seguridad). 5. *Cobertura del contexto de uso*: La capacidad de satisfacer las necesidades del usuario en diversas condiciones de uso.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación se enfoca en el estudio de herramientas de software utilizadas para la enseñanza de la programación para lo cual se hace foco en:

- Diseñar y aplicar una estrategia de búsqueda sistemática en bases de datos académicas reconocidas, siguiendo lineamientos metodológicos para estudios de mapeo sistemático.
- Seleccionar y analizar los estudios identificados, extrayendo datos clave sobre las herramientas de software usadas para la enseñanza de la programación, sus características técnicas y su aplicación pedagógica.
- Clasificar las herramientas encontradas en el punto anterior según criterios como, por ejemplo, tipo de herramienta, paradigma de programación, lenguaje de programación, nivel de enseñanza, entre otros.
- Identificar enfoques y métricas existentes en la literatura que se relacionen con la evaluación de la calidad de dichas herramientas, especialmente aquellas alineadas con normas internacionales como la ISO 25000.
- Elaborar un marco de referencia conceptual y descriptivo sobre el uso de herramientas de software usadas en la enseñanza de la programación, que sirva como base para futuras investigaciones orientadas a evaluar su calidad en uso.

Este trabajo contribuye al campo de la Ingeniería de Software aplicada a la educación proporcionando una base para la selección crítica de herramientas de software para la enseñanza de la programación en los diferentes niveles educativos.

3. RESULTADOS OBTENIDOS

Se realizó una revisión de la literatura siguiendo los siguientes pasos del mapeo sistemático [13], cuyos resultados se encuentran disponibles en [14].

Cuenta con la elaboración de un diseño del protocolo que incluye la definición de las preguntas de investigación (RQ), estrategia de búsqueda, criterios de inclusión y exclusión, procedimientos de selección y extracción de datos y esquema de clasificación y análisis.

Las preguntas planteadas en [14] fueron:

- RQ1: ¿Qué herramienta/s software se utilizan para la enseñanza de la programación?
- RQ2: ¿Qué tipo de licencia poseen estas herramientas?
- RQ3: ¿Presenta alguna evaluación de calidad? ¿Cuáles fueron, qué métodos se siguieron?

- RQ4: ¿En qué ámbito se realizó la evaluación?
- RQ5: ¿A qué paradigmas hacen referencia?
- RQ6: ¿Qué enfoque o estrategia educativa se utilizó?
- RQ7: ¿Presenta alguna vinculación con herramientas de IA generativa?

Análisis de los resultados

De los resultados obtenidos para cada una de las preguntas formuladas anteriormente surgieron las siguientes conclusiones:

RQ1: Las herramientas de software que más se utilizan para la enseñanza de la programación son Scratch, Blockly, Python, Java, C++, Arduino, LEGO, App Inventor, Alice, y Greenfoot.
RQ2: Respecto del licenciamiento de las herramientas de software para la enseñanza de la programación más populares son mayoritariamente de código abierto. Las profesionales/industriales tienden a ser propietarias.
RQ3: Con relación a la calidad del software usado para la enseñanza de la programación, pocas investigaciones la evalúan formalmente. En aquellas donde se evidencia que existe evaluación, se enfocan principalmente en los atributos, tales como, la eficacia (cumplimiento de tareas), satisfacción del usuario y facilidad de aprendizaje. Considerando el concepto de calidad del software, en este estudio se adopta el concepto de calidad del producto software, entendido en el sentido de la familia ISO/IEC 25000 como el conjunto de atributos intrínsecos del software (p. ej. funcionalidad, rendimiento, usabilidad, seguridad, mantenibilidad, portabilidad) que permiten caracterizar su calidad. Este enfoque se centra en las propiedades del producto en sí mismo y no en los resultados derivados de su uso en un contexto específico. Por lo tanto, el análisis se orienta a identificar las características del software que se consideran relevantes para su aplicación para la enseñanza de la programación, sin abordar, la evaluación de los resultados obtenidos por los usuarios finales. Las evaluaciones de calidad del software son especialmente comunes en entornos universitarios.
RQ4: Los ámbitos educativos en los que se realizaron las evaluaciones comprenden: primaria, secundaria, superior/universitaria y también en el ámbito de formación docente.
RQ5: Entre los paradigmas de programación predominantes se observa: orientado a objetos, procedural y basado en eventos.
RQ6: En relación con el enfoque o estrategia educativa utilizadas, se observa que la estrategia más mencionada es la gamificación / aprendizaje basado en juegos, seguida por entornos visuales y Aprendizaje Basado en Problemas (ABP), lo que indica un mayor interés académico en enfoques activos y lúdicos para enseñar programación.

RQ7: Se evidencia una tendencia emergente hacia la integración de herramientas de IA generativa en la enseñanza de programación. La baja frecuencia de publicaciones sobre IA generativa en este conjunto de datos no implica ausencia de relevancia, sino que refleja un desfase temporal entre la rápida evolución tecnológica y el ritmo de publicación académica y sistematización de evidencias incluidas en la revisión.

4. RESULTADOS ESPERADOS

En base a la información relevada se continuará con la elaboración de una taxonomía de herramientas, según criterios técnicos y pedagógicos, que permitirá seguir avanzando hacia la evaluación de la calidad de las herramientas de software utilizadas para la enseñanza de la programación en los distintos niveles educativos.

5. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan dos docentes investigadores y una tesista de posgrado de la Maestría en Ingeniería de Software en la Facultad de Informática, UNLP, que actualmente realiza su trabajo final integrador para obtener el grado de Especialista.

6. REFERENCIAS

[1] Theodoropoulos, A., & Lepouras, G. (2020). “Digital game-based learning and computational thinking in P-12 education: A systematic literature review on playing games for learning programming”. En K. Graziano (Ed.), Research Anthology on Computational Thinking, Programming, and Robotics in the Classroom (pp. 159–183). IGI Global. Available: https://doi.org/10.4018/978-1-7998-4576- 8.ch007

[2] McGill, M. M., & Decker, A. (2020). “Construction of a taxonomy for tools, languages, and environments across computing education”. En Proceedings of the 2020 ACM Conference on International Computing Education Research (ICER ’20) (pp. 124–135). Association for Computing Machinery. Available: https://doi.org/10.1145/3372782.3406258

[3] Rodrigues, G., Monteiro, A. F., & Osório, A. (2022). “Introductory programming in higher education: A systematic literature review”. En Proceedings of the Third International Computer Programming Education Conference (ICPEC 2022). Open Access Series in Informatics (OASIcs), 102, 4:1–4:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. Available: https://doi.org/10.4230/OASIcs.ICPEC.2022.4

[4] Lytvynova, S. H., Rashevska, N. V., & Proskura, S. L. (2024). “The use of artificial intelligence in teaching students programming languages”. En Proceedings of the IX International Workshop on Professional Retraining and Life-Long Learning using ICT: Person-oriented Approach (3L-Person 2024) (co-ubicado con la 19a International Conference on ICT in Education, Research, and Industrial Applications), 10–29. CEUR-WS (ISSN 1613-0073). Disponible en https://ceur-ws.org/Vol-3781/paper01.pdf

[5] Callejas-Cuervo, M., & Alarcón-Aldana, A. C. (2017). “Modelos de calidad del software: un estado del arte”. Entramado, 13(1), 236–250. Disponible en https://doi.org/10.18041/entramado.2017v13n1.25464

[6] International Organization for Standardization. (2005). ISO/IEC 25000:2005 Software engineering — Software product Quality Requirements and Evaluation (SQuaRE). Available: https://www.iso.org/standard/35734.html

[7] Segura, A., Vidal-Castro, C., & Prieto, M. E. (2008, octubre). “Evaluación de la calidad del software para el aprendizaje”. En A. B. Gil G., J. Á. Velázquez Iturbide, & F. J. García-Peñalbo (Eds.), Actas del X Simposio Internacional de Informática Educativa (SIIE 2008) (pp. 59–64). Ediciones Universidad de Salamanca. Disponible: https://doi.org/10.13140/2.1.2023.5528

[8] Kirkwood, A., & Price, L. (2015). “Achieving improved quality and validity: Reframing research and evaluation of learning technologies”. European Journal of Open, Distance and E-learning, 18(1). Available: https://www. eurodl. org/ index. php?p= speci al& sp= artic les& inum= 6& artic le= 672& artic le= 678

[9] Kuz, Antonieta y Ariste María Cecilia. “Análisis y revisión de softwares educativos para el aprendizaje de la programación en entornos lúdicos”. Tecné, Episteme y Didaxis: TED. 2022; (52):117-136. [fecha de Consulta 5 de octubre de 2025]. ISSN: 2665-3184. Disponible en: https://www.redalyc.org/articulo.oa?id=614274349008

[10] Denning, P. J., & Tedre, M. (2021). “Computational Thinking: A Disciplinary Perspective. Informatics in Education”, 20(3), 361-390. DOI: 10.15388/infedu.2021.21.

[11] ISO25000:2005 Software Engineering, Software product Quality Requirements and Evaluation (SQuaRE). Technical report, International Organization for Standardization.

[12] ISO 25010:2005, Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models. Technical report International Organization for Standardization.

[13] Kitchenham, B. A., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report, EBSE-2007-01. Disponible en: https://legacyfileshare.elsevier.com/promis_misc/525444systematicreviewsguide.pdf

[14] Medina, Y. “Evaluación de herramientas software para la enseñanza de la programación desde una perspectiva de Ingeniería de Software: un estudio de mapeo sistemático.” En Actas de las Jornadas de Calidad de Software y Agilidad 2025 (JCSA 2025), Rcia., Argentina. UTN., en prensa.

INTEGRACIÓN DE TECNOLOGÍAS EMERGENTES PARA FORTALECER EL APRENDIZAJE COLABORATIVO Y EL DESARROLLO DE COMPETENCIAS

Carlos Neil, Marcelo De Vincenzi,
Nelson Garrido, Fabian Maffei, Silvia Poncio, Marcelo Monferrato, Julian Rodriguez Escobedo

Universidad Abierta Interamericana. Facultad de Tecnología Informática.
Centro de Altos Estudios en Tecnología Informática. Buenos Aires. Argentina
{carlos.neil, medevincenzi, silvia.poncio, marcelo.monferrato}@uai.edu.ar
{nelson.garrido, fabian.maffei, julianmartin.rodriguezescobedo}@alumnos.uai.edu.ar

RESUMEN

Las instituciones educativas enfrentan la necesidad de adaptar sus modelos pedagógicos para formar graduados competentes frente a los desafíos globales. En este trabajo, se establecen las bases conceptuales que coadyuven a una mejora de la calidad educativa bajo el enfoque del Aprendizaje Centrado en el Estudiante (ACE), fundamentado en el constructivismo y la colaboración en un entorno de nuevas tecnologías emergentes. La irrupción de las nuevas tecnologías, todas teñidas de Inteligencia Artificial (IA) ha transformado el escenario educativo al redefinir la distinción entre el proceso de aprendizaje (actividades mentales) y el producto observable (resultados entregables). La correlación que tradicionalmente se asumía entre proceso y producto desafía esa lógica, reduciendo el valor del producto como indicador de aprendizaje genuino. Por lo tanto, para evitar que el uso indiscriminado de las tecnologías emergentes obstaculice el desarrollo de habilidades cognitivas, se debe garantizar tanto el alineamiento constructivo como el uso responsable de herramientas tecnológicas.

Palabras clave: aprendizaje centrado en el estudiante; desarrollo de competencias; evaluación; tecnologías emergentes

CONTEXTO

El proyecto de investigación se desarrolla en la Facultad de Tecnología Informática de la UAI, donde se implementan las estrategias didácticas y herramientas derivadas de las investigaciones en asignaturas de grado y posgrado.

1. INTRODUCCIÓN

Este trabajo es una continuación del [1] donde se focalizó en el impacto de la inteligencia artificial en la definición, desarrollo y evaluación de competencias. En esta nueva propuesta, el proyecto busca establecer una conexión coherente entre los conceptos fundamentales que sustentan la calidad educativa, enmarcados en la integración de tecnologías emergentes para fortalecer el aprendizaje colaborativo y el desarrollo de competencias

Como eje rector de la propuesta, coloca el faro en el Aprendizaje Centrado en el Estudiante (ACE) que posiciona al alumno en el epicentro del proceso educativo. Sustentado teóricamente en el constructivismo, el ACE promueve la construcción activa del conocimiento a través del trabajo en equipo y la colaboración. En este marco, se reflexiona sobre el impacto de las nuevas tecnologías en educación, distinguiendo entre su uso por parte de docentes y estudiantes, y advirtiendo sobre los posibles efectos negativos que un uso indiscriminado puede tener en el desarrollo de procesos cognitivos.

Para operacionalizar el ACE, se propone la integración de un modelo por competencias en el diseño curricular, haciendo hincapié en la aplicación práctica y la preparación del estudiante para situaciones reales. En este esquema, el uso de rúbricas, lineamientos para el aprendizaje y listas de verificación derivados de las mismas actúan como guías esenciales en el proceso de enseñanza aprendizaje. Como concepto integrador, la definición, el desarrollo y la evaluación de competencias se estructuran bajo el principio del alineamiento constructivo, garantizando una coherencia necesaria entre los resultados de aprendizaje, las actividades de enseñanza y aprendizaje y las estrategias de evaluación.

Finalmente, se aborda el desafío de integrar las tecnologías emergentes no solo como una herramienta de automatización, sino como un recurso que, correctamente gestionado, posee la capacidad de potenciar el enfoque constructivista y el modelo de competencias en un entorno educativo en constante transformación.

1.1. Tecnologías emergentes y educación

Los avances tecnológicos están generando transformaciones profundas en el ámbito

educativo, redefiniendo paradigmas de aprendizaje y enseñanza [2]. Estas tecnologías no solo ofrecen acceso a una cantidad sin precedentes de información, sino que también introducen herramientas innovadoras que transforman la manera en que las concebimos [3] creando experiencias adaptativas que no estaban disponibles en los ámbitos educativos previos [4]. La incorporación de tecnologías emergentes en la educación nos interpela a reconsiderar los métodos tradicionales tanto en la enseñanza como en la evaluación que resultan inadecuados frente a las nuevas dinámicas educativas.

Si concebimos el aprendizaje como una transformación de las estructuras cognitivas [5], es importante distinguir entre su *proceso* y el *producto* resultante. El primero, entendido como las actividades mentales que el estudiante desarrolla, constituyen la base del aprendizaje significativo [6]. El *producto*, por otro lado, se manifiesta visible en los comportamientos observables de los estudiantes. Si no se distinguen ambos, existe el riesgo de centrarse exclusivamente en los resultados finales (*productos*) y descuidar los procesos de análisis y las estrategias cognitivas y metacognitivas (*procesos*) que condujeron a esos resultados. La transformación de las estructuras cognitivas depende de un proceso activo y significativo en donde los docentes deben utilizar estrategias de aprendizaje efectivas (*proceso*) y no solo evaluar los resultados (*producto*).

El uso de las tecnologías emergentes debe diferenciarse según el rol que desempeñan los actores principales: docentes y estudiantes. El proceso de aprendizaje se construye mediante la participación del estudiante y el inevitable uso de la tecnología plantea desafíos en relación con el equilibrio entre facilitar el aprendizaje y la adquisición de resultados educativos sin esfuerzo. La tecnología pueden aliviar la carga cognitiva del estudiante, pero también se corre el riesgo de limitar su participación en el proceso de aprendizaje, sin obviar los temas éticos vinculados [7].

Por otro lado, en las múltiples actividades docentes, que implican mucho esfuerzo y tiempo en su diseño y desarrollo, el uso de la tecnología sí abre un abanico amplio de posibilidades que van desde la evaluación eficiente y objetiva mediante la automatización de procesos de corrección, pasando por el diseño y la personalización curricular y llegando al desarrollo de sistemas de gestión del aprendizaje que ayuden a identificar lagunas de conocimiento y propongan estrategias didácticas.

1.2. Enfoque constructivista

Desde la perspectiva constructivista el aprendizaje es un proceso activo donde los estudiantes construyen su conocimiento a través de experiencias contextualizadas. Fue propuesto inicialmente por Piaget [8] a partir de la teoría del desarrollo cognitivo, posteriormente expandida por Vygotsky [9]. Por otro lado, Ausubel [6], enfatizó en el aprendizaje significativo, donde los nuevos conocimientos se relacionan con los previos del estudiante, utilizando organizadores y materiales que faciliten la asimilación de la información y su vínculo con otros conceptos.

Otras perspectivas también interactúan con las tecnologías emergentes. El conductismo, a través de herramientas, como los sistemas de aprendizaje adaptativo, que utilizan la IA para ofrecer refuerzos inmediatos y personalizados [4]. El conectivismo, con su enfoque en el aprendizaje a través de redes y conexiones, encuentra en la IA una poderosa herramienta para organizar información y fomentar interacciones en entornos digitales [10].

1.3. Aprendizaje centrado en el estudiante

El ACE es un enfoque central en la educación [11], se basa en teorías constructivistas que proponen que los estudiantes deben construir y reconstruir el conocimiento para aprender de forma efectiva. Es también un proceso de cambio cualitativo, reforzando y empoderando al estudiante, desarrollando sus habilidades críticas [12].

1.4. Trabajo colaborativo

El trabajo colaborativo, fundamentado en el concepto de zona de desarrollo próximo [13] ha experimentado un notable impulso. Una forma de implementarlo es mediante el aprendizaje basado en problemas que fomenta el pensamiento complejo y la reflexión grupal y cooperativa para tomar decisiones sobre problemas reales y relevantes dentro del ámbito profesional en el que participa y se forma el estudiante [14]. La IA puede potenciar el trabajo colaborativo y el aprendizaje basado en problemas al crear entornos virtuales donde los estudiantes colaboren para resolver problemas reales. Asistentes inteligentes pueden facilitar la comunicación y coordinación entre los miembros del grupo, mientras que algoritmos de IA pueden generar problemas adaptados al nivel de los estudiantes, promoviendo el pensamiento crítico y complejo.

1.5. Competencias y resultados de aprendizaje

El concepto de competencia es amplio; algunos la definen en términos de rendimiento y habilidades adquiridas a través de la formación, otros adoptan una perspectiva más holística que engloba el conocimiento, la comprensión, las habilidades, las destrezas y las actitudes. Los resultados de aprendizaje, más específicos, son formulaciones que indican lo que se espera que los estudiantes conozcan, sean capaces de hacer o demostrar al finalizar un período formativo [15]. [16].

La tecnología puede contribuir al modelo por competencias al facilitar tanto el desarrollo como la evaluación de manera personalizada. Sistemas basados en IA pueden considerar a las habilidades, conocimientos y actitudes de los estudiantes en tiempo real, alineándolos con los resultados de aprendizaje esperados mediante el análisis de la matriz de competencias, identificando áreas de mejora y distribuyendo las responsabilidades de manera más eficiente en los espacios curriculares.

1.6. Evaluación y lineamientos para el aprendizaje utilizando rúbricas

Las rúbricas brindan criterios de calidad relacionados con las competencias y son un instrumento válido para su evaluación [17]. La autoevaluación basada en su uso incorpora los criterios que los estudiantes deben aplicar para valorar su propio trabajo [18]. Las rúbricas, por otro lado, no se limitan únicamente a la evaluación, también tienen un componente instructivo transformándose en una estrategia de aprendizaje donde el estudiante no resuelve un problema y luego se autoevalúa, sino que lo desarrolla con las consideraciones establecidas en los criterios de evaluación, devenidos en lineamientos para el aprendizaje, que le permite reflexionar y actuar en consecuencia [19]. Estas tecnologías pueden enriquecer el uso de rúbricas al automatizar la retroalimentación basada en los criterios de evaluación, ayudando a los estudiantes a reflexionar y mejorar en tiempo real. Además, la IA puede utilizar las rúbricas como herramienta instructiva, adaptando las actividades de aprendizaje a las necesidades del estudiante.

Se está desarrollando una aplicación (https://case.uai.edu.ar/) para la gestión de rúbricas que optimiza el proceso de diseño mediante la personalización de criterios, la exportación multiformato y un repositorio colaborativo para reutilizar recursos. La aplicación se desarrolla mediante una arquitectura de microservicios y un sistema de almacenamiento híbrido (SQL y NoSQL).

1.7. Estilos de aprendizaje

Las tecnologías emergentes pueden favorecer el aprendizaje colaborativo enriquecido (ACE) al habilitar plataformas capaces de adaptar contenidos, actividades y retroalimentación al ritmo y a las preferencias de aprendizaje del estudiantado. Estos entornos, apoyados en técnicas de inteligencia artificial, permiten identificar patrones de interacción y ofrecer ajustes pedagógicos oportunos, fortaleciendo los procesos de personalización sin desvincularlos de las decisiones didácticas. En esta línea, los estudios desarrollados en el marco del proyecto muestran que la aplicación de técnicas de inteligencia artificial orientadas a la detección de estilos de aprendizaje contribuye a mejorar los procesos de adaptación pedagógica cuando dichas tecnologías se integran al diseño didáctico como soporte de la mediación docente y no como sustituto de esta [20], [21].

1.8. Competencias Blandas e IA

La convergencia entre el *Business Intelligence* (BI) y la Inteligencia Artificial (IA) permite un diagnóstico predictivo y profundo de las competencias blandas en el mercado laboral. Mientras que el BI organiza los datos y la Gestión del Conocimiento los transforma en activos intangibles compartidos, la IA potencia este ecosistema mediante el procesamiento de grandes volúmenes de información no estructurada y la detección de patrones de talento emergentes. esta integración tecnológica permite a las organizaciones trascender la gestión de datos convencional para convertirse en entidades de aprendizaje continuo, donde el conocimiento fluye de manera automatizada para optimizar la toma de decisiones estratégicas y la creación de valor [22]. El avance de la IA resalta la necesidad de fomentar habilidades blandas como factor diferenciador en el entorno profesional. El presente artículo analiza el rol estratégico de las universidades en el fortalecimiento de estas competencias, concluyendo que la formación integral requiere un compromiso compartido entre la institución académica y la participación proactiva del alumnado [23].

1.9. Alineamiento constructivo

El alineamiento constructivo es un enfoque pedagógico reconocido por su énfasis en la

necesaria coherencia que debe existir entre los resultados de aprendizaje, el proceso educativo y la evaluación [24]. En esta línea, es fundamental que los resultados de aprendizaje, las actividades de enseñanza, así como los métodos de evaluación se definan de manera coherente para maximizar el aprendizaje. La IA puede coadyuvar a alineamiento constructivo mediante la sugerencia de actividades y evaluaciones que estén directamente vinculadas con los resultados de aprendizaje esperados, asegurando que cada uno de los componentes del proceso educativo estén alineado entre sí.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación se desarrolla en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la UAI y se enmarca en la línea de investigación *Sociedad del conocimiento y tecnologías aplicadas a la educación*.

3. RESULTADOS OBTENIDOS / ESPERADOS

3.1. Resultados obtenidos

En el marco del proyecto de investigación se han presentado, en 2025, los siguientes artículos científicos:
Automation of learning outcome generation using a Large Language Model [25]; Learning Outcome Generation using LLM: Design and Validation [26]; Diseño de artefactos pedagógicos mediante IA: un enfoque basado en LLM y RAG [27]. Además, las siguientes defensas de tesis doctorales en el Doctorado en informática (UAI). Silvia Poncio. Diagnóstico de competencias blandas mediante herramientas de business intelligence y Nestor Balich. Gemelos digitales híbridos para la enseñanza de robótica. Un entorno distribuido en tiempo real.

3.2. Resultados esperados

En 2026, se desarrollarán trabajos que se verán reflejados en artículos científicos, tesis y trabajos finales de posgrado.

A partir de la aplicación del proceso de desarrollo conducido por rúbricas [28]: 1) el diseño de mapas conceptuales (*Rubric-Driven Concept Maps Design*); 2) la escritura de resultados de aprendizaje (*Rubric-Driven Learning Outcomes Writing*) que amplía [29] y, 3) el diseño de rúbricas (*Rubric-Driven Rubric Design*) que amplía [30]. Como parte de tesis y trabajos finales de posgrado, se desarrollarán: 4) *Percepción de estilos de aprendizaje y personalización con Inteligencia Artificial*. 5) análisis del uso de IA desde la perspectiva del estudiante. 6) el rediseño de UAIRubric (https://case.uai.edu.ar/) y 7) el análisis de estilos de aprendizaje utilizando herramientas de IA.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está compuesto estudiantes de posgrado de la UAI, que están desarrollando sus respectivas tesis y trabajos finales en temas relacionados con el proyecto: Nelson Garrido, Fabian Maffei, Edgardo Go, Marcelo Monferrato, Julián Escobedo y los estudiantes de grado, Facundo Romeu y Ángelo Perrotta.

5. BIBLIOGRAFÍA

[1] C. Neil, C. G., Battaglia, N., De Vincenzi, , F., Garrido N., M, Maffei; Poncio, Monferrato, "El impacto de IA en la definición, desarrollo y evaluación de competencias," in *XXVII Workshop de Investigadores en Ciencias de la Computación (WICC).*, 2025.

[2] M. M. O. Fernández and A. A. Ferreiro, "Las tecnologías digitales en el entrenamiento de las funciones ejecutivas: una revisión sistemática de literatura," *RiiTE Rev. Interuniv. Investig. en Tecnol. Educ.*, pp. 120–136, 2023.

[3] C. Beaulac and J. S. Rosenthal, "Predicting university students' academic success and major using random forests," *Res. High. Educ.*, vol. 60, pp. 1048–1064, 2019.

[4] H. Peng, S. Ma, and J. M. Spector, "Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment," *Smart Learn. Environ.*, vol. 6, no. 1, pp. 1–14, 2019.

[5] D. B. Novak, J. D. y Gowin, *Aprendiendo a Aprender*. Martínez Roca, 1988.

[6] D. Ausubel, J. Novak, and H. Hanesian, "Significado y aprendizaje significativo," *Psicol. Educ. un punto vista cognoscitivo*, vol. 1, no. 2, pp. 53–106, 1976.

[7] A. Nguyen, H. N. Ngo, Y. Hong, B. Dang, and B.-P. T. Nguyen, "Ethical principles for artificial intelligence in education," *Educ. Inf. Technol.*, vol. 28, no. 4, pp. 4221–4241, 2023,

[8] L. S. Vygotsky, *The collected works of LS Vygotsky: Problems of the theory and history of psychology*, vol. 3. Springer Science & Business Media, 1987.

[9] J. Piaget, M. Cook, and W. W. Norton, "The

origins of intelligence in children, vol. 8 New York," *NY Int. Univ. Press. Sch.*, 1952.
[10] G. Siemens, "Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning," *Online] retrieved from http//www. idtl. org/Journal/Jam _05/article01. html*, 2005.
[11] M. Mendoza and M. Rodríguez, "Aprendizaje centrado en el estudiante desde la planificación en investigación. CIENCIAMATRÍA, 6 (10), 560-572." 2019.
[12] J. J. C. Cantu, E. D. C. Ramirez, M. E. L. Almeida, E. S. I. Mendo, and J. M. Vergaray, "El aprendizaje autónomo en educación superior. Revisión sistemática," *Salud, Cienc. y Tecnol.*, no. 3, p. 391, 2023.
[13] C. N. Suárez Guerrero, "La Zona de Desarrollo Próximo: categoría pedagógica para el análisis de la interacción en contextos de virtualidad," *Pixel-Bit*, 2004.
[14] S. L. Olivares Olivares and Y. Heredia Escorza, "Desarrollo del pensamiento crítico en ambientes de aprendizaje basado en problemas en estudiantes de educación superior," *Rev. Mex. Investig. Educ.*, vol. 17, no. 54, pp. 759–778, 2012.
[15] J. H. P. Prieto, "Secuencias didácticas: Aprendizaje y evaluación de competencias en educación superior," *Bordon*, vol. 63, no. 1. pp. 77–92, 2011.
[16] M. Neil, C. G., Battaglia, N., & De Vincenzi, "Diseño de competencia conducido por rúbricas," in *XXIX Congreso Argentino de Ciencias de la Computación*, 2023.
[17] M. P. García Sanz, "La evaluación de competencias en Educación Superior mediante rúbricas: un caso práctico," *Rev. Electrónica Interuniv. Form. del Profr.*, vol. 17, no. 1, 2014.
[18] E. Panadero, J. A. Tapia, and J. A. Huertas, "Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education," *Learn. Individ. Differ.*, vol. 22, no. 6, pp. 806–813, 2012.
[19] N. Battaglia, C. G. Neil, and M. De Vincenzi, "Software Engineering Competence-Based Learning in Collaborative Virtual Environments," in *EDUNINE 2021 - 5th IEEE World Engineering Education Conference: The Future of Engineering Education: Current Challenges and Opportunities, Proceedings*, 2021.
[20] F. Maffei, C. Neil, and N. Battaglia, "Estilos de aprendizajes personalizados con inteligencia artificial. Un mapeo sistemático de la literatura," in *XXIX Congreso Argentino de Ciencias de la Computación*, 2023.
[21] C. Maffei, F., Neil, "Teorías de aprendizaje en el contexto de la Inteligencia Artificial," in *Congreso Nacional de Ingeniería Informática / Sistemas de Información 12º CoNaIISI. Universidad Tecnológica de Catamarca.*, 2024.
[22] S. V. Poncio, C. N. Cuña Giménez, N. Cardú, and G. Ruiz, "Data Warehouse Soft Skills: modelo diagnóstico para la observación de la actitud empática," in *XX Congreso de Tecnología en Educación y Educación en Tecnología (TE & ET)(UNDAV, 12 y 13 de junio de 2025)*, 2025.
[23] A. Olivera, "Desarrollo de Habilidades Blandas ante el avance de la Inteligencia Artificial," 2023.
[24] J. B. Biggs, "Calidad del aprendizaje universitario," *Educ. Siglo XXI*, vol. 22, p. 272, 2004.
[25] N. Garrido and C. Neil, "Automation of learning outcome generation using a Large Language Model," *JAIIO, Jornadas Argentinas de Informática*, vol. 11, no. 8, pp. 30–33, 2025.
[26] C. Garrido, N.; Neil, "Impact of Artificial Intelligence on the Learning Experience: A Systematic Literature Mapping," in *54JAIIO Jornadas Argentinas de Informática e Investigación Operativa*, 2025.
[27] C. P. Nelson Garrido, Carlos Neil, "Diseño de artefactos pedagógicos mediante IA: un enfoque basado en LLM y RAG," in *XXXI Congreso Argentino de Ciencias de la Computación*, 2025.
[28] C. Neil, N. Battaglia, and M. De Vincenzi, "Rubric-Driven Competency Development: A Case Study BT - Computer Science – CACIC 2023," P. Pesado, W. Panessi, and J. M. Fernández, Eds., Cham: Springer Nature Switzerland, 2024, pp. 350–365.
[29] C. G. Neil, N. Battaglia, and M. De Vincenzi, "La matriz de competencias como herramienta para orientar la escritura de resultados de aprendizaje," in *XVIII Congreso Nacional de Tecnología en Educación y Educación en Tecnología-TE&ET*, 2023.
[30] C. Neil, N. Battaglia, and M. E. D. V. Zemborain, "Marco metodológico para el diseño de rúbricas analíticas," *Edutec, Rev. Electrónica Tecnol. Educ.*, no. 80, 2022.

La IA Generativa en la enseñanza introductoria de la programación: posibles transformaciones en la práctica docente

Ana M. Piccin[1], Graciela D. S. Hadad[1,2], Macarena Gorgal[1]

1 Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano
2 Instituto de Ingeniería y Nuevas Tecnologías, Universidad Nacional del Oeste
ana.piccin@comunidad.ub.edu.ar, graciela.hadad@comunidad.ub.edu.ar, macarena.gorgal@comunidad.ub.edu.ar

RESUMEN

La Inteligencia Artificial Generativa (GenAI) ha irrumpido en la enseñanza a futuros profesionales, en particular, de las carreras relacionadas con la informática. Publicaciones en congresos y revistas especializadas muestran el creciente interés académico por estudiar, comprender y describir las nuevas formas que va adquiriendo la enseñanza tras el advenimiento de la GenAI. En las carreras de informática, la enseñanza introductoria de la programación aparece como susceptible de transformaciones dramáticas: las demandas del mercado laboral exigirán otras competencias, probablemente de nuevas categorías; se deberán renovar objetivos pedagógicos; los docentes deberán seleccionar herramientas didácticas acordes a su enfoque de la enseñanza; los trabajos prácticos deberán ser rediseñados; los métodos de evaluación tradicionales ya están siendo cuestionados. El objetivo principal de este proyecto es revelar el impacto de dichas tecnologías en tres dimensiones centrales: las prácticas docentes, el aprendizaje de los estudiantes y las estrategias de evaluación.

***Palabras clave*:** Inteligencia Artificial Generativa, Enseñanza de programación, Educación universitaria, Prácticas docentes, Evaluación.

CONTEXTO

Este proyecto se inscribe en la línea de investigación institucional de la Universidad de Belgrano "Abordaje de problemas de la educación" y corresponde al área Teoría de la Computación, Programación y Lenguajes de la Facultad de Ingeniería y Tecnología Informática.

Dentro de esta línea institucional, en la Facultad de Ingeniería y Tecnología Informática, se encuentra en ejecución una línea específica para el desarrollo de herramientas para la enseñanza en Ciencias de la Computación (CS). Dentro de esta línea se encuentra en desarrollo un proyecto orientado a la enseñanza de la IA "Framework para la construcción de estrategias para el desarrollo de competencias en IA" (FEDAI), en el que se encuadra el trabajo que se presenta.

Los proyectos correspondientes a esta línea de investigación se caracterizan por estar diseñados de modo de facilitar la incorporación de alumnos de las carreras de grado en cualquiera de sus etapas. En todos los casos los proyectos culminan en la construcción de modelos didácticos sobre los que se construyen aplicaciones informáticas.

1. INTRODUCCIÓN

En 2022, la publicación de ChatGPT marcó un avance significativo en el uso de chatbots de Inteligencia Artificial (IA), con impacto notable en campos como la formación de profesionales de informática. Sin embargo, aún no habría un acuerdo generalizado acerca de cómo los estudiantes integran los resultados de las consultas que proponen a las IA generativas (GenAI) en sus procesos de aprendizaje [1, 2, 3].

La IA generativa presenta desafíos y oportunidades para la educación informática y, en particular para la enseñanza introductoria de la programación, que exigirán estrategias pedagógicas actualizadas, centradas en nuevas competencias [2, 3]. En esta transformación didáctica el estudiante es colocado en el centro del proceso de aprendizaje, lo que implica una redefinición inevitable del papel del profesor y, por lo tanto, de la práctica docente [6]. Es sin dudas un cambio copernicano.

La IA ha impuesto un nuevo paradigma en el diseño de aplicaciones informáticas y esto, a su vez, ha despertado la preocupación con respecto a cómo enseñarla, en particular, en los cursos de educación superior. Esto ha puesto en el eje de la formación de profesionales de la computación la reformulación de contenidos y de las prácticas de enseñanza [7].

Algunos autores esperan que, con la generación de código impulsada por IA, la programación en lenguaje natural y el aprendizaje colaborativo, pueda crearse una educación en programación más inclusiva y eficaz para preparar a la próxima generación de programadores [8]. Sin embargo, dado que los modelos de IA generativa son capaces de generar soluciones a problemas típicos de los cursos introductorios de programación, surgen preocupaciones sobre una posible dependencia excesiva y un uso indebido por parte de los estudiantes [8].

Las competencias desarrolladas en los cursos introductorios de programación son consideradas un fundamento esencial para las competencias informáticas más avanzadas; por tanto, deberían proveer habilidades básicas requeridas para cursos posteriores dentro de los planes de estudios de las carreras de informática [9, 7].

Alineado con dicha premisa, este proyecto está dirigido al estudio de la aplicación de herramientas GenAI en la enseñanza introductoria de la programación. Se considerarán para este caso las instancias de: presentación de contenidos que el docente hace a los alumnos; la producción de trabajos prácticos; y la evaluación de aprendizajes.

Por un lado, "…la aplicación de la GenAI en la enseñanza facilita un aprendizaje activo y personalizado durante la generación de código, la resolución interactiva de problemas y proyectos colaborativos" [8], y "…permite enfrentar a los alumnos a tareas auténticas y desafiantes que requieren una integración de habilidades cognitivas y técnicas, fomentando el aprendizaje autónomo y colaborativo" [7], por otro, en lo que respecta a las evaluaciones, el hecho de que los estudiantes puedan utilizar fácilmente ChatGPT y herramientas similares para completar una amplia gama de tareas con un esfuerzo mínimo de su parte preocupa esencialmente a los docentes [8].

Si bien se desconoce a ciencia cierta el impacto que la llegada de la GenAI pueda tener en la evaluación del aprendizaje y la integridad académica, es posible anticipar que los métodos de evaluación deberán centrarse más en evaluar el pensamiento crítico y los procesos de aprendizaje que los productos finales [7].

El desarrollo de competencias apela a abordajes didácticos que promuevan la reflexión sobre los procesos cognitivos personales [8]. En [9] se describen cuatro abordajes de la enseñanza de la programación. Tres de ellos, con diferentes matices con respecto a la organización del trabajo en el aula, promueven la introspección, el análisis de los propios procesos internos, y recurren a la evaluación durante el proceso. Este método surge como un recurso idóneo para corroborar los conocimientos y competencias adquiridas por los alumnos cuando se incorporan herramientas GenAI a los cursos de programación. Sin embargo, los resultados en [10] muestran que el enfoque de enseñanza depende fuertemente de las concepciones didácticas del docente y que aún se puede encontrar en las aulas quienes siguen adhiriendo a métodos conductistas.

El hecho de que las herramientas de GenAI brinden orientación en tiempo real y

personalizada permitiría que los estudiantes con menos experiencia completen tareas de programación que de otro modo les resultarían difíciles. “Sin embargo, aún no está claro si GenAI realmente “igualan las condiciones” o si la dependencia de estas herramientas afecta el aprendizaje profundo y el pensamiento crítico” [11].

La integración exitosa de la GenAI en la enseñanza de la programación dependería de estrategias didácticas intencionales, evaluaciones diseñadas concienzudamente y procesos estructurados de integración [12]. No obstante esta afirmación, serán las concepciones de los docentes las que influirán en sus decisiones con respecto a qué incluyen, qué excluyen, cómo transforman sus prácticas, qué estrategias pedagógicas son las adecuadas para enseñar la materia a sus alumnos [10, 13]. Sin dudas, la incorporación de la GenAI en la enseñanza introductoria de la programación en las carreras de informática implicará la adaptación y transformación de los métodos didácticos tradicionales. Un estudio sobre las concepciones docentes con respecto a la aplicación de la GenAI en el aula permitiría diseñar planes de capacitación docente.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Objetivo:

Describir cómo el uso de GenAI transforma la dinámica pedagógica en la enseñanza de la programación, identificando beneficios, riesgos y desafíos en la formación de competencias fundamentales.

Metodología:

Para el estudio de las concepciones didácticas se requiere un proceso de indagación y revelación de impresiones y percepciones que fundamentan las decisiones didácticas de los docentes que participen en la investigación. La codificación de los relatos de los docentes, en términos de la teoría de la didáctica general y, en particular, de los modelos de enseñanza [13,12], permitiría describir la manera en que los docentes enfocan la enseñanza introductoria [14,15].

Se propone un enfoque mixto que combina encuestas y entrevistas a docentes y estudiantes, análisis comparativo de cohortes con y sin acceso a GenAI, y estudios de caso en instituciones que ya integran estas herramientas. Además, se incluirá un análisis de percepciones éticas vinculadas a autonomía, plagio y responsabilidad académica.

3. RESULTADOS ESPERADOS

Se obtendrá una clasificación de los principales usos de GenAI en programación (asistente de sintaxis, generador de código, tutor de debugging), la identificación de riesgos (dependencia, pérdida de autonomía, dilemas éticos) y oportunidades (creatividad, motivación, personalización del aprendizaje), así como recomendaciones prácticas para docentes sobre integración responsable de estas tecnologías.

Las categorías resultantes del estudio de concepciones didácticas con respecto a la aplicación de herramientas GenAI en la enseñanza introductoria de la programación facilitarán el diseño de planes de capacitación docente.

4. FORMACIÓN DE RECURSOS HUMANOS

En el presente proyecto se está formando en investigación a una docente de la carrera de Ingeniería en Informática, y a una alumna que está desarrollando su trabajo final de carrera.

El diseño del proyecto permite la incorporación paulatina de estudiantes a lo largo de sus diferentes etapas, como auxiliares de investigación. Alumnos de la materia Habilitación Profesional II participarán en el diseño de aplicaciones didácticas. Los investigadores que integran el equipo son docentes de las carreras de informática de la universidad. Esta circunstancia facilita la transferencia de conocimientos hacia los alumnos.

5. BIBLIOGRAFÍA

[1] Marina Lepp, Joosep Kaimre (2025) *Does generative AI help in learning programming: Students' perceptions, reported use end relation to performance*, Computers in Human Behaviour Reports, 18 (2025),

[2] Paez, N., Marcet, D., & Cozzi, G. (2025). *Uso de la inteligencia artificial en la enseñanza de la programación: un estudio preliminar*. JAIIO, Jornadas Argentinas De Informática, 11(8), 74-79.https://revistas.unlp.edu.ar/JAIIO/article/view/19934

[3] Vera Paredes, D. A., Franco Arias, O. O., & Córdova Martínez, L. C. (2025). *Impacto de la inteligencia artificial en el aprendizaje de la programación informática en los estudiantes universitarios*. Ciencia Y Educación, 6(3), 33 - 50. Elsevier.

[4] Amruth N. Kumar, Rajendra K. Raj, Sherif G. Aly, Monica D. Anderson, Brett A. Becker, Richard L. Blumenthal, Eric Eaton, Susan L. Epstein, Michael Goldweber, Pankaj Jalote, Douglas Lea, Michael Oudshoorn, Marcelo Pias, Susan Reiser, Christian Servin, Rahul Simha, Titus Winters, and Qiao Xiang. (2023). Computer Science Curricula 2023. ACM Press, IEEE Computer Society Press and AAAI Press. DOI: https://doi.org/10.1145/3664191

[5] Abdulla, S et al. 2024. *Using ChatGPT in Teaching Computer Programming and Studying its Impact on Students Performance*, Electronic Journal of e-Learning, 22(6), pp 66-81, ISSN 1479-4403 66

[6] Llerena-Izquierdo, J.; Mendez-Reyes, J.; Ayala-Carabajo, R.; Andrade-Martinez, C. (2024) *Innovations in Introductory Programming Education: The Role of AI with Google Colab and Gemini.* Educ. Sci. 2024, 14, 1330. https://doi.org/10.3390/educsci14121330

[7] Piccin, A.M.; Hadad, Graciela D.S.; Angeleri, P. (2024) *Hacia la construcción de estrategias para la enseñanza de inteligencia artificial en carreas de informática*, XXVI Workshop de Investigadores en Ciencias de la Computación (WICC) (Puerto Madryn, 18 y 19 de abril de 2024) ISBN: 978-987-8352-57-2 Pp: 332-336

[8] Agbo, F.J et al. (2025) *Computing education using generative artificial intelligence tools: A systematic literature review*; Computers and Education Open Volume 9, December 2025.

[9] Piccin, A.M.; Hadad, G.D.S., et al. (2024) *Concepciones de Docentes de Cursos Introductorios de Programación en relación con las Competencias enunciadas en Estándares*, EJS 23 (2) 2024 p. 111-131 112

[10] Piccin, A.M. (2021) *Enseñanza de la programación: las concepciones didácticas de los docentes de los cursos introductorios*. Autora: XIII Workshop de Investigadores en Ciencias de la Computación 335 15 y 16 de abril de 2021 RedUNCI- UNdeC. ISBN: 978-987-24611-3-3

[11] Houda Bouamor, Gabriela Gongora-Svartzman, Larry Heiman, Shihong Huang (2025) *Evaluating GenAI's Effectiveness for Students with Varied Programming Backgrounds in a Software Development Course,* SIGCSE TS 2025, February 26-March 1, 2025, Pittsburgh, PA, USA

[12] Jemimah Nathaniel et al. (2025) *"Literature Review on the Integration of Generative AI in Programming Education",* International Journal of Artificial Intelligence in Education (2025) 35:2724–2755 https://doi.org/10.1007/s40593-025-00524-3

[13] Camilloni, A.R.W.: *Los profesores y el saber didáctico*. En: Camilloni, A. (comp.) El saber didáctico. Paidós, Buenos Aires (2008).

[14] Joyce, B.R., Weil, M., Calhoum, E.: Models of Teaching. 9th. ed. Pearson (2017)

[15] Gimeno Sacristán, J., Pérez Gómez, A. I.: Comprender y transformar la enseñanza. Ediciones Morata, Madrid (2002).

La integración de la Inteligencia Artificial en la formación del Ingeniero en Sistemas: desafíos, oportunidades y transformaciones en la práctica docente

Gabriela Caliva[1]

[1] Facultad de Tecnología Informática, Universidad Abierta Interamericana, Rosario, Argentina.

gabrielalorena.caliva@uai.edu.ar

RESUMEN

La creciente presencia de la Inteligencia Artificial Generativa (IAG) en la industria tecnológica está transformando las competencias requeridas para el ejercicio profesional del Ingeniero en Sistemas. Sin embargo, su incorporación en la formación universitaria continúa siendo parcial y dispareja.

Este trabajo presenta una síntesis de una línea de investigación orientada a la innovación educativa en la carrera de Ingeniería en Sistemas de la Universidad Abierta Interamericana (UAI). Desde una mirada situada, se combinan aportes teóricos con un relevamiento exploratorio realizado con estudiantes de segundo año, en el marco de la asignatura Ingeniería de Requerimientos.

Los resultados evidencian un uso extendido de herramientas de IA en contextos académicos, una valoración positiva respecto de su aporte al aprendizaje y preocupaciones vinculadas a la dependencia, la confiabilidad de la información y el uso ético. A partir de estos hallazgos, se reflexiona sobre las transformaciones que demanda la práctica docente universitaria y se proponen líneas de acción orientadas a fortalecer la formación del Ingeniero en Sistemas frente a este escenario tecnológico.

Palabras clave: Inteligencia Artificial Generativa; Educación en Ingeniería; Práctica docente universitaria; Ingeniería de Requerimientos.

CONTEXTO

La Inteligencia Artificial (IA), y en particular la Inteligencia Artificial Generativa (IAG), se ha convertido en un componente estructural del ecosistema tecnológico contemporáneo. Herramientas basadas en modelos generativos, como ChatGPT, Copilot o Gemini, están modificando de manera acelerada los procesos de desarrollo de software, análisis de información y resolución de problemas complejos, redefiniendo competencias profesionales y roles laborales en el ámbito de la ingeniería.

En este escenario, la formación universitaria de los Ingenieros en Sistemas enfrenta el desafío de integrar tecnologías que evolucionan a gran velocidad y que ya forman parte de las prácticas cotidianas de los estudiantes. El estudiantado no es ajeno a este fenómeno: utiliza IA para estudiar, programar, analizar problemas y resolver ejercicios, incluso cuando su uso no siempre es abordado de manera explícita en las asignaturas.

La presente línea de investigación se desarrolla en el contexto de la carrera de Ingeniería en Sistemas de la Universidad Abierta Interamericana (UAI), a partir de una experiencia situada en la asignatura Ingeniería de Requerimientos, correspondiente al segundo año de la carrera. Esta materia resulta particularmente pertinente para explorar el uso de IA como herramienta de apoyo en procesos de elicitación, análisis y validación de requerimientos, así como en la simulación de escenarios y stakeholders.

1. INTRODUCCIÓN

La incorporación acelerada de la Inteligencia Artificial Generativa en la industria del software está produciendo transformaciones significativas en los modos de trabajo de los profesionales de sistemas y, en consecuencia, en los procesos de formación universitaria. Herramientas basadas en IA no solo automatizan tareas, sino que también funcionan como asistentes cognitivos capaces

de intervenir en actividades de análisis, diseño y toma de decisiones.

En este contexto, las carreras de Ingeniería en Sistemas se enfrentan al desafío de revisar sus propuestas formativas, tanto en términos de contenidos como de estrategias didácticas. La presencia extendida de la IA en las prácticas cotidianas de los estudiantes plantea interrogantes sobre el rol del docente, los criterios de evaluación y las competencias que resultan necesarias para una formación profesional pertinente.

Este trabajo se propone abordar estas problemáticas desde una perspectiva situada, analizando el impacto del uso de Inteligencia Artificial Generativa en la formación del Ingeniero en Sistemas y las transformaciones que ello demanda en la práctica docente universitaria.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La línea de investigación presentada se orienta a analizar la integración pedagógica de la Inteligencia Artificial Generativa en la formación del Ingeniero en Sistemas. En particular, se estructura en torno a los siguientes ejes:

- Uso de la IA como herramienta de apoyo en los procesos de enseñanza y aprendizaje en carreras de ingeniería.
- Identificación de competencias emergentes asociadas al trabajo con IA, tales como la formulación de prompts, la validación crítica de resultados y el uso responsable de la tecnología.
- Análisis de las transformaciones del rol docente universitario en contextos educativos mediados por IA.

La línea surge a partir de experiencias concretas de aula y se proyecta como un espacio de articulación entre docencia, investigación e innovación educativa.

3. RESULTADOS

- Resultados esperados

A partir de esta línea de trabajo se espera contribuir a la mejora de las prácticas docentes mediante la integración planificada de la IA en las asignaturas de la carrera, el desarrollo de competencias emergentes en los estudiantes y la generación de insumos para la reflexión institucional sobre la actualización curricular y la formación docente.

- Resultados obtenidos

El relevamiento exploratorio se realizó mediante una encuesta online, anónima y voluntaria, aplicada a estudiantes de segundo año de la carrera de Ingeniería en Sistemas de la Universidad Abierta Interamericana, en el marco del cursado de la asignatura Ingeniería de Requerimientos. La encuesta fue respondida por 31 estudiantes y permitió relevar usos, valoraciones y percepciones vinculadas al empleo de herramientas de Inteligencia Artificial Generativa en contextos académicos.

El análisis del relevamiento exploratorio realizado a estudiantes de segundo año de la carrera de Ingeniería en Sistemas muestra un uso extendido de herramientas de Inteligencia Artificial Generativa en contextos académicos. La totalidad de los estudiantes encuestados declaró haber utilizado IA, siendo ChatGPT, Copilot y Gemini las herramientas más mencionadas.

Los principales usos identificados incluyen la explicación de contenidos no comprendidos, la búsqueda de información, la generación de ideas, la escritura y corrección de trabajos y la resolución de ejercicios de programación. En el contexto específico de la asignatura Ingeniería de Requerimientos, los estudiantes destacan el uso de la IA para la generación y reformulación de requerimientos, la validación de ideas y la simulación de stakeholders o escenarios de elicitación.

Las valoraciones generales sobre el uso de la IA son mayoritariamente positivas. Los estudiantes señalan como ventajas la mejora en la comprensión de conceptos, la agilización de procesos, la posibilidad de obtener distintas perspectivas sobre un mismo problema y el apoyo en la redacción y organización de ideas. Al mismo tiempo, emergen preocupaciones vinculadas a la dependencia excesiva, la reducción del esfuerzo personal, la confiabilidad de las

respuestas y aspectos éticos asociados al uso responsable de la tecnología.

Asimismo, se evidencia una expectativa clara respecto al rol docente, especialmente en relación con la enseñanza explícita de estrategias de uso de la IA, la formulación de prompts efectivos, la validación crítica de la información generada y la incorporación de propuestas didácticas basadas en simulaciones y casos prácticos.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación contribuye a la formación de recursos humanos tanto en el nivel de grado como en el ámbito de la docencia universitaria. En el plano estudiantil, promueve el desarrollo de competencias vinculadas al uso crítico, reflexivo y responsable de la Inteligencia Artificial Generativa, fortaleciendo habilidades como el análisis de requerimientos, la validación de información y la toma de decisiones fundamentadas.

En el ámbito docente, la línea impulsa procesos de reflexión y actualización sobre las prácticas de enseñanza en contextos mediados por tecnologías emergentes, favoreciendo el intercambio de experiencias y la incorporación de estrategias didácticas innovadoras. Asimismo, se proyecta como un espacio de articulación con futuras producciones académicas y de posgrado vinculadas a la temática.

5. REFERENCIAS

Andreoli, S., Aubert, E., Cherbavaz, M. C., & Perillo, L. (2024). Entre humanos y algoritmos: percepciones docentes sobre la exploración con inteligencia artificial generativa en la enseñanza del nivel superior. Revista de Educación Superior, 53(2), 45–62.

García Cabrera, L., Ruano Ruano, I., Balsas Almagro, J. R., & Quesada Real, F. J. (2025). Uso de inteligencia artificial generativa en educación de ingeniería: Propuestas de integración. En Jornadas de Automática (Vol. 46). Universidad de Jaén.

Hidalgo Abdala, M. F. (2024). Usos pedagógicos de la inteligencia artificial generativa (IAG) en la docencia universitaria de grado. Revista Latinoamericana de Tecnología Educativa, 23(1), 67–83.

Lion, C., & Maggio, M. (2019). Desafíos para la enseñanza universitaria en los escenarios digitales contemporáneos. Cuadernos de Investigación Educativa, 10(1), 13–25.

Martínez-Valdez, R. I., Pérez-Martínez, K. S., & Cuevas-Gutiérrez, A. I. (2024). Motivos, satisfacción y frecuencia de uso de la inteligencia artificial generativa en estudiantes universitarios de ingeniería y administración. Revista Mexicana de Investigación Educativa, 29(101), 455–478.

Moyano, E., Jerez, M., Romano, L., Villarreal, M., Moncho, M., Villanueva, T., Rigoni, B., Andrade, D., & Iost, D. (2021). Diseño de estrategias didácticas para la enseñanza de ingeniería de requerimientos. Revista Iberoamericana de Educación en Ingeniería, 15(2), 33–48.

Sánchez Mendiola, M., & Carbajal Degante, E. (2023). La inteligencia artificial generativa y la educación universitaria: ¿Salió el genio de la lámpara? Perfiles Educativos, 45(Especial), 70–86.

Modelo curricular AI-Native con Trazabilidad N4 para la Formación Universitaria en Programación

Alberto Cortez[1,2], Carlos Martinez[1], Mariana Brachetta[1,2], Juan Sarmiento[1], Juan Robledo[1], Claudia Naveda[1], Bruno Roberti[1], Daniela Carbonari[1]

[1]UTN Facultad Regional Mendoza, Ingeniería en Sistemas de Información
Laboratorio de Auditoría y Seguridad de TIC (AusegTics) / Laboratorio de Integración de Tecnologías Aplicadas a Prototipos de Software (LITAPS)
[2]Universidad Champagnat

{alberto.cortez, carlos.martinez, mariana.brachetta, juan.sarmiento, juan.robledo, claudia.naveda, bruno.roberti, daniela.carbonar} @ docentes.frm.utn.edu.ar

RESUMEN

La presencia de sistemas de Inteligencia Artificial generativa está modificando profundamente el aprendizaje de la programación y la resolución algorítmica. La posibilidad de obtener código funcional sin evidencia del razonamiento subyacente afecta la validez de los modelos educativos tradicionales. En este trabajo se presenta una propuesta de diseño del Modelo Curricular AI-Native con Trazabilidad Cognitiva N4, que articula un marco teórico de cognición distribuida, una taxonomía cognitiva específica (TC-AIN), una arquitectura institucional interoperable y un sistema de trazabilidad organizado en cuatro niveles (N1–N4). El modelo aspira a permitir la reconstrucción verificable del proceso cognitivo del estudiante en escenarios híbridos humano–IA, asegurando la transparencia, la auditoría y la evaluación auténtica del aprendizaje.

Palabras clave: *IA Generativa, Cognición Distribuida, Programación, Evaluación, Auditoría Algorítmica, Educación Superior.*

CONTEXTO

La Universidad Tecnológica Nacional – Facultad Regional Mendoza (UTN-FRM), a través del AuSegTics y LITAPS, y la Universidad Champagnat por medio de su Facultad de Informática y Diseño y el Instituto de Investigación de la UCh, han manifestado interés en desarrollar una línea conjunta de trabajo orientada a la transformación digital y la innovación tecnológica, promoviendo una cooperación interinstitucional basada en el trabajo colaborativo, el intercambio académico y el desarrollo de proyectos con impacto regional. El proyecto surge de la constatación empírica de que los instrumentos evaluativos habitualmente utilizados en espacios curriculares vinculados a la enseñanza de la programación —como exámenes parciales, trabajos prácticos y proyectos integradores— fueron concebidos bajo el supuesto de que el código entregado por el estudiante constituye evidencia directa de su proceso de razonamiento, un supuesto que el uso de IA generativa invalida en su base epistemológica.

La IA generativa produce soluciones completas que pueden ser entregadas sin que haya mediado un proceso de construcción de conocimiento y comprensión conceptual por

parte del alumno. Ma et al. demuestran que los estudiantes la utilizan predominantemente para corregir código, evitando procesos cognitivos superiores como planificación y evaluación [1], lo que genera delegación epistémica, dependencia cognitiva y pérdida de agencia. Esta situación configura un desafío estructural para la calidad educativa. El contexto institucional refuerza la pertinencia del proyecto: la UTN Facultad Regional Mendoza ofrece las carreras de Ingeniería en Sistemas de Información y Tecnicatura Universitaria en Programación, mientras que la Universidad Champagnat dicta la Licenciatura en Sistemas de Información, conformando en conjunto un ecosistema académico consolidado, con un cuerpo docente calificado e infraestructura tecnológica adecuada. La articulación entre ambas universidades incorpora, además, perspectivas complementarias provenientes de las ciencias de la educación y de la ética aplicada.

1. INTRODUCCIÓN

Diversos marcos conceptuales permiten comprender la profundidad del fenómeno. La Cognición Distribuida de Hutchins [2] sostiene que el pensamiento se distribuye entre el ser humano y las herramientas que utiliza; en el contexto educativo, cuando un estudiante interactúa con IA generativa, el proceso cognitivo resultante es inherentemente híbrido. La Mente Extendida de Clark y Chalmers [3] radicaliza esta posición al postular que los sistemas de IA participan constitutivamente de los procesos cognitivos, obligando a redefinir qué entendemos por "pensar" y "aprender" cuando el entorno incluye agentes artificiales capaces de generar soluciones de forma autónoma.

Desde la perspectiva normativa, los estándares internacionales convergen en exigir trazabilidad, transparencia y responsabilidad humana. La Recomendación de UNESCO sobre la Ética de la IA [4] establece principios de proporcionalidad y rendición de cuentas aplicables a contextos educativos. El marco IEEE de Diseño Éticamente Alineado [5] prioriza el bienestar humano y la transparencia algorítmica. La norma ISO/IEC 23894 [6] proporciona un marco de gestión de riesgos de IA adaptable al ámbito educativo. Frente a esta convergencia, la educación superior requiere mecanismos concretos para registrar y auditar el proceso cognitivo del estudiante en interacción con IA, superando la dicotomía entre "uso permitido" y "uso prohibido".

La literatura reciente aborda esta problemática desde perspectivas complementarias. Parra [7] advierte que los filtros epistémicos del estudiante se debilitan ante respuestas de IA que aparentan corrección. Levchuk [9] implementa un tutor inteligente basado en IA generativa para programación, aunque limitado a una lógica tutorial sin arquitectura curricular institucional. Sayad [10] argumenta que la facilidad para obtener respuestas erosiona el pensamiento crítico y la reflexión autónoma. Gil et al. [11] destacan que la trazabilidad debe entenderse como dimensión ética y formativa, no solo como mecanismo de control técnico. Gallardo et al. [12] proponen criterios de transparencia evaluativa desde la Educación 5.0. Pardines Murcia [13] enfatiza la verificación de la trazabilidad en recursos educativos generados por IA. Lion y Lippenholtz [14] proponen una auditoría educativa sostenida desde la cognición distribuida. Naharros Pérez [15] ofrece directrices éticas para IA en contextos profesionales. Vidarte et al. [16] proponen una alfabetización digital crítica que articule ética y competencias algorítmicas. Porcelli [17] reflexiona sobre autoría y responsabilidad en contextos mediados por IA. Medici [18] argumenta que la trazabilidad debe garantizar

la soberanía epistemológica de las instituciones educativas sobre sus propios datos y procesos formativos.

En conjunto, estos estudios coinciden en la urgencia de redefinir los marcos epistemológicos, éticos y evaluativos ante el uso de la IA generativa en educación, aunque la mayoría se limita al diagnóstico conceptual o a soluciones parciales. En este sentido el Modelo AI-Native con Trazabilidad Cognitiva N4 propuesto en el presente trabajo aspira a ofrecer una propuesta integral que aborda dimensiones vinculadas al rediseño curricular y la trazabilidad del proceso cognitivo en entornos mediados por IA, articulando una arquitectura tecnológica interoperable y protocolos de gobernanza.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Este trabajo se enmarca en la línea de investigación sobre rediseño curricular y evaluación en entornos mediados por IA generativa. El foco principal es el diseño, desarrollo y validación del modelo AI-Native con Trazabilidad Cognitiva N4 para carreras universitarias que involucran la enseñanza de programación y la resolución algorítmica. La denominación "AI-Native" expresa que el modelo concibe la IA como condición estructural del entorno formativo, desde la cual se diseña la totalidad de la propuesta curricular, pedagógica, evaluativa y tecnológica.

El componente central es el Sistema de Trazabilidad Cognitiva N4, que permite reconstruir la trayectoria cognitiva completa del estudiante mediante cuatro niveles complementarios:

N1 – Comprensión del Problema. Registra cómo el estudiante interpreta el enunciado, identifica entidades del dominio y activa conocimientos previos antes de codificar. Evita la ilusión de competencia descrita por Sweller [8]: sin registro explícito de la comprensión inicial, es imposible distinguir entre un estudiante que comprende el problema y delega selectivamente a la IA, y otro que simplemente copia un prompt y acepta la primera respuesta generada.

N2 – Razonamiento Explicado. Captura las decisiones estratégicas durante la resolución: elección de algoritmos, planificación de la estructura del programa y justificación de las decisiones de diseño. Según Ma et al. [1], esta es la fase más erosionada por la IA generativa, ya que los estudiantes delegan las decisiones de orden superior reteniendo sólo tareas mecánicas de corrección sintáctica. El nivel N2 exige que el estudiante articule las razones de sus decisiones antes de implementarlas en código.

N3 – Evaluación y Auditoría Cognitiva. Incluye verificación conceptual del código, análisis de errores y auditoría explícita de las respuestas de IA cuando fueron consultadas. Responde a los riesgos epistémicos señalados por Parra [7], quien advierte que los filtros epistémicos se debilitan cuando la IA produce respuestas que aparentan corrección y completitud. El nivel N3 requiere que el estudiante identifique posibles errores, sesgos o limitaciones en las sugerencias recibidas.

N4 – Integración Humano–IA y Gobernanza. Registra los prompts enviados a sistemas de IA, las intenciones cognitivas que motivaron cada consulta, las respuestas recibidas y las decisiones del estudiante respecto a aceptar, modificar o rechazar cada sugerencia con su justificación. Se alinea con los principios de transparencia de IEEE [5] e ISO/IEC 23894 [6], cerrando el ciclo de

auditoría al hacer visible la interacción completa entre razonamiento humano y generación artificial.

Se propone operativizar la trazabilidad cognitiva mediante un rastro digital verificable - Cognitive Trace Record (CTR) que codifica estados cognitivos, decisiones y eventos de interacción en formato estructurado y auditable. El CTR registra cada transición entre niveles N1–N4, permitiendo reconstruir la secuencia temporal completa del proceso de resolución. Los niveles funcionan como una ontología interoperable con los componentes tecnológicos institucionales: AI-Gateway (intermedia interacciones con modelos de lenguaje), VectorDB (almacena CTR con búsqueda semántica), Git (versiona artefactos de código) y n8n (orquesta auditorías automatizadas). Esta arquitectura habilita auditorías que detectan patrones de delegación cognitiva excesiva y generan alertas ante desviaciones significativas.

Además, el modelo propone incorporar una Taxonomía Cognitiva -AI-Native (TC-AIN)-, que distingue seis niveles de actividad cognitiva en programación con IA: reconocimiento, reproducción, aplicación guiada, adaptación crítica, diseño autónomo e integración gobernada. La TC-AIN incorpora explícitamente la interacción con IA como variable constitutiva de cada nivel, evaluando no solo la complejidad cognitiva de la tarea sino la calidad de la colaboración humana–IA.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto se encuentra en fase de diseño y prototipado. Como resultado obtenido, se ha completado el diseño curricular de una Tecnicatura Universitaria en Programación y Sistemas con Inteligencia Artificial Agéntica, que operacionaliza los principios del modelo AI-Native en una estructura académica formal destinada a su presentación.

Se ha elaborado la taxonomía cognitiva TC-AIN con sus seis niveles y se ha formalizado la estructura del Cognitive Trace Record (CTR) como artefacto central del sistema de trazabilidad. Asimismo, se ha definido la arquitectura tecnológica institucional (AI-Gateway, VectorDB, sistema de versionado, orquestador de auditorías) y sus interfaces de interoperabilidad.

Como resultados esperados, se anticipan mejoras en tres dimensiones. En la evaluativa, la trazabilidad permitirá fundamentar calificaciones en evidencia verificable del proceso cognitivo, superando los instrumentos tradicionales que solo evalúan el producto final. En la pedagógica, la explicitación del pensamiento promoverá la metacognición, el pensamiento crítico y la autorregulación del aprendizaje. En la institucional, el modelo aspira a ofrecer una arquitectura interoperable y replicable con protocolos de gobernanza alineados con estándares de UNESCO, IEEE e ISO/IEC 23894, contribuyendo a un marco de referencia para la educación en programación en la era de la IA generativa.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente proyecto contribuye a la formación de recursos humanos mediante un esquema colaborativo que integra docentes-investigadores, graduados, estudiantes de grado y becarios de ambas universidades.

Se prevé impacto directo en:

- Formación de docentes y actualización de contenidos en las cátedras: Desarrollo de

Software, Seguridad en los Sistemas de Información y Gestión Gerencial.
- Capacitación de alumnos de grado mediante participación en modelado, desarrollo, integración y testing del prototipo.
-Becarios de investigación con responsabilidades técnicas específicas.
-Apoyo a Trabajos Finales de Grado y proyección a tesis de posgrado vinculadas a la temática.
-Vinculación y transferencia hacia pares de investigación y profesionales del sector (software/salud digital).

5. BIBLIOGRAFÍA

[1] B. Ma et al., "Scaffolding Metacognition in Programming Education: Understanding Student-AI Interactions" arXiv preprint arXiv:2511.04144, 2025.

[2] E. Hutchins, Cognition in the Wild. Cambridge, MA: MIT Press, 1995.

[3] A. Clark and D. Chalmers, "The Extended Mind" Analysis, vol. 58, no. 1, pp. 7-19, 1998.

[4] UNESCO, Recommendation on the Ethics of Artificial Intelligence, 2021.

[5] IEEE, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition, 2019.

[6] ISO/IEC 23894:2023, Artificial Intelligence — Risk Management.

[7] D. E. R. Parra , "Uso de la inteligencia artificial en educación universitaria: Comprensión desde los filtros epistémicos en el proceso de reducción de la evidencia cualitativa," Dialnet, vol. 32, no. 24, 2024.

[8] J. Sweller, P. Ayres, and S. Kalyuga, Cognitive Load Theory. New York, NY: Springer, 2011.

[9] O. Levchuk, "Diseño y evaluación de un tutor inteligente basado en IA Generativa para la adquisición de habilidades de programación" tesis de maestría, CICESE, 2024.

[10] A. L. V. Sayad, Inteligencia artificial y pensamiento crítico. Camino para la educación. Bogotá: Corporación Universitaria Minuto de Dios – UNIMINUTO, 2024.

[11] C. R. Gil, J. M. G. Vilela, and M. V. Martínez, Epistemología, Ética, Educación e Inteligencia Artificial. Editorial Alema, 2024.

[12] A. R. Gallardo, J. R. H. Morales, and S. S. Carrillo, "Impacto de la Inteligencia Artificial Generativa en la Educación 5.0" 2024.

[13] C. E. Pardines Murcia, "Uso de la inteligencia artificial para la generación de recursos y unidades didácticas," Trabajo de Fin de Máster, Universidad Miguel Hernández de Elche, 2024.

[14] C. Lion and B. Lippenholtz, Experimentar con IA: Notas para educadores alertas. Buenos Aires: Tilde editora, 2025.

[15] S. Naharros Pérez, "Elaboración de una guía de buenas prácticas para el empleo de la inteligencia artificial para profesionales TIC" Trabajo de Fin de Grado, Universitat Politècnica de València, RIUNet, 2024.

[16] J. A. B. Vidarte, P. A. B. Zelada, and R. Carranza, "La IA y la Educación Superior: Hacia una Alfabetización Digital Crítica y Ética" 2024.

[17] A. M. Porcelli, "La inteligencia artificial y la robótica: dilemas éticos y jurídicos". Derecho global. Estudios sobre derecho y justicia, vol. 6, no. 16, pp. 49-105, 2020.

[18] A. Medici, "Capitalismo de vigilancia, colonialismo y extractivismo digital: desafíos regulatorios para el derecho crítico" REDHES, Revista de Derechos Humanos y Estudios Sociales, 202

Pensamiento computacional, pensamiento crítico y programación en la universidad: Metodologías, herramientas y recursos para su enseñanza

Gladys N. Dapozo, Raquel H. Petris, Cristina L. Greiner, Ana M. Company, María C. Espíndola, Yanina Medina, Silvana Armana, María Isabel Sánchez, Jeremías González

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste. 9 de Julio Nº 1449. Corrientes. Argentina.
{gndapozo, cgreiner, acompany, mcespindola, yanina}@exa.unne.edu.ar, {rpetris, silvana.armana}@comunidad.unne.edu.ar, marita@unne.edu.ar, jeremiasgonzalez7464@gmail.com

RESUMEN

La Ingeniería de Software comprende todas las etapas del desarrollo de sistemas de información, siendo la programación su herramienta central. El pensamiento computacional resulta esencial para este campo, ya que impulsa la abstracción, la lógica y la creatividad. En el contexto educativo, la incorporación de la IA generativa plantea el reto de reforzar el pensamiento crítico, entendido como la capacidad de analizar, evaluar y reflexionar en la toma de decisiones. Paralelamente, el desarrollo de la industria del software exige promover vocaciones en TIC, con especial atención a la inclusión de las mujeres, a fin de incrementar su representación en el sector. Esta línea de investigación tiene como propósito investigar y aplicar estrategias educativas innovadoras que integren herramientas de enseñanza de la programación e IA generativa. Se busca potenciar el pensamiento computacional y crítico en el aprendizaje de la programación, a la vez, fomentar el interés de jóvenes, y en particular motivar la participación de más mujeres en carreras tecnológicas.

Palabras clave: Enseñanza de la programación en la universidad, IA generativa. Pensamiento crítico.

CONTEXTO

Las líneas de I/D corresponden al PI 25F007 “Pensamiento computacional, pensamiento crítico y programación en la universidad: Metodologías, herramientas y recursos para su enseñanza”, aprobado por la SGCyT-UNNE para el periodo 2026-2029.

1. INTRODUCCIÓN

Este proyecto de investigación se orienta al estudio y aplicación experimental de estrategias educativas que incorporen modelos, metodologías y recursos innovadores que contribuyan a promover el pensamiento computacional a fin de favorecer el aprendizaje de la programación, habilidad clave para la construcción de sistemas informáticos.

La preocupación inicial que dio origen a estas líneas de investigación, estaba dada por el bajo número de graduados en las carreras de las áreas STEM (Science, Technology, Engineering and Mathematics), en particular en las carreras de Informática, en las que se observaban además preocupantes índices de deserción y desgranamiento. Así también, la baja participación de las mujeres en áreas de ciencia y tecnología. En particular, en las carreras de informática las mujeres representan tan sólo un 18% de los estudiantes, generando tanto una profunda inequidad en la distribución del ingreso y del capital intelectual, como privando al sector de la mirada de más de la mitad de la población. El estudio “Y las mujeres… ¿Dónde están? de la Fundación Sadosky [1], señala varias causas que alejan a las mujeres de la informática.

En el plano académico, las investigaciones del área dan cuenta de que el aprendizaje de la programación inicial conlleva cierto grado de dificultad para los estudiantes universitarios, que se refleja en los índices de desgranamiento en los primero años de las carrera de Informática.

Por otra parte, cabe destacar la dinámica que afecta el proceso de enseñanza de este tópico central en la formación de informáticos. Como una de las consecuencias de la pandemia, las carreras de Informática del país incrementaron notablemente su matrícula. Este fenómeno produjo una importante masividad de los cursos de los primeros años, afectando la enseñanza de la programación que requiere, por sus características, un seguimiento personalizado. Esta situación dio lugar a la implementación de nuevas estrategias pedagógicas, incorporando métodos y recursos de educación a distancia. En forma casi simultánea, se dio también la aparición de la IA generativa. Estas dos novedades generan un panorama más complejo para la enseñanza de la programación en la actualidad.

En este nuevo escenario, la enseñanza de la programación se ve atravesada por la IA. En un trabajo reciente [2] realizado en el marco de la primera asignatura de enseñanza de programación de la Licenciatura en Sistemas de Información de la UNNE (LSI) se detectó que el 83% de los estudiantes usa herramientas de IA en el proceso de aprendizaje de los conceptos básicos de programación. Los docentes observan muchas situaciones en las que los alumnos presentan el código de una solución tomado de alguna herramienta de IA, que si bien cumplen con lo exigido por la cátedra, no entienden la solución que proponen, esto es, no hay aprendizaje por parte del alumno.

Esta propuesta de investigación busca promover el pensamiento computacional porque es una habilidad primordial para la programación, pero a la vez fomentar también el pensamiento crítico, buscando que el alumno reflexione acerca de sus actitudes y resultados, especialmente en actividades asociadas con la IA.

La formación en programación en la universidad no solo implica la adquisición de destrezas técnicas, sino también el desarrollo de competencias cognitivas superiores que permitan a los estudiantes analizar, cuestionar, evaluar y tomar decisiones fundamentadas en entornos complejos.

En este sentido, el pensamiento crítico, entendido como la capacidad de razonar de manera lógica, reflexiva y autónoma, se complementa con el pensamiento computacional, que aporta herramientas para la resolución de problemas de forma estructurada, creativa y eficiente. La vinculación de ambas dimensiones resulta estratégica para formar profesionales capaces de diseñar soluciones tecnológicas responsables, evaluar críticamente los impactos de sus decisiones, y aplicar el razonamiento computacional en contextos interdisciplinarios y sociales.

La IA y la educación

El uso de aplicaciones basadas en IA generativa se ha incrementado en todos los ámbitos de la sociedad. El uso de enfoques de aprendizaje automático (*machine learning,* ML), que constituyen la base de muchas aplicaciones de IA, es hoy común en actividades cotidianas como las compras, las redes sociales y los servicios de entidades financieras, así como en áreas como la salud y la educación. En el ámbito educativo es necesario adoptar una mirada crítica frente al contenido generado por sistemas de IA, ya que tanto docentes como estudiantes deben ser conscientes de que los resultados producidos por estas tecnologías, en especial por la IA generativa, pueden contener imprecisiones y errores. Además, dichos resultados pueden estar cargados de sesgos, favoreciendo perspectivas mayoritarias y dominantes, y esta conciencia crítica suele estar ausente tanto en el profesorado como en el estudiantado [3].

La irrupción de la IA en las universidades presenta un gran dilema en los estudiantes: para qué aprender o memorizar o razonar si esto se puede obtener desde un dispositivo móvil con algoritmos de IA. Esto genera preocupación en los educadores dado que estas herramientas pueden resolver eficazmente tareas y problemas de exámenes, pero no producen conocimiento en los estudiantes [4].

Gracias a la IA generativa, el ámbito de la educación se encuentra en la cúspide de una revolución. Entre las ventajas de esta tecnología se puede mencionar: el aprendizaje personalizado, que permite adaptar los itinerarios educativos para que se ajusten a las necesidades individuales de los estudiantes; los sistemas de tutoría virtual pueden contribuir en áreas con escasez de docentes proporcionando retroalimentación instantánea, asegurando el progreso continuo de los estudiantes; la creación de contenidos fomenta la creatividad entre los estudiantes; la traducción en tiempo real ha experimentado notables avances en precisión; el procesamiento del lenguaje natural ha aumentado las interacciones entre humanos y computadoras, haciéndolas más intuitivas [5].

Sin embargo, este inmenso poder implica una gran responsabilidad. Una mayor dependencia de la IA podría llevar a un estancamiento cognitivo en los estudiantes, frustrando el propósito mismo de la educación. Desde el punto de vista ético, los modelos de IA, que reflejan los datos con los que se entrenan, pueden perpetuar inadvertidamente los sesgos sociales. Es crucial garantizar que estos modelos sean equitativos y no profundicen aún más las divisiones sociales. Por último, la cuestión de la autenticidad sigue siendo importante. En un mundo donde distinguir entre contenido generado por humanos y por IA se vuelve cada vez más difícil, garantizar la confianza y la transparencia es primordial. En conclusión, el futuro iluminado por la IA generativa es a la vez prometedor y desconcertante. A medida que la IA generativa continúe reformulando los paradigmas de enseñanza, aprendizaje y comunicación, nuestra responsabilidad colectiva es garantizar que su recorrido esté anclado en la ética, la equidad y la excelencia [6].

En este contexto, los educadores tienen la tarea crucial de preparar a los estudiantes para un futuro donde la IA será omnipresente, no solo en industrias impulsadas por la tecnología, sino en diversos sectores. Para lograr esto, los educadores deben dotar a los estudiantes con un conjunto integral de habilidades que incluya principios y conocimientos de programación tradicionales, además de la capacidad de aprovechar las herramientas de IA de vanguardia [7].

Por otra parte, diversos estudios dan cuenta de que el aprendizaje de la programación conlleva cierto grado de dificultad para los estudiantes. En [8] se señala que la investigación existente ha llevado a muchas discusiones e ideas sobre la mejor manera de enseñar la programación introductoria, y de las dificultades que enfrentan los estudiantes en este tipo de cursos. Entre estas, el hecho de que los estudiantes consideran a la programación como una actividad puramente técnica en lugar de un conjunto de habilidades combinadas de resolución de problemas. Por lo tanto, la mayoría de los estudiantes que se inician en la programación tiende a desarrollar un conocimiento superficial y no crean estrategias de resolución de problemas mediante el uso de construcciones de programación.

Estudios recientes [9] indican que: (a) el pensamiento computacional promueve el desarrollo sostenible a través de la educación STEM y de prácticas de enseñanza innovadoras; (b) constituye una habilidad fundamental para los procesos de resolución de problemas que evolucionan con el progreso tecnológico; (c) su desarrollo es una preocupación global, no limitada al nivel de desarrollo de un país; y (d) su introducción en una etapa temprana brinda oportunidades para el avance de grupos vulnerables.

A lo largo de las últimas dos décadas, el pensamiento computacional (PC) y el pensamiento crítico (PCr) han confluido en la agenda educativa como capacidades transversales para aprender, resolver problemas y participar responsablemente en entornos digitales altamente complejos. En términos clásicos, el PCr se define como un razonamiento reflexivo y sensato orientado a decidir qué creer o qué hacer, que combina

habilidades de analizar, inferir, evaluar, favoreciendo a la apertura mental, búsqueda de razones [10]. Vinculado con la IA [11] ofrece un modelo didáctico para el pensamiento crítico con IA configurado por seis dimensiones, aspectos o competencias específicas de aprendizaje con la IA: Interrogación, Comparación, Diálogo crítico, Verificación, Reelaboración personal y Reflexión.

Lejos de ser dominios independientes, PC y PCr se complementan en al menos tres sentidos:

Primero, ambos comparten un núcleo analítico: descomponer/abstraer (PC) exige evaluar la relevancia de la información y las suposiciones (PCr).

Segundo, el diseño de algoritmos requiere justificar pasos y evaluar evidencias sobre su corrección y eficiencia, procesos típicamente críticos.

Tercero, la transferencia de soluciones computacionales a contextos reales demanda juicios éticos y epistémicos (calidad de datos, sesgos, consecuencias), foco tradicional del PCr y hoy imprescindible en educación digital [12].

La investigación empírica reciente respalda esta sinergia. Una revisión de estudios experimentales muestra que intervenciones de PC/codificación mejoran resultados cognitivos de los estudiantes, incluidas habilidades de resolución de problemas y, en algunos diseños, componentes próximos al razonamiento crítico, aunque con ciertas diferencias según el contexto, la problemática y la integración curricular [13].

En síntesis, el PC aporta métodos estructurados para diseñar y probar soluciones; el PCr aporta criterios normativos y epistémicos para juzgar su validez, fiabilidad e impacto. Integrarlos favorece aprendizajes transferibles y responsables: estudiantes capaces de modelar problemas, justificar decisiones y auditar soluciones técnicas a la luz de evidencia y valores.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para cumplir con estos objetivos se propone la metodología de investigación-acción. Esta metodología propone cuatro fases:

- Diagnóstico y reconocimiento de la situación inicial.
- Desarrollo de un plan de acción para mejorar aquello que ya está ocurriendo.
- Actuación para poner el plan en práctica y la observación de sus efectos en el contexto que tiene lugar.
- Reflexión en torno a los efectos como base para una nueva planificación.

Las estrategias educativas destinadas a los estudiantes universitarios, que incluyen recursos tecnológicos y didácticos, se aplicarán al interior de las asignaturas de programación: Algoritmos y Estructuras de Datos I, Algoritmos y Estructuras de Datos II y Programación Orientada a Objetos de la carrera Licenciatura en Sistemas de Información de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la UNNE, que están a cargo de las docentes investigadoras que integran el proyecto.

2. RESULTADOS ESPERADOS

Para cumplir los objetivos del proyecto se realizarán las siguientes actividades.

1. Diseñar, implementar y analizar estrategias educativas orientadas a fortalecer la integración del pensamiento crítico y del pensamiento computacional en los procesos de enseñanza y aprendizaje de la programación.
2. Estudiar herramientas software para la enseñanza de la programación, incluyendo herramientas de IA generativa, estableciendo criterios para su selección, evaluación y uso que contribuyan a la calidad de los procesos de enseñanza-aprendizaje y a la innovación pedagógica.
3. Promover vocaciones en el ámbito de las TIC entre los jóvenes, con especial atención a la inclusión de las mujeres, a fin de incrementar su representación en el sector.

Al integrar el pensamiento crítico, se busca formar estudiantes con capacidad para cuestionar, analizar alternativas y evaluar consecuencias, contribuyendo así a una educación universitaria más integral, pertinente y comprometida con la sociedad del conocimiento.

3. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan 5 docentes investigadoras, tres tesistas de posgrado que desarrollan sus planes de tesis en temas vinculados con el proyecto y un estudiante de grado que lleva adelante su Proyecto Final de Carrera de la Licenciatura en Sistemas de Información de la UNNE.

4. REFERENCIAS

[1] M. Zukerfeld, M. F. Botta, L. Dughera, and G. Yansen, ¿Y las mujeres… dónde están? Fundación Dr. Manuel Sadosky / Program.AR, Ministerio de Ciencia, Tecnología e Innovación Productiva, 2013.

[2] Dapozo, G., Greiner, C., Petris, R., Company, A.M., Espíndola, M.C. (2024). Uso de IA en programación inicial desde la perspectiva de los estudiantes. In JADiCC 2024, disponible en: https://www.unirioeditora.com.ar/producto/reflexiones-practicas-didactica-las-ciencias-la-computacion/

[3] O. Almatrafi, A. Johri, H. Lee, A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023), Computers and Education Open,Volume 6, 2024, https://doi.org/10.1016/j.caeo.2024.100173.

[4] S. Lau, P. J. Gou, “From "Ban It TillWe Understand It" to "Resistance is Futile": How University Programming Instructors Plan to Adapt as More Students Use AI Code Generation and Explanation Tools such as ChatGPT and GitHub Copilot”, in Proceedings of the 2023 ACM Conference on International Computing Education Research, New York, NY, USA, 2023.

[5] F. J. García-Peñalvo, "Generative Artificial Intelligence: New Scenarios in Teaching, Learning, and Communication," VIII Congreso Internacional de Estudios sobre Medios de Comunicación. Universidad Complutense de Madrid, España, 6 de septiembre de 2023. Available from: https://bit.ly/3sGUA3Y.

[6] D. Sun, A. Boudouaia, C. Zhu, “Would ChatGPT-facilitated programming mode impact college students’ programming behaviors, performances, and perceptions? An empirical study”. International Journal of Educational Technology in Higher Education, vol 2, N° 14, 2024.

[7] Zviel-Girshin, R. (2024). The Good and Bad of AI Tools in Novice Programming Education. Education Sciences, 14(10), 1089.
ttps://doi.org/10.3390/educsci14101089

[8] C. Kazimoglu, M. Kiernan, L. Bacon y L. Mackinnon, “A serious game for developing computational thinking and learning introductory computer programming,” Procedia - Social and Behavioral Sciences 47, pp. 1991–1999, 2012.

[9] R. Tariq, B. M. Aponte Babines, J. Ramírez, I. Alvarez-Icaza y F. Naseer, "Computational thinking in STEM education: current state-of-the-art and future research directions," *Frontiers in Computer Science*, vol. 6, p. 1480404, Jan. 2025, doi: 10.3389/fcomp.2024.1480404

[10] R. H. Ennis, "The Nature of Critical Thinking: An Outline of Critical Thinking Dispositions and Abilities," *University of Illinois*, 2011. [Online]. Available: https://education.illinois.edu/docs/default-source/faculty-documents/robert-ennis/thenatureofcriticalthinking_51711_000.pdf?utm_source=chatgpt.com

[11] M. Area-Moreira (2025). Luces y sombras de la IA en la educación superior. Didáctica para el pensamiento crítico. RIULL Repositorio Institucional de la Universidad de La Laguna, España. Disponibvle en: http://riull.ull.es/xmlui/handle/915/40470

[12] UNESCO (2023). Informe GEM 2023: Tecnología en la educación. Disponible en: https://www.unesco.org/gem-report/sites/default/files/medias/fichiers/2023/07/2023reportflyer_SP.pdf

[13] C. Montuori, F. Gambarota, G. Altoè, and B. Arfé, "The cognitive effects of computational thinking: A systematic review and meta-analytic study," Computers & Education, vol. 210, p. 104961, Mar. 2024, doi: 10.1016/j.compedu.2023.104961.

Predicción del riesgo de abandono de estudiantes en las carreras de Ingeniería de la UTN FRC

Analía Guzmán, Mario Alejandro Garcia, María Alejandra Jewsbury, Claudia Castro, Andrea Fabiana Delgado, María Fernanda Giubergia, Silvia Graciela Socolovsky, Gabriel Martinez Ocampo

Universidad Tecnológica Nacional - Facultad Regional Córdoba,
Maestro Marcelo López esq. Cruz Roja Argentina – Córdoba

aguzman@frc.utn.edu.ar; mgarcia@frc.utn.edu.ar; ajewsbury@gmail.com; ingclaudiacastro@gmail.com; andreafdelgado@gmail.com; fernanda.giubergia@gmail.com; silviagsocolovsky@gmail.com; gbye.martinez@gmail.com

RESUMEN

Este trabajo presenta una línea de investigación que aborda la problemática de la deserción universitaria en las carreras de Ingeniería de la Facultad Regional Córdoba (FRC) de la Universidad Tecnológica Nacional (UTN). El estudio se centra en el desarrollo de un modelo de predicción basado en inteligencia artificial y aprendizaje automático (*machine learning*) para identificar a estudiantes en riesgo de abandono. Para ello, se estudian y seleccionan algoritmos destinados a analizar registros históricos, incluyendo calificaciones, asistencia y datos demográficos, con el fin de descubrir patrones y factores de riesgo asociados. El alcance del proyecto contempla la construcción de un modelo de predicción preciso y el desarrollo de una herramienta de visualización que integre un Sistema de Alerta Temprana (SAT). Estos instrumentos permitirán a los gestores académicos implementar intervenciones personalizadas, tales como tutorías y programas de mentoría, optimizando la asignación de recursos institucionales. Mediante la integración de estas tecnologías, se busca fortalecer las trayectorias académicas y mejorar sustancialmente las tasas de retención estudiantil en la educación superior tecnológica.

Palabras clave: *inteligencia artificial, predicción, deserción, abandono.*

CONTEXTO

El presente trabajo forma parte del proyecto de investigación y desarrollo (PID) homologado por la Secretaría de Investigación, Desarrollo y Posgrado de la Universidad Tecnológica Nacional (UTN), reconocido con el código: COTEEC533. En el mismo se reúnen integrantes del grupo GIEIMeT (Grupo de Investigación en Enseñanza de la Ingeniería Mediada por Tecnología) y del grupo GIA (Grupo de Inteligencia Artificial), ambos pertenecientes a la UTN FRC, para abordar la problemática de la deserción universitaria en la FRC, desde el aporte que puede realizar la Inteligencia Artificial en el campo de la educación.

La motivación de realizar esta investigación surge ya que los modelos de aprendizaje automático se utilizan cada vez más para predecir y detectar la deserción en carreras de grado universitarias. Estos modelos aprovechan datos históricos de estudiantes, como calificaciones, asistencia, datos demográficos y espacios de aprendizaje en línea, para identificar patrones que puedan indicar un riesgo de abandono, los cuales

pueden servir para mejorar las tasas de retención y así garantizar que más estudiantes puedan completar sus estudios con éxito.

1. INTRODUCCIÓN

La problemática de la deserción universitaria cobró relevancia política en Latinoamérica a mediados de la década de los 90, impulsada por recomendaciones de organismos internacionales sobre la eficiencia del gasto público, respondiendo frecuentemente a determinantes exógenos y paradigmas de gestión pública global [1]. En Argentina, la Ley de Educación Superior de 1995 estableció el marco para garantizar la igualdad de oportunidades y la permanencia, creando la CONEAU para monitorear datos estadísticos de ingreso y egreso de las instituciones [2]. Dentro de este contexto, la definición estándar describe la deserción como el abandono de los estudios antes de finalizar el programa académico, concepto ampliamente aceptado en la literatura científica por consenso [3, 4, 5]. Desde una perspectiva operativa, se categoriza bajo el concepto de abandono a los estudiantes inscritos que, tras iniciar el cursado, no registran reinscripción, no mantienen la regularidad ni se han graduado al momento del relevamiento [6]. Es fundamental considerar que la decisión de abandonar no es un evento instantáneo, sino un proceso de conflictividad prolongada en el tiempo y el espacio [7]. Este fenómeno es transversal a diversas carreras e instituciones, con tasas de permanencia en primer año que oscilan globalmente entre el 50% y el 60% [8]. En el caso de la Facultad Regional Córdoba, estudios previos del equipo de investigación muestran resultados que coinciden con estas tendencias, observándose una tasa de abandono cercana al 44% en estudiantes de los primeros años de las carreras de ingeniería, asociada principalmente a dificultades académicas y organizativas [9].

Para analizar el fenómeno, se identifican cinco enfoques principales: psicológico, sociológico, económico, organizacional e interaccional [10]. Los factores que condicionan la trayectoria incluyen atributos individuales como aptitudes y habilidades, experiencias educativas previas y el perfil familiar, los cuales influyen directamente en el compromiso del estudiante con la institución y su objetivo de graduarse.

Actualmente, existen diversas herramientas tecnológicas para la predicción del riesgo de deserción [11, 12]. Entre ellas, los modelos de aprendizaje automático, como redes neuronales o árboles de decisión, permiten identificar a estudiantes con mayor probabilidad de abandono basándose en datos históricos. El análisis predictivo utiliza técnicas estadísticas para hallar patrones en series temporales y agrupamientos, mientras que los SAT emplean indicadores como calificaciones bajas o asistencia irregular para generar intervenciones personalizadas. Por su parte, la minería y la ciencia de datos permiten extraer conocimiento "oculto" y tomar decisiones sobre grandes volúmenes de datos [13, 14, 15]. Estas herramientas han sido aplicadas con resultados positivos en diversos trabajos académicos tanto en Argentina como en el resto del mundo [16, 17, 18, 19, 20].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta investigación se sustenta en la sinergia entre los grupos GIEIMeT (Enseñanza de la Ingeniería) y GIA (Inteligencia Artificial) de la UTN FRC. El trabajo trasciende la analítica de datos educativa para posicionarse en las siguientes áreas:

- *IA Aplicada a la Educación: Integración de herramientas de aprendizaje automático para resolver problemas críticos de gestión académica, como la retención estudiantil.*
- *Transferibilidad del Conocimiento: Exploración de cómo los modelos desarrollados pueden adaptarse a otros niveles educativos o sectores sociales.*
- *Modelos Predictivos en Poblaciones de Riesgo: Estudio del comportamiento humano ante el abandono de planes, con*

potencial impacto en ámbitos de salud (tratamientos médicos) o motivación laboral.

- *Fortalecimiento Interdisciplinar: Desarrollo de nuevas metodologías de investigación que permiten a docentes y becarios abordar problemáticas regionales e internacionales desde la ciencia de datos.*

El proyecto permite predecir el abandono de estudiantes de las carreras de Ingeniería en la Facultad Regional Córdoba, esto es importante porque le permitirá a la institución contar con información para tomar medidas proactivas para retener a los estudiantes, mejorar las tasas de retención y promover el éxito académico y personal de los estudiantes. Dichas medidas proactivas podrían:

- *Facilitar la intervención temprana de la institución con acciones que puedan proporcionar a estudiantes en riesgo de abandono, el apoyo necesario para mejorar sus posibilidades de éxito académico.*
- *Optimizar los recursos, ya que con la identificación temprana de estudiantes en riesgo de deserción se podría asignar de manera más eficiente sus recursos y servicios de apoyo, focalizándose en los estudiantes que más los necesitan. Favorecer la vinculación del estudiante con la institución fortaleciendo lazos de pertenencia.*

En resumen, el proyecto permite a la institución de educación superior contar con herramientas concretas para determinar políticas de retención y acompañamiento de población estudiantil, permite anticiparse a las necesidades de atención de estudiantes adquiriendo los recursos necesarios en tiempo y forma, por ejemplo determinar cantidad de estudiantes que requieren recursar una materia o estudiantes que requieren excepciones de cursado.

La presencia de estudios de este tipo permite que otras instituciones educativas, de otros niveles, puedan apropiarse de este conocimiento para conocer sus poblaciones en estado temprano de abandono.

El impacto de este modelo se puede trasladar, si se encuentran particularidades semejantes, al estudio de otras poblaciones que abandonan sus planes. Por ejemplo, enfermos con riesgo de abandonar su tratamiento de salud, o trabajadores en riesgo de abandonar su motivación por el trabajo.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto tiene como objetivo general desarrollar un modelo predictivo, basado en Inteligencia Artificial, que identifique a estudiantes de las carreras de Ingeniería de la Facultad Regional Córdoba en riesgo de abandono de sus estudios.

Para ello se han identificado los siguientes objetivos particulares:

1. *Obtener nuevo conocimiento sobre las variables asociadas a la deserción y creación y validación de características (features) complejas asociadas también a la deserción universitaria;*
2. *Construir un modelo predictivo que puedan predecir con precisión qué estudiantes tienen mayor probabilidad de abandonar sus estudios en el futuro cercano, utilizando variables como calificaciones, asistencia, datos demográficos, comportamientos académicos, entre otros;*
3. *Desarrollar un herramienta de visualización para la toma de decisiones, que incluya un SAT que notifique a los asesores académicos y administradores cuando un estudiante sea identificado como en riesgo de deserción, permitiendo intervenciones oportunas;*
4. *Publicar los resultados de los modelos para diseñar intervenciones personalizadas y recursos adicionales que aborden las necesidades específicas de los estudiantes en riesgo, como tutorías, asesoramiento académico o programas de mentoría.*

Al finalizar el proyecto, se esperan los siguientes resultados:

- *Un modelo de predicción de individuos de una población estudiantil específica y que se encuentran en riesgo de abandono de sus estudios a partir de la valoración de variables comprobables empíricamente.*
- *Una herramienta de visualización para la toma de decisiones, que incluya SAT que permita observar los datos de los estudiantes que muestran signos de riesgo de abandonar sus estudios antes de que ocurra la deserción.*

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por docentes de la UTN que se desempeñan en la enseñanza de distintas ingenierías y también participan en la producción de conocimiento desde otros grupos y proyectos de investigación.

Se reúnen integrantes del grupo GIEIMeT (Grupo de Investigación en Enseñanza de la Ingeniería Mediada por Tecnología) y del grupo GIA (Grupo de Inteligencia Artificial), ambos pertenecientes a la Facultad Regional Córdoba de la Universidad Tecnológica Nacional, para abordar la problemática de la deserción universitaria en la Facultad Regional Córdoba, desde el aporte que puede realizar la Inteligencia Artificial en el campo de la educación.

En el transcurso de la investigación, se propone formar becarios de la Facultad Regional Córdoba, con estudiantes y graduados, lo que permitirá comenzar la formación de investigadores en el área de estudio.

Este nuevo proyecto permitirá fortalecer la interdisciplina otorgando aprendizajes nuevos, como es la Inteligencia artificial, para la incursión en otros temas de investigación, donde los docentes investigadores explorarán las problemáticas y las posibles soluciones a las mismas, permitiendo ampliar sus posibilidades, en cuanto a líneas de investigación de interés, tanto local como regional o internacional, y acceder a nombres y trabajos de posibles colaboradores en sus respectivos campos de estudio.

5. BIBLIOGRAFÍA

[1] Bentancur, Nicolás. "Gobiernos, Banco Mundial y universidades: el legado de una década de políticas universitarias en América Latina", Pensamiento Universitario, nro. 11, pp. 7-15, 2004.

[2] LEY de educación superior. Ley Nº 24.521, Sancionada: Julio 20 de 1995.

[3] Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45(1), 89-125.

[4] Bean, J.P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. Research in Higher Education, 12(2), 155-187.

[5] Cabrera, A.F. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. Journal of Higher Education, 64(2), 123-139.

[6] Maya, M., González, R., Aguirre, G., Juárez, J., Fernández, Y., & Abdala, V. (2019). O abandono nos cursos de Ciências Biológicas da Universidade Nacional de Tucumán. Década 1991-2001. Revista iberoamericana de educación superior, 10(28), 131-156.

[7] Jewsbury, A. e I. Haefeli (2000). "Análisis de la deserción en universidades públicas argentinas", en V Congreso Internacional del CLAD, Santo Domingo, CLAD.

[8] Sanseau, María; Sánchez Cestona, Sanseau; Calio, Sabrina. (2023). "Permanencia de las y los estudiantes en la universidad". Editado por Mirta Caucia. - 1a ed. - Ciudad Autónoma de Buenos Aires: CONEAU

Comisión Nacional de Evaluación y Acreditación Universitaria. ISBN 978-987-3765-75-9.

[9] Giubergia, María Fernanda; Jewsbury, Alejandra; Socolovsky, Silvia Graciela (2025). "Una aproximación al estudio del abandono en la universidad desde la perspectiva del retraso académico", en 13° Congreso Nacional de Ingeniería Informática/Sistemas de información, Universidad Tecnológica Nacional, Facultad Regional Córdoba, Córdoba, Argentina.

[10] Tinto, V. (1993). Leaving College: Rethinking the Causes and Cures of Student Attrition, Chicago, University of Chicago Press.

[11] Márquez Vera, C. (2015). Predicción del fracaso y abandono escolar mediante técnica de minería de datos.

[12] Cruz, E., González, M., & Rangel, J. C. (2022). Técnicas de machine learning aplicadas a la evaluación del rendimiento ya la predicción de la deserción de estudiantes universitarios, una revisión. Prisma Tecnológico, 13(1), 77-87.

[13] D. M. West and J. R. Allen. (2020). How artificial intelligence is transforming the world.

[14] F. Galbusera, G. Casaroli and T. (2019). Bassani Artificial intelligence and machine learning in spine research. JOR SPINE, 2(1), p.e1044.[15] Hastie, T., Tibshirani, R., Friedman, J. H. (2023). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (Second Edition).

[16] Istvan, R. M., Lasagna, V., Bacigalupe, M. D. L. Á., & Rivero, J. (2022). Sistema informático de predicción de deserción estudiantil universitaria basado en un modelo de indicadores de regulación del aprendizaje, en entornos educativos mediados por TIC. In XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza).

[17] Valía, L., Rostagno, J., Moine, J. M., Bigatti, C., Riva, F. M., & Amar, E. (2017, August). Minería de datos aplicada a la educación: modelo de deserción universitaria en la Universidad Tecnológica Nacional, Facultad Regional Rosario. In XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires).

[18] Aulck, L., Velagapudi, N., Blumenstock, J., West, J. (2016). Predicting student dropout in higher education. In: 2016 ICML Workshops #Data4Good Machine Learning, New York, vol. abs/1606.06364, pp. 16–20.

[19] Jadri´c, M., Gara Çca, Ç Z., Cuku Çsi´c, M. (2010). Student dropout analysis with application of data mining methods. Manag. J. Contemp. Manag. Issues 15(1), 31–46.

[20] Kotsiantis, S.B., Pierrakeas, C.J., Pintelas, P.E. (2003). Preventing student dropout in distance learning using machine learning techniques. In: Palade, V., Howlett, R.J., Jain, L. (eds.) KES 2003. LNCS (LNAI), vol. 2774, pp. 267–274. Springer, Heidelberg. https://doi.org/10.1007/978-3-540-45226-3 37.

PUNTO TECNOLÓGICO PARA EL DESARROLLO DE ESTRATEGIAS DE ACCESIBILIDAD COMUNICACIONAL

María Andrea Guisen[1,2] [0000-0001-8704-0897], **Abner Safir Vasquez Yrigoin**[1] [0009-0009-8941-3816], **Iván Capomasi**[1] [0009-0002-8127-3424], **Marcos Banducci**[1] [0009-0004-7567-1375], **Pedro López, Mauro Alejandro Soto**[3] [0000-0003-3581-0153], **Nadia Carolina Ksybala**[4] [0009-0008-6044-497]

[1]Centro de Altos Estudios en Tecnología Informática
Facultad de Ingeniería en Sistemas Informáticos – Universidad Abierta Interamericana (UAI), Argentina
{mariaandrea.guisen; pedro.lopez}@uai.edu.ar
{abnersafir.vasquezyrigoin, ivan.capomasi, marcos.banducci}@alumnos.uai.edu.ar
[2]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
[3]Universidad Nacional de Salta, maurosotoal@gmail.com
[4]Universidad Nacional de la Patagonia Austral, nadia.carolina.ksybala@gmail.com

RESUMEN

Este trabajo presenta EntteR, un punto tecnológico para la accesibilidad comunicacional orientado al desarrollo experimental de estrategias mediadas por tecnología para personas con condiciones neurológicas. El proyecto adopta un enfoque de investigación aplicada y se estructura a partir de un modelo cíclico incremental, en el cual la fase de diseño y desarrollo tecnológico constituye el eje central desde las Ciencias de la Computación. Se conformó un piso tecnológico integrado por recursos de hardware, software, periféricos de interacción y elementos de ergonomía, que habilita procesos iterativos de evaluación, selección y configuración situada de tecnologías según los requerimientos de cada usuario. Entre 2024 y 2025 se desarrollaron diez estrategias de accesibilidad comunicacional, materializadas en compilaciones funcionales de tecnologías que combinan dispositivos móviles, configuraciones de accesibilidad, comunicación aumentativa y tecnologías de interacción alternativa, incluyendo seguimiento ocular. Como resultado del segundo ciclo de implementación, se destaca el desarrollo del periférico alternativo "DeVo" (Dedo Conductivo), diseñado para ampliar las modalidades de interacción con pantallas táctiles capacitivas. Los resultados evidencian el potencial del enfoque para generar soluciones computacionales personalizadas y replicables en contextos de diversidad funcional.

Palabras claves: accesibilidad comunicacional, interacción humano–computador, diseño centrado en el usuario, *assistive technologies* (tecnologías de apoyo), comunicación aumentativa y alternativa.

CONTEXTO

El proyecto tiene su origen en el año 2023, a partir de una convocatoria del Ministerio de Ciencia, Tecnología e Innovación de la Nación Argentina destinada al financiamiento de Proyectos Especiales de Innovación Social (PEIS), orientados a fortalecer la articulación entre el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y organizaciones de la sociedad civil. En este marco se conformó una red interdisciplinaria[1] que permitió el desarrollo del proyecto "Punto tecnológico para la accesibilidad de personas con síndrome de Rett y otras condiciones neurológicas", implementado entre noviembre de 2023 y febrero de 2025. Como resultado de este primer ciclo, el proyecto fue reconocido

[1] El núcleo asociativo lo conformó el CONICET y la Fundación Síndrome de Rett Argentina, con la participación de la cátedra "Patología y terapéutica fonoaudiológica en la parálisis cerebral" de la Esc. de Fonoaudiología (Fc. de Ciencias Médicas, UNR), el Centro de Desarrollo e Investigación en Tecnologías Especiales (CeDITE, UTN–FRRo), el Equipo Neuro (centro de rehabilitación y estimulación temprana), y el programa "Educación y Sociedad. Hacia una mayor inclusión educativa" de la UNL.

con el sello de calidad PDTS (Proyecto de Desarrollo Tecnológico y Social), lo que da cuenta de la solidez del enfoque metodológico y del impacto social alcanzado.

A partir de 2025, y tras la radicación de la dirección del proyecto en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI), se dio inicio a un segundo ciclo de implementación, financiado y acreditado por la UAI. Esta nueva etapa se orienta a profundizar el desarrollo tecnológico, con énfasis en soluciones computacionales, dispositivos físicos y prototipos funcionales, así como en la integración de recursos humanos en formación del campo de las Ciencias de la Computación. En este contexto, el proyecto consolida su carácter como punto tecnológico para la accesibilidad, articulando investigación aplicada, desarrollo experimental y transferencia al medio.

1. INTRODUCCIÓN

El acceso a la comunicación constituye un derecho humano fundamental y una condición habilitante para la participación social, la autodeterminación y el ejercicio de otros derechos [1]. No obstante, en gran parte de los contextos latinoamericanos, su ejercicio se encuentra condicionado por las limitadas posibilidades de acceso a tecnologías de apoyo y por la concentración de recursos técnicos y profesionales.

En este contexto, las personas con condiciones neurológicas conforman un grupo particularmente afectado, en tanto la afectación del sistema nervioso central se traduce en dificultades motoras y del habla que restringen el desarrollo funcional del lenguaje expresivo y, en consecuencia, el acceso a la comunicación [2].

En respuesta a las demandas vinculadas con la accesibilidad comunicacional de este grupo poblacional, el proyecto tuvo como objetivo general constituir un dispositivo científico-tecnológico para el desarrollo de estrategias de accesibilidad comunicacional, que abarcan el diseño y desarrollo de soluciones informáticas destinadas a optimizar las condiciones de acceso a la comunicación.

El proyecto adopta un enfoque de investigación aplicada y desarrollo experimental, integrando principios de la Investigación e Innovación Responsable [3] y el Enfoque Sociotécnico [4].

Desde el punto de vista metodológico, se apoya en el **"**Modelo cíclico incremental para el desarrollo de estrategias de accesibilidad mediadas por tecnología", validado en el primer ciclo de implementación. Dicho modelo estructura el proceso de desarrollo en cuatro fases interrelacionadas, de las cuales la segunda, orientada al diseño y desarrollo tecnológico, constituye el eje central desde el campo de las Ciencias de la Computación.

En esta fase, el trabajo se realiza desde una lógica ingenieril, que articula análisis de requerimientos, diseño, desarrollo, implementación y testeo mediante un proceso cíclico, iterativo y experimental. Las decisiones técnicas se sustentan en principios de Interacción Humano–Computador (HCI) [5] y Diseño Centrado en el Usuario (UCD) [6], permitiendo modelar formas de interacción alternativas y configurar sistemas de accesibilidad comunicacional *ad hoc,* ajustados a las necesidades de cada usuario.

Las estrategias de accesibilidad comunicacional desarrolladas se materializan en compilaciones funcionales de tecnologías, que integran dispositivos de apoyo a la ergonomía, dispositivos móviles —tablets, teléfonos móviles y notebooks— y periféricos de interacción tradicionales, adaptados o alternativos, destinados a ampliar o sustituir modalidades de entrada convencionales. En este conjunto adquieren especial relevancia las *eye tracking technologies* (tecnologías de seguimiento ocular), que combinan hardware y software para habilitar el acceso a interfaces digitales mediante la mirada [7].

Las compilaciones incluyen la customización de funcionalidades de accesibilidad de los sistemas operativos**,** la selección y configuración de aplicaciones de Comunicación Aumentativa y Alternativa y, de manera complementaria, el uso de *serious games* orientados al desarrollo de competencias digitales.

Las tecnologías empleadas pueden ser de libre acceso, adquiridas y configuradas, o desarrolladas específicamente cuando las soluciones disponibles no responden a los requerimientos identificados, priorizando enfoques *low cost* y tecnologías *open source* siempre que sea posible.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Accesibilidad**:** concebida como propiedad emergente de sistemas sociotécnicos y decisiones de diseño.

Interacción Humano–Computador (HCI)**:** análisis y diseño de modalidades de interacción e interfaces en contextos situados.

Eye tracking technologies: desarrollo y aplicación de modalidades no convencionales, con énfasis en seguimiento ocular.

Assistive technologies: diseño de dispositivos y sistemas para la mediación de la interacción con entornos digitales y físicos.

Fabricación digital e impresión 3D**:** prototipado iterativo de dispositivos personalizados mediante impresión 3D.

Comunicación Aumentativa y Alternativa (CAA)**:** integración de herramientas para la comunicación asistida.

Vinculación científico-tecnológica y Ciencia Tecnología y Sociedad (CTS)**:** análisis de procesos de co-construcción de conocimiento situado.

3. RESULTADOS OBTENIDOS / ESPERADOS

En relación con el propósito de constituir un punto tecnológico para la accesibilidad comunicacional, durante el período 2024–2025 —que comprende el primer ciclo de implementación y el inicio del segundo— se logró consolidar y poner en funcionamiento "EntteR: Estrategias de Accesibilidad" como un dispositivo científico-tecnológico operativo, con capacidad de intervención situada y potencial de replicabilidad.

En este marco, se generaron las condiciones técnicas, éticas y metodológicas necesarias para el desarrollo de estrategias de accesibilidad comunicacional mediadas por tecnología. Esto incluyó la adquisición de un piso tecnológico integrado por recursos de hardware, software, periféricos de interacción y elementos de ergonomía, que opera como base para la evaluación, selección y configuración situada de tecnologías, así como la validación de instrumentos de recolección de información y la elaboración de documentación técnica orientada al registro sistemático de los procesos de desarrollo.

A partir de estas condiciones, entre 2024 y 2025 se desarrollaron diez estrategias de accesibilidad comunicacional destinadas a niños, algunos de ellos con Síndrome de Rett, jóvenes y una persona adulta con condiciones neurológicas. Las estrategias fueron diseñadas de manera situada, atendiendo a objetivos vinculados al acceso a la comunicación mediada por tecnología y a la adecuación de las modalidades de interacción a las posibilidades de cada usuario. Para cada caso se elaboró una compilación funcional de tecnologías, que constituyó el sistema de accesibilidad comunicacional del destinatario (Figura 1).

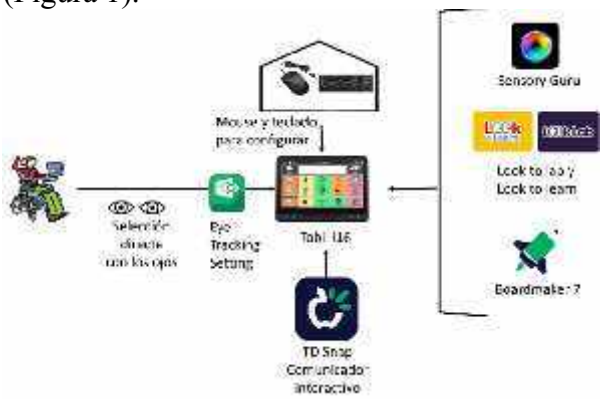


Figura 1. Compilación funcional de tecnologías de una destinataria del proyecto usuaria de Comunicación Aumentativa y Alternativa, y *eye tracking technologies*.

Estas acciones se complementaron con cinco instancias de capacitación comunitaria dirigidas a: una el equipo completo de EntteR, dos a estudiantes universitarios, una a fonoaudiólogas y médicos orientados al trabajo en discapacidad, y una a terapeutas y familiares (Figura 2).



Figura 2. Taller para la adaptación de juguetes para niños de la primera infancia con discapacidad motriz en la Fc. de Humanidades y Ciencias de la UNL, dirigida a terapeutas y familiares.

En el plano de la sistematización y generación de conocimiento, por cada estrategia desarrollada se elaboró un Informe Técnico que documenta el proceso de intervención, las tecnologías empleadas y los resultados alcanzados. Asimismo, se consolidó el "Modelo cíclico incremental para el desarrollo de estrategias de accesibilidad mediadas por tecnología", acompañado por una matriz de variables e indicadores para el análisis sistemático de los resultados.

Asimismo, los avances y resultados obtenidos fueron presentados en tres congresos científicos, con publicación de trabajos en *proceedings* [8] [9], y un workshop.

Se realizaron también dos acciones de divulgación de la ciencia y la técnica (Figura 3), y se registraron cinco participaciones en medios de comunicación masivos, entre radio y televisión, además de una nota periodística [10], contribuyendo a la visibilización pública de las problemáticas abordadas y de las estrategias desarrolladas.



Figura 3. Participación en la Semana de la Ciencia, la tecnología y la Innovación con lugar en la Fc. de Cs. Económicas de la UNER dirigida a estudiantes de escuela media.

Como proyección para el año 2026, se prevé avanzar en el desarrollo de un motor de síntesis de voz infantil, de tono masculino y con acento argentino, compatible con entornos abiertos de Comunicación Aumentativa y Alternativa. Asimismo, se contempla profundizar la articulación con el sector privado con el objetivo de subsidiar estrategias de accesibilidad comunicacional destinadas a personas que las requieran y no cuenten con los recursos necesarios para afrontarlas.

Desarrollo del dispositivo DeVo

El desarrollo del periférico alternativo "DeVo" (Dedo Conductivo) se inscribe en el segundo ciclo de implementación del proyecto y emerge del proceso experimental habilitado por el piso tecnológico conformado para el desarrollo de estrategias de accesibilidad comunicacional. El dispositivo surge a partir de un caso de uso concreto, representativo de un grupo poblacional de personas que presentan dificultades en la flexión de los dedos meñique, anular, medio y pulgar, lo que interfiere en el uso preciso del dedo índice, principal modalidad de señalamiento en la interacción con pantallas táctiles. En este contexto, DeVo fue concebido como una extensión funcional del dedo índice, que posibilita la interacción manteniendo la mano abierta, sin requerir movimientos de flexión fina.

Tras el análisis de requerimientos ergonómicos y funcionales, se optó por una estructura compuesta por tres segmentos articulados mecánicamente, diseñada para integrar un sistema conductor interno y garantizar una interacción eficaz con pantallas táctiles capacitivas. El soporte físico consta de tres falanges independientes impresas en 3D utilizando filamento PETG, articuladas mediante un núcleo de alambre galvanizado y consolidadas con adhesivo de poliuretano, lo que permite una configuración postural ajustable a las necesidades del usuario.

En el interior de la estructura impresa se integra un material textil conductor, cuya selección respondió a las condiciones de uso previstas, caracterizadas por el contacto prolongado con la piel y la exposición a

transpiración y salivación, particularmente relevantes en usuarios con condiciones neurológicas. Ensayos preliminares realizados con distintos materiales conductores, entre ellos cobre, evidenciaron procesos de oxidación y pérdida de estabilidad eléctrica bajo estas condiciones. En función de estos resultados, se optó por el uso de hilo conductor de acero inoxidable AISI 316L, cosido sobre una cinta textil flexible mediante un patrón denso, lo que provee continuidad eléctrica, flexibilidad mecánica y resistencia a la corrosión. Este conjunto textil constituye la superficie de contacto con el dedo del usuario y se fija a la estructura impresa mediante un adhesivo compatible, preservando sus propiedades eléctricas y estructurales.

La falange distal presenta una terminación geométrica en forma de "hongo", diseñada para cumplir una doble función: actuar como trampa mecánica que asegura la fijación del material conductor y facilitar la correcta distribución de la carga capacitiva en el extremo del dispositivo. Sobre esta base se aplica una punta conductiva elaborada a partir de una matriz de caucho de silicona con base platino dopada con grafito en polvo de alta pureza, cuya formulación permite emular la capacitancia de la yema del dedo humano.

En el marco del prototipado, fue necesario diseñar y fabricar un molde específico para la conformación de la punta conductiva. En esta etapa se desarrolló un molde unitario, adecuado para la validación funcional del dispositivo, previendo en instancias posteriores la fabricación de un molde múltiple que permita la producción simultánea de un mayor número de piezas (Figura 4).



Figura 4. A la izquierda se presenta el molde utilizado para la fabricación de la punta conductiva. A la derecha el modelo de DeVo con todas las partes ensambladas.

Como resultado, el dispositivo posibilita un accionamiento confiable de pantallas táctiles capacitivas —en tablets y smartphones— así como de teclados físicos, ofreciendo una respuesta táctil precisa y estable, acorde a los requerimientos de interacción identificados.

4. FORMACIÓN DE RECURSOS HUMANOS

En 2024, se desarrollaron dos pasantías de grado (Comunicación Social, UNR) orientadas a la gestión de la comunicación del proyecto, así como una práctica profesional supervisada (Fonoaudiología, UNR) vinculada a la elaboración de informes técnicos para la sistematización de estrategias de accesibilidad comunicacional.

En 2025, se incorporaron tres estudiantes de Ingeniería en Sistemas Informáticos (UAI) al desarrollo tecnológico del dispositivo DeVo.

5. BIBLIOGRAFÍA

[1] Organización de las Naciones Unidas. (2006). *Convención sobre los derechos de las personas con discapacidad*. Naciones Unidas.

[2] Palacios-Lenis, J., Bonilla-Villegas, D., Rodríguez-Campos, D., Ortega-Ledesma, M. F., & Argüello-Vélez, P. (2022). Implementación del Sistema Pictográfico de Comunicación en un deportista con parálisis cerebral espástica: Reporte de un caso. *Revista de Investigación en Logopedia, 12*(2), e79173.

[3] Mezzadra, S., & Tan, W. (2024). Social Innovation in Inclusive Design: Engaging Stakeholders in Participatory Methodologies. *Journal of Inclusive Practices, 12*(1), 45–67.

[4] Thomas, H. (2022). Para un desarrollo inclusivo y sustentable hace falta tener alguna idea contrahegemónica. *Trilogía Ciencia, Tecnología, Sociedad, 14*(28), e2536.

[5] Cook, A. M., & Polgar, J. M. (2020). *Assistive Technologies: Principles and Practice* (5th ed.). Elsevier.

[6] Fakhimi, M. M., Hughes, A., & Gustavson, A. M. (2025). Designing for Dignity: Empowering Life Through Synergistic and Integrated Design Solutions for Aging in Place. *Preprints.org*.

[7] Bhattacharya, U., et al. (2023). Why Don't They Talk to Our Daughter? Eye Tracking AAC and Medical Communication in Rett Syndrome. *Language and Health* 1 (1): 32–43

[8] Guisen, M. A., Soto, M. A., Ksybala, N. C., & Brasca, R. (2024). Proyecto EntteR: punto tecnológico para el desarrollo de estrategias de accesibilidad de personas con síndrome de Rett y otras condiciones neurológicas. *JAIIO, Jornadas Argentinas de Informática*, *10*(4), 60-71.
[9] Guisen, M. A. (en prensa). The EntteR Project: Women-Led Innovation in Assistive Technologies for Gender-Focused Communicational Accessibility Strategies. *Actas de la Latin American Women in Computing Conference (LAWCC), CLEI 2025*.
[10] Consejo Nacional de Investigaciones Científicas y Técnicas. (2025, 23 de enero). *Especialistas del CONICET lideran un proyecto para mejorar la accesibilidad comunicacional de personas con diferentes afecciones neurológicas*.

Sistema de Tutorías Híbridas para el acompañamiento de estudiantes en el inicio de la Carrera Ingeniería en Sistemas de Información

Roberto M. Muñoz, Analía Guzmán, Claudia Castro,
Lorena Genoud, Martín G. Casatti

GIEIMeT - Grupo de Investigación en Enseñanza de la Ingeniería Mediada por Tecnologías
Facultad Regional Córdoba – Universidad Tecnológica Nacional

{rmunoz; aguzman; cscastro; lgenoud; mcasatti }@frc.utn.edu.ar

RESUMEN

El proyecto investiga el diseño e implementación de un sistema de tutorías híbridas y asistido por inteligencia artificial, para fortalecer el acompañamiento académico en la educación superior. La línea de investigación aborda los desafíos persistentes de retención y egreso, específicamente en los tramos iniciales de la formación universitaria, mediante dispositivos que integran tecnología y pedagogía, dado que las estrategias tradicionales de tutoría suelen presentar limitaciones en escalabilidad y personalización. Se propone un modelo híbrido, que combina la labor de docentes tutores con asistentes automatizados, capaces de brindar orientación inmediata y personalizada, permite además monitorear el desempeño académico, anticipar espacios y situaciones de riesgo y ofrecer recomendaciones adaptadas al perfil de cada estudiante. La incorporación de tecnológica se concibe como un recurso complementario, bajo supervisión docente, que permite preservar la dimensión humana del vínculo pedagógico. El estudio contempla una fase de validación en entornos reales de implementación, con el objetivo de analizar el impacto del sistema en las trayectorias académicas y generar lineamientos para su escalabilidad institucional. La investigación fortalece la formación de nuevos investigadores y aporta evidencia empírica y criterios de diseño pedagógico en el uso ético y significativo de las tecnologías emergentes en el acompañamiento académico universitario.

Palabras clave: *Tutorías Híbridas, Inteligencia Artificial, Acompañamiento Académico, Chatbots Educativos.*

CONTEXTO

El presente trabajo forma parte del Proyecto de Investigación y Desarrollo (PID) homologado por la Secretaría de Investigación, Desarrollo y Posgrado de la Universidad Tecnológica Nacional (UTN), reconocido con el código: COTEEC1190. El mismo se desarrolla en el Grupo de Investigación en Enseñanza de la Ingeniería Mediada por Tecnologías (GIEIMeT), de la Facultad Regional Córdoba de la UTN (FRC-UTN).

El proyecto está enfocado en el diseño, desarrollo, implementación y evaluación de un sistema de tutorías híbridas asistido por inteligencia artificial que, en articulación con tutorías docentes, que ya están implementadas, pueda brindar acompañamiento académico y personalizado a estudiantes de primer año de la carrera de Ingeniería en Sistemas de Información de la FRC–UTN. El sistema combina un componente automatizado, basado en un chatbot educativo inteligente, con el acompañamiento de docentes tutores,

articulando ambos dispositivos en una propuesta pedagógica integrada.

El alcance contempla la implementación piloto del sistema en una asignatura del primer año, para validar su funcionamiento en un entorno real y obtener evidencia empírica sobre su impacto. También incluye la integración del sistema con plataformas institucionales existentes, como Moodle, y el diseño de instrumentos de seguimiento y evaluación.

Finalmente, el proyecto prevé la elaboración de lineamientos para su escalabilidad, considerando su posible aplicación en otras asignaturas del mismo nivel u otras carreras de la FRC-UTN.

.

1. INTRODUCCIÓN

La retención y el egreso en la educación superior constituyen desafíos persistentes que exigen respuestas institucionales, ante la influencia de diversos factores académicos, personales y contextuales en la continuidad de los estudios. En el contexto argentino, según datos de la CONEAU de 2022, el 68% de los estudiantes universitarios no logra completar su formación, mientras que apenas el 29,5% egresa en el tiempo previsto por el plan de estudios, lo que evidencia la necesidad de implementar dispositivos de acompañamiento más efectivos [1]. Esta situación cobra especial relevancia en las carreras vinculadas a la ingeniería y la tecnología debido a su papel estratégico en la formación de capital humano para el desarrollo económico del país. Para abordar las dificultades académicas, actitudinales y motivacionales de los estudiantes se presentan las actividades de apoyo tutorial, como un mecanismo fundamental para brindar contención y evitar el alejamiento de las carreras. Estas tutorías se definen como apoyos que abarcan aspectos académicos, metodológicos y administrativos con el fin de mejorar el rendimiento y facilitar la adaptación a la vida universitaria [2]. En este sentido, un programa de tutorías enfocado en el primer año resulta esencial, ya que es la etapa donde se registra la mayor deserción, debido a los retos que implica la transición institucional y la organización del tiempo.

La FRC-UTN cuenta con una trayectoria importante de trabajo en la implementación de tutorías para los primeros años de sus carreras. En 2014, el Departamento de Ingeniería en Sistemas de Información lanzó el Programa de Tutorías de Inicio de Carrera dentro del marco de PROMINF, estableciendo un espacio de acompañamiento voluntario liderado por docentes y estudiantes avanzados [3]. Tras mantenerse vigente hasta 2017, el programa recibió continuidad entre 2018 y 2021 a través de la Secretaría de Planeamiento Académico. En 2022, se formalizó esta iniciativa mediante la Resolución N° 1507/2022, asignando su dependencia a la Secretaría Académica de la Facultad Regional. Actualmente, el programa busca garantizar la igualdad de condiciones para el acceso y permanencia en ingeniería, reducir el desgranamiento y mejorar el desempeño académico mediante el apoyo en el cursado y la articulación directa con las cátedras. Este esfuerzo refleja un interés sostenido, por más de una década, en el desarrollo de políticas institucionales de acompañamiento.

En este contexto, la inteligencia artificial (IA) es concebida por la UNESCO como máquinas capaces de imitar funcionalidades de la inteligencia humana, tales como el razonamiento y el aprendizaje [4]. Su integración en los procesos de enseñanza abre perspectivas transformadoras, aunque exige un abordaje responsable frente a sus desafíos asociados [5]. En el ámbito universitario, la IA potencia el aprendizaje personalizado mediante sistemas de tutoría inteligente y chatbots educativos, herramientas que facilitan la detección temprana de necesidades de apoyo y promueven la inclusión [6][7]. Ante la creciente adopción de asistentes como ChatGPT, resulta imperativo establecer criterios éticos y pedagógicos que orienten su uso [8]. Las tutorías asistidas por IA se presentan así como una estrategia innovadora, para enriquecer las trayectorias estudiantiles, ofreciendo un acompañamiento más accesible

y sostenible [9]. No obstante, el rol del docente permanece como un pilar irremplazable, para garantizar la supervisión ética y el desarrollo del pensamisento crítico, asegurando que la IA actúe como un recurso complementario [10].

Finalmente, en consonancia con las tendencias internacionales, Argentina ha manifestado un creciente interés en la incorporación de sistemas de tutoría asistidos por inteligencia artificial en el ámbito universitario. Diversas instituciones han desarrollado experiencias relevantes con enfoques diferenciados, entre las que se destacan el chatbot *SintIA* de la Universidad de Buenos Aires, orientado a la personalización de la enseñanza en asignaturas específicas [11], y la implementación de chatbots en la Universidad Nacional de Córdoba para la automatización y gestión de consultas académicas y administrativas [12].

En el ámbito de las universidades privadas, la Universidad Blas Pascal lanzó *Chat UBP*, un tutor virtual disponible las 24 horas y validado por docentes, con foco en el acompañamiento académico personalizado [13]. Por su parte, dentro de la Universidad Tecnológica Nacional, la Facultad Regional La Plata incorporó el *Tutor Virtual UTN* como dispositivo de apoyo a estudiantes ingresantes [14], mientras que la Facultad Regional Buenos Aires implementó el asistente *Cel.IA Recomendador*, orientado al soporte técnico del campus virtual [15].

En esta misma línea de innovación, la Universidad Tecnológica Nacional – Facultad Regional Córdoba prevé poner en marcha, a partir del ciclo lectivo 2026, el proyecto UTN-AVE (Asistente Virtual Educativo) en el ámbito de la carrera de Ingeniería en Sistemas de Información. Esta iniciativa integra inteligencia artificial generativa con el objetivo de ofrecer acceso oportuno a información confiable y acompañar procesos de aprendizaje más activos y autónomos.

En conjunto, estas iniciativas evidencian un compromiso creciente de las instituciones de educación superior por integrar tecnologías emergentes en sus dispositivos de acompañamiento académico, destacando la necesidad de avanzar hacia modelos pedagógicamente integrados que articulen la automatización con la mediación docente y el fortalecimiento del vínculo educativo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el presente trabajo se explorarán las siguientes líneas de investigación, las que pueden ser adaptadas o adecuadas de acuerdo a los hallazgos producidos del propio trabajo de investigación:

- *Modelos de intervención pedagógica híbrida: Investigación sobre la articulación entre el acompañamiento humano (docentes y tutores pares) y el soporte automatizado de inteligencia artificial para personalizar la retroalimentación académica.*

- *Análisis contextual y estado del arte en tutorías inteligentes: Estudio de la literatura científica sobre sistemas de tutoría universitaria mediadas por IA y análisis histórico de las políticas de acompañamiento institucional en la UTN-FRC.*

- *Desarrollo y entrenamiento de asistentes educativos automatizados: Investigación sobre la creación de agentes conversacionales capaces de facilitar conceptos complejos y automatizar el seguimiento personalizado del progreso estudiantil.*

- *Interoperabilidad tecnológica en entornos de aprendizaje: Estudio de la integración técnica mediante APIs y plugins, para conectar sistemas inteligentes con plataformas institucionales de gestión educativa como Moodle.*

- *Evaluación del impacto y eficacia pedagógica: Análisis cualitativo y*

cuantitativo del uso de asistentes inteligentes en entornos reales de cursado, para validar su influencia en la retención y el rendimiento académico.

- *Estrategias de escalabilidad y transferencia institucional: Investigación sobre la creación de lineamientos y guías metodológicas que permitan replicar el modelo de tutoría híbrida en otros niveles curriculares y especialidades de ingeniería.*

El desarrollo de este sistema de tutorías híbridas generará aportes significativos en múltiples dimensiones:

- *Fortalecimiento de la Retención: Proporcionar una herramienta proactiva para la detección temprana de dificultades, facilitando intervenciones que mejoren el éxito y la permanencia estudiantil.*
- *Eficiencia y Sostenibilidad: Optimizar los recursos institucionales al complementar el soporte humano con tecnología, ofreciendo una alternativa flexible ante las limitaciones presupuestarias de las tutorías presenciales.*
- *Formación de Recursos Humanos: Permitir la especialización de docentes en tecnologías emergentes y la formación de becarios alumnos y graduados en el área de investigación.*
- *Vínculo Institucional: Reforzar el sentido de pertenencia de los estudiantes al ofrecer una experiencia educativa más dinámica, cercana y con soporte constante.*
- *Ética y Supervisión Docente: Garantizar que la integración de la IA se realice bajo una guía pedagógica responsable, preservando el rol insustituible del docente en la formación crítica.*

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto tiene como objetivo general diseñar e implementar un sistema de tutorías híbridas asistido por inteligencia artificial que, en articulación con tutorías docentes, brinde acompañamiento académico personalizado a estudiantes en el inicio de la carrera de Ingeniería en Sistemas de Información de la UTN–FRC.

En este marco, se espera alcanzar los siguientes resultados:

- *La definición y validación de un modelo de intervención tutorial híbrido que combine tutorías humanas e inteligentes, orientado a la personalización de recomendaciones, explicaciones y retroalimentación académica.*
- *El desarrollo de un prototipo funcional de chatbot educativo, con capacidades de seguimiento académico personalizado.*
- *La integración efectiva del sistema con plataformas institucionales existentes, particularmente Moodle.*
- *La obtención de evidencia empírica, a partir de una prueba piloto, sobre la eficacia del sistema desde la perspectiva de estudiantes y docentes.*
- *La elaboración de recomendaciones para la escalabilidad del modelo en otros niveles de la carrera y en otras carreras de la Facultad.*

4. FORMACIÓN DE RECURSOS HUMANOS

En el equipo de investigación participan docentes de la carrera de Ingeniería en Sistemas de Información, con la dirección del Secretario Académico de la Facultad Regional Córdoba.

En el transcurso de la investigación, se propone formar becarios de la Facultad Regional Córdoba, con alumnos y graduados, lo que permite comenzar la formación de investigadores en el área de estudio.

Los docentes podrán abordar la problemática de las tutorías híbridas en la Facultad Regional Córdoba, desde el aporte que puede realizar la Inteligencia Artificial en el campo de la educación, siendo central el rol docente, no sólo como guía y acompañante pedagógico del estudiante, sino también como actor clave en el diseño, validación y evaluación de las estrategias tutoriales mediadas por tecnología. Sus experiencias y conocimientos, del contexto académico, permite integrar los recursos de IA de manera crítica, asegurando que el sistema responda a las necesidades reales de enseñanza y aprendizaje, y que se preserve la dimensión humana del vínculo educativo.

5. BIBLIOGRAFÍA

[1] Sanseau, M., Sánchez Cestona, J., & Calio, S. (2023). Permanencia de las y los estudiantes en la Universidad. CONEAU - Comisión Nacional de Evaluación y Acreditación Universitaria: CABA.

[2] Haefeli, I., Muñoz, R., & Odetti, M. A. (2015). Programa de Tutorías Para Primer Año: Una experiencia con la intervención de monitores alumnos. Facultad Regional Córdoba - Universidad Tecnológica Nacional.

[3] Marciszack, M., Muñoz, R., & Odetti, M. A. (2018) Proyecto Tutoría: Programa de retención-recuperación de estudiantes avanzados de la carrera de Ingeniería en Sistemas de Información-Plan 2008. VI CONAIISI, Universidad CAECE – Mar del Plata, Buenos Aires, Argentina. 2018

[4] Hu, X., Neupane, B., Echaiz, L. F., Sibal, P., & Rivera Lam, M. (2019). Steering AI and advanced ICTs for knowledge societies: A Rights, Openness, Access, and Multi-stakeholder Perspective. UNESCO Publishing.

[5] Unesco. (2023). Oportunidades y desafíos de la era de la inteligencia artificial para la educación superior: una introducción para los actores de la educación superior.

[6] González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. Applied sciences, 11(12), 5467).

[7] Cruz, E. T. (2023). Impacto de la Inteligencia Artificial en la educación universitaria (Doctoral dissertation, Doctoral dissertation, Universidad Nacional del Altiplano). https://downloads.editoracientífica.com.br/articles/230513147.pdf.

[8] Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. Ieee Access, 8, 75264-75278.

[9] Martínez, R. L. I., Morales, J. L. C., & González, M. N. P. (2023). Inteligencia artificial en la educación. Revista Digital de Tecnologías Informáticas y Sistemas.

[10] Guerschberg, L., & Gutierrez, Y. E. (2024). Tutoría con Inteligencia Artificial Generativa en la Educación Superior: Oportunidades y Desafíos en el Aprendizaje. Ciencia Latina Revista Científica Multidisciplinar, 8(5), 9960-9975.

[11] https://www.educrear.com.ar/es/innovacion-tecnologia/ubachatbox-tecnologia-innovacion-universidad.

[12] https://fcefyn.unc.edu.ar/facultad/secretarias/secretaria-deinformatica/inauguramos-un-chatbot-en-el-aula-virtual/.

[13] https://www.ubp.edu.ar/novedades/chat-ubp-la-universidad-blaspascal-lanza-el-primer-tutor-academico-virtual-para-acompanar-a-losestudiantes-las-24-horas/.

[14] https://portal.amelica.org/ameli/journal/266/2663842004/html/?utm_source=chatgpt.com.

[15] https://blog.centrodeelearning.com/cel-ia-elegir-capacitacionesinteligencia-artificial/.

Innovación en Sistemas de Software

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

Arquitecturas de Software Modulares y Energéticamente Eficientes para IoT

Moreno, Juan Pablo[1]; Baigorria, Lorena[2]; Montejano, Germán[2]; Heredia, Diego[1]; Baquinsay, Manuel[1]; Ferraro, Matìas[1]; Figueroa, Matìas[1]

[1]Grupo IoT/Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca
[2]Facultad de Ciencias Físico Matemáticas y Naturales/Universidad Nacional de San Luis
Maximio Victoria Nº 55 - C.P: 4700 - San Fernando del Valle de Catamarca.
Tel: +54 (03834) 435112
jpmoreno@tecno.unca.edu.ar, lorenabaigorria@gmail.com

RESUMEN

La línea de Investigación y Desarrollo presentada se orienta al diseño de arquitecturas innovadoras de software embebido para nodos IoT (Internet of Things - Internet de las Cosas) de bajo consumo aplicados al monitoreo ganadero en entornos rurales. El eje central del trabajo es la adaptación y optimización de la administración de energía mediante un enfoque modular de firmware, desacoplado del hardware específico, que permita extender significativamente la vida útil de los dispositivos, mejorar su mantenibilidad y facilitar su portabilidad.

La propuesta surge como evolución de desarrollos previos validados en campo por la Universidad Nacional de Catamarca (UNCA) en colaboración con el Instituto Nacional de Tecnología Agropecuaria (INTA), donde se evidenció que las arquitecturas monolíticas limitan la aplicación de técnicas avanzadas de gestión dinámica de energía. En este contexto, la línea de I+D propone integrar principios de Ingeniería de Software, como modularidad, bajo acoplamiento y reutilización, con técnicas de Dynamic Power Management (DPM), permitiendo el apagado selectivo e inteligente de periféricos.

Se espera que los resultados contribuyan tanto al avance académico en innovación de sistemas de software embebido como al desarrollo de soluciones IoT más sostenibles, escalables y replicables en ámbitos rurales y productivos.

Palabras clave: Internet de las Cosas, Software Embebido, Gestión Energética, Arquitectura Modular, Sistemas IoT de Bajo Consumo.

CONTEXTO

La presente línea de Investigación se encuentra inserta en el proyecto interdisciplinario "Adaptación de la administración de energía para una mayor vida útil de los nodos IoT de seguimiento de ganado", presentado en el marco del Plan de Investigación, Desarrollo e Innovación (PIDI 2025).

El proyecto es coordinado por la Facultad de Tecnología y Ciencias Aplicadas (FTyCA) de la UNCA, a través del Grupo de Investigación en IoT (GIIOT), y se desarrolla en colaboración con el INTA. La acreditación y financiamiento corresponden a la Universidad Nacional de Catamarca, con aportes complementarios en infraestructura, conocimiento aplicado y validación experimental por parte del INTA.

1. INTRODUCCIÓN

La innovación en sistemas de software para dispositivos IoT constituye un área estratégica de investigación debido a su creciente impacto en sectores productivos, industriales y ambientales. En particular, las aplicaciones IoT en entornos rurales presentan desafíos

específicos asociados a la autonomía energética, la confiabilidad operativa y la dificultad de mantenimiento, lo que exige soluciones de software embebido altamente eficientes y adaptables.

Diversos estudios han demostrado que la gestión energética es uno de los factores críticos en la vida útil de los nodos, especialmente en dispositivos alimentados por baterías y desplegados en escenarios de acceso limitado (Anastasi et al., 2009; Dargie, 2012). Técnicas como Duty Cycling, Dynamic Voltage and Frequency Scaling (DVFS) y Dynamic Power Management (DPM) permiten reducir significativamente el consumo energético; sin embargo, su adopción efectiva depende en gran medida de la arquitectura del firmware.

En muchos desarrollos reales, el firmware presenta un diseño monolítico y fuertemente acoplado al hardware, lo que dificulta la incorporación de nuevas políticas energéticas, la reutilización de componentes y la evolución del sistema. Si bien existen sistemas operativos embebidos como FreeRTOS, TinyOS, etc., su sobrecarga en memoria y complejidad limita su aplicación en plataformas de ultra bajo consumo.

En este contexto, la línea de investigación propuesta aborda la problemática desde la perspectiva de la Ingeniería de Software, planteando el diseño de una arquitectura modular y ligera para nodos IoT sin sistema operativo, basada en interfaces abstractas para los subsistemas de sensores, comunicación, almacenamiento y energía. Este enfoque permite integrar decisiones energéticas inteligentes, mejorar la mantenibilidad del software y facilitar la portabilidad a nuevas plataformas.

La investigación se apoya en antecedentes teóricos y experimentales en redes de sensores inalámbricos, consumo energético en LPWAN y diseño de arquitecturas reutilizables (Pottie & Kaiser, 2000; Bouguera et al., 2018; Gamma et al., 1995), y se valida en un caso de aplicación real como lo es el monitoreo ganadero mediante nodos IoT de largo alcance.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La línea de investigación se estructura sobre los siguientes ejes:

- Diseño de arquitecturas modulares de firmware para sistemas IoT de bajo consumo.
- Gestión dinámica de energía (DPM) aplicada a nodos sensores sin sistema operativo.
- Optimización del consumo energético en dispositivos IoT con tecnologías LPWAN.
- Aplicación de principios de Ingeniería de Software en sistemas embebidos.
- Evaluación experimental y comparativa de arquitecturas monolíticas versus modulares.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los resultados obtenidos se destacan:

- Validación en campo de nodos IoT ganaderos con una autonomía promedio de 20 días.
- Medición experimental del consumo energético de sensores y periféricos críticos.
- Identificación de limitaciones estructurales del firmware monolítico actual.

Como resultados esperados se plantea:

- Incrementar la autonomía de los nodos en al menos un 50%.

- Reducir el consumo energético promedio mediante apagado selectivo de periféricos.
- Mejorar la mantenibilidad y reutilización del firmware.
- Obtener un modelo arquitectónico replicable para otros sistemas IoT.

4. FORMACION DE RECURSOS HUMANOS

La línea de I+D es desarrollada por un equipo interdisciplinario de la Facultad de Tecnología y Ciencias Aplicadas (FTyCA-UNCA), integrado por docentes-investigadores y recursos humanos en formación, con antecedentes en Ingeniería de Software, Electrónica y Sistemas de Información.

El equipo está integrado por:

- Dos investigadores con título de Doctor, con experiencia en proyectos de investigación y desarrollo tecnológico.
- Un docente maestrando, que participa activamente en las tareas de investigación, aportando a su formación de posgrado en el marco de la línea.
- Docentes universitarios en etapa inicial de investigación, que se incorporan progresivamente a actividades de I+D, fortaleciendo su perfil académico y científico.
- Un alumno avanzado de la carrera Ingeniería en Electrónica, que participa en tareas de diseño, implementación y validación experimental del sistema, en el marco de una beca de investigación, contribuyendo a su formación temprana en actividades científicas y tecnológicas.

La línea se articula con acciones sistemáticas de formación de recursos humanos, promoviendo la participación de docentes y estudiantes en actividades de investigación aplicada, así como el desarrollo de tesis de posgrado y tesinas de grado vinculadas a sistemas embebidos, Internet de las Cosas (IoT) y eficiencia energética, tanto en curso como en proyección.

5. BIBLIOGRAFÍA

Anastasi, G., Conti, M., Di Francesco, M., & Passarella, A. (2009). Energy conservation in wireless sensor networks: A survey. Ad Hoc Networks. doi:10.1016/j.adhoc.2008.06.003.

Dargie, W. (2012). Dynamic Power Management in Wireless Sensor Networks: State-of-the-Art. IEEE Sensors Journal. doi:10.1109/JSEN.2011.2174149.

Bouguera, T., et al. (2018). Energy Consumption Model for Sensor Nodes Based on LoRa and LoRaWAN. Sensors. doi:10.3390/s18072104.

Pottie, G. J., & Kaiser, W. J. (2000). Wireless integrated network sensors. Communications of the ACM. doi:10.1145/332833.332838

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). Design Patterns: Elementos de software orientado a objetos reutilizables. Addison-Wesley. ISBN: 0-201-63361-2.

Blockchain y Ciudades Inteligentes: Posibilidades de Implementación en la Gestión Pública Local en Argentina

Rocío A. Rodríguez[1,2,3], Pablo M. Vera[1,2,3], Santiago Zavatti[1], Nicolas Repetti[1], Claudia G. Alderete[1,2,3], Dora Mendoza[1,4]

[1]Centro de Altos Estudios en Tecnología Informática (CAETI)
Facultad de Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

[2]Facultad Regional Haedo (FRH)
Universidad Tecnológica Nacional (UTN)
Paris 522, Haedo, Buenos Aires, Argentina

[3]Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT)
Universidad Nacional de La Matanza (UNLaM)
Florencio Varela 1903, San Justo, Buenos Aires, Argentina

[4]Laboratorio de Investigación y Desarrollo Tecnológico en Energías Renovables
Sede Regional Orán
Universidad Nacional de Salta, Argentina

{rocioandrea.rodriguez, pablomartin.vera, claudia.alderete}@uai.edu.ar
{nicolas.repetti, SantiagoNahuel.Zavatti} @alumnos.uai.edu.ar
dmendoza412@oran.unsa.edu.ar

RESUMEN

Las ciudades inteligentes constituyen un marco estratégico para el desarrollo urbano sostenible, apoyado en el uso de tecnologías de la información y el software como habilitadores de nuevos servicios y formas de gestión pública. En trabajos previos, el grupo de investigación ha abordado distintos aspectos vinculados al análisis y evaluación de modelos de Smart Cities, considerando sus dimensiones, métricas y planes de acción en el contexto latinoamericano.

En esta nueva etapa, se propone explorar la incorporación de tecnologías emergentes de software, particularmente blockchain, como una nueva línea de investigación aplicada a gobiernos locales. El objetivo del trabajo es analizar de manera conceptual el potencial de blockchain como infraestructura para la validación, trazabilidad y transparencia de la información municipal, en articulación con los distintos ejes del modelo de ciudad inteligente. Asimismo, se mencionan otras tecnologías en estudio, como los sistemas conversacionales (chatbots), orientadas a mejorar la interacción entre municipios y ciudadanía.

Palabras clave: Ciudades Inteligentes, TIC, Blockchain

CONTEXTO

Esta línea de investigación y desarrollo (I+D) forma parte de los proyectos radicados en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). En este proyecto participan docentes y alumnos, de tres sedes de la universidad sede Centro, Castelar (ambas en la provincia de Buenos Aires) y sede Rosario. El proyecto es continuación de una línea de investigación

previa, cuenta con financiamiento asignado y una duración de 2 años.

1. INTRODUCCIÓN

"Una blockchain no es otra cosa que una base de datos que se halla distribuida entre diferentes participantes, protegida criptográficamente y organizada en bloques de transacciones relacionados entre sí matemáticamente. Expresado de forma más breve, es una base de datos descentralizada que no puede ser alterada" [1].

Blockchain permite registrar información de forma inmutable, verificable y descentralizada. Eso hace que sea muy potente para garantizar que un dato es auténtico y no fue alterado. Las propiedades significativas son:

- Inmutabilidad: una vez que la información se registra, no se puede modificar sin dejar rastro.
- Trazabilidad: se puede ver el historial completo de un dato.
- Descentralización: no depende de una única autoridad que "certifique".
- Consenso: varios nodos validan que la información sea correcta antes de aceptarla.

Es importante entender que, por lo general, blockchain no almacena el dato completo por cuestiones de costos y privacidad; lo habitual es guardar un hash del dato (huella digital), el dato original queda fuera del blockchain. Si el dato cambia, aunque sea un único carácter, el hash cambia y se detecta fraude.

Existen diversas plataformas blockchain que pueden ser utilizadas según el contexto institucional, desde redes públicas con soporte de contratos inteligentes (como Ethereum) hasta soluciones privadas orientadas a entornos organizacionales (como Hyperledger Fabric).

Las ventajas frente a sistemas tradicionales son las que se muestran en la tabla 1.

Tabla 1. Comparación entre sistemas

Sistema Tradicional	Blockchain
Autoridad central	Confianza distribuida
Base de datos editable	Registro inmutable
Difícil auditoría	Auditoría transparente
Vulnerable a corrupción	Alta resistencia a manipulación

Cabe destacar también las limitaciones de blockchain

- No garantiza que el dato original sea verdadero (solo que no fue modificado)
- Costos y escalabilidad
- Complejidad técnica y legal
- Privacidad, mal diseñada puede ser un problema

Es importante comprender que blockchain valida integridad, no veracidad.

En Argentina la Secretaría de Modernización de la Nación (Presidencia de la Nación Argentina) ha establecido un modelo de ciudad inteligente [2], [3]. Este modelo surge del análisis de otros modelos previos existentes: ISO 37.120 [4]; ONU ITU [5]; European SCI [6]; ONU Habitat [7], [8].

El modelo establece 5 dimensiones (ver figura 1).

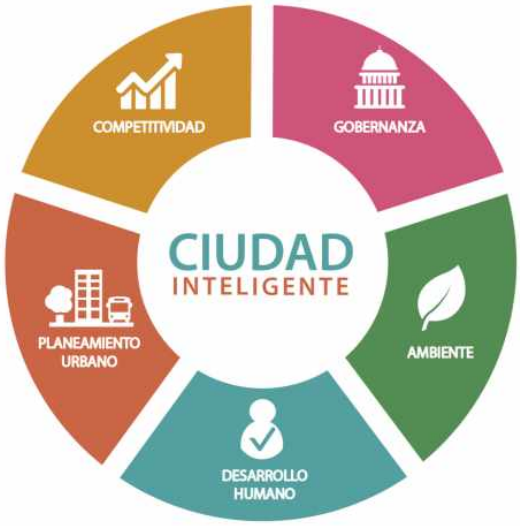


Figura 1. Dimensiones del Modelo de Ciudad Inteligente - Argentina

Todas las dimensiones del modelo pueden beneficiarse de implementar blockchain. En el modelo de ciudad inteligente, la tecnología blockchain funciona como una capa transversal de validación y trazabilidad de la información, fortaleciendo la gobernanza, la transparencia y la confianza en las distintas dimensiones.

A continuación, se explicitan usos posibles en cada una de las dimensiones:

1. Gobernanza:
    a) Validación de documentos públicos: Ordenanzas, Decretos, Resoluciones, Licitaciones. El registro en blockchain del hash del documento, fecha de publicación, dependencia emisora; beneficia la transparencia e integridad.

    b) Presupuesto participativo: Registro de propuestas, Registro de votaciones, Registro de resultados.
    Blockchain: No vota por sí sola, Garantiza que los resultados no se alteren

    c) Auditoría y control: Tribunal de Cuentas / Concejo Deliberante, Acceso a un registro inmutable
    Blockchain actúa como libro de actas digital.

2. Ambiente:
    a) Monitoreo ambiental confiable: Calidad del aire, Ruido urbano, Agua
    Los datos de sensores se almacenan en sistemas municipales se sellan con hash en blockchain. Nadie puede alterar mediciones después.

    b) Gestión de residuos: Registro de recolección, Trazabilidad del reciclaje, Certificación de tratamiento final.
    Blockchain: Garantiza que el proceso declarado ocurrió realmente

3. Desarrollo humano:
    a) Certificados y constancias: Cursos municipales, Capacitaciones, Oficios
    Blockchain: Certifica autenticidad, Evita falsificación, Facilita verificación por terceros
    b) Programas sociales: Asignación de beneficios, Seguimiento de cumplimiento
    Blockchain: No publica datos personales, Registra reglas y entregas

4. Planeamiento Urbano:
    a) Expedientes urbanos: Permisos de obra, Planos aprobados, Cambios de zonificación. Blockchain registra: Versiones, Fechas, Aprobaciones
    Evita modificaciones retroactivas.
    b) Obras públicas: Pliegos, Certificados de avance, Actas de inspección
    Blockchain: Da trazabilidad, Facilita auditorías, Reduce conflictos legales
    c) Datos urbanos abiertos: Publicación de datasets, Garantía de que no fueron manipulados

5. Competitividad
    a) Simplificación de trámites: Habilitaciones comerciales, Permisos, Certificados
    Blockchain: Reduce intermediarios, Acelera validaciones
    b) Compras y contrataciones públicas
    Smart contracts simples: Pago contra cumplimiento, Condiciones verificables.
    Blockchain: Aumenta confianza de proveedores.
    c) Ecosistema emprendedor: Validación de proyectos, Certificación de incubadoras municipales.
    Blockchain: Trazabilidad de fondos públicos

En la tabla 2, se resume la aplicación de Blockchain en cada dimensión del modelo.

Tabla 2. Aplicación de blockchain en cada Eje del Modelo de Ciudad Inteligente

Dimensión	Aplicación de Blockchain
Gobernanza	Transparencia, integridad documental
Ambiente	Datos confiables y auditables
Desarrollo humano	Certificación e inclusión digital
Planeamiento urbano	Trazabilidad de decisiones
Competitividad	Confianza, eficiencia, simplificación

En los últimos años, académicamente, se ha mostrado un interés creciente en el estudio de blockchain más allá de las criptomonedas, evaluando su potencial para transformar servicios públicos y mejorar la transparencia y confianza en la administración pública. Artículos recientes han propuesto estrategias de uso de blockchain en notificaciones certificadas de gobiernos locales y han analizado barreras organizativas para su adopción en el sector público. Además, se han desarrollado marcos conceptuales que integran blockchain con otras tecnologías para reforzar la integridad de datos en el contexto de ciudades inteligentes [9], [10], [11].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Analizar el uso potencial de tecnologías emergentes de software, particularmente blockchain, en el contexto de ciudades inteligentes.
- Identificar oportunidades de aplicación de blockchain en los distintos ejes del modelo de Smart City para gobiernos locales.
- Explorar el uso de blockchain como mecanismo de validación y trazabilidad de información municipal.
- Explorar el uso de sistemas conversacionales (chatbots) como herramientas de software para la interacción entre gobiernos locales y la ciudadanía.

3. RESULTADOS OBTENIDOS/ESPERADOS

El equipo de investigación cuenta con una trayectoria consolidada en el área de ciudades inteligentes, reflejada en numerosos trabajos publicados. En esta etapa, se propone explorar nuevas tecnologías emergentes de software, como blockchain, así como la posible incorporación de herramientas de inteligencia artificial, tales como sistemas conversacionales (chatbots), en el contexto de plataformas municipales. Si bien parecen ser dos temas separados pueden en un futuro integrarse teniendo un chatbot que consulta información certificada en blockchain.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está formado por 5 docentes todos con estudios de posgrado realizados ó en curso: 2 de ellos son doctores en Ciencias Informática graduados en la Universidad Nacional de La Plata (UNLP), 2 con maestría y actualmente realizando el Doctorado en Informática en la Universidad Abierta Interamericana (UAI) y 1 realizando la Maestría en Gestión Educativa (UAI).

Este proyecto también cuenta con la participación de alumnos de grado y posgrado de la UAI que no están vinculados con actividades docentes (actualmente en el proyecto colaboran 4 alumnos).
En esta área se encuentran en realización 1 tesis de doctorado (UNLP), 1 tesis de doctorado (UAI) y 1 tesina de grado (UAI), siendo directores de dichas tesis miembros del equipo de investigación.

5. BIBLIOGRAFIA

[1] Preukschat, Alex (2017). Blockchain. *La Revolución Industrial de Internet*.

[2] Castella L. (2017). “La Importancia de un Modelo de Planificación Estratégica

para el Desarrollo de Ciudades Inteligentes". Secretaría de Modernización, Presidencia de la Nación. JATIC 2017

[3] Palacio Cortés C. (2019). "Gobernanza: base en la Estrategia de Ciudades Inteligentes. El caso Argentino de País Digital". Dirección de Investigación, Innovación y Control Subsecretaría de Modernización de la Nación, Presidencia de la Nación. Smart City Expo Curitiba 2019

[4] ISO (2014), Desarrollo de Comunidades Sostenibles y Resilientes: Indicadores Globales para los servicios de las ciudades y la calidad de vida.

[5] ITU-T (2015), Grupo de Estudio 5, Informe Técnico sobre Ciudades Inteligentes y Sostenibles.

[6] Giffinger, Rudolf et al. (2007-2015), European Smart Cities, Universidad Tecnológica de Viena

[7] ONU Hábitat III (2015), United Nations Conference on Housing and Sustainable Urban Development, ONU.

[8] ONU Hábitat III, Objetivo 11 (2015) United Nations Conference on Housing and Sustainable Urban Development, ONU.

[9] Zabala-Vargas, S., Álvarez-Pizarro, Y., Sánchez-Galvis, I., & Rubio-Vásquez, K. (2024). Blockchain-Based Strategy to Optimize Certified Notifications from Government Entities. *Administrative Sciences*, *14*(9), 195.

[10] Murano, A., Bruno, B., & Vespri, V. (2025). Blockchain barriers in public organizations: A systematic literature review. *Socio-Economic Planning Sciences*, 102364.

[11] Islam, R., Bose, R., Roy, S., Khan, A. A., Sutradhar, S., Das, S., ... & AlZubi, A. A. (2025). Decentralized trust framework for smart cities: a blockchain-enabled cybersecurity and data integrity model. *Scientific Reports*, *15*(1), 23454.

De ERC-721 a Soulbound: La Ingeniería de CertChain

Transformando un Estándar de Propiedad en una Prueba de Identidad Académica Intransferible

Esp. Ing. Carolina Canessa | Mgter. Marcelo Palma | Ing. Luis Canessa

Licenciatura en Sistemas de Información — Facultad de Informática y Diseño

Universidad Champagnat - Mendoza, Argentina

{canessacarolina, palmamarcelo, canessaluis}@uch.edu.ar

RESUMEN

CertChain es un proyecto de investigación y desarrollo orientado a la emisión y validación de certificaciones académicas digitales mediante tecnología blockchain. El trabajo aborda un desafío técnico central: la adaptación del estándar ERC-721, originalmente diseñado para activos transferibles, hacia un modelo de credenciales académicas intransferibles mediante la implementación del patrón Soulbound. La arquitectura propuesta integra contratos inteligentes en redes EVM-compatible (Moonbeam y Polygon), almacenamiento híbrido on-chain/off-chain con IPFS y mecanismos de control de revocación institucional. Los resultados alcanzados incluyen el desarrollo de un prototipo funcional validado en entorno de pruebas, análisis de costos de gas y confirmación de la viabilidad técnica y económica del modelo. Como líneas de continuidad, se prevé la ejecución de un piloto institucional, la incorporación de identidad digital soberana y la ampliación hacia microcredenciales, contribuyendo a la modernización de los sistemas de certificación universitaria en Argentina.

Palabras clave: blockchain; certificación académica; ERC-721; Soulbound Tokens; contratos inteligentes; credenciales digitales; identidad digital soberana.

CONTEXTO

Los sistemas de certificación académica en Argentina se sustentan mayormente en infraestructuras centralizadas que requieren validaciones presenciales, documentación física y procedimientos administrativos de alta carga burocrática. Estas dinámicas generan demoras, costos operativos significativos y limitaciones para la verificación remota, especialmente en escenarios internacionales o interinstitucionales. Asimismo, la dependencia de repositorios cerrados incrementa la exposición a riesgos de falsificación, manipulación documental o pérdida de información ante fallas técnicas o administrativas.

En este marco, la tecnología blockchain se presenta como una infraestructura capaz de aportar inmutabilidad, trazabilidad y verificabilidad descentralizada en la gestión de credenciales digitales. Diversos estudios han señalado su potencial en el ámbito educativo; sin embargo, la adopción efectiva requiere resolver tensiones entre los estándares tecnológicos existentes y la naturaleza jurídica de los títulos académicos,

que deben estar vinculados de manera permanente e intransferible a la identidad del egresado.

Este escenario configura el problema central que motiva la presente línea de investigación: diseñar un modelo de certificación académica digital que combine garantías criptográficas, viabilidad económica y adecuación institucional.

1. INTRODUCCIÓN

CertChain surge como una aplicación descentralizada (dApp) orientada a la emisión y validación de certificaciones académicas digitales mediante contratos inteligentes en redes EVM-compatible. La arquitectura integra frontend web, backend de lógica de negocio, almacenamiento descentralizado en IPFS y smart contracts desarrollados en Solidity.

El desafío técnico abordado en este trabajo radica en la adaptación del estándar ERC-721, originalmente concebido para activos digitales transferibles, hacia un esquema de credenciales académicas intransferibles. Mientras que un NFT tradicional representa propiedad sobre un bien digital, un título universitario constituye una acreditación personal que no puede ser comercializada ni transferida sin desnaturalizar su significado jurídico y académico.

En este sentido, la pregunta de investigación que guía el artículo es: ¿cómo transformar un estándar de representación de propiedad digital en un mecanismo de identidad académica verificable e intransferible?

El presente trabajo se focaliza específicamente en la ingeniería de contratos inteligentes necesaria para implementar el patrón Soulbound, detallando las decisiones técnicas adoptadas, las modificaciones sobre ERC-721 y el análisis de métricas de gas en la red seleccionada. De este modo, se profundiza y amplía la línea de investigación previamente iniciada [1], concentrándose en el aporte técnico que posibilita la materialización del modelo propuesto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1. El Problema con ERC-721 Estándar

El estándar ERC-721 [2] establece una interfaz para tokens no fungibles que incluye funciones de transferencia como transferFrom y safeTransferFrom. Estas funciones permiten que el propietario actual de un NFT lo transfiera a otra dirección, facilitando la creación de mercados secundarios de activos digitales [3].

Sin embargo, esta característica resulta problemática para la certificación académica. Si utilizáramos un ERC-721 estándar sin modificaciones, un egresado podría teóricamente "vender" su título universitario o transferirlo a otra wallet. Esto contradice el principio fundamental de soberanía digital en educación: la credencial debe estar intrínsecamente ligada a la identidad del individuo que completó los estudios, no ser un activo comercializable.

2.2. Implementación de Soulbound Tokens (SBTs)

La solución implementada se basa en el concepto de Soulbound Tokens propuesto por Buterin et al. [4], definidos como tokens no transferibles que representan compromisos o credenciales ligadas a una identidad.

Técnicamente, modificamos el contrato UCHCertification.sol (heredado de la librería OpenZeppelin [5]) para realizar un override total de las funciones de transferencia. Cualquier llamada a transferFrom o

safeTransferFrom ejecuta una operación revert, bloqueando la transacción a nivel de EVM y garantizando la intransferibilidad. El mapeo _owners[tokenId] vincula matemáticamente la credencial a la wallet del egresado y no puede ser modificado post-mint.

Aunque la arquitectura actual utiliza un bloqueo forzoso (hard revert), el diseño contempla la futura adopción del estándar EIP-5192 (Minimal Soulbound NFTs), lo que permitiría a las interfaces de billeteras detectar estandarizadamente el estado de "bloqueo" del token.

Adicionalmente, se implementó una función revokeCertificate(uint256 tokenId) protegida por control de acceso (onlyRole(ADMIN_ROLE)). Si bien reconocemos que esto introduce un vector de centralización, es un requisito legal indispensable para anular títulos en casos de fraude académico. Se investiga actualmente la transición de este rol administrativo hacia un esquema de gobernanza multifirma (MultiSig) para mitigar riesgos de abuso unilateral.

2.3. Arquitectura Tecnológica Chain-Agnostic

CertChain fue diseñado con una arquitectura modular que permite deployments en múltiples blockchains EVM-compatible sin cambios significativos en el código:

• Moonbeam (Polkadot Ecosystem): Seleccionado como blockchain principal para títulos de grado debido a su finalidad determinística (~12s), interoperabilidad con el ecosistema Polkadot vía XCM [6], y gobernanza on-chain robusta. Costo promedio por mint: ~$0.05 USD.

• Polygon PoS: Utilizado para microcredenciales y certificados de menor criticidad. Ofrece finalidad probabilística (~2s), costos ultra-bajos (~$0.001 USD por transacción), y alta adopción empresarial.

La modularidad se logra mediante abstracciones en el backend (Python) que encapsulan las diferencias entre redes (RPC endpoints, chain IDs, tiempos de confirmación) mientras el contrato Solidity permanece idéntico.

2.4. Almacenamiento Híbrido On-Chain/Off-Chain

Para optimizar costos y mantener la verificabilidad del documento original, CertChain implementa un modelo híbrido:

• On-Chain (Blockchain): Se almacenan únicamente los metadatos críticos (hash SHA-256 del documento, dirección del egresado, fecha de emisión, institución emisora, estado de revocación).

• Off-Chain (IPFS): El documento PDF completo se almacena en InterPlanetary File System [7], generando un CID (Content Identifier) determinístico. Este CID se referencia en el token URL del NFT.

Esta arquitectura permite que cualquier verificador: (1) descargue el PDF desde IPFS usando el CID, (2) calcule su hash SHA-256, (3) compare con el hash on-chain para garantizar integridad.

3. RESULTADOS OBTENIDOS / ESPERADOS

3.1. Resultados Alcanzados

• Prototipo Funcional Completo: Sistema validado exitosamente en ambiente de pruebas (Moonbase Alpha testnet) con ciclo completo de emisión, verificación y revocación de certificados. Interacción correcta entre todos los módulos de la arquitectura.

• Estabilidad de Contratos Inteligentes: El contrato UCHCertification.sol fue desplegado

y probado extensivamente en Moonbase Alpha. Se confirmó el correcto funcionamiento del bloqueo de transferencias (revert en 100% de intentos de transferFrom/safeTransferFrom), la capacidad de revocación administrativa, y la resistencia a reentrancy attacks.

• Viabilidad Económica Confirmada: El análisis de costos en Moonbeam mainnet indica un promedio de ~$0.05 USD por certificado emitido (costo de gas), excluyendo el almacenamiento en IPFS que se gestiona mediante nodos propios o servicios de pinning de bajo costo. Para una universidad con 500 graduados/año, esto representa aproximadamente $25 USD/año en costos de red, una reducción drástica comparada con los procesos de certificación tradicional.

3.2. Líneas Futuras

• Piloto Institucional: Ejecutar una implementación piloto con la primera cohorte de graduados de la Universidad Champagnat (estimado: 50-100 usuarios) para recoger métricas empíricas de adopción, usabilidad y satisfacción.

• Extensión a Microcredenciales: Ampliar la funcionalidad hacia certificados intermedios (cursos, diplomados, talleres) utilizando el deployment en Polygon PoS para minimizar costos en emisiones de alto volumen.

• Integración de Identidad Digital Soberana (DID): Explorar la incorporación de W3C Decentralized Identifiers (DIDs) y Verifiable Credentials para interoperabilidad global con sistemas educativos internacionales y plataformas de empleo (LinkedIn, Indeed).

• Pruebas de Seguridad y Resiliencia: Realizar auditoría formal del smart contract con herramientas como Slither y Mythril, y pruebas de stress testing para validar la capacidad del sistema ante emisión masiva de certificados.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto CertChain ha servido como plataforma de formación en tecnologías emergentes para la comunidad académica de la Universidad Champagnat:

• Equipo de Investigación: 3 investigadores (1 Especialista en Ingeniería, 1 Magíster en Tecnología Informática, 1 Ingeniero) con dedicación parcial al proyecto.

• Estudiantes de Grado: 8 alumnos de la carrera de Ingeniería en Sistemas participaron en prácticas profesionales supervisadas, trabajando en diferentes módulos del sistema (frontend, backend, smart contracts, testing).

• Áreas de Capacitación: Los participantes adquirieron competencias en desarrollo blockchain (Solidity, Web3.js, Ethers.js), arquitecturas descentralizadas (IPFS, nodos blockchain), frameworks modernos (Vue.js 3 con Composition API), y buenas prácticas de seguridad en smart contracts (auditoría con herramientas estáticas, patrones como Checks-Effects-Interactions).

• Difusión Científica: El proyecto generó 1 artículo presentado en EnIDI 2025 (San Rafael, Mendoza) titulado "La Blockchain y la Certificación de Documentos en Argentina: Análisis del Proyecto CertChain", y 1 presentación técnica en DevConnect enfocada en la ingeniería de Soulbound Tokens.

• Integración Académica: El material técnico del proyecto fue incorporado como caso de estudio en la Diplomatura en LegalTech de la Universidad Nacional de La Plata, específicamente en el módulo de tokenización de activos del mundo real (RWA) y contratos inteligentes.

5. BIBLIOGRAFÍA

[1] C. Canessa, M. Palma, and L. Canessa, "La Blockchain y la Certificación de

Documentos en Argentina: Análisis del Proyecto CertChain," in Actas de EnIDI 2025, San Rafael, Mendoza, Argentina, 2025.

[2] W. Entriken, D. Shirley, J. Evans, and N. Sachs, "EIP-721: Non-Fungible Token Standard," Ethereum Improvement Proposals, no. 721, Jan. 2018. [Online]. Available: https://eips.ethereum.org/EIPS/eip-721

[3] A. M. Antonopoulos and G. Wood, Mastering Ethereum: Building Smart Contracts and DApps. Sebastopol, CA, USA: O'Reilly Media, 2018.

[4] V. Buterin, E. G. Weyl, and P. Ohlhaver, "Decentralized Society: Finding Web3's Soul," SSRN Electron. J., May 2022, doi: 10.2139/ssrn.4105763.

[5] OpenZeppelin, "OpenZeppelin Contracts Documentation - ERC721," 2024. [Online]. Available: https://docs.openzeppelin.com/contracts/4.x/erc721

[6] G. Wood, "Polkadot: Vision for a heterogeneous multi-chain framework," Parity Technologies, White Paper, 2016. [Online]. Available: https://github.com/polkadot-io/polkadot-white-paper/blob/master/PolkaDotPaper.pdf.

[7] J. Benet, "IPFS - Content Addressed, Versioned, P2P File System," arXiv preprint arXiv:1407.3561, 2014.

[8] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," Bitcoin.org, 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf.

[9] P. De Filippi and S. Hassan, "Blockchain technology as a regulatory technology: From code is law to law is code," First Monday, vol. 21, no. 12, Dec. 2016. [Online]. Available: https://firstmonday.org/ojs/index.php/fm/article/view/7113

[10] A. Grech and A. F. Camilleri, "Blockchain in Education," JRC Science for Policy Report, European Commission, Luxembourg, 2017, doi: 10.2760/60649.

Diseño de un Metamodelo Conceptual de Infraestructura Pública Digital basado en Ingeniería de Software

Mauro Cambarieri, Alejandra Viadana, Héctor Ruiz, Bruno Gorosito, Carlos Lugani, Nicolás Martin

Universidad Nacional de Rio Negro. Sede Atlántica
Laboratorio de Informática Aplicada
{mcambarieri, caviadana, hruiz ,bjgorosito, clugani, namartin}@unrn.edu.ar

RESUMEN

La Infraestructura Pública Digital (Digital Public Infrastructure, DPI) se ha consolidado en los últimos años como un concepto estratégico para la transformación digital de los gobiernos. Iniciativas impulsadas por organismos internacionales como el Programa de las Naciones Unidas para el Desarrollo (PNUD), el Banco Mundial y el G20 destacan la necesidad de construir DPI como bienes públicos digitales basados en principios de interoperabilidad, estándares abiertos, reutilización, apertura y gobernanza. Sin embargo, gran parte de estas propuestas se presentan en un nivel macro o político, sin una formalización conceptual que permita su traducción sistemática en arquitecturas de software y prácticas de ingeniería.

Este trabajo propone una línea de investigación orientada al desarrollo de un metamodelo conceptual de Infraestructura Pública Digital, concebido como un puente entre los marcos estratégicos de Gobierno Digital y las decisiones técnicas de diseño e implementación. El metamodelo integra enfoques de ingeniería de software centrados en el dominio, como el Diseño Dirigido por el Dominio (DDD), la Ingeniería Dirigida por Modelos (MDE) y las Líneas de Productos de Software (LPS), alineándose con arquitecturas basadas en microservicios, APIs y patrones promovidos por iniciativas como GovStack[1].

El objetivo es desarrollar un marco de trabajo que permita describir, analizar, evaluar y gobernar infraestructuras públicas digitales en distintos niveles de gobierno, sin imponer arquitecturas tecnológicas específicas. Se propone, en consecuencia, un artefacto conceptual que permita describir, analizar y evolucionar soluciones de DPI de manera sistemática, reutilizable y adaptable a diversos contextos institucionales.

Palabras clave: Infraestructura Pública Digital; Metamodelo; Ingeniería de Software.

CONTEXTO

La transformación digital del Estado se asienta en un escenario global caracterizado por aceleración tecnológica, creciente demanda de los servicios públicos, transparencia, eficiencia y la necesidad de sostenibilidad institucional. En este contexto, los gobiernos ya no enfrentan únicamente el desafío de "digitalizar" procesos administrativos, sino de rediseñar la forma en que producen valor público.

Durante las primeras etapas del gobierno electrónico, la incorporación de tecnologías digitales se orientó principalmente a la automatización de trámites y a la provisión de portales informativos. Posteriormente, el concepto de Gobierno Digital amplió esta visión, incorporó un enfoque estratégico orientado al uso de tecnologías digitales para crear valor público, promover la interoperabilidad entre las agencias de gobierno y avanzar hacia una gestión centrada en el ciudadano (BID, 2022). La Organización para la Cooperación y el Desarrollo Económicos (OCDE) ha definido el Gobierno Digital como el uso estratégico de tecnologías digitales para crear valor público dentro de un ecosistema que involucra múltiples actores institucionales y sociales (OCDE, 2014). En los últimos años, la noción de DPI ha emergido como una evolución conceptual que propone estructurar el entorno digital del Estado a partir de bloques fundamentales reutilizables, tales como identidad digital, pagos e intercambio de datos. Estos bloques, concebidos como bienes públicos digitales, permiten habilitar múltiples servicios sobre una base común, reduciendo la fragmentación tecnológica y favoreciendo la interoperabilidad (Clastornik Taube & Muente, 2024).

Organismos internacionales (PNDU, Banco mundial, G20) coinciden en que una DPI efectiva debe basarse en principios de apertura, modularidad, reutilización, seguridad y gobernanza, permitiendo que

[1] https://www.govstack.global/

distintos actores —gubernamentales, privados y de la sociedad civil— puedan innovar sobre una base común. No obstante, la ausencia de modelos conceptuales compartidos dificulta la alineación entre políticas públicas, marcos normativos y arquitecturas de software, generando soluciones fragmentadas, con alta dependencia tecnológica y limitada capacidad de evolución.

Desde la Ingeniería de Software aplicada al sector público, surge la necesidad de formalizar estos conceptos en modelos que permitan reducir la brecha entre el nivel estratégico del gobierno digital y el nivel técnico de implementación, especialmente en contextos caracterizados por cambios regulatorios, organizacionales y tecnológicos constantes. Esta brecha entre el nivel estratégico de la DPI y su materialización técnica constituye el problema central que motiva la presente investigación.

Este proyecto busca diseñar un metamodelo conceptual de Infraestructura Pública Digital que identifique los bloques de construcción, defina las interacciones entre actores del ecosistema y establezca un marco de evaluación de madurez para su implementación en diferentes niveles de gobierno.

1. INTRODUCCIÓN

La Infraestructura Pública Digital (Digital Public Infrastructure, DPI) se define como un enfoque de digitalización centrado en la creación de "bloques de construcción digitales fundamentales diseñados para el beneficio público". Esta ha emergido como el paradigma dominante para comprender la transformación digital del sector público. Según la definición adoptada por los líderes del G20 en la Declaración de Nueva Delhi de 2023(G20, 2023), DPI se conceptualiza como "un conjunto de sistemas digitales compartidos, seguros e interoperables, construidos sobre tecnologías abiertas, para proporcionar acceso equitativo a servicios públicos y/o privados a escala social" (Ardic Alper et al., 2023). Esta definición establece tres pilares tecnológicos fundamentales —identificación digital, interoperabilidad de datos y métodos de pago— que han sido adoptados como marco de referencia por los países de la región.

DPI puede entenderse como una capa de software intermedia dentro del ecosistema digital del Estado, situada por encima de la infraestructura física (conectividad, dispositivos, centros de datos, nube) y por debajo de las aplicaciones y plataformas de servicios específicos como muestra la Figura 2 (Clark,J et al,2025).

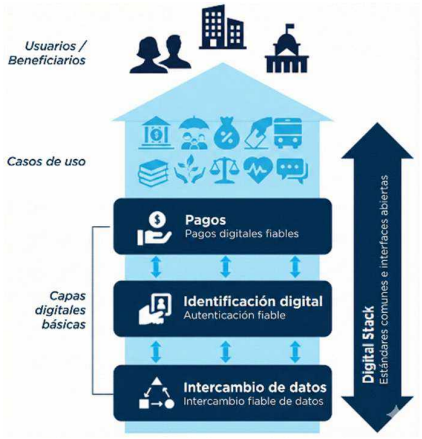


Figura 2. Fuente: World Bank (2022).

Diversos organismos internacionales han propuesto definiciones de DPI, evidenciando tanto convergencias conceptuales como matices estratégicos.

El Grupo de Trabajo sobre Economía Digital del G20 (Digital Economy Working Group, DEWG) define la DPI como un conjunto de sistemas digitales compartidos que deben ser seguros e interoperables, y que pueden construirse sobre estándares y especificaciones abiertas para ofrecer y garantizar acceso equitativo a servicios públicos y/o privados a escala social. Estos sistemas deben estar gobernados por marcos legales aplicables y reglas habilitantes que promuevan el desarrollo, la inclusión, la innovación, la confianza y la competencia, respetando los derechos humanos y las libertades fundamentales (G20, 2023a).

La iniciativa GovStack, junto con la Digital Public Goods Alliance (DPGA), concibe a DPI como soluciones y sistemas que habilitan la provisión efectiva de funciones y servicios esenciales en los sectores público y privado. Esto incluye, entre otros elementos, identidad digital y mecanismos de verificación, registro, sistemas de pago, intercambio de datos y sistemas de información, incluyendo aquellos específicos de sectores como salud o educación. Asimismo, reconocen que la DPI de un país puede estar compuesta por implementaciones múltiples, tanto propietarias como de código abierto, incluyendo bienes públicos digitales (GovStack, 2022).

La OCDE define la DPI como sistemas digitales comunes y fundacionales que habilitan la provisión de servicios en la era digital. Estos elementos pueden ser desarrollados por el sector público, el sector privado o de manera colaborativa, con el objetivo de facilitar la provisión y el acceso a servicios a través de sectores e incluso de fronteras. La OCDE enfatiza que la DPI debe basarse en estándares y ser reutilizable (OCDE, 2024).

Finalmente, el Centro de Infraestructura Publica Digital (Center for Digital Public Infrastructure, CDPI) propone una definición más orientada al enfoque estratégico, caracterizando la DPI como un método para abordar problemas

socioeconómicos a gran escala mediante la combinación de intervenciones tecnológicas, gobernanza público-privada e innovación de mercado (CDPI, 2024).

Los sistemas identificados como "DPI principales" incluyen:

1. Identidad digital y firmas electrónicas: Permiten la autenticación remota segura y validan legalmente las transacciones en línea.
2. Pagos digitales: Facilitan transferencias de dinero seguras, instantáneas y sin efectivo entre personas, gobiernos y empresas.
3. Intercambio de datos: Permiten el flujo y la reutilización optimizada de datos de manera confiable entre diferentes entidades, protegiendo la información personal.

La distinción entre infraestructura digital tradicional y DPI constituye uno de los desarrollos conceptuales más relevantes en la evolución del gobierno digital. Mientras que la infraestructura digital tradicional se enfocaba predominantemente en la provisión de hardware, conectividad y sistemas de información aislados, el DPI adopta un paradigma de "bloques constructivos" (building blocks) que pueden ser reutilizados, combinados e innovados por múltiples actores del ecosistema digital (World Bank,2022). La siguiente figura muestra la distinción:

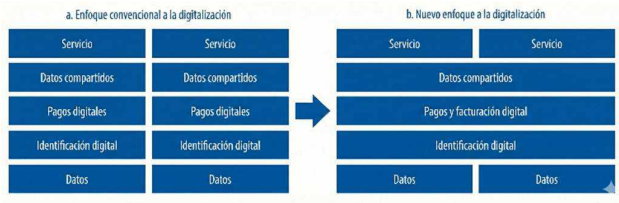


Fig. 1. Enfoque tradicional comparado con DPI

La infraestructura digital tradicional en la región se caracterizaba por sistemas monolíticos desarrollados de manera independiente para cada servicio público, con baja interoperabilidad, alta redundancia y arquitecturas propietarias que generaban dependencia de proveedores específicos. Este enfoque, predominante durante las décadas de 1990 y 2000, produjo silos de información gubernamental que replicaban la fragmentación del Estado físico en el entorno digital. Los ciudadanos enfrentaban experiencias inconsistentes, múltiples credenciales para diferentes servicios, y procesos que requerían presentar repetidamente la misma información a diferentes agencias. (Desai, V. etal, 2023).

En los últimos años, la DPI ha sido promovida por organismos internacionales como un habilitador estructural del desarrollo económico y social. El Banco Mundial ha señalado que los países que logran articular sistemas de identidad digital, pagos interoperables e intercambio seguro de datos pueden acelerar la inclusión financiera, mejorar la eficiencia del gasto público y ampliar el acceso a servicios esenciales. Sin embargo, en América Latina persisten desafíos significativos que obstaculizan la consolidación de estas infraestructuras (Porteous, D. et al, 2024).

En primer lugar, se observa una marcada fragmentación tecnológica e institucional. Las soluciones digitales suelen implementarse de manera aislada, bajo lógicas sectoriales o proyectos de corto plazo, sin una arquitectura común que garantice interoperabilidad y reutilización. En segundo lugar, existe una ausencia de estandarización formal de los bloques de construcción que componen la DPI, lo que dificulta su replicabilidad y escalabilidad entre niveles de gobierno (municipal, provincial y nacional). Finalmente, los países carecen de instrumentos conceptuales y métricas específicas que permitan evaluar el grado de madurez de su infraestructura digital pública, limitando la planificación estratégica y la comparación interjurisdiccional.

Si bien la literatura y los marcos internacionales —incluyendo los desarrollos analíticos de la Organización para la Cooperación y el Desarrollo Económicos (OCDE)— reconocen la relevancia de la transformación digital del Estado, persiste un vacío metodológico: no se dispone de un metamodelo conceptual que formalice los componentes estructurales de la DPI, modele las interacciones del ecosistema y permita evaluar su evolución de manera sistemática. Este vacío constituye el problema central que aborda la presente línea de investigación.

En el ámbito del Gobierno Digital, el diseño de sistemas de información enfrenta desafíos particulares asociados a la complejidad normativa, la diversidad de actores involucrados y la necesidad de garantizar continuidad, transparencia y confianza pública. En este contexto, enfoques tradicionales de desarrollo de software resultan insuficientes para abordar la escala y criticidad de las plataformas que sustentan servicios públicos esenciales.

La Ingeniería de Software ha desarrollado, en las últimas décadas, enfoques que priorizan la comprensión del dominio y la separación de responsabilidades, como el Diseño Dirigido por el Dominio, cuyo principio central consiste en alinear el diseño del software con el conocimiento del dominio, promoviendo una colaboración estrecha y continua entre expertos del dominio y desarrolladores (Evans & Vernon, 2023). La Ingeniería Basada en Modelos (Model Driven Engineering, MDE) constituye un paradigma de desarrollo que sitúa a los modelos como artefactos centrales del proceso de software. Su objetivo principal es elevar el nivel de abstracción, reducir errores y aumentar la productividad mediante la automatización de tareas repetitivas (France & Rumpe, 2010). Por su parte, las Líneas de Productos de Software, proponen un cambio de paradigma desde el desarrollo de aplicaciones individuales hacia la

construcción de familias de productos relacionados, mediante la identificación explícita de elementos comunes y variabilidades (Parra et al., 2014). Y finalmente, la arquitectura de software centrada en el dominio, como la Arquitectura Hexagonal (Cockburn,2023), permite el desacoplamiento entre la lógica de negocio y las tecnologías subyacentes, lo que establece criterios para la integración de tecnologías emergentes, en consonancia con los principios de la DPI. Estos enfoques permiten capturar explícitamente el conocimiento del dominio, gestionar la variabilidad y reducir el acoplamiento entre la lógica de negocio y las decisiones tecnológicas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo parte de la hipótesis de que la Infraestructura Pública Digital (DPI) puede beneficiarse significativamente de los enfoques centrados en el dominio si estos se articulan mediante un metamodelo conceptual que permita describir formalmente sus componentes, relaciones y reglas de evolución. Un metamodelo de DPI no pretende imponer una arquitectura tecnológica específica, sino ofrecer un marco de referencia común que facilite el diseño, la evaluación y la gobernanza de infraestructuras públicas digitales en distintos niveles de gobierno.

En este contexto, el objetivo general consiste en diseñar un metamodelo conceptual de Infraestructura Pública Digital que estandarice los bloques de construcción, defina las interacciones entre los actores del ecosistema y establezca un marco de evaluación de madurez para su implementación en niveles municipal, provincial y nacional. A partir de esta premisa, se definen los siguientes objetivos específicos:

- O1: Integrar enfoques de Software de Dominio Específico (Domain-Specific Software, DSW), proponiendo que la tecnología no sea un accesorio, sino un reflejo fiel de la lógica del Estado.
- O2: Identificar y catalogar los bloques de construcción (building blocks) transversales y específicos de dominio.
- O3: Modelar las interacciones y flujos de datos entre los actores del ecosistema (Estado, ciudadanos, empresas y organismos de control).
- O4: Definir un Índice de Madurez de DPI (DPI-MI) con indicadores claros para cada nivel.
- O5: Validar el metamodelo mediante un caso de estudio en un dominio específico.
- O6: Difundir conocimiento sobre el metamodelo mediante documentos, estándares, buenas prácticas y servicios de asistencia técnica.
- O7: Proponer un artefacto conceptual que permita describir, analizar y evolucionar soluciones de DPI de manera sistemática, reutilizable y adaptable a distintos contextos institucionales.

La propuesta metodológica se estructura en cinco fases (I–V), organizadas de manera secuencial e iterativa, que permiten avanzar desde la fundamentación conceptual hasta la validación empírica y la transferencia de resultados.

Fase I – Fundamentación conceptual y delimitación del dominio (O1 y O7)

Tiene como propósito establecer el marco teórico del metamodelo, integrando enfoques de Software de Dominio Específico (DSW) al contexto de Gobierno Digital y DPI. Se realiza una revisión sistemática de literatura sobre DPI, arquitectura empresarial y modelado centrado en el dominio; se identifican los principios estructurales de la DPI (identidad, pagos, interoperabilidad y datos); se delimita el dominio administrativo estatal (procesos, reglas normativas, entidades y relaciones); y se define un meta-lenguaje conceptual que represente formalmente la lógica administrativa. La fase concluye con una validación preliminar con expertos. El resultado es un artefacto conceptual formal que vincula la lógica administrativa con su representación digital y funciona como marco descriptivo, analítico y evolutivo, reutilizable en distintos contextos institucionales.

Fase II – Estandarización de bloques de construcción (O2)

Busca identificar, clasificar y formalizar los building blocks que estructuran la DPI. Se realiza un relevamiento comparativo de componentes transversales (identidad digital, interoperabilidad, firma digital, pagos, infraestructura de datos) y de bloques específicos por dominio (salud, políticas sociales, gestión tributaria). Se definen atributos estructurales como criticidad regulatoria, dependencia, escalabilidad y reutilización, construyendo una taxonomía formal integrada al metamodelo. El resultado es una matriz estandarizada de componentes DPI comparables y reutilizables entre áreas y jurisdicciones.

Fase III – Modelado del ecosistema e interoperabilidad (O3)

Se orienta a representar formalmente las interacciones entre actores y los flujos de datos. Se identifican actores multinivel (Estado, ciudadanos, empresas, organismos de control), se modelan relaciones estructurales y eventos clave (registro, autorización, pago, notificación), y se establecen principios de gobernanza de datos. El resultado es un modelo sistémico que explicita dependencias e interoperabilidad dentro del ecosistema DPI.

Fase IV – Diseño del Índice de Madurez DPI (DPI-MI) (O4)

Tiene como finalidad desarrollar un instrumento de evaluación del grado de desarrollo de la DPI. Se definen dimensiones (gobernanza, interoperabilidad, estandarización, cobertura y escalabilidad), niveles progresivos de madurez e indicadores cuantitativos y cualitativos, junto con una metodología de aplicación diagnóstica. El resultado es un instrumento operativo de evaluación comparativa del nivel de desarrollo de DPI.

Fase V – Validación y transferencia (O5 y O6)

Se selecciona un dominio específico para aplicar el metamodelo, mapear procesos reales, identificar building blocks y evaluar madurez mediante el DPI-MI. Los resultados permiten ajustar iterativamente el modelo. En paralelo, se elaboran estándares conceptuales, buenas prácticas y publicaciones académicas. Esta fase consolida la validación empírica y la transferibilidad del metamodelo como marco de referencia para la implementación sistemática de DPI.

3. RESULTADOS OBTENIDOS/ ESPERADOS

El metamodelo conceptual propuesto contribuye a la formalización teórica de la DPI, integrando principios de Ingeniería de Software con marcos estratégicos de Gobierno Digital. Proporciona un lenguaje común para describir infraestructuras digitales públicas, reduce ambigüedades conceptuales y habilita evaluación comparativa de madurez. Su carácter no prescriptivo permite adaptabilidad tecnológica sin perder coherencia estructural.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se encuentra formado por cuatro investigadores formados, dos investigadores en formación y dos alumnos avanzados de la carrera Licenciatura en Sistemas. En su marco se desarrollarán prácticas profesionales supervisadas, becas de formación práctica y se producirán dos trabajos finales de carrera de grado y la elaboración del plan de trabajo de la carrera de Doctorado (UNRN-Sede Atlántica).

5. BIBLIOGRAFÍA

Ardic Alper, O. P., Galicia Rabadan, G. A., Marin, G., Natarajan, H., Piveteau, T. Y., Ramteke, N. C., & Sarkar, A. (2023). World Bank Group. http://documents.worldbank.org/curated/en/099092023121016458/P178703046f82d07c0bbc60b5e474ea7841

Clark, J., Marin, G., Ardic Alper, O. P., & Galicia Rabadan, G. A. (2025). Digital public infrastructure and development: A World Bank Group approach (Digital Transformation White Paper, Vol. 1). World Bank Group.

Clastornik Taube, J., & Muente, A. (2024). La transformación digital se transforma: Tendencias y avances en la transformación pública digital y su impacto en las adquisiciones de tecnologías y servicios digitales. https://doi.org/10.18235/0013259

Cockburn, A. (2023). Hexagonal architecture. https://www.alistair.cockburn.us

Desai, V., Marskell, J., Marin, G., & Varghese, M. (2023, 15 de marzo). How digital public infrastructure supports empowerment, inclusion, and resilience. World Bank Blogs: Digital Development. https://blogs.worldbank.org/digital-development/how-digital-public-infrastructure-supports-empowerment-inclusion-and-resilience

Evans, E., & Vernon, V. (2023). Domain-driven design distilled. Addison-Wesley.

France, R., & Rumpe, B. (Eds.). (2010). Model-driven software development. Springer.

G20. (2023, 19 de agosto). Digital Economy Ministers Meeting: Outcome document and chair's summary. Bengaluru, Karnataka. https://g7g20-documents.org/database/document/2023-g20-india-sherpa-track-digitaleconomy-ministers-ministers-language-g20-digital-economy-ministers-meeting-outcome-document-and-chair-summary#section-3

GovStack Community of Practice. (2022). GovStack definitions: Understanding the relationship between digital public infrastructure, building blocks & digital public goods. https://www.digitalpublicgoods.net/DPI-DPG-BB-Definitions.pdf

OECD (2014). Recommendation of the Council on Digital Government Strategies. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0406

OECD Publishing. (2024). Developing scalable and secure digital public infrastructure. En Government at a glance: Latin America and the Caribbean 2024. https://www.oecd-ilibrary.org/governance/government-at-a-glance-latin-america-and-the-caribbean-2024_4abdba16-en

Parra, C., et al. (2014). An SOA approach for automating software product lines. En Proceedings of the 29th Annual ACM Symposium on Applied Computing.

Porteous, D. J., Rifon Perez, A., Appaya, S., Martinezgarza, R., Albornoz, N., & Zaballos, A. G. (2024). Unlocking the potential of digital public infrastructure (DPI) in Latin America and the Caribbean (LAC): A region-specific perspective – A World Bank–Inter-American Development Bank technical note. World Bank Group. http://documents.worldbank.org/curated/en/099072524161522274

World Bank. (2022). A digital stack for transforming service delivery: ID, payments, and data sharing. https://documents1.worldbank.org/curated/en/099755004072288910/pdf/P1715920edb5990d60b83e037f756213782.pdf

World Bank. (2024). DPI thinking. CDPI. https://docs.cdpi.dev/the-dpi-wiki/dpi-overview

Fortaleciendo la Conexión entre Proyectos de Software y el Bienestar Comunitario

Mg. Rodolfo Bertone [1], C.C Marcelo Taruschio [2], Lic. Diego Luna[3], Ing. Franco Quintero[4], Santiago Mariano Bustos[5], Juan Santiago Cazón[6], Valentino Creuso[7], Nicolás Federico Haluska[8], Franco Santiago Timpanaro[9], Jonatan Scaglia[10], Luciano Signorelli[11]

Facultad de Ciencias Exactas e Ingeniería. Universidad Católica de La Plata.
[1]rodolfo.bertone@ucalp.edu.ar [2]marcelo.taruschio@ucalp.edu.ar [3]diego.luna@ucalp.edu.ar
[4]franco.quintero@ucalp.edu.ar [5]smbust38@yahoo.com [6]juanscazon@gmail.com
[7]valentinocreuso@gmail.com [8]nicohaluska@gmail.com [9]francotimpanaro20@gmail.com
[10]jonatanscaglia@gmail.com [11]luchitosignorelli@gmail.com

Resumen

La Usina de Ideas constituye un dispositivo académico institucional orientado a la generación sistemática, evaluación estructurada y desarrollo progresivo de soluciones tecnológicas con potencial de transferencia efectiva al medio socio-productivo. Integrada en el ámbito de la Facultad de Ciencias Exactas e Ingeniería, articula de manera orgánica las funciones sustantivas de investigación, extensión y vinculación tecnológica bajo un modelo operativo basado en proyectos con orientación aplicada.

Durante el período 2025 se consolidaron líneas estratégicas centradas en el desarrollo de sistemas basados en software con impacto ambiental, organizacional y cultural, priorizando problemáticas concretas del entorno territorial. Entre los principales desarrollos se encuentran: (i) un sistema de gestión de residuos sólidos urbanos no habituales con geolocalización y planificación dinámica; (ii) una plataforma digital para la administración integral de patrimonio museístico; y (iii) una solución tecnológica para la optimización de espacios de trabajo compartido mediante algoritmos de asignación eficiente de recursos.

El modelo metodológico adoptado integra prácticas contemporáneas de ingeniería de software, metodologías ágiles, control de versiones distribuido, integración y despliegue continuo, junto con marcos conceptuales de innovación abierta y triple hélice. Este enfoque permite combinar rigurosidad técnica, escalabilidad arquitectónica y formación avanzada de recursos humanos.

La experiencia acumulada demuestra la viabilidad de un esquema replicable que articula producción tecnológica, formación profesional y transferencia efectiva, fortaleciendo el rol de la universidad como actor estratégico en los sistemas regionales de innovación.

Palabras clave: innovación aplicada; transferencia tecnológica; desarrollo colaborativo; gestión ambiental; patrimonio digital; ingeniería de software; impacto social.

1. Introducción

En el escenario contemporáneo de transformación digital acelerada, caracterizado por la convergencia de tecnologías emergentes, automatización de procesos y creciente digitalización de servicios públicos y privados, las instituciones universitarias enfrentan el desafío de trascender el modelo tradicional de generación de conocimiento teórico para avanzar hacia esquemas integrados de producción científica aplicada. En este contexto, la articulación entre investigación, extensión y transferencia tecnológica se configura

como un eje estratégico para el desarrollo institucional y territorial.

La Usina de Ideas surge como respuesta estructurada a esta necesidad, constituyéndose en un entorno formalizado para la incubación, diseño, desarrollo y validación de proyectos tecnológicos con aplicabilidad concreta. Su concepción organizacional promueve la interacción sistemática entre docentes investigadores, estudiantes avanzados y actores externos —municipios, organizaciones culturales, cooperativas y empresas— favoreciendo la convergencia entre demandas sociales reales y capacidades académicas disponibles.

Desde una perspectiva epistemológica, el proyecto se inscribe en el paradigma de innovación abierta, donde el conocimiento no se genera ni se utiliza de manera aislada, sino que circula dinámicamente entre múltiples actores que aportan saberes complementarios. Asimismo, adopta el modelo de triple hélice universidad–empresa–Estado, el cual enfatiza la co-creación de soluciones con impacto territorial mediante la cooperación interinstitucional.

El presente trabajo describe la estructura conceptual, metodológica y operativa del proyecto, detalla los desarrollos en curso y analiza sus resultados preliminares, destacando su potencial de replicabilidad y escalabilidad en otros contextos institucionales. [1][2][3]

2. Marco Conceptual

2.1 Innovación y Sistemas de Conocimiento [4][5][6]

La literatura contemporánea en economía de la innovación sostiene que los sistemas nacionales y regionales de innovación dependen de la interacción efectiva entre actores institucionales, productivos y académicos. La transferencia tecnológica no constituye un fenómeno lineal donde la universidad “produce” conocimiento y el sector productivo lo “consume”, sino un proceso complejo de retroalimentación, adaptación y aprendizaje mutuo.

En este sentido, la universidad asume un rol ampliado que supera la docencia y la investigación básica, incorporando funciones de vinculación estratégica, desarrollo tecnológico y generación de soluciones aplicadas orientadas a la resolución de problemáticas específicas del territorio.

2.2 Innovación Abierta y Triple Hélice

El enfoque de innovación abierta propone que las organizaciones deben aprovechar flujos internos y externos de conocimiento para acelerar el desarrollo tecnológico y mejorar su capacidad de adaptación. Esto implica compartir información, colaborar con actores externos y generar ecosistemas de innovación distribuidos.

Por su parte, el modelo de triple hélice conceptualiza la innovación como resultado de la interacción dinámica entre universidad, sector productivo y sector público. Cada actor cumple un rol diferenciado pero complementario: la universidad aporta conocimiento científico y capacidades técnicas; el sector productivo contribuye con orientación al mercado y sostenibilidad económica; el Estado provee marcos regulatorios, financiamiento y legitimidad institucional.

La Usina de Ideas operacionaliza estos marcos teóricos mediante proyectos concretos que integran actores municipales, instituciones culturales y organizaciones intermedias, generando espacios de co-diseño y validación participativa.

2.3 Transferencia Tecnológica y Formación Profesional

La transferencia tecnológica se concibe no sólo como la cesión de resultados de investigación, sino como un proceso continuo de adaptación contextual,

implementación progresiva y mejora incremental basada en retroalimentación de usuarios reales. Implica acompañamiento técnico, documentación estructurada y generación de capacidades locales para garantizar sostenibilidad.

Simultáneamente, el modelo fortalece la formación de estudiantes, quienes participan activamente en todas las etapas del ciclo de vida del software, adquiriendo experiencia en entornos que simulan condiciones profesionales reales.

3. Metodología

El modelo operativo se estructura en fases iterativas que permiten combinar planificación estructurada con flexibilidad adaptativa:

1. Ideación y definición del problema mediante sesiones de brainstorming guiado y análisis de actores.
2. Análisis de factibilidad técnica, económica y social.
3. Diseño arquitectónico y modelado conceptual y lógico de datos.
4. Desarrollo incremental bajo metodologías ágiles.
5. Validación funcional con usuarios reales o simulados.
6. Documentación técnica y académica.

Se adoptan prácticas contemporáneas de ingeniería de software que garantizan calidad, trazabilidad y escalabilidad:

- Control de versiones distribuido: permite que múltiples desarrolladores trabajen simultáneamente sobre el mismo proyecto sin pérdida de información, registrando cada cambio realizado y posibilitando revertir versiones anteriores si es necesario.
- Integración continua (CI): consiste en la automatización de la compilación y pruebas del software cada vez que se incorpora nuevo código, reduciendo errores acumulativos.
- Despliegue continuo (CD): automatiza la publicación de nuevas versiones en entornos de prueba o producción.
- Pruebas automatizadas: incluyen pruebas unitarias (validan funciones individuales), pruebas de integración (verifican interacción entre módulos) y pruebas funcionales.
- Arquitectura basada en microservicios: cuando la escala lo requiere, el sistema se divide en servicios independientes que se comunican mediante APIs, facilitando mantenimiento y escalabilidad.
- Principios de arquitectura limpia: separación clara entre lógica de negocio, infraestructura y presentación, lo que mejora la mantenibilidad del sistema.

Este enfoque no sólo garantiza calidad técnica, sino que permite generar productos mínimos viables (MVP) con proyección de crecimiento estructural.

4. Desarrollos en Curso

4.1 Sistema de Gestión de Residuos Sólidos Urbanos No Habituales

El sistema implementa una plataforma web colaborativa orientada a optimizar la recolección de residuos voluminosos o no convencionales (electrodomésticos, escombros, muebles).

Funcionalidades principales

- Registro ciudadano con autenticación segura.
- Geolocalización automática mediante APIs cartográficas.
- Clasificación tipológica de residuos.
- Panel de planificación dinámica para operadores municipales.
- Sistema de notificaciones multicanal.
- Módulo de trazabilidad integral.

La planificación dinámica se basa en algoritmos de optimización de rutas, que permiten asignar recorridos eficientes reduciendo tiempos y costos logísticos.

Impacto esperado

- Disminución de microbasurales.
- Mejora en eficiencia logística.
- Integración con actores de economía circular.
- Generación de datos estadísticos para políticas públicas.

El MVP desarrollado permitió validar tiempos de respuesta, escalabilidad básica y estabilidad del sistema.

4.2 Plataforma de Gestión Patrimonial Museística

Se desarrolla como solución SaaS orientada a instituciones culturales con limitaciones presupuestarias y necesidad de digitalización progresiva.

Componentes principales

- Inventario estructurado con metadatos normalizados.
- Gestión de exposiciones.
- Repositorio multimedia.
- Gestión de artistas y colecciones.
- Almacenamiento de modelos 3D.
- Generación automática de fichas técnicas.

El uso de metadatos normalizados facilita interoperabilidad con otros sistemas y repositorios digitales, promoviendo preservación y difusión del patrimonio.

La arquitectura contempla escalabilidad horizontal, lo que permite aumentar capacidad agregando servidores adicionales sin modificar la lógica central del sistema.

4.3 Sistema de Gestión de Espacios de Trabajo Compartido

La solución integra herramientas de reserva, asignación inteligente de recursos y analítica de uso en tiempo real.

Características técnicas

- Motor de reglas configurable para asignación automática.
- Panel de métricas con indicadores de ocupación.
- Interfaz centrada en experiencia de usuario.
- API de integración externa.

El motor de reglas permite establecer criterios como prioridad institucional, horarios pico o necesidades específicas de equipamiento, optimizando utilización de infraestructura.

5. Resultados Preliminares

Durante el segundo cuatrimestre de 2025 se logró:

- Desarrollo de un MVP funcional del sistema de residuos.
- Validación técnica interna mediante pruebas estructuradas.
- Documentación formal con estándares académicos.
- Presentación de avances en eventos científicos y tecnológicos.
- Conformación de equipos interdisciplinarios estables.

El modelo demuestra capacidad de generar resultados concretos en ciclos académicos anuales, combinando productividad tecnológica y formación avanzada.

6. Formación de Recursos Humanos

El proyecto funciona como laboratorio de aprendizaje avanzado, integrando teoría y práctica.

Los estudiantes participan en:

- Relevamiento de requerimientos reales.
- Modelado de datos y diseño arquitectónico.
- Programación backend y frontend.
- Implementación de pruebas automatizadas.
- Documentación científica.
- Presentación en congresos.

Esta experiencia fortalece competencias técnicas (programación, arquitectura), metodológicas (gestión ágil, trabajo colaborativo) y comunicacionales (redacción científica y exposición oral).

7. Discusión

La experiencia confirma que la integración estructurada entre docencia, investigación y extensión no sólo es viable, sino estratégicamente necesaria en el contexto de transformación digital. El modelo permite reducir la brecha entre conocimiento académico y aplicación práctica.

No obstante, la sostenibilidad a largo plazo dependerá de:

- Institucionalización de financiamiento.
- Formalización de métricas de impacto.
- Consolidación de alianzas externas.
- Generación de marcos normativos internos.

La replicabilidad requiere sistematización metodológica y documentación de buenas prácticas.

8. Conclusiones

La Usina de Ideas constituye un modelo replicable de innovación aplicada universitaria que integra:

- Desarrollo tecnológico con impacto social.
- Transferencia efectiva.
- Formación avanzada de recursos humanos.
- Producción científica documentada.
- Articulación interinstitucional estratégica.

El desafío futuro radica en escalar las soluciones desarrolladas, consolidar redes interinstitucionales y fortalecer indicadores de impacto que permitan medir resultados en términos sociales, ambientales y económicos.

Referencias

[1] Freeman, C. (1995). "The 'National System of Innovation' in historical perspective." *Cambridge Journal of Economics, 19*(1), 5-24.

[2] Von Hippel, E. (2005). *Democratizing Innovation*. MIT Press.

[3] Chesbrough, H. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business Press.

[4] Etzkowitz, H., & Leydesdorff, L. (2000). "The dynamics of innovation: From National Systems and 'Mode 2' to a Triple Helix of university–industry–government relations." *Research Policy, 29*(2), 109-123.

[5] Perkmann, M., & Walsh, K. (2007). "University–industry relationships and open innovation: Towards a research agenda."*International Journal of Management Reviews, 9*(4), 259-280.

[6] Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). "Links and impacts: The influence of public research on industrial R&D." *Management Science, 48*(1), 1-23.

[8] United Nations. (2020). The UN E-Government Survey 2020: Digital Government in the Decade of Action for Sustainable Development. United Nations. Recuperado de https://www.un.org/development/desa/dpad/publication/un-e-government-survey-2020/

Infraestructura Blockchain Interoperable: Integración de Soluciones Web 3.0 y Web 2.0 mediante API REST Chain-Agnostic con Soporte de Privacidad

Alberto M. Palma; Javier Rosenstein; Carolina Canessa
Licenciatura en Sistemas de Información — Facultad de Informática y Diseño
Universidad Champagnat, Mendoza, Argentina
palmamarcelo@uch.edu.ar, rosensteinjavier@uch.edu.ar, canessacarolina@uch.edu.ar

RESUMEN

El presente trabajo se enmarca en el proyecto de investigación "Infraestructura Blockchain Interoperable: Integración de Soluciones Web 3.0 y Web 2.0" (2024–2026), cuyo objetivo es reducir la fricción tecnológica en la adopción de blockchain por organizaciones Web 2.0 mediante una arquitectura de abstracción chain-agnostic.

Se diseña e implementa ChainBridge API, una API REST interoperable que permite interactuar con múltiples blockchains EVM-compatible de manera transparente, desacoplando la lógica de negocio de la infraestructura subyacente.

Como aporte central, se formaliza el Dilema Privacidad-Interoperabilidad (DPI), que evidencia la tensión entre auditabilidad empresarial y anonimato del usuario en middleware blockchain.

Se propone una taxonomía de operaciones por sensibilidad (T0, T1, T2) y se implementa un Privacy Mode con batching transaccional y wallets dedicadas por aplicación. La validación se realiza mediante dos casos de uso contrastantes: CertChain (transparencia) y CriptoVendimia (anonimato). Los resultados preliminares demuestran la viabilidad técnica de una arquitectura interoperable con privacidad selectiva, aportando una contribución tanto aplicada como teórica en el diseño de infraestructuras Web 3.0 integradas con entornos Web 2.0.

Palabras clave: interoperabilidad blockchain; API REST; arquitectura chain-agnostic; privacidad; middleware Web 3.0; votación descentralizada; integración Web 2.0/Web 3.0.

CONTEXTO

El trabajo se desarrolla en el marco de la línea de investigación sobre infraestructura blockchain interoperable radicada en la Licenciatura en Sistemas de Información de la Universidad Champagnat. La creciente adopción de tecnologías Web 3.0 por parte de organizaciones tradicionales enfrenta un obstáculo estructural: la fragmentación del ecosistema blockchain y la complejidad técnica asociada a la gestión multi-cadena.

La interoperabilidad limitada entre redes genera silos de datos y dependencia de implementaciones específicas, mientras que la integración con aplicaciones Web 2.0 existentes exige arquitecturas que desacoplen la lógica de negocio de la infraestructura descentralizada. Esta problemática se agrava cuando los casos de uso presentan requerimientos heterogéneos, como transparencia total en sistemas de certificación académica o anonimato en sistemas de votación.

En este escenario, surge la necesidad de diseñar una capa intermedia que permita abstraer la complejidad técnica de múltiples blockchains, manteniendo compatibilidad con los principios de descentralización y, al mismo tiempo, incorporando mecanismos de privacidad adecuados al dominio de aplicación.

1. INTRODUCCIÓN

Para abordar esta problemática se diseñó e implementó ChainBridge API, una API REST interoperable y chain-agnostic que permite a aplicaciones Web 2.0 interactuar con múltiples blockchains EVM-compatible mediante una interfaz unificada. La arquitectura adopta un patrón Adapter que desacopla la lógica de

negocio de las particularidades de cada red, posibilitando la ejecución de operaciones multi-cadena mediante la simple parametrización del identificador de red.

Durante el proceso de diseño e integración se identificó una tensión estructural entre interoperabilidad y privacidad, particularmente al canalizar aplicaciones con requerimientos de anonimato a través de una API que, por diseño, incorpora mecanismos de logging y auditoría. Esta observación condujo a la formalización del Dilema Privacidad-Interoperabilidad (DPI), que constituye el principal aporte teórico del proyecto.

El presente trabajo describe el diseño arquitectónico de ChainBridge API, la formalización del DPI y la implementación de un modo de privacidad selectiva (Privacy Mode) que incorpora batching transaccional y aislamiento criptográfico mediante wallets dedicadas por aplicación. La validación conceptual y técnica se realiza a través de dos dominios contrastantes —certificación académica y votación descentralizada— permitiendo evaluar la robustez de la solución frente a requerimientos funcionales divergentes.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El proyecto articula cuatro líneas de investigación interrelacionadas:

2.1 Interoperabilidad blockchain mediante abstracción chain-agnostic

Se investiga la viabilidad de una capa de abstracción REST que unifique la interacción multi-cadena. La innovación central es el Contract Registry Dinámico, un módulo que mapea nombres lógicos de contratos (ej. "Votación2026") con sus direcciones físicas en cada red desplegada.

La API implementa una arquitectura hexagonal donde los adaptadores de cada blockchain (BaseEVMAdapter) inyectan las dependencias específicas de la red. Esto permite que una aplicación consumidora ejecute la misma operación de negocio cambiando únicamente el parámetro chainId en el header de la petición HTTP. Las pruebas preliminares sugieren que esta abstracción reduce significativamente el tiempo de desarrollo de clientes Web 2.0, al eliminar la necesidad de gestionar librerías complejas como Ethers.js o Web3.js en el frontend.

2.2 Privacidad en middleware blockchain (Dilema Privacidad-Interoperabilidad)

Al intentar canalizar el caso de uso de votación (CriptoVendimia) a través de ChainBridge API, se identificó que las propiedades de auditoría de la API (logging de parámetros de transacción, granularidad transaccional unitaria, wallet compartida) entran en conflicto con los requerimientos de anonimato. Se formalizó este conflicto como el Dilema Privacidad-Interoperabilidad (DPI), se identificaron tres puntos de fuga de privacidad específicos, y se propuso una taxonomía de operaciones blockchain por sensibilidad (T0-Transparente, T1-Pseudónima, T2-Anónima). El riesgo de re-identificación en datasets anonimizados, demostrado por Narayanan y Shmatikov [5], fundamenta la necesidad de esta clasificación.. Esta línea constituye la contribución teórica principal del proyecto.

2.3 Votación descentralizada con privacidad desde el diseño (Privacy by Design)

Se diseña CriptoVendimia como aplicación Web 2.0 pura (React + Firebase) que consume ChainBridge API. A diferencia de sistemas de votación electrónica tradicionales como Helios [6], que requieren que el votante realice operaciones criptográficas complejas en el navegador, nuestra propuesta delega la complejidad en la API sin comprometer el anonimato.

El sistema implementa un esquema de Identidad Digital Soberana Simplificada mediante un hash derivado de la autenticación Firebase. Para garantizar el secreto del voto, se desarrolló un mecanismo de batching con aleatorización Fisher-Yates [7] ejecutada en un entorno de ejecución controlado off-chain (Cloud Function serverless aislada). Esto previene el análisis de correlación temporal (timing attacks),

alcanzando un nivel de privacidad práctica suficiente para el dominio de aplicación, sin requerir la sobrecarga computacional de las pruebas de conocimiento cero (ZK-SNARKs) [8]. A diferencia del enfoque propuesto por Groth [9], que ofrece garantías criptográficas formales, nuestra solución prioriza la viabilidad de implementación en un contexto académico con recursos limitados.

Este enfoque se alinea estrictamente con los principios de Privacy by Design postulados por Cavoukian [10], específicamente el principio de "Privacidad incorporada en el diseño" (Privacy embedded into design), donde el anonimato no es una funcionalidad añadida, sino la arquitectura base del módulo de votación.

2.4 Resolución del DPI: Privacy Mode en APIs blockchain

Se propone y se implementa una extensión de ChainBridge API (v2.0 → v2.1) con tres componentes: (a) un endpoint de transacciones batch que agrupa N operaciones en una sola transacción on-chain, (b) un modo de privacidad (Privacy Mode) que controla selectivamente qué parámetros se almacenan en los logs de transacción, y (c) un sistema de wallets dedicadas por aplicación (Application Signers) que provee aislamiento criptográfico entre casos de uso. La extensión mantiene compatibilidad total hacia atrás.

3. RESULTADOS OBTENIDOS / ESPERADOS

Resultados obtenidos:

R1 — Diseño completo de ChainBridge API v2.1.

Se completó el diseño técnico de la API con especificación de 18 endpoints REST, 5 modelos de datos MongoDB (Chain, Contract, ApiKey, TransactionLog, AppSigner), patrón Adapter con BaseEVMAdapter y 3 adaptadores específicos, autenticación por API Key + JWT con tres niveles de permisos, y documentación OpenAPI auto-generada. La extensión v2.1 agrega el endpoint de batch con Privacy Mode y el sistema de Application Signers sin modificar la API existente.

R2 — Formalización del Dilema Privacidad-Interoperabilidad.

Se identificaron formalmente tres puntos de fuga de privacidad: (PF-1) almacenamiento de parámetros en logs de auditoría, (PF-2) correlación temporal por granularidad transaccional unitaria, y (PF-3) trazabilidad de *wallet* compartida. Se demostró que los requisitos de auditabilidad empresarial y privacidad del usuario son contradictorios bajo una política uniforme de logging.

Como aporte teórico principal, se definió una **Taxonomía de Operaciones Blockchain por Sensibilidad (TOBS)**, clasificando las transacciones en tres niveles:

- **Nivel T0 (Transparente):** Operaciones públicas donde la trazabilidad es deseable (ej. emisión de certificados en CertChain). Se audita el 100% de los datos.
- **Nivel T1 (Pseudónima):** Operaciones que requieren ocultar la identidad civil pero permiten trazabilidad de la wallet (ej. pagos DeFi). Se aplica ofuscación parcial de logs.
- **Nivel T2 (Anónima):** Operaciones críticas donde no debe existir vínculo entre el actor y la acción (ej. voto en CriptoVendimia). Se requiere *batching*, *mixing* de transacciones y *wallet* dedicada (Application Signer aislada).

R3 — Arquitectura de CriptoVendimia como caso de validación.

Se diseñó la arquitectura completa de un sistema de votación descentralizada que opera como app Web 2.0 pura consumidora de ChainBridge API. El diseño demuestra que la abstracción chain-agnostic es viable para dominios con requisitos heterogéneos (transparencia vs. anonimato) cuando la API incorpora mecanismos de privacidad selectiva.

Resultados esperados (en curso):

E1 — Implementación y deploy.

Desarrollo completo de ChainBridge API v2.1 desplegada en Railway con las tres cadenas operativas, y de CriptoVendimia como segundo

caso de uso funcional. Se espera completar la implementación durante el primer semestre de 2026.

E2 — Validación empírica de anonimato. Verificación exhaustiva de todos los puntos de almacenamiento (Firestore, MongoDB, blockchain pública) para confirmar que no existe cadena de correlación reconstructible entre identidad del votante y candidata elegida.

E3 — Métricas comparativas.

Medición y publicación de métricas de latencia, throughput y costo de gas de la API bajo dos tipos de carga: certificaciones (baja frecuencia, transparentes) vs. votación (alta frecuencia en ventanas cortas, anónimas), sobre las tres blockchains.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto involucra un equipo de 8 integrantes (1 director, 7 investigadores), todos vinculados a la Licenciatura en Sistemas de Información de la Universidad Champagnat. La formación de recursos humanos se articula en tres niveles:

Formación en investigación: Los integrantes del equipo participan en todas las fases del proyecto (diseño, implementación, validación, documentación), adquiriendo competencias en metodología de investigación aplicada, redacción académica y presentación de resultados en congresos.

Formación técnica especializada: El proyecto provee formación práctica en tecnologías emergentes: desarrollo de APIs REST con TypeScript/Fastify, smart contracts con Solidity/Hardhat, arquitecturas cloud-native con Railway/MongoDB Atlas, y criptografía aplicada (hashing, cifrado AES-256-GCM, firma ECDSA).

Vinculación con tesis de grado: Se han identificado al menos tres temas de tesis de grado derivados del proyecto: (a) la implementación del módulo de Privacy Mode en ChainBridge API, (b) el sistema de votación CriptoVendimia como caso de integración Web 2.0/Web 3.0, y (c) el análisis comparativo de rendimiento multi-chain.

5. BIBLIOGRAFÍA

[1] R. Belchior, A. Vasconcelos, S. Guerreiro, and M. Correia, "A Survey on Blockchain Interoperability: Past, Present, and Future Trends," ACM Computing Surveys, vol. 54, no. 8, pp. 1-41, 2021.

[2] X. Xu, I. Weber, and M. Staples, Architecture for Blockchain Applications. Cham, Switzerland: Springer, 2019.

[3] V. Buterin, "A Next-Generation Smart Contract and Decentralized Application Platform," Ethereum Whitepaper, 2014. [Online]. Available: https://ethereum.org/whitepaper/

[4] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, Design Patterns: Elements of Reusable Object-Oriented Software. Reading, MA: Addison-Wesley, 1994.

[5] A. Narayanan and V. Shmatikov, "Robust De-anonymization of Large Sparse Datasets," in Proc. IEEE Symposium on Security and Privacy, Oakland, CA, USA, 2008, pp. 111-125.

[6] B. Adida, "Helios: Web-based Open-Audit Voting," in Proc. 17th USENIX Security Symposium, San Jose, CA, USA, 2008, pp. 335-348.

[7] R. A. Fisher and F. Yates, Statistical Tables for Biological, Agricultural and Medical Research. London, UK: Oliver & Boyd, 1938.

[8] E. B. Sasson et al., "Zerocash: Decentralized Anonymous Payments from Bitcoin," in Proc. IEEE Symposium on Security and Privacy, San Jose, CA, USA, 2014, pp. 459-474.

[9] J. Groth, "Efficient Maximal Privacy in Board-Room Voting and Anonymous Broadcast," in Financial Cryptography and Data Security, FC 2004, Lecture Notes in Computer Science, vol. 3110. Springer, Berlin, Heidelberg, 2004, pp. 90-104.

[10] A. Cavoukian, "Privacy by Design: The 7 Foundational Principles," Information and Privacy Commissioner of Ontario, Canada, 2009.

INNOVACIÓN EN SISTEMAS DE SOFTWARE PARA AUTOMATIZACIÓN, EFICIENCIA ENERGÉTICA E INTEGRACIÓN DE ENERGÍAS RENOVABLES EN EL CENTRO DE CÓMPUTOS DE LA UTN LA PLATA

Martínez, Nicolás Agustín[1,2]; Maccarone, José[1]; Lasagna, Valeria[2]; Istvan, Romina[2]; Ramírez, José[2]

[1] Centro de Investigación de Codiseño Aplicado - Área Energía Eléctrica (CODAPLI)

[2] Laboratorio de Ingeniería en Sistemas de Información (LINES)

Facultad Regional La Plata, Universidad Tecnológica Nacional

Av. 60 s/n° esquina 124, CP 1900, La Plata, Buenos Aires, Argentina

martinez.nicolas@alu.frlp.utn.edu.ar; {macarone; valerial; ristvan}@frlp.utn.edu.ar; jramirez@alu.frlp.utn.edu.ar

RESUMEN

Se presenta una línea de Investigación y Desarrollo orientada al diseño e implementación de una plataforma software-hardware para la modernización del Centro de Cómputo de la Facultad Regional La Plata de la Universidad Tecnológica Nacional.

La propuesta integra monitoreo térmico distribuido mediante microcontroladores y sensores ambientales, supervisión de continuidad eléctrica y calidad de potencia mediante analizadores de redes, y desarrollo de mecanismos progresivos de automatización para climatización y respaldo energético.

El proyecto contempla una arquitectura modular de adquisición de datos, telemetría, analítica, alarmas y control, con énfasis en trazabilidad, confiabilidad operativa y seguridad. A partir de la generación sistemática de datos históricos, se desarrollarán modelos de eficiencia energética, caracterización de capacidad instalada y estudios prospectivos para la eventual incorporación de generación fotovoltaica.

La línea de I/D propone un enfoque interdisciplinario entre ingeniería eléctrica y sistemas de información, orientado a obtener resultados medibles en disponibilidad, estabilidad térmica y optimización energética, además de sentar bases replicables para otros centros de cómputo universitarios.

Palabras clave: Data Centers; Eficiencia energética; IoT industrial; Ciudades Inteligentes; Control energético.

CONTEXTO

La presente línea de Investigación y Desarrollo se integra al proyecto institucional de modernización tecnológica del Centro de Cómputo de la Universidad Tecnológica Nacional – Facultad Regional La Plata.

El proyecto es coordinado conjuntamente por el Centro de Investigación de Codiseño Aplicado – Área Energía Eléctrica (CODAPLI) y el Laboratorio de Ingeniería en Sistemas de Información (LINES), articulando docentes-investigadores y estudiantes de las carreras de Ingeniería en Energía Eléctrica e Ingeniería en Sistemas de Información. Esta colaboración promueve un enfoque interdisciplinario orientado a la automatización, el análisis de datos y la gestión energética de infraestructura crítica.

La iniciativa se desarrolla en el marco de las actividades institucionales de investigación y desarro-

llo de la Facultad, con el objetivo de fortalecer la infraestructura universitaria mediante soluciones tecnológicas innovadoras y replicables.

El proyecto es acreditado por la Universidad Tecnológica Nacional – Facultad Regional La Plata y se financia mediante recursos institucionales y dedicación docente-investigador, contemplándose la posibilidad de futuras vinculaciones con programas de promoción científico-tecnológica.

1. INTRODUCCIÓN

Los centros de cómputo constituyen infraestructuras críticas cuyo funcionamiento continuo resulta esencial para instituciones académicas y de investigación. La creciente digitalización de procesos administrativos, actividades de docencia virtual, almacenamiento de datos científicos y servicios en la nube ha incrementado de manera sostenida la demanda computacional y, consecuentemente, el consumo energético asociado a estas instalaciones.

Diversos estudios recientes destacan que la eficiencia energética en data centers se ha convertido en un eje central de investigación, tanto desde la perspectiva de métricas de evaluación como de estrategias de optimización térmica y eléctrica [4, 5]. En particular, la medición sistemática del desempeño energético y la caracterización del entorno térmico resultan fundamentales para comprender el comportamiento operativo y detectar oportunidades de mejora [1]. La literatura enfatiza que una gestión térmica inadecuada no solo incrementa el consumo energético, sino que también afecta la confiabilidad y la vida útil del equipamiento.

En este contexto, el monitoreo en tiempo real y el sensado distribuido han demostrado ser herramientas efectivas para la evaluación dinámica de eficiencia y control térmico [3]. Tecnologías basadas en redes inalámbricas de sensores, como LoRaWAN y arquitecturas IoT, han sido aplicadas con éxito para el monitoreo ambiental en data centers, permitiendo obtener mapas térmicos detallados y detectar anomalías operativas [2]. Estas aproximaciones habilitan la transición desde esquemas de supervisión reactiva hacia modelos de gestión basados en datos.

Asimismo, la literatura especializada señala la necesidad de integrar métricas energéticas, calidad de potencia y análisis de capacidad instalada en un enfoque sistémico que contemple tanto eficiencia operativa como planificación futura [4]. La disponibilidad de datos históricos confiables constituye un requisito clave para evaluar escenarios de expansión y potencial incorporación de generación renovable en infraestructuras tecnológicas.

No obstante, gran parte de las soluciones documentadas se orientan a grandes data centers corporativos o de hiperescala, mientras que los centros de cómputo universitarios de escala media presentan características particulares: crecimiento incremental, restricciones presupuestarias y limitada automatización integrada. Esta situación plantea la necesidad de desarrollar enfoques modulares, interdisciplinarios y replicables que combinen ingeniería eléctrica, sistemas de información y analítica de datos aplicada a infraestructura crítica.

En este marco, la presente línea de Investigación y Desarrollo se orienta a contribuir al estudio y modernización de centros de cómputo universitarios mediante la integración de monitoreo térmico distribuido, caracterización eléctrica y análisis energético basado en datos, promoviendo una gestión más eficiente, medible y sustentable.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto se estructurará en líneas de investigación y desarrollo articuladas entre sí, orientadas a diseñar, implementar y validar un sistema integral de observabilidad, automatización y optimización energética para el Centro de Cómputo. Cada línea combinará desarrollo tecnológico, generación sistemática de datos medibles y validación progresiva de resultados, asegurando trazabilidad, replicabilidad y fundamento técnico para futuras decisiones de inversión.

2.1 Observabilidad térmica y modelado ambiental

Se desarrollará una arquitectura de sensado distribuido destinada a caracterizar el comportamiento térmico del Centro de Cómputo. Para ello, se prevé la instalación estratégica de nodos basados en microcontroladores ESP32, ubicados en zonas frontal y posterior de racks, a distintas alturas, en proximidad a equipos de climatización y en un punto central de referencia.

Los nodos registrarán temperatura y humedad relativa con frecuencia configurable, generando una base histórica que permitirá:

- Construir un mapa térmico dinámico del recinto.
- Identificar gradientes verticales y acumulaciones de calor.
- Detectar ineficiencias en la distribución de aire frío.
- Analizar correlaciones entre carga computacional y comportamiento térmico.

En una primera etapa, el sistema operará exclusivamente en modo observacional, con el objetivo de establecer líneas base térmicas y validar la coherencia y estabilidad de las mediciones. Los datos obtenidos constituirán el insumo para las etapas posteriores de automatización y optimización.

2.2 Automatización térmica inteligente

Sobre la base de los modelos térmicos construidos, se desarrollará un sistema de control activo de climatización mediante emisión de señales infrarrojas hacia los equipos existentes, sin modificar su infraestructura interna.

El sistema implementará:

- Reglas de control con histéresis configurable.
- Validación cruzada entre sensores.
- Umbrales de seguridad térmica.
- Registro completo de cada acción ejecutada.
- Modos de operación manual, asistido y automático.

Inicialmente se prevé una etapa de control asistido, donde el sistema emitirá recomendaciones de ajuste. Posteriormente, y tras validar su desempeño en distintos escenarios, podrá evolucionar hacia automatización parcial bajo reglas predefinidas.

El impacto será evaluado mediante comparación cuantitativa de indicadores antes y después de la implementación.

2.3 Continuidad y calidad del suministro eléctrico

Se investigará la estabilidad y calidad del suministro eléctrico mediante un sistema de monitoreo seguro de tensión en el tablero general. Se registrarán:

- Cortes totales de red.
- Microcortes
- Variaciones fuera de tolerancia.
- Duración y frecuencia de eventos.

En relación con el grupo electrógeno, se diseñará una lógica de arranque automatizado escalonado con:

- Verificación efectiva de ausencia de red.
- Retardos configurables para evitar falsas activaciones.

- Enclavamientos eléctricos y lógicos.
- Registro detallado de intentos y estados.

Asimismo, mediante analizadores de redes se medirán potencia activa y reactiva por fase, factor de potencia, corriente RMS, distorsión armónica total (THD) y perfiles de carga horarios y semanales, generando métricas objetivas de calidad de servicio eléctrico.

2.4 Caracterización de capacidad y eficiencia energética

Se realizará un relevamiento integral del sistema eléctrico existente, incluyendo identificación de líneas dedicadas, capacidad nominal de protecciones, sección de conductores y distribución de cargas por fase.

A partir de las mediciones obtenidas se evaluará:

- La relación entre capacidad instalada y capacidad utilizada.
- Márgenes disponibles para expansión de infraestructura.
- Desbalances entre fases.
- Necesidad potencial de corrección del factor de potencia.
- Riesgos de sobrecarga futura.

El análisis histórico permitirá desarrollar modelos de consumo energético, curvas de demanda diaria y estacional, y estimaciones de potencia pico y media.

2.5 Modelado prospectivo y pre-factibilidad fotovoltaica

Utilizando los datos consolidados, se modelarán escenarios de incorporación futura de generación fotovoltaica. Se estimarán potencias requeridas, perfiles de demanda coincidente y potencial de autoconsumo. Esta línea no contemplará instalación física inmediata, sino la generación de evidencia técnica cuantitativa que fundamente decisiones estratégicas futuras y complete el análisis integral del sistema energético.

2.6 Plataforma de integración, auditoría y replicabilidad

Todas las mediciones y acciones proyectadas se integrarán en una plataforma centralizada de almacenamiento y visualización de datos, que garantizará:

- Auditoría completa de eventos.
- Trazabilidad de decisiones automatizadas.
- Exportación de datos para análisis académico.
- Generación automática de reportes técnicos.
- Construcción de indicadores comparativos antes/después.

Se priorizarán estructuras de datos normalizadas y metodologías documentadas que permitan la replicabilidad del modelo en otros centros de cómputo universitarios. En conjunto, estas líneas de investigación y desarrollo permitirán abordar de manera integral la estabilidad térmica, la continuidad eléctrica y la eficiencia energética del Centro de Cómputo, articulando observación sistemática, automatización progresiva y análisis prospectivo con fundamento cuantitativo.

3. RESULTADOS ESPERADOS

La presente línea de Investigación y Desarrollo se encuentra en etapa inicial de ejecución, por lo que no se reportan aún resultados empíricos. Se establecen, no obstante, metas verificables que orientarán su desarrollo durante el período de implementación.

En una primera fase se prevé la puesta en operación de la infraestructura de sensado térmico distribuido y monitoreo eléctrico, con el objetivo de obtener una línea base ambiental y energética del Centro de Cómputo. Esta caracterización incluirá indicadores tales como temperatura por zona,

gradientes térmicos, consumo por franja horaria, factor de potencia y registro de eventos eléctricos.

En una segunda etapa se proyecta la validación de mecanismos de control asistido y automatización parcial de climatización y respaldo eléctrico, con evaluación comparativa antes/después. Se espera mejorar la estabilidad térmica, reducir la variabilidad entre zonas, identificar desbalances eléctricos y construir métricas de eficiencia energética adaptadas al contexto institucional.

Asimismo, el histórico de datos permitirá modelar escenarios de expansión de infraestructura y analizar la pre-factibilidad técnica de incorporación futura de generación fotovoltaica. Desde el punto de vista académico, se prevé la formación en investigación aplicada, la generación de documentación técnica replicable y la producción de indicadores que respalden decisiones estratégicas institucionales.

4. FORMACIÓN DE RECURSOS HUMANO

La línea de Investigación y Desarrollo se desarrolla en el marco del trabajo conjunto entre el Centro de Investigación de Codiseño Aplicado – Área Energía Eléctrica (CODAPLI) y el Laboratorio de Ingeniería en Sistemas de Información (LINES), con la participación de docentes-investigadores y estudiantes avanzados de grado. El equipo se estructura de la siguiente manera:

- Directores e investigadores responsables, a cargo de la conducción científica y técnica del proyecto.
- Docentes-investigadores participantes, que intervienen en los ejes de automatización, monitoreo eléctrico y analítica de datos.
- Estudiantes avanzados de grado, quienes participan en tareas de diseño, implementación y validación de los desarrollos tecnológicos en el marco del proyecto.

Además constituye el marco académico de un Trabajo Final de Grado actualmente en desarrollo y se proyecta como base para futuras instancias de formación en investigación aplicada.

5. BIBLIOGRAFÍA

[1]. Z. G. M. Setyo, H. B. Rijal, N. Aqilah y N. Abdullah, "*Energy Efficiency Measurement Method and Thermal Environment in Data Centers — A Literature Review*", *Energies*, vol. 18, no. 14, 3689, 2025.

[2]. T. Polonelli, D. Brunelli, A. Bartolini y L. Benini, "*A LoRaWAN Wireless Sensor Network for Data Center Temperature Monitoring",* arXiv, preprint arXiv:1902.09400, 2019.

[3]. B. Lajevardi, K. R. Haapala y J. F. Junker, "*Real-time monitoring and evaluation of energy efficiency and thermal management of data centers"*, *Journal of Manufacturing Systems*, vol. 37, pp. 511--516, 2015.

[4] X. Shao, Z. Zhang, P. Song, Y. Feng y X. Wang, "*A review of energy efficiency evaluation metrics for data centers"*, *Energy and Buildings*, vol. 271, 112308, 2022.

[5] S. Long, Y. Li, J. Huang, Z. Li y Y. Li, "*A review of energy efficiency evaluation technologies in cloud data centers"*, *Energy and Buildings*, vol. 260, 111848, 2022.

Integración de asistencia inteligente a la terapia del habla

Iván Oscar Antonioli[1]; Alicia Mon[1]; Antonio Piccinno[2]
[1]Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903 - San Justo (CP 1754)
Buenos Aires, Argentina
[2] Università degli studi di Bari Aldo Moro
Piazza Umberto I - 70121
Bari , Italy
amon@unlam.edu.ar; iantonioli@alumno.unlam.edu.ar; antonio.piccinno@uniba.it

RESUMEN

El presente artículo expone los resultados parciales de un proyecto de investigación en curso sobre la integración de un Asistente Inteligente, específicamente Alexa smart assistant (Amazon, s.f.) a una aplicación web enfocada en la terapia del habla en chicos de entre 4 y 6 años.

La aplicación web desarrollada en la Universidad de Bari Aldo Moro permite que los logopedas, o terapistas del habla puedan desarrollar y monitorear ejercicios resueltos por sus pacientes, ayudándolos a tratar sus problemas lingüísticos. Para mejorar la experiencia del usuario y reducir sus esfuerzos cognitivos al interactuar con el sistema se busca agregar al sistema ya desarrollado una integración con un asistente inteligente del hogar.

Palabras clave: *Asistente Inteligente; Terapias del habla.*

1. CONTEXTO

El estudio que se presenta esta desarrollado en el marco del proyecto TNE-DeSK (Developing shared knowledge) (TNE DESK, s. f.). financiado por el Plan Nacional de Recuperación y Resiliencia (PNRR) (PNRR, 2021), con el apoyo de la Unión Europea en el marco del programa NextGenerationEU (Comisión Europea, s.f.). Este proyecto es liderado por la Universidad de Bari Aldo Moro, Italia en tanto que la Universidad Nacional de La Matanza (UNLaM) aporta su colaboración con la participación de estudiantes de Ingeniería Informática realizando una estancia de investigación a partir de enero de 2026 en la Universidad de Bari (TNE DESK, s. f.).

2. INTRODUCCIÓN

La terapia del habla es un campo de la medicina que estudia y trata discapacidades lingüísticas, específicamente aquellas que afectan las habilidades del paciente al pronunciar palabras, o al comunicarse con otras personas (Butcher, 1987).

Estos problemas lingüísticos suelen estar presentes durante la infancia, pudiendo dificultar el desarrollo de las habilidades sociales y comunicativas, así como la interacción con otras personas en las actividades diarias. Por esto, las estrategias de tratamientos clínicos recomiendan tratar los síntomas lo antes posible, en las etapas tempranas del desarrollo de una persona (Law, 2003).

El tratamiento de la terapia del habla aborda diferentes áreas, para problemas relacionados a la producción del sonido tales como distorsiones y sustituciones, se trabaja sobre la articulación. En el área de la fluidez, se tratan los problemas que impiden hablar de corrido, ya sea por repetición de palabras, pausas entre ellas, o la fusión de diferentes palabras (Nicklaus Children's Hospital, s.f.).

Durante el tratamiento también se abordan las dificultades relacionadas al lenguaje receptivo y expresivo, en el primero al

entender el lenguaje oral, mientras que el segundo al momento de expresar ideas, pensamientos o sentimientos de forma verbal o escrita. A su vez, el procedimiento aborda algunos aspectos de la voz y la motricidad oral, centrándose en problemas tales como el tono, volumen y calidad de la voz, mediante ejercicios que permiten identificar y tratar debilidades motoras (Nicklaus Children's Hospital, s.f.).

Además, el tratamiento aborda el área del juego, buscando desarrollar habilidades relacionadas con la resolución de conflictos y pensamiento simbólico, los cuales son indispensables para el desarrollo del lenguaje (Nicklaus Children's Hospital, s.f.).

En la actualidad, existen algunas aplicaciones destinadas a la terapia del habla como, por ejemplo, "Speech blubs" (Speech Blubs, s.f.), que busca ayudar a los infantes aplicando la técnica de imitación, la cual se basa en escuchar y repetir palabras. También existen aplicaciones destinadas a adultos, como es el caso de "Constant therapy health" (Constant Therapy Health, s.f.), que ayuda a quienes viven con afecciones neurológicas o se están recuperando de accidentes cerebrovasculares.

Para colaborar en los tratamientos de estos problemas lingüísticos en el marco del proyecto se desarrolló "e-SpeechT" (Barletta et al., 2023), que constituye una aplicación web que proporciona funcionalidades para 3 categorías de usuarios:

1. Logopedas: son usuarios conformados por el personal médico especializado en la temática que define los tratamientos personalizados para ayudar a los pacientes con dificultades en el habla.
2. Pacientes: son los usuarios destinatarios de la aplicación, constituidos por los infantes de entre 4 y 8 años que tienen alguna discapacidad lingüística.
3. Tutores: son los usuarios conformados por los padres o tutores de los niños que acompañan en sus tratamientos y hacen un seguimiento sobre las sesiones de terapia que se llevan a cabo (Barletta et al., 2023).

La aplicación se centra en el aprendizaje basado en juegos, los aspectos y estética de la interacción también siguen ese principio (Di Bitonto et al., 2014)

La aplicación web solicita para iniciar sesión que el paciente seleccione un terapista del habla y contraseña para poder acceder al sistema, la contraseña consiste en una imagen de dibujo animado a seleccionar entre los disponibles, esto lo debe hacer el paciente junto a su tutor y una vez seleccionado tendrá acceso al sistema (Barletta et al., 2023).

Hay 3 categorías de ejercicios dentro de la aplicación que se pueden llevar a cabo durante las sesiones de terapia:

Naming images: A la izquierda de la pantalla se mostrará una imagen y a la derecha dos íconos, un micrófono para que el paciente grabe un audio diciendo la palabra; y el segundo ícono, un altavoz, permite dar una pista en forma de audio sobre la imagen.

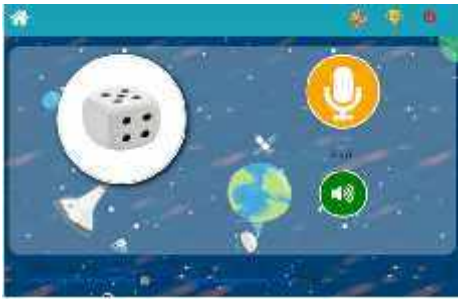


Gráfico 1: Ejercicio Naming Images. Tomado de Barletta et al. (2023).

Minimum pair recognition: Consiste en dos imágenes a los costados de la pantalla y un icono de altavoz en el centro, se deberá escuchar el audio y seleccionar la imagen correspondiente a la palabra enunciada.



Gráfico 2: Ejercicio Minimum pair recognition. Tomado de Barletta et al. (2023).

Repetition of Words: En la pantalla se verán dos íconos en el centro, un micrófono y un altavoz. Debajo habrá 3 palabras que se deberán pronunciar en el orden correcto. Para completar satisfactoriamente el ejercicio, el paciente podrá escucharla secuencia de palabras oprimiendo el icono del altavoz y luego presionando el icono del micrófono deberá grabar un audio diciendo la secuencia de palabras. A su vez presionando cada una de las palabras, aparecerá una foto asociada en forma de pista (Barletta et al., 2023).

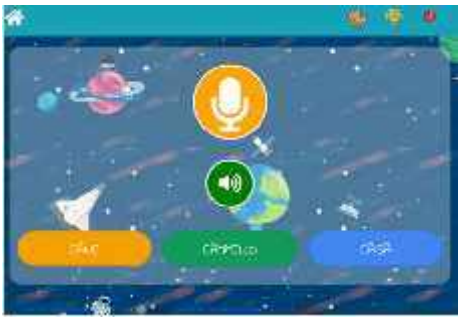


Gráfico 3: Ejercicio Repetition of Words. Tomado de Barletta et al. (2023).

Como se puede ver en las imágenes, las actividades están diseñadas siguiendo el principio de aprendizaje basado en juegos para captar la atención de los pacientes y mejorar su experiencia. También cabe destacar que los ejercicios tienen relación con las áreas de la fluidez y la voz, presentes en los tratamientos de terapia del habla.

3. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto en curso que se expone en el presente artículo se propone como objetico integrar la web ya desarrollada con una aplicación que funcione como un Asistente Inteligente de modo tal de mejorar la experiencia de los pacientes al interactuar con la aplicación y asistir a las sesiones de terapia. Los avances realizados a la actualidad proponen usar el sistema web integrándolo con Alexa (Amazon, s.f.). Esto permitirá generar recordatorios de sesiones de terapia próximas y cuando es hora de asistir a dicha sesión Alexa reproduciría un audio indicando que se debe asistir a una sala específica de la casa. A su vez consulta al paciente si ya está en la sala correspondiente para que diga una frase de invocación, permitiendo que la sesión comience (Calvano et al., 2024).

Como otro objetivo de aporte a la aplicación existente, se planea realizar una mejora del sistema agregando la geolocalización de un dispositivo previamente configurado para que se inicie automáticamente la sesión, solo si se encuentra dicho dispositivo a una distancia determinada y hay una sesión programada para ese horario (Calvano et al., 2024).

4. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos hasta el momento, utilizando la aplicación web y Alexa se basan en una prueba realizada con usuarios reales por parte de la Universidad de Bari Aldo Moro. En la prueba participaron 10 niños, 2 observadores y 1 facilitador, estos últimos tenían como objetivo guiar a los chicos durante la prueba y explicarles cómo se debía llevar a cabo. El rango etario de los niños que participaron fue de 4 a 8 años. Se notó una diferencia entre las habilidades de lectura/escritura de los chicos involucrados debido a la diferencia de edad y el consecuente desarrollo de dichas habilidades (Barletta et al., 2024).

Usuarios Facilitadores	Usuarios Observadores	Usuarios Paciente
1	2	10

Gráfico 4: Cantidad de usuarios involucrados. Fuente: elaboración propia.

Cabe destacar que a los niños más grandes les fue más fácil el uso de los dispositivos porque estaban familiarizados. Los más chicos tenían dificultades para enunciar la frase de invocación que iniciaba el sistema. Esto confirmó la necesidad de configurar la aplicación de distinta forma dependiendo de la edad de los pacientes involucrados (Barletta et al., 2024).

Para recopilar información acerca del estudio se desarrollaron dos tipos cuestionarios. El primero consistía en una serie de 18 preguntas enfocadas a la experiencia del tutor y el paciente, el responsable de completarla fue el observador simulando el rol de tutor. El segundo estaba compuesto por una única pregunta que se realizaba a los chicos para conocer cómo se sintieron durante la experiencia, el formato era una escala de satisfacción.



Gráfico 5: Distribución de preguntas relacionadas a la experiencia de cada usuario. Fuente: Elaboración propia



Gráfico 6: Escala de satisfacción. Tomado de Barletta et al. (2024).

En base a estos resultados se está desarrollando una integración con Alexa, como asistente inteligente que busca facilitar el entendimiento por parte del usuario paciente, adaptándose a sus expectativas y necesidades, las cuales estarán delimitadas por su edad. Para lograr esto se propone el desarrollo de una interfaz gráfica y de voz en Alexa que facilite el inicio de una sesión de terapia.

Cuando Alexa anuncie que es hora de una sesión programada, la idea es ofrecer distintas formas de iniciar dicha sesión, el usuario podrá hacerlo mediante comandos de voz, por medio de la interfaz gráfica, o si es posible se activará automáticamente al detectar que el usuario está cerca. La primera opción consistirá en un comando que el usuario debe decir para iniciar la aplicación web en el dispositivo seleccionado. En el segundo escenario, si el usuario no logra invocar a Alexa, se busca guiarlo mediante la interfaz gráfica para que seleccione su contraseña, un dibujo animado, y con esa interacción se ejecutaría la aplicación para asistir a la sesión. Finalmente, en el último escenario se busca que Alexa reconozca que el usuario se encuentra cerca, mediante la conexión con algún dispositivo compatible, y por consiguiente ejecutar la aplicación.

De esta forma, agregando distintas opciones de inicio al momento de tener una sesión de terapia, se podrá guiar a todos los usuarios, incluyendo a aquellos que hasta el momento les resultaba difícil utilizar la aplicación debido a su edad.

5. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación GIS, del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, colabora en el proyecto a partir de la formación de estudiantes avanzados de Ingeniería Informática.

Los proyectos precedentes del grupo de investigación han generado diversos resultados científicos y académicos publicados en diferentes congresos nacionales e internacionales y la interdisciplinariedad de este ha facilitado el desarrollo de diversos Proyectos Finales de Ingeniería, así como de Tesis de Doctorado.

En ese contexto, uno de los autores del trabajo se encuentra desarrollado una Estancia de Investigación en la Universidad de Bari, participando activamente en la consecución del presente proyecto.

6. BIBLIOGRAFÍA

- Amazon. (s.f.). Alexa. https://alexa.amazon.com/
- Butcher, P., Elias, A., Raven, R. and Yeatman, J., and Littlejohns, D.:Psychogenic voice disorder unresponsive to speech therapy: Psychological characteristics and cognitive-behaviour therapy in Children: A Systematic Review of Features and Challenges; In International Journal of language and speech disorders, 1987. https://doi.org/10.3109/13682828709088690
- Barletta, V., Calvano, M., Curci, A., and Piccinno, A. (2023). A protocol to assess usability and feasibility of e-speechT, a web-based system supporting speech therapies.
- Barletta, V., Calvano, M., Curci, A., and Piccinno, A. (s.f.). A new interactive paradigm for speech therapy.
- Barletta, V., Calvano, M., Curci, A., Pagano, A., and Piccinno, A. (2024). Evaluation of ”Speech System” and ”Skill”: An Interaction Paradigm for Speech Therapy.
- Calvano, M., Curci, A., Pagano, A., and Piccinno, A. (2024). Speech Therapy Supported by AI and Smart Assistants.
- Comisión Europea. (s.f.). NextGenerationEU. https://next-generation-eu.europa.eu/
- Constant Therapy Health. (s.f.). Constant Therapy. https://constanttherapyhealth.com/es/constant-therapy/
- Di Bitonto, P., Corriero, N., Pesare, E., Rossano, V., and Roselli, T. (2014). Training and learning in e-health using the gamification approach: The trainer interaction. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8515 LNCS(PART 3):228 – 237.
- Italia Domani. (2021, 30 de abril). Piano Nazionale di Ripresa e Resilienza. Presidenza del Consiglio dei Ministri. www.italiadomani.gov.it
- Laviano, Pier Luigi. (2020). *Re-design and evolution of a system supporting speech therapists* [Master’s degree Thesis in Computer Science]. University of Bari Aldo Moro.
- Law, J.: Speech and language therapy intervention for children with primary speech and language delay or disorder: Cochrane Developmental, Psychosocial and learning Problems Group. Cochrane Database Syst Rev. 2003;2003(3):CD004110
- Nicklaus Children’s Hospital. (s.f.). Terapia del habla. https://www.nicklauschildrens.org/tratamientos/terapia-del-habla
- Speech Blubs. (s.f.). Speech Blubs: Language Therapy. https://speechblubs.com/
- TNE DESK. (s.f.). Scopri. https://www.tne-desk.org/#scopri

Integración de la privacidad desde el diseño y los datos abiertos para fortalecer las plataformas digitales públicas

Roxana Martínez, Ramiro Macías, Agustín Aponte

Secretaría de Investigación y Transferencia Científica
Universidad Siglo 21 (S21)
Calle De los Latinos 8555, Córdoba, Argentina

roxana-martinez@mi.21.edu.ar; roxana.martinez@ues21.edu.ar
{jjrmac, agustin.aponte}@gmail.com

RESUMEN

La creciente digitalización de los servicios públicos implica un amplio procesamiento de datos ciudadanos a través de plataformas gubernamentales digitales, lo que plantea nuevos desafíos en términos de transparencia, calidad de los datos y protección de la privacidad. Este trabajo presenta una evolución de investigaciones previas sobre datos abiertos y transparencia gubernamental, incorporando una perspectiva de ingeniería orientada a la privacidad desde la etapa de diseño para el diseño y la evaluación de plataformas digitales públicas.

Proponemos integrar métricas de calidad de datos abiertos con mecanismos de protección de la privacidad y prácticas de seguridad observables en sistemas públicos. Al aplicar este marco a plataformas digitales públicas argentinas seleccionadas, identificamos brechas recurrentes y oportunidades de mejora. Los resultados esperados buscan contribuir al desarrollo de mejores prácticas para plataformas digitales públicas más transparentes, seguras y confiables, promoviendo tanto la reutilización de datos abiertos como la protección de la información sensible de los ciudadanos.

Una contribución distintiva de este trabajo reside en conectar las iniciativas de transparencia basadas en datos abiertos con prácticas de ingeniería orientadas a la privacidad, una integración que a menudo se aborda por separado en los esfuerzos de digitalización del sector público. Al combinar la evaluación de la calidad de los datos, los principios de privacidad desde el diseño y los mecanismos prácticos de evaluación, la propuesta proporciona un marco replicable que apoya tanto la transparencia como la gestión responsable de datos. Esta doble perspectiva permite que las plataformas digitales públicas evolucionen hacia modelos que no solo promueven la apertura y la rendición de cuentas, sino que también garantizan la protección de los datos de los ciudadanos, fomentando la confianza a largo plazo en los servicios gubernamentales digitales.

Palabras Claves: Open Data; Privacy-by-Design; Public Digital Platforms; Data Quality.

CONTEXTO

Este trabajo forma parte del proyecto "*Procesamiento de Datos Públicos para Promover la Transparencia en el Gobierno Abierto*", iniciado en diciembre de 2023. El proyecto pertenece a la línea de investigación Transformación Tecnológica, dentro del área de Ciencias Aplicadas de la Secretaría de Investigación y Transferencia Científica. Sus actividades están directamente relacionadas con los programas de grado de Ingeniería y Sistemas de la Universidad Siglo 21 (S21). El proyecto cuenta con el financiamiento y la evaluación de la Secretaría de Investigación y Transferencia Científica e incluye la participación de profesores y estudiantes de grado del programa de Informática.

Este trabajo se enmarca en una línea de investigación centrada en el análisis y la mejora de las prácticas de transparencia y gestión de datos en entornos digitales públicos. Investigaciones previas han examinado los desafíos relacionados con los datos abiertos gubernamentales, incluyendo cuestiones de

calidad, accesibilidad, interoperabilidad y estandarización de los datos en los sistemas de información pública. Estos estudios muestran que, si bien la apertura de datos ha avanzado en los últimos años, persisten importantes brechas en cuanto a la fiabilidad de los datos, su potencial de reutilización y las prácticas de gobernanza de la información.

Las investigaciones más recientes han incorporado enfoques tecnológicos emergentes y nuevos marcos de evaluación destinados a fortalecer los mecanismos de transparencia y rendición de cuentas en las plataformas digitales gubernamentales. Sin embargo, muchas de estas iniciativas se centran principalmente en la disponibilidad y apertura de la información, sin abordar adecuadamente la protección de los datos personales y sensibles procesados por los servicios digitales públicos.

En este contexto, el presente trabajo representa una continuación y evolución de estas investigaciones al integrar prácticas de ingeniería orientadas a la privacidad con enfoques de evaluación de la calidad de los datos abiertos. La propuesta busca contribuir al desarrollo de plataformas digitales públicas que promuevan simultáneamente la transparencia, la reutilización de datos y la gestión responsable de la información, impulsando así procesos sostenibles de transformación digital en el sector público.

1. INTRODUCCION

La transformación digital del sector público ha impulsado el desarrollo de plataformas digitales gubernamentales destinadas a mejorar la prestación de servicios y facilitar la interacción entre el Estado y la ciudadanía. Estas plataformas procesan cada vez más grandes volúmenes de datos personales y administrativos, lo que plantea nuevos retos en materia de transparencia, seguridad y protección de datos [1], [2]. Paralelamente, las políticas de Gobierno Abierto han promovido la publicación de datos públicos con el objetivo de fortalecer la rendición de cuentas, fomentar la participación ciudadana y fomentar la reutilización de la información para generar valor social y económico [3], [4].

Sin embargo, varios estudios recientes muestran que los datos abiertos no siempre se traducen en una mayor transparencia efectiva, debido a problemas relacionados con la calidad de los datos [5], su actualización, interoperabilidad y facilidad de reutilización [6], [7]. En muchos casos, los portales de datos abiertos presentan inconsistencias estructurales, conjuntos de datos incompletos o documentación insuficiente, lo que limita su uso por parte de ciudadanos, investigadores y organismos de control [8]. Además, la estandarización insuficiente en la publicación de datos dificulta la comparación y el análisis a nivel regional o nacional [9].

Estas limitaciones se ven agravadas por un desafío creciente: la necesidad de proteger la privacidad y los datos personales en las plataformas digitales públicas. La expansión de los servicios digitales gubernamentales ha incrementado la exposición de información sensible, generando riesgos asociados con violaciones de datos, configuraciones de seguridad inadecuadas y prácticas insuficientes de minimización de datos [10], [11]. Estudios recientes destacan que muchas plataformas públicas priorizan la disponibilidad de la información y la digitalización de trámites sin incorporar adecuadamente los principios de protección de datos desde las primeras etapas del diseño del sistema [12].

En este contexto, el enfoque de privacidad por diseño ha cobrado relevancia como estrategia para integrar los requisitos de privacidad y seguridad en el proceso de ingeniería de software, promoviendo la protección de datos desde la concepción misma de los sistemas digitales [13]. Sin embargo, la literatura reciente indica que la implementación práctica de este enfoque en plataformas gubernamentales aún presenta desafíos, especialmente en contextos donde las políticas de datos abiertos coexisten con la necesidad de salvaguardar información sensible [14].

Además, investigaciones recientes en gobierno digital muestran que la confianza pública en los servicios digitales depende tanto de la transparencia como de la percepción de la

protección de datos personales [15][16]. En consecuencia, es necesario avanzar hacia modelos que integren simultáneamente mecanismos de apertura de información pública y prácticas técnicas para salvaguardar la privacidad, evitando que ambas dimensiones se desarrollen de forma aislada. Ante este escenario, este artículo propone una integración entre las métricas de calidad de datos abiertos y las prácticas de ingeniería orientadas a la privacidad, permitiendo la evaluación de plataformas digitales públicas desde una doble perspectiva: transparencia y protección de datos. Este enfoque busca contribuir al desarrollo de plataformas gubernamentales más confiables, seguras y reutilizables, apoyando procesos sostenibles de transformación digital en el sector público.

Por tanto, esta investigación sugiere que la evolución de las políticas de datos abiertos requiere avanzar hacia modelos que integren simultáneamente apertura, calidad de los datos y protección de la privacidad, permitiendo a las plataformas digitales públicas promover la transparencia sin comprometer la seguridad de la información de los ciudadanos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto se enmarca en el tema de investigación Transformación Tecnológica, dentro del área de Ciencias Aplicadas de la Secretaría de Investigación y Transferencia Científica. En el marco de las actividades de I+D+i (Investigación, Desarrollo e Innovación), el trabajo se centra en el estudio y la mejora de los procesos de gestión, calidad y protección de datos en plataformas digitales públicas, en línea con las iniciativas de Gobierno Abierto y la transformación digital del sector público.

Las principales áreas de investigación del proyecto son: Evaluar la calidad de los conjuntos de datos públicos disponibles, considerando criterios de completitud, consistencia, actualización y potencial de reutilización; Analizar los riesgos asociados a la exposición de datos personales o sensibles en plataformas digitales públicas; Integrar prácticas de ingeniería orientadas a la privacidad en el desarrollo y la evaluación de servicios digitales gubernamentales; Proponer directrices técnicas y metodológicas que equilibren la apertura de datos públicos con la protección de la información personal; Diseñar recursos y experiencias de formación que permitan transferir los resultados del proyecto a estudiantes de ingeniería y sistemas.

Problemática:

En el contexto actual de transformación digital, los gobiernos están aumentando progresivamente la disponibilidad de servicios y datos en plataformas digitales públicas. Sin embargo, persisten desafíos relacionados con la calidad, la interoperabilidad y la reutilización de los datos abiertos, lo que limita su impacto en términos de transparencia y control ciudadano. Simultáneamente, la digitalización de los servicios públicos implica el procesamiento de grandes volúmenes de datos personales, lo que genera riesgos relacionados con la seguridad de la información y la protección de la privacidad. En muchos casos, los sistemas públicos priorizan la disponibilidad de la información sin incorporar mecanismos adecuados de protección de datos desde la etapa de diseño de la plataforma.

Esta situación genera tensión entre los objetivos de apertura y transparencia, por un lado, y la necesidad de salvaguardar la información personal y sensible, por otro. La falta de estándares uniformes para la publicación y gestión de datos también dificulta la interoperabilidad entre agencias y la reutilización efectiva de la información. Por consiguiente, es necesario desarrollar enfoques que permitan la evaluación y mejora simultáneas de la calidad de los datos abiertos y la protección de la privacidad en las plataformas digitales públicas, promoviendo ecosistemas de datos más fiables y sostenibles.

Pregunta de investigación del proyecto:

¿Cómo puede la integración de prácticas de evaluación de la calidad de los datos abiertos y

enfoques de ingeniería orientados a la privacidad mejorar la transparencia y la gestión responsable de la información en las plataformas digitales gubernamentales?

3. RESULTADOS OBTENIDOS Y ESPERADOS

Enfoque del Proyecto

El proyecto busca evaluar y mejorar la transparencia y la gestión responsable de los datos públicos en las plataformas digitales gubernamentales mediante la aplicación de enfoques técnicos y metodológicos que equilibren los datos abiertos con la protección de la privacidad.

Como parte de los resultados esperados, el proyecto busca fortalecer las políticas de datos abiertos mediante la evaluación de la calidad, la disponibilidad y la reutilización de la información publicada, a la vez que promueve mecanismos que reduzcan la exposición de datos personales o sensibles en plataformas públicas. Este enfoque está estrechamente vinculado con áreas como la ingeniería de software, la arquitectura de software y la gobernanza de datos

Objetivo principal:

Promover las mejores prácticas para la publicación, gestión y reutilización de datos públicos, fomentando simultáneamente la transparencia gubernamental y la protección de la información personal en plataformas digitales públicas.

Objetivos específicos esperados

Identificar los riesgos asociados a la exposición de datos personales en plataformas digitales públicas.

Integrar prácticas de ingeniería orientadas a la privacidad en la evaluación de los servicios digitales gubernamentales.

Proponer directrices técnicas y metodológicas para equilibrar la apertura de datos con la protección de la información sensible.

Desarrollar prototipos o herramientas de apoyo para evaluar la calidad de los datos y los riesgos de exposición de la información.

Diseñar recursos de capacitación y transferencia tecnológica para estudiantes y profesores participantes en el proyecto.

Metodología y Técnicas

Para lograr estos resultados, se adoptó una metodología mixta que combina técnicas cualitativas y cuantitativas. En primer lugar, se realizó una revisión bibliográfica de la literatura científica, los marcos regulatorios y los portales gubernamentales de datos abiertos, identificando los criterios y métricas utilizados para evaluar la calidad y la reutilización de la información pública. Posteriormente, se realizó un análisis comparativo de las plataformas digitales gubernamentales, evaluando la disponibilidad de datos, las actualizaciones, la documentación, los formatos de publicación y el potencial de reutilización, así como los posibles riesgos asociados a la exposición de información personal o sensible.

Además, se analizaron estudios de caso sobre la implementación de prácticas de protección de datos y privacidad en servicios digitales públicos para identificar enfoques aplicables al contexto regional. Como parte del trabajo, se planea desarrollar herramientas o prototipos conceptuales para evaluar la calidad de los datos y los riesgos de exposición de la información en plataformas públicas.

Resultados Obtenidos

Como parte de los avances en esta línea de investigación, se han publicado publicaciones en congresos científicos que analizan la evolución de las iniciativas de transparencia en la región.

Entre los resultados obtenidos se encuentran estudios centrados en el análisis y la evaluación de herramientas digitales gubernamentales diseñadas para promover la transparencia y el acceso a la información pública. Estos estudios identifican avances en la disponibilidad de servicios digitales, así como limitaciones relacionadas con la escalabilidad, la calidad de los datos y la facilidad de reutilización de la información. Asimismo, estudios recientes realizados por el equipo han permitido diagnosticar la disponibilidad y la calidad de los datos abiertos en diversas agencias gubernamentales. Estos estudios han identificado fortalezas en las políticas de acceso a la información pública, pero también dificultades relacionadas con la

falta de estandarización y homogeneidad en la publicación de datos, lo que limita su reutilización y comparabilidad a nivel nacional.

Estos resultados constituyen la base de la presente investigación, que busca avanzar hacia un modelo que integre simultáneamente la transparencia, la calidad de los datos y la protección de la privacidad en las plataformas digitales públicas.

4. RECURSOS HUMANOS

Este proyecto está integrado por dos profesores de grado con varios años de experiencia docente en diversas universidades de Argentina y más de 20 años de experiencia en el campo de las Tecnologías de la Información. Además, cuentan con títulos de posgrado: uno es profesora investigadora, doctora en Ciencias de la Computación por la Universidad Nacional de La Plata (UNLP), actual directora del Proyecto de Investigación, y tiene una maestría en Tecnologías de la Información por la Universidad Abierta Interamericana (UAI). El proyecto también incluye a otro profesor investigador, con una maestría en Tecnología Educativa por la Universidad Abierta Interamericana (UAI). El equipo también incluye estudiantes de grado y graduados de las carreras de Ingeniería y Sistemas de la Universidad Siglo 21, específicamente de la carrera de Informática.

5. REFERENCES

[1] Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government information quarterly*, 37(3), 101493.

[2] Mu, R., & Wang, H. (2022). A systematic literature review of open innovation in the public sector: Comparing barriers and governance strategies of digital and non-digital open innovation. *Public Management Review*, 24(4), 489-511.

[3] Barati, M. (2023). Open government data programs and information privacy concerns: a literature review. *arXiv preprint arXiv:2312.10096.*

[4] Kempeneer, S., & Wolswinkel, J. (2023). Rethinking open government data for citizen participation. An introduction to a special issue. *Information Polity*, 28(2), 163-173.

[5] Martínez, M. R. (2022). *Quality metrics for validating governmental public open datasets* (Doctoral dissertation, Universidad Nacional de La Plata).

[6] Quarati, A. (2023). Open government data: usage trends and metadata quality. *Journal of Information Science*, 49(4), 887-910.

[7] Kumar, V., Chandrappa, & Harinarayana, N. S. (2025). Exploring dimensions of metadata quality assessment: a scoping review. *Journal of Librarianship and Information Science*, 57(3), 661-673.

[8] Talukder, K. A. (2025). The impact of data-driven decision support systems on governance and policy implementation in US institutions. *ASRC Procedia: Global Perspectives in Science and Scholarship*, 1(01), 994-1030.

[9] Soylu, A., Corcho, Ó., Elvesæter, B., Badenes-Olmedo, C., Yedro-Martínez, F., Kovacic, M., ... & Roman, D. (2022). Data quality barriers for transparency in public procurement. *Information*, 13(2), 99.

[10] Andrade, V. C., Gomes, R. D., Reinehr, S., Freitas, C. O. D. A., & Malucelli, A. (2022, November). Privacy by design and software engineering: A systematic literature review. In *Proceedings of the XXI Brazilian Symposium on Software Quality* (pp. 1-10).

[11] Tunca, M. (2024, December). Privacy by Design: A Systematic Literature Review of European and British Regulatory Perspectives for Software and Information Engineering. In *Proceedings of the 2024 13th International Conference on Software and Information Engineering* (pp. 54-63).

[12] Iwaya, L. H., Babar, M. A., & Rashid, A. (2023). Privacy engineering in the wild: Understanding the practitioners' mindset, organizational aspects, and current practices. *IEEE Transactions on Software Engineering*, 49(9), 4324-4348.

[13] Del-Real, C., De Busser, E., & van den Berg, B. (2025). A systematic literature review of security and privacy by design principles, norms, and strategies for digital technologies. *International Review of Law, Computers & Technology,* 39(3), 374-405.

[14] Martínez, R. (2025). Digital Governance and Personal Data Protection: Proposals for an Ethical Public Transformation through the Privacy-by-Design Approach. *Salud, Ciencia y Tecnología-Serie de Conferencias*, 4, 3.

[15] Ernest, E., Osakwe, J., & Nhamu, I. (2023, November). A Framework for Data Protection and Privacy in Data Service Centres in the Public Sector: A Case Study of a Security Organisation in a Developing Countries. In *Proceedings of the International Conference on Information Systems and Emerging Technologies* (ICISET).

[16] Marwenny, E., Syafwar, R., & Yani, P. (2024). Personal Data Protection in Public Services. *Jurnal Ilmiah Ekotrans & Erudisi*, 4(2), 74-85.

MIGEM: MÓDULO INTERACTIVO PARA LA GESTIÓN DE EMERGENCIAS MÉDICAS

Barbaro, Cristian Rodrigo[1]; Iglesias, Agustín[2]; Boyé, Mauricio[3]; Giménez, Fernando Marcos[4]; Guaresti, Germán[5], Britos, Paola[6]

[1]crbarbaro@unrn.edu.ar Hospital Zonal Bariloche / CePIIeS - Universidad Nacional de Río Negro; [2]aiglesia@unrn.edu.ar Hospital Zonal Bariloche / CePIIeS - Universidad Nacional de Río Negro; [3]mauriboye@gmail.com Ingeniería en Computación - Sede Andina - Universidad Nacional de Rio Negro; [4]gimenezfernandomarcos@gmail.com Ingeniería en Computación - Sede Andina - Universidad Nacional de Rio Negro; [5]ggguaresti@unrn.edu.ar CePIIeS - Universidad Nacional de Río Negro; [6]pbritos@unrn.edu.ar , Ingeniería en Computación - Sede Andina - Universidad Nacional de Rio Negro.

RESUMEN

MIGEM es un sistema interactivo de asistencia clínica diseñado para mejorar la toma de decisiones en emergencias médicas mediante sensores biomédicos, algoritmos expertos e inteligencia artificial. El proyecto surge ante la necesidad de estandarizar intervenciones críticas en contextos con escasez de especialistas, alta carga asistencial y variabilidad en la calidad de atención. MIGEM integra monitoreo fisiológico, alarmas inteligentes, soporte farmacológico y guías de acción en tiempo real, funcionando tanto en entornos hospitalarios como prehospitalarios. Actualmente se encuentra en TRL 3, con una aplicación funcional y algoritmos validados mediante datos simulados. La línea de I/D se orienta a avanzar hacia la integración de hardware biomédico real, validación en laboratorio y posterior implementación en zonas rurales de Río Negro. El proyecto combina innovación tecnológica, impacto sanitario y potencial de transferencia a sistemas de salud nacionales e internacionales.

Palabras clave: emergencias médicas, inteligencia artificial, sensores biomédicos, soporte clínico, salud digital.

CONTEXTO

La línea de investigación se inserta en el desarrollo del proyecto MIGEM, coordinado por un equipo interdisciplinario de la Universidad Nacional de Río Negro, el Hospital Zonal Bariloche y el Centro Patagónico Interdisciplinario de Investigaciones en Salud (CePIIeS). El proyecto se orienta a resolver problemas críticos de la atención de emergencias en zonas urbanas y rurales, especialmente en regiones con baja disponibilidad de especialistas.

1. INTRODUCCIÓN

La medicina de emergencias enfrenta un escenario crítico a nivel nacional e internacional. Los equipos médicos de urgencias trabajan bajo **alta carga asistencial**, **presión emocional constante**, **fatiga del personal**, **escasez de especialistas** y **variabilidad en la calidad de las intervenciones**. Estos factores incrementan el riesgo de errores clínicos y comprometen la seguridad del paciente, especialmente en situaciones de alta complejidad como paro cardiorrespiratorio, shock o trauma.

En Argentina, esta problemática se profundiza en **zonas rurales y remotas**, donde la disponibilidad de profesionales capacitados es limitada. La provincia de Río Negro, por ejemplo, presenta amplias áreas con baja densidad poblacional y recursos sanitarios insuficientes. Aunque existen sistemas provinciales de emergencias (como SIARME), su cobertura es desigual y no logra garantizar tiempos de respuesta adecuados ni soporte especializado continuo.

A nivel global, la pandemia de COVID-19 evidenció la fragilidad de los sistemas de salud y aceleró la adopción de tecnologías de soporte clínico, simulación y automatización. Sin embargo, las soluciones disponibles actualmente son **fragmentadas**: monitores que solo registran datos, aplicaciones móviles sin lógica clínica, dispositivos de alerta sin integración biomédica o sistemas de triaje sin soporte terapéutico.

En este contexto, existe una **brecha tecnológica y asistencial**: no hay herramientas que integren monitoreo, análisis inteligente y asistencia clínica automatizada en tiempo real para apoyar decisiones críticas.

La falta de herramientas de soporte clínico en emergencias genera consecuencias directas:

- Aumento en la frecuencia de error humano en escenarios de alta presión.
- Desigualdad en el acceso a atención de calidad entre zonas urbanas y rurales.
- Sobrecarga del personal, que debe tomar decisiones complejas en segundos.
- Variabilidad en la calidad de las maniobras, especialmente en equipos sin especialistas.
- Limitaciones en la formación continua del personal de salud.

Resolver este problema tiene impacto en múltiples niveles:

- Impacto sanitario: Mejora la seguridad del paciente, reduce complicaciones y mortalidad en emergencias tiempo-dependientes, estandariza la atención incluso en contextos con poca experiencia clínica.
- Impacto social: Aumenta la equidad en el acceso a atención de calidad, fortalece la capacidad de respuesta en zonas vulnerables.
- Impacto educativo: Permite entrenar personal de salud mediante simulación y retroalimentación inteligente.
- Impacto económico: Reduce costos asociados a errores, traslados innecesarios y complicaciones, impulsa el desarrollo de tecnología médica nacional.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el proyecto se enmarcan los siguientes ejes: a) Integración de sensores biomédicos, b) Algoritmos clínicos y soporte de decisión, c) Inteligencia artificial y capacidades predictivas, d) Interfaz interactiva y experiencia de usuario y e) Simulación clínica y formación médica.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se ha desarrollado una aplicación funcional con algoritmos clínicos validados mediante datos simulados, lo cual permite demostrar la viabilidad del concepto y su comportamiento ante distintos escenarios médicos críticos. El

equipo ha definido la arquitectura del sistema, seleccionado los sensores y plataformas embebidas, y cuenta con una integración parcial del software clínico orientado a emergencias.

La próxima fase del proyecto es la integración de sensores (ECG, SpO_2, etc.) con el sistema embebido y la validación del prototipo en entorno de laboratorio con señales fisiológicas reales, validación en laboratorio, ensayos simulados, implementación piloto en zonas rurales y finalmente, expansión institucional.

4. FORMACIÓN DE RECURSOS HUMANOS

Equipo compuesto por 2 estudiantes avanzados de Ingeniería en Computación, 2 médicos especialistas en emergencias médicas y los directores de la carrera de Medicina y de Ingeniería en Computación. Cabe destacar que el director de la carrera de Medicina está especializado en gestión sanitaria en educación médica.

5. BIBLIOGRAFÍA

Bates, D. W., & Gawande, A. A. (2003). Improving safety with information technology. *New England Journal of Medicine, 348*(25), 2526–2534. https://doi.org/10.1056/NEJMsa020847

Bender, D., & Sartipi, K. (2013). HL7 FHIR: An agile and RESTful approach to healthcare information exchange. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, 326–331. https://doi.org/10.1109/CBMS.2013.6627810

Carayon, P., Schoofs Hundt, A., Karsh, B.-T., Gurses, A. P., Alvarado, C. J., Smith, M., & Brennan, P. F. (2006). Work system design for patient safety: The SEIPS model. *Quality and Safety in Health Care, 15*(Suppl 1), i50–i58. https://doi.org/10.1136/qshc.2005.015842

Desautels, T., Calvert, J., Hoffman, J., Jay, M., Kerem, Y., Shieh, L., Shimabukuro, D., Chettipally, U., Feldman, M. D., Barton, C., Wales, D. J., & Das, R. (2016). Prediction of sepsis in the intensive care unit with minimal electronic health record data: A machine learning approach. *JMIR Medical Informatics, 4*(3), e28. https://doi.org/10.2196/medinform.5909

Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., Sam, J., & Haynes, R. B. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *JAMA, 293*(10), 1223–1238. https://doi.org/10.1001/jama.293.10.1223

Henry, K. E., Hager, D. N., Pronovost, P. J., & Saria, S. (2015). A targeted real-time early warning score (TREWScore) for septic shock. *Science Translational Medicine, 7*(299), 299ra122. https://doi.org/10.1126/scitranslmed.aab3719

International Electrotechnical Commission. (2020). *IEC 62366-1:2015+AMD1:2020 Medical devices—Application of usability engineering to medical devices*. IEC.

International Medical Device Regulators Forum. (2017). *Software as a Medical Device (SaMD): Clinical evaluation (IMDRF/SaMD WG/N41FINAL:2017)*. https://www.imdrf.org

International Organization for Standardization. (2019). *ISO 14971:2019 Medical devices—Application of risk management to medical devices*. ISO.

Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C., & Faisal, A. A. (2018). The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine, 24*(11), 1716–1720. https://doi.org/10.1038/s41591-018-0213-5

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Royal College of Physicians. (2017). *National Early Warning Score (NEWS2): Standardising the assessment of acute-illness severity in the NHS*. RCP.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206–215. https://doi.org/10.1038/s42256-019-0048-x

Shillan, D., Sterne, J. A. C., Champneys, A., & Gibbison, B. (2019). Use of machine learning to analyse routinely collected intensive care unit data: A systematic review. *Critical Care, 23*, 284. https://doi.org/10.1186/s13054-019-2564-9

Smith, G. B., Prytherch, D. R., Meredith, P., Schmidt, P., & Featherstone, P. I. (2013). The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care admission, and death. *Resuscitation, 84*(4), 465–470. https://doi.org/10.1016/j.resuscitation.2012.12.016

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digital Medicine, 3*, 17. https://doi.org/10.1038/s41746-020-0221-y

U.S. Food and Drug Administration. (2023). *Software as a Medical Device (SaMD)*. https://www.fda.gov

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. WHO.

Propuesta Arquitectónica para Interoperabilidad en Salud mediante SSI, IPS-FHIR y ZK-Rollups

Carlos Martinez, Mariana Brachetta, Julio Monetti, Alberto Cortez

Laboratorio de Integración de Tecnologías Aplicadas a Prototipos de Software (LITAPS), Ingeniería en Sistemas de Información, UTN Facultad Regional Mendoza, Mendoza, Argentina

{carlos.martinez, mariana.brachetta, julio.monetti, alberto.cortez}@ docentes.frm.utn.edu.ar

RESUMEN

La interoperabilidad en salud digital, basada en Health Level Seven (HL7) Fast Healthcare Interoperability Resources (FHIR), está limitada por modelos de identidad centralizados que fragmentan la información. Esta fragmentación obstaculiza el intercambio del Resumen Internacional del Paciente (IPS), incrementando la dependencia de intermediarios y el riesgo de inconsistencias en escenarios regionales e internacionales.

La investigación explora una arquitectura de Identidad Auto-Soberana (SSI) para que el paciente comparta de forma segura y verificable su IPS o recursos HL7 FHIR específicos. Mediante consentimiento granular y mecanismos criptográficos, se garantiza la autenticidad e integridad de la información clínica entre distintas jurisdicciones.

La solución integra estándares W3C, como DIDs y Verifiable Credentials (VCs), con perfiles IHE (PDQm, MHD y SVCM). El marco de confianza se ancla en Ethereum (Capa 1) y utiliza zkSync (Capa 2, ZK-rollup) para ejecutar operaciones de identidad —emisión, verificación y revocación— con escalabilidad y eficiencia de costos.

El objetivo principal es validar una arquitectura de interoperabilidad que posibilite una gestión segura de la identidad del paciente, preserve la privacidad y la soberanía de los datos mediante un modelo SSI basado en credenciales verificables, y promueva la eficiencia de costos y la escalabilidad requeridas para su implementación en escenarios interjurisdiccionales.

Palabras clave: Interoperabilidad en Salud, Identidad Auto-Soberana (SSI), HL7 FHIR, Resumen Internacional del Paciente (IPS), Verifiable Credentials (VCs), Ethereum, zkSync, ZK-Rollups.

CONTEXTO

El presente trabajo se desarrolla en el Laboratorio de Integración de Tecnologías Aplicadas a Prototipos de Software (LITAPS), en el Departamento de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional – Facultad Regional Mendoza (UTN-FRM). Se enmarca en la línea del laboratorio orientada al diseño y validación de prototipos de software para problemáticas de alto impacto en el dominio sanitario, con foco en interoperabilidad, arquitectura de sistemas y tecnologías emergentes. En particular, este artículo sintetiza avances de una línea de I+D sobre interoperabilidad en salud digital e Identidad Auto-Soberana (SSI).

1. INTRODUCCIÓN

En los últimos años, la interoperabilidad en salud digital ha cobrado relevancia estratégica por el aumento de la movilidad de los pacientes, la continuidad del cuidado y la necesidad de intercambio oportuno de información clínica entre instituciones. La adopción de estándares como HL7 FHIR ha impulsado interfaces modernas para representar, consultar e intercambiar recursos clínicos de forma estructurada, y ha fortalecido perfiles orientados al intercambio transfronterizo, como el International Patient Summary (IPS), concebido como un conjunto mínimo de datos clínicos esenciales para escenarios interjurisdiccionales [1], [2].

Sin embargo, persisten brechas estructurales asociadas a modelos de gestión y validación de identidad centralizados y a esquemas de intercambio donde el paciente no controla de forma directa la portabilidad de su información. En la práctica, esto contribuye a la fragmentación de registros, dificulta la conciliación clínica entre efectores y traslada la verificación a intermediarios institucionales, muchas veces incumpliendo el marco regulatorio o con limitaciones en este sentido. El problema se evidencia especialmente en contextos regionales e internacionales, donde la validación de inmunizaciones y el intercambio del IPS pueden volverse lentos o inconsistentes, incrementando costos operativos, duplicación de registros y riesgos de errores de identificación o de reconciliación de datos.

Para abordar estas limitaciones, investigaciones recientes proponen la Identidad Auto-Soberana (SSI) como un enfoque centrado en la persona, en el cual el control de las credenciales se desplaza a una wallet digital y la confianza se sustenta en estándares abiertos. En este paradigma, los W3C Decentralized Identifiers (DIDs) posibilitan la identificación descentralizada de personas e instituciones, mientras que las W3C Verifiable Credentials (VCs) habilitan credenciales firmadas criptográficamente, verificables y a prueba de manipulaciones, reduciendo la dependencia de verificaciones sincrónicas con el emisor en cada transacción [3], [4]. En el ámbito sanitario, se han reportado marcos SSI aplicados a credenciales y a la gestión de tarjetas de salud electrónicas, con mejoras en privacidad e interoperabilidad frente a modelos convencionales [5]. Asimismo, otros trabajos exploran plataformas y modelos SSI para salud centrada en el paciente, apoyados en blockchain y orientados al cumplimiento de regulaciones de protección de datos [6], [7], [8].

A diferencia de estos aportes —que suelen concentrarse en el modelo SSI y su aplicación a credenciales sanitarias—, la presente propuesta se enfoca en una arquitectura de referencia para el intercambio interjurisdiccional de información clínica conforme a HL7 FHIR, particularmente el IPS, integrando SSI con perfiles IHE que operacionalizan interoperabilidad técnica y semántica: PDQm, MHD y SVCM [9], [10], [11]. Asimismo, se aborda explícitamente un desafío crítico para su adopción generalizada: la necesidad de reducir costos operativos y promover escalabilidad.

En este marco, el uso de contratos inteligentes como ancla de confianza aporta trazabilidad e inmutabilidad para registros críticos, como emisores autorizados y políticas de revocación, mientras que la incorporación de una solución de Capa 2 permite coordinar operaciones frecuentes con mayor eficiencia y menor costo transaccional, preservando las garantías de seguridad de la Capa 1. Concretamente, se propone anclar la confianza en Ethereum (Capa 1) [12] y emplear zkSync (Capa 2, ZK-rollup) [13] para soportar operaciones de identidad, incluyendo emisión, verificación y revocación.

En este contexto, la contribución de esta línea de investigación consiste en proponer y validar una arquitectura de referencia que integra SSI (DIDs/VCs) con HL7 FHIR e IHE (PDQm, MHD, SVCM), habilitando un intercambio clínico centrado en el paciente, donde este pueda compartir, bajo consentimiento granular, su IPS o recursos FHIR específicos de forma segura y selectiva entre efectores de salud provinciales, nacionales o internacionales, preservando privacidad y soberanía de los datos, con escalabilidad y eficiencia de costos.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La línea de trabajo se sitúa en la intersección de la interoperabilidad en salud digital y la Identidad Auto-Soberana (SSI), priorizando el uso de estándares abiertos y tecnologías blockchain para la gestión descentralizada de identidades. El marco metodológico adoptado es DSRM [15], abarcando desde la identificación del problema hasta la evaluación del prototipo.

Los ejes principales del trabajo son:

- Diseñar y validar una arquitectura basada en W3C DIDs y Verifiable Credentials [3], [4] para el intercambio seguro de recursos HL7 FHIR e IPS [1], [2].
- Integrar perfiles IHE (PDQm, MHD y SVCM) para asegurar la interoperabilidad técnica y semántica en entornos móviles y regionales [9], [10], [11].
- Desarrollar el componente Identity Connector como puente de integración entre sistemas de salud legados y ecosistemas SSI [8].
- Evaluar la escalabilidad y eficiencia de costos de las operaciones de identidad mediante el uso de redes de Capa 2 (zkSync) sobre Ethereum [12], [13].
- Validar funcionalmente la portabilidad del IPS y la verificación criptográfica de certificados de inmunización en escenarios de movilidad interjurisdiccional [14].

3. RESULTADOS OBTENIDOS Y ESPERADOS

3.1 Resultados obtenidos

Bajo el marco de la DSRM [15], se completó el diseño y desarrollo del artefacto, materializado en dos hitos principales:

A. Arquitectura de Referencia Multicapa: Una solución con privacidad por diseño (PHI *off-chain*) que integra una *wallet* digital para el paciente (Holder), servicios de interoperabilidad clínica basados en HL7 FHIR/IPS e IHE (Issuer), el componente Identity Connector como puente hacia el ecosistema SSI, y un anclaje de confianza en Ethereum (L1) con escalabilidad en zkSync (L2) [3], [4], [8], [12], [13].

B. Laboratorio de Pruebas (Interoperabilidad): Se ha desplegado un entorno basado en el *IPS National Backend* de RACSEL-LACPASS sobre infraestructura contenerizada [14]. Este laboratorio integra núcleos clínicos FHIR, terminología normalizada (SNOMED CT), mediación para perfiles IHE MHD y servicios para la generación de Enlaces de Salud Verificables (VHL) [2], [10], [14].

3.2 Resultados esperados

En las fases subsiguientes, se prevé obtener evidencia empírica mediante la ejecución de

un prototipo integrado extremo a extremo. Los resultados esperados incluyen:

1. Validación técnica: confirmación de la integridad criptográfica y verificación de estado sobre infraestructura L1/L2 [12], [13].
2. Métricas de desempeño: medición de latencia, tasas de éxito y viabilidad económica del modelo mediante la optimización de costos en zkSync [13].
3. Escenarios de movilidad: validación funcional de la portabilidad del IPS para continuidad asistencial y verificación de certificados sanitarios con minimización de datos [1], [2], [6], [8].

Se espera que esta arquitectura aporte evidencia sobre la viabilidad de integrar SSI en redes de escalado para mitigar la fragmentación de datos, estableciendo una base técnica eficiente para la soberanía del paciente en entornos interjurisdiccionales.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto se desarrolla en el Laboratorio de Integración de Tecnologías Aplicadas a Prototipos de Software (LITAPS) de la UTN-FRM (Ing. en Sistemas de Información) y contribuye a la formación de recursos humanos mediante un esquema colaborativo que integra docentes investigadores, graduados, estudiantes de grado y becarios. La línea fortalece capacidades en interoperabilidad en salud (HL7 FHIR/IHE), arquitectura de software, seguridad y privacidad, e identidad auto-soberana (SSI: DIDs/VCs) con tecnologías blockchain, aplicadas al diseño y validación de prototipos.

Se prevé impacto directo en:

- Formación de docentes y actualización de contenidos en las cátedras: Desarrollo de Software, Paradigmas de Programación, Seguridad en los Sistemas de Información y Sistemas de Gestión.
- Capacitación de alumnos de grado mediante participación en modelado, desarrollo, integración y testing del prototipo.
- Becarios de investigación con responsabilidades técnicas específicas.
- Apoyo a Trabajos Finales de Grado y proyección a tesis de posgrado vinculadas a la temática.
- Vinculación y transferencia hacia pares de investigación y profesionales del sector (software/salud digital).

5. BIBLIOGRAFÍA

[1] HL7 International, HL7 FHIR: *International Patient Summary (IPS) Implementation Guide, HL7*. [Online]. Available: https://hl7.org/fhir/uv/ips/. [Accessed: 03-Feb-2026].

[2] HL7 International, *FHIR Release 4 (R4) Specification,* HL7. [Online]. Available: https://hl7.org/fhir/R4/. [Accessed: 03-Feb-2026].

[3] World Wide Web Consortium (W3C), *Decentralized Identifiers (DIDs) v1.0: Core Architecture, Data Model, and Representations,* W3C Recommendation, 2022. [Online]. Available: https://www.w3.org/TR/did-core/. [Accessed: 03-Feb-2026].

[4] World Wide Web Consortium (W3C), *Verifiable Credentials Data Model v1.1,* W3C Recommendation, 2022. [Online]. Available: https://www.w3.org/TR/vc-data-model/. [Accessed: 03-Feb-2026].

[5] D. Salah, S. Mnasri, and H. Idoudi, *Self-Sovereign Digital Identity in Blockchain-Based Systems for E-Health Cards Management*, Procedia Computer Science, vol. 270, pp. 985–993, 2025, doi: 10.1016/j.procs.2025.09.219.

[6] P. Bai, S. Kumar, G. Aggarwal, M. Mahmud, O. Kaiwartya, y J. Lloret, Self-sovereignty identity management model for smart healthcare system, Sensors (Basel), vol. 22, núm. 13, p. 4714, 2022. doi: 10.3390/s22134714.

[7] A. G. de Siqueira, A. F. da Conceição, and V. Rocha, *User-Centric Health Data Using Self-Sovereign Identities*, arXiv:2107.13986, 2021, doi: 10.48550/arXiv.2107.13986.

[8] D. T. Harrell et al., "Technical Design and Development of a Self-Sovereign Identity Management Platform for Patient-Centric Health Care Using Blockchain Technology," Blockchain in Healthcare Today, vol. 5, Art. no. 196, 2022, doi: 10.30953/bhty.v5.196. [Online]. Available: https://blockchainhealthcaretoday.com/index.php/journal/article/view/196. [Accessed: 03-Feb-2026].

[9] Integrating the Healthcare Enterprise (IHE), *Patient Demographics Query for Mobile (PDQm)*, IHE IT Infrastructure Technical Framework. [Online]. Available: https://profiles.ihe.net/ITI/PDQm/. [Accessed: 03-Feb-2026].

[10] Integrating the Healthcare Enterprise (IHE), *Mobile Access to Health Documents (MHD)*, IHE IT Infrastructure Technical Framework. [Online]. Available: https://profiles.ihe.net/ITI/MHD/. [Accessed: 03-Feb-2026].

[11] Integrating the Healthcare Enterprise (IHE), *Sharing Valuesets, Codes, and Maps (SVCM)*, IHE IT Infrastructure Technical Framework. [Online]. Available: https://profiles.ihe.net/ITI/SVCM/. [Accessed: 03-Feb-2026].

[12] Ethereum Foundation, *Introduction to Ethereum.* [Online]. Available: https://ethereum.org/en/what-is-ethereum/. [Accessed: 03-Feb-2026].

[13] Matter Labs, *zkSync Documentation.* [Online]. Available: https://docs.zksync.io. [Accessed: 03-Feb-2026].

[14] Red de América Latina y el Caribe de Cooperación en Salud Digital (RACSEL-LACPASS), "IPS-national-backend," GitHub repository. [Online]. Available: https://github.com/RACSEL/IPS-national-backend. [Accessed: 03-Feb-2026].

[15] T. Haryanti, N. A. Rakhmawati, A. P. Subriadi, and A. Tjahyanto, The Design Science Research Methodology (DSRM) for Self-Assessing Digital Transformation Maturity Index in Indonesia, in Proc. 2022 IEEE 7th Int. Conf. on Information Technology and Digital Applications (ICITDA), Yogyakarta, Indonesia, pp. 1–7, 2022, doi: 10.1109/ICITDA55840.2022.9971171.

SEGUIMIENTO AUTOMÁTICO DE ACTIVIDADES EN ADULTOS MAYORES MEDIANTE GEOFENCING Y PROCESAMIENTO LOCAL

Esteban A. Carnuccio, Pablo M. Vera, Rocío A. Rodríguez

Centro de Altos Estudios en Tecnología Informática (CAETI)
Universidad Abierta Interamericana (UAI)

Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT)
Universidad Nacional de La Matanza (UNLaM)

Florencio Varela 1703, San Justo, Buenos Aires, Argentina

{estebanandres.carnuccio} @alumnos.uai.edu.ar
{pablomartin.vera, rocioandrea.rodriguez} @uai.edu.ar

RESUMEN

El incumplimiento de actividades planificadas, como turnos médicos o actividades sociales, constituye un problema frecuente en adultos mayores y genera preocupación en su entorno familiar. En este trabajo se presenta el diseño e implementación de un sistema stand-alone compuesto por un smartphone Android y un smartwatch con Wear OS, cuyo procesamiento se realiza de forma local sin depender de servicios externos. Mediante geolocalización y geofencing, el sistema permite configurar citas asociadas a una fecha, horario y ubicación determinada. Para verificar la asistencia se desarrolló un algoritmo propio basado en una máquina de estados que detecta manualmente eventos de ingreso y permanencia en las áreas definidas. El resultado se notifica automáticamente mediante SMS a los contactos configurados. De esta forma, se pretende brindar una herramienta que permita el seguimiento de actividades realizadas por adultos mayores en forma remota.

Palabras clave: Adultos mayores, Android, Geolocalización, Geofencing, Wear Os

CONTEXTO

Esta línea de investigación y desarrollo (I+D) forma parte de los proyectos de la Facultad de Tecnología Informática radicados en el Centro de Altos Estudios en Tecnología Informática (CAETI) sede Ciudad Autónoma de Buenos Aires de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado.

1. INTRODUCCIÓN

Según la Organización Mundial de la Salud (OMS), el proceso de envejecimiento puede considerarse desde los 45 años, etapa denominada envejecimiento temprano [1]. Sin embargo, a partir de los 65 años suele observarse un descenso más acelerado de la capacidad funcional, acompañado por un deterioro físico y neurocognitivo progresivo, a este grupo etario la OMS clasifica como adultos mayores [2]. En esta etapa, muchas personas comienzan a requerir distintos niveles de asistencia para el desarrollo de sus actividades cotidianas.

En este sentido, el grado de dependencia

funcional puede evaluarse mediante instrumentos como el índice de Barthel, el cual permite clasificar el nivel de autonomía de una persona en función de su capacidad para realizar actividades básicas de la vida diaria. Cuando el adulto mayor presenta dificultades en estas tareas, suele necesitar el acompañamiento de familiares, allegados o cuidadores [3]. No obstante, este acompañamiento permanente puede impactar negativamente en su percepción de autonomía, afectando su estado de ánimo, privacidad y calidad de vida [4].

Por otra parte, cuando el adulto mayor conserva un alto grado de independencia, puede surgir una situación inversa: la preocupación de sus familiares ante la falta de información sobre sus actividades cotidianas, tales como asistir a consultas médicas o realizar caminatas Si bien la autonomía constituye un factor central para el bienestar y la calidad de vida [5], esta misma independencia puede generar intranquilidad en el entorno cercano, que busca garantizar la seguridad sin vulnerar la libertad personal.

En este contexto, uno de los problemas frecuentes en la edad adulta es la desorientación espacial, junto con olvidos vinculados a actividades planificadas. Por ejemplo, puede ocurrir que la persona no asista a una cita médica programada o que olvide recoger a sus nietos en un horario determinado. Asimismo, puede presentarse la dificultad para reconocer el entorno, incluso en lugares previamente conocidos, lo que incrementa el riesgo de ingresar a zonas potencialmente peligrosas para su integridad física. Situaciones como transitar por veredas en mal estado pueden derivar en caídas, generando lesiones en tejidos blandos o fracturas que impactan significativamente en la autonomía y en la vida diaria del adulto mayor. Estos eventos no solo afectan a la persona, sino que también producen una carga emocional, psicológica y económica en sus familiares, quienes deben intensificar el cuidado y asumir posibles gastos asociados.

A nivel tecnológico, se observa una limitada integración de soluciones que combinen de manera automática la detección temprana de desorientación, la ocurrencia de caídas y el seguimiento del cumplimiento de actividades planificadas. En muchos casos, estas funcionalidades se encuentran implementadas de forma independiente, con escasa automatización e integración entre módulos. Asimismo, diversos sistemas presentan deficiencias de usabilidad, ofreciendo alertas poco adaptadas a las capacidades cognitivas y físicas de los adultos mayores. A esto se suma, en ciertos desarrollos, una fuerte dependencia de infraestructuras externas, lo que condiciona el funcionamiento del sistema a la disponibilidad de conectividad y puede implicar costos adicionales. Finalmente, numerosas soluciones comerciales corresponden a sistemas cerrados que limitan su escalabilidad y la incorporación de nuevas funcionalidades.

En este contexto, surge la necesidad de diseñar una solución integrada que permita monitorear de forma no invasiva la ubicación del adulto mayor, detectar eventos relevantes como caídas o desorientación espacial, y verificar el cumplimiento de actividades planificadas, preservando al mismo tiempo su autonomía. Asimismo, resulta deseable que dicha solución reduzca la dependencia de infraestructuras externas y ofrezca un diseño flexible que facilite su ampliación y adaptación a distintos escenarios de uso.

Por este motivo, en este trabajo de investigación se presenta un sistema orientado a abordar estas limitaciones, integrando mecanismos de geolocalización, detección de eventos y notificación automática dentro de una arquitectura que prioriza la autonomía del usuario y la asistencia contextual a sus familiares.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto de investigación busca:

- Hacer uso de la geolocalización y de las técnicas de Geofencing para el monitoreo de adultos mayores
- Construir una aplicación de monitoreo no invasiva en Android y Wear OS.
- Realizar el seguimiento de cumplimiento de actividades planificadas

3. RESULTADOS OBTENIDOS/ESPERADOS

En este trabajo se desarrolló un sistema de monitoreo orientado a adultos mayores, el cual se encuentra conformado por dos subsistemas principales: un smartphone con sistema operativo Android y un smartwatch con Wear OS.

El sistema fue diseñado con el objetivo de no depender de servicios externos para el procesamiento de los datos ni para la generación de notificaciones mediante plataformas de terceros. Por este motivo, el procesamiento se realiza localmente en el propio dispositivo, funcionando bajo el enfoque de un sistema stand-alone. Dicho procesamiento se lleva a cabo en el smartphone, debido a su mayor capacidad de cómputo y disponibilidad energética en comparación con los relojes inteligentes. En consecuencia, las notificaciones de alertas ante eventos de emergencia se realizan mediante mensajes de texto (SMS) enviados a los contactos responsables del adulto mayor. No obstante, la aplicación del smartphone utiliza funcionalidades provistas por las APIs de Google Maps con el fin de descargar y visualizar mapas de manera gráfica, permitiendo al usuario observar la ubicación geográfica correspondiente.

El desarrollo se orientó específicamente al uso de smartwatches que no poseen conectividad 4G ni GSM. Por tal motivo, el smartphone cumple además el rol de gateway del sistema, ya que es el encargado de gestionar la comunicación externa y de enviar los mensajes SMS a los contactos definidos.

En este sentido, cada subsistema realiza determinadas tareas, las cuales son mencionadas a continuación.

El smartphone permite realizar las siguientes tareas al usuario:

- Configurar áreas de geofence básicas.
- Establecer prioridad de las áreas
- Configurar Zonas seguras
- Configurar citas de asistencias
- Establecer contacto de notificación
- Visualizar zonas de geofences
- Mostrar notificaciones push de alertas
- Enviar SMS
- Configurar de forma personalizada al sistema

Por otro lado, el smartwatch le permite al usuario recibir notificaciones personalizadas, que fueron adaptadas a las necesidades de los adultos mayores. Asimismo, es el dispositivo encargado de detectar y gestionar eventos de caídas que el adulto pueda llegar a sufrir durante su vida diaria. En este sentido, las funcionalidades implementadas en el reloj fueron descritas en [6].

En la **Fig. 1** se puede visualizar la pantalla principal de la aplicación llamada *AbuMonitor*. En esta imagen se puede visualizar las diferentes opciones que puede realizar el usuario administrador del sistema.



Fig. 1 Pantalla Inicial de AbuMonitor

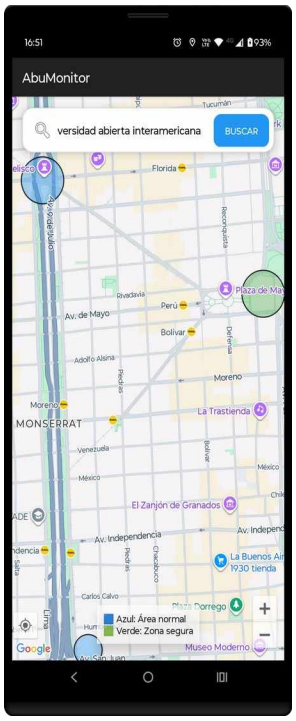


Fig. 2 Interfaz para definir áreas de Geofence

En el sistema se pueden definir tres tipos diferentes de áreas de geofence.

- Áreas Normales
- Áreas de Zonas Seguras
- Áreas correspondientes a citas de asistencia

Las áreas normales son zonas de geofencing que el administrador crea y configura para que se genere una notificación según el nivel de prioridad establecido por el usuario. En este sentido, si el área es de prioridad baja, se notificará únicamente a los contactos del adulto mayor vía SMS cuando este ingrese, salga o permanezca en el área durante un tiempo determinado. En las áreas de prioridad media, además del envío de SMS a los contactos, se alertará al adulto mayor en su smartphone mediante notificaciones push. En cambio, en las áreas de prioridad alta se generarán notificaciones vía SMS a los contactos, notificaciones push en el smartphone del adulto mayor y, adicionalmente, se mostrará una alerta en su smartwatch.

Por otra parte, se encuentran las áreas de zonas seguras. Estas son áreas en las que el adulto mayor debe ingresar o salir dentro de un rango horario específico. Si el adulto sale fuera del horario configurado, el sistema lo considera un evento de advertencia y notifica vía SMS a los contactos sobre dicho suceso. Este tipo de áreas se configuran por defecto con prioridad baja.

En la Fig. 2, puede visualizarse un ejemplo de áreas de geofence normales y de zonas seguras. Las primeras se representan en color azul, mientras que las segundas se muestran en color verde.

Finalmente, el tercer tipo de área de geofence corresponde a las citas de asistencia. Estas permiten realizar un seguimiento del cumplimiento de actividades planificadas que deben llevarse a cabo en una fecha y horario determinados, por ejemplo, cuando el adulto mayor tiene un turno médico, asiste a una obra de teatro o juega al ajedrez con sus amigos en un día específico. Por lo tanto, el sistema notificará a los contactos tanto en caso de inasistencia como de asistencia a la cita. La verificación de dicha asistencia se realiza mediante un algoritmo propio que utiliza las áreas de geofence asociadas a cada evento.



Fig. 3 Recordatorio de cita en smartwatch



Fig. 4 Recordatorio de cita en Smartphone

Al crear una cita de asistencia, el usuario debe indicar el horario y día de la cita, además debe ingresar cuanto tiempo antes se debe emitir un

recordatorio de la cita al adulto mayor, por lo que en el momento indicado el sistema generará un recordatorio tanto en el smartwatch como en el smartphone, tal como se muestra en la Fig. 3 y Fig. 4. Es importante mencionar que las notificaciones en el reloj fueron adaptadas para que sean más fáciles de utilizar por los adultos mayores, tal como se explican en [6].

El sistema mantiene un historial con el estado de cada cita de asistencia, tal como puede observarse en la Fig. 5.

En el mapa las áreas de geofence asignadas a citas de asistencia, se identifican por el color amarillo, tal como se muestra en la Fig. 6.



Fig. 5 Estados de las citas de asistencia

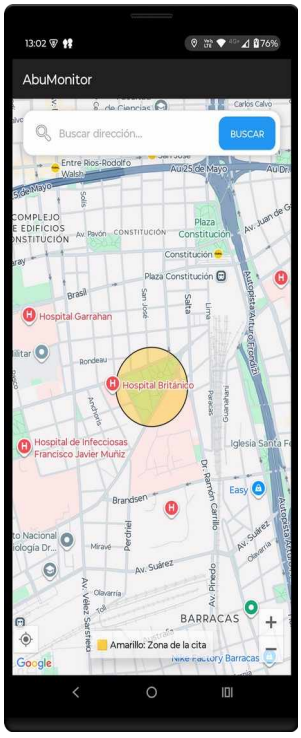


Fig. 6 Área de cita de asistencia

En este sentido, cuando se crea un evento de este tipo, su estado inicial se muestra como "*cita pendiente*". Posteriormente, el sistema actualiza automáticamente su estado a "*cita asistida*" cuando el adulto mayor ingresa al área correspondiente en la fecha y hora indicadas por el administrador, y luego sale de dicha zona tras haber permanecido en ella al menos un minuto por defecto, que podría llegar a ser configurable en un futuro.

Por otro lado, el sistema marca la cita como "*cita no asistida*" si, transcurrido un tiempo posterior al horario de inicio establecido, el adulto mayor no ingresa ni sale del área indicada por el administrador. En ambos casos, ya sea que el adulto mayor asista o no a la cita programada, el sistema notifica mediante SMS a los contactos sobre lo ocurrido.

Desafíos encontrados durante el desarrollo

Inicialmente, el sistema de geofencing había sido desarrollado utilizando las APIs de los servicios ofrecidos por Google Play, en conjunto con el sistema operativo, para gestionar las áreas definidas. En consecuencia, la detección de los eventos de geofence se realizaba a través de dichas herramientas. Al delegar esta funcionalidad a los servicios de PlayServices, el sistema optimizaba el consumo de recursos, ya que la aplicación no necesitaba estar constantemente verificando la ubicación geográfica. En su lugar, la Api le informaba cuando se producían un evento de este tipo, minimizando así el consumo energético. No obstante, tras una actualización de estos servicios, la detección dejó de funcionar correctamente, generándose demoras y retrasos en la identificación de los eventos.

Por este motivo, fue necesario implementar la detección de manera manual mediante el desarrollo de un algoritmo propio encargado de gestionar estos eventos. Para ello, se diseñó una máquina de estados que verifica, en intervalos de tiempo determinados, si el adulto mayor se encuentra dentro o fuera de un área de geofence previamente definida en el sistema.

Si bien este algoritmo puede implicar un mayor consumo de recursos en comparación con la solución basada en las APIs originales, presenta la ventaja de realizar el procesamiento de geofencing de forma local y con mayor precisión, reduciendo la dependencia de servicios externos y mejorando la confiabilidad del sistema.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 5 personas, docentes de grado, postgrado y alumnos.

En el área de dispositivos móviles se encuentran en realización, 2 tesis de maestría y 1 tesina de grado en la UAI (Universidad Abierta Interamericana).

5. BIBLIOGRAFÍA

[1] Europaem, “The ageing brain: how to maintain cognitive functions into old age,” *europaem.eu/en/the-ageing-brain-how-to-maintain-cognitive-functions-into-old-age/*, 2025. .

[2] OMS, *Informe Mundial sobre el envejecimiento y la salud*. Estados Unidos: OMS, 2015.

[3] R. V. Rosa, *Manual de cuidados generales para el adulto mayor disfuncional o dependiente*, Literatura. México, 2014.

[4] A. Ardila, D. Orejana, and L. Silva, “Relación del nivel de dependencia y la depresión en los adultos mayores mayores de un hogar vida de florida blanca, Santander en el año,” Universidad Cooperativa de Colombia, 2018.

[5] T. M. Celeiro and M. P. Galizzi, “Calidad de vida en adultos mayores entre 70 a 85 años institucionalizados y no institucionalizados de la ciudad de Nogoyá,” Universidad Católica Argentina, 2019.

[6] P. M. Vera, E. A. Carnuccio, and R. A. Rodríguez, “Sistema de Alerta de Geofencing para el monitoreo de adultos mayores basados en Android y Wear OS,” in *Libro de Actas - XXVII Workshop de Investigadores en Ciencias de la Computación WICC 2025*, Mendoza, 2025, pp. 502–506.

Sistema de respuesta a emergencias y gestión de riesgos en Ciudades Inteligentes debido al cambio climático

Carlos Lugani, Mauro Cambarieri, Diego Birochio, Guillermo Malpeli, María Parrotti

Universidad Nacional de Rio Negro. Sede Atlántica
Laboratorio de Informática Aplicada / Laboratorio de Investigación y Conservación de la Biodiversidad
{clugani, mcambarieri, dbirochio, gmalpeli}@unrn.edu.ar
mparrotti@gmail.com

RESUMEN

Este proyecto tiene como propósito establecer los factores que intervienen en la gestión de respuesta ante múltiples incidentes con origen en el cambio climático dentro del concepto de Ciudades Inteligentes. La sociedad debe entender las situaciones que podrían ocurrir promoviendo la educación de los ciudadanos y las iniciativas para proteger la vida de las personas, el medio ambiente donde se establecen las ciudades y las mismas ciudades ante una emergencia.

Se pretende definir que tipo de incidentes o desastres son posibles y el nivel de preparación real y el necesario ante estas situaciones que pudiera sufrir una ciudad. Asimismo se debe analizar la interrelación de ocurrencias (se presenta la figura de efecto dominó) que se pueden sumar determinado que un incidente posterior también se produzca y así escalando en forma sucesiva. Se destaca la existencia de incidentes debido al cambio climático como ser: inundaciones, olas de calor, tormentas, sequías, incendios, aludes, deterioro de la calidad del aire, perdida de la diversidad urbana, colapso de drenajes, cortes de energía, daños en vias de transporte o en sistemas de transporte, entre otras situaciones, que a su vez pueden ocurrir en forma relacionada o no.

En este proyecto se enumerarán factores intervinientes, se analizará el posible impacto teniendo en cuenta ocurrencias históricas en ciudades (aun no siendo Ciudades Inteligentes). Asimismo se contextualizarán las acciones que han sido llevadas a cabo en el pasado para analizar si han sido exitosas y prever los mejores planes de acción para Ciudades Inteligentes y cómo el desarrollo de capacidades ayudarían a gestionar incidentes en especial teniendo en cuenta los habitantes mas vulnerables.

Palabras clave: Ciudades Inteligentes, Emergencias, Riesgos, Cambio Climático

CONTEXTO

En el contexto mundial la gestión de múltiples incidentes relacionados con el cambio climático es de importancia crítica por las siguientes razones:

Vulnerabilidad Creciente de las Ciudades

Las ciudades concentran más del 55% de la población mundial y son responsables del 75% de las emisiones globales de gases de efecto invernadero. Paradójicamente, también son las más vulnerables a los impactos del cambio climático. Las Ciudades Inteligentes, a pesar de sus avances tecnológicos, enfrentan riesgos amplificados debido a su dependencia de sistemas interconectados que pueden generar efectos cascada durante desastres.

La investigación de Seong y Jiao (2024) [1] revela que las ciudades inteligentes priorizan enfoques metodológicos para adaptar tecnologías a la resiliencia ante desastres, pero rara vez abordan

aspectos sociológicos relacionados con el medio ambiente, la economía y la gobernanza.

Efecto Dominó de Incidentes

Los sistemas urbanos de ingeniería (transporte, telecomunicaciones, agua y energía) demuestran altos niveles de interconectividad lo que los hace particularmente afectados a fallas en serie. Un incidente inicial puede desencadenar una serie de eventos secundarios que amplifican el impacto más allá del evento original. El caso del huracán Katrina en Nueva Orleans / Estados Unidos en al año 2005, ejemplifica este fenómeno: la ruptura de diques provocó inundaciones que colapsaron redes eléctricas, sistemas de agua y el orden social, creando un efecto dominó significativo y devastador para la ciudad.

Según datos de NOAA (2024) [1], entre los años 2018 al 2024 hubo apenas 18 días en promedio entre desastres con un alto costo asociado en Estados Unidos, comparado con 82 días en la década de 1980. Estos intervalos más cortos significan menos tiempo y recursos disponibles para responder, recuperarse y prepararse para eventos futuros, produciendo impactos en cascada particularmente mas difíciles de sobrellevar para los habitantes más vulnerables.

Protección de Poblaciones Vulnerables

La investigación de Goonetilleke et al. (2024) [2]. enfatiza que la confianza comunitaria, la brecha digital y la sensibilidad cultural son variables contextuales críticas para ejecutar efectivamente la gestión de desastres basada en la comunidad con tecnologías inteligentes. Los asentamientos informales y/o barrios marginales, típicamente situados en áreas periféricas, ambientalmente no adecuadas y carentes o con faltas significativas de infraestructura; soportan la mayor carga de desastres relacionados con el clima, son menos propensos a recibir ayuda durante la ocurrencia de incidentes, así como posteriormente de ocurridos.

De acuerdo estudios de Naciones Unidas [3] para al año 2040, más de dos mil millones de personas que actualmente viven en ciudades podrían estar expuestas a un aumento adicional de temperatura de al menos 0.5°C. Las ciudades necesitaran una inversión estimada de $4.5 a $5.4 billones anuales para construir y mantener sistemas resilientes al clima, pero el financiamiento actual es de apenas $831 mil millones, solo una fracción del monto requerido.

La siguiente figura pretende esquematizar la interrelación de los diferentes factores de los cuales se van a analizar sus interrelaciones.



Como antecedentes del grupo de investigación sobre la temática de Ciudades Inteligentes se mencionan los siguientes proyectos han sido desarrollados o se están realizando en el Laboratorio de Informática Aplicada:

- PI 40-C-875 “Herramientas Informáticas de Dominio Específico para el Desarrollo de Servicios Digitales Innovadores para Comunidades Urbanas y Rurales en el Marco de Ciudades y Regiones Inteligentes”
- PI 40-C-797 - “Identificación de personas y dispositivos IoT en ciudades inteligentes”
- Proyecto de Investigación bajo Res. 1296/2024 - Esquema de información para municipios con capacidad de Ciudad Inteligente Sostenible.

Este proyecto pretende crear un esquema de trabajo y plantear como debería ser un sistema para la gestión de incidentes que podrían afectar a las Ciudades Inteligentes, que incluya la definición de roles y responsabilidades, la estandarización de procesos y acciones, así como establecer que

[1] https://www.noaa.gov/ National Oceanic and Atmospheric – Administración Nacional Oceánica y Atmosferica – Departamento de Comercio de Estados Unidos

herramientas serían necesarias para la adecuada gestión de incidentes.

1. INTRODUCCIÓN

La gestión de incidentes climáticos en Ciudades Inteligentes se desarrolla en varios dominios disciplinarios:

Resiliencia Urbana y Adaptación Climática

Una revisión sistemática [4] de 548 artículos publicada en 2025 sobre ciudades inteligentes resilientes revela que las ciudades resilientes incorporan capacidades adaptativas en sistemas cívicos para mantener la continuidad operacional durante eventos disruptivos. Esto requiere planes de preparación bien desarrollados, respuestas de emergencia coordinadas y protocolos de recuperación eficientes. A diferencia de revisiones previas que abordan ciudades inteligentes y resiliencia por separado, la investigación propuesta analiza la convergencia de ambos términos.

Este proyecto identifica una paradoja: la existencia y dependencia las tecnologías y la aplicación en Ciudades Inteligentes, si bien mejoran la vida de las personas, también introducen vulnerabilidades debido a la mayor interdependencia de la tecnología y de las infrestructuras codependientes y a las disrupciones que pudieran ocurror debido a desastres naturales. Por lo que se detecta que esta situación no ha sido suficientemente abordada en la literatura existente.

Innovaciones Tecnológicas para Resiliencia ante Desastres

Singh (2023) [5] clasifica las tecnologías emergentes y disruptivas utilizadas para mejorar la resiliencia ante desastres en Ciudades Inteligentes. A pesar de los avances, las Ciudades Inteligentes enfrentan peligros amplificados: su mayor dependencia de sistemas y redes interconectadas las hace susceptibles a riesgos más allá de los convencionales, conduciendo a efectos en cascada.

Las tecnologías clave incluyen: Internet de las Cosas para monitoreo en tiempo real, inteligencia artificial, gemelos digitales para simulación de escenarios, sistemas de alerta temprana basados en sensores y plataformas Web y aplicaciones móviles para información ciudadana.

Marco de Sendai y Gestión Integral del Riesgo

El Marco de Sendai para la Reducción del Riesgo de Desastres 2015-2030[2] proporciona directrices globales para la gestión holística del riesgo ante desastres en todos los niveles. Desde el año 2015, el número de países con estrategias nacionales de reducción de riesgo con perspectiva en desastres se ha más que duplicado. En el año 2024, 131 países reportaron tener estrategias nacionales y locales, comparado con 57 en 2015.

El estudio de Samir et al. (2023) [5] propone la Ciudad Resiliente Inteligente como enfoque para mejorar la reducción del riesgo de desastres, integrando tecnologías de ciudad inteligente con principios del Marco de Sendai. El enfoque presenta la importancia de políticas de uso de suelo flexibles, inversiones dirigidas y toma de decisiones impulsada por la comunidad.

Sistemas de Alerta Temprana y Gemelos Digitales

La investigación sobre modelado de ciudades 3D, sistemas de alerta temprana y gemelos digitales (Sensors, 2023) [6] destaca la creación de tecnología para construir resiliencia climática. Se pueden utilizar gemelos digitales para integrar datos de sensores en tiempo real con modelos virtuales 3D, permitiendo adquirir información actualizada y responder efectivamente a eventos climáticos extremos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Existen varios aspectos iniciales, intermedios y finales a definir, la visión de Ciudades Inteligentes determina cómo se pueden aprovechar los beneficios de comunicaciones, tecnología, participación ciudadana, rol de la gestión del estado y empresas. Se analizaran los relevamientos

[2] https://aidmi.org/sendai-framework-for-disaster-risk-reduction-2015-2030/?gad_source=1&gad_campaignid=20719207678&gbraid=0AAAAAp90Ml7arE3M1goq_CvcmMAGMLNFz&gclid=CjwKCAiAkvDMBhBMEiwAnUA9BWlqm3gv1ENDFa5RXLr_q-LCl4tknT_n6vrbyjN2KBSxJgsfzrypKxoCy1gQAvD_BwE

necesarios y posibles respuestas eficientes de acuerdo al tiempo y de qué forma se puede reducir el impacto de ocurrencias.

Se tendrá en cuenta un análisis de riesgos e impactos, para determinar planes de acción en forma individual, relacionados y en conjunto, que puedan ser revisados en forma dinámica por un organismo que tenga injerencia sobre las acciones previas y posteriores a incidentes.

Objetivo General: definir un esquema de interrelaciones que permitirá observar e identificar riesgos de desastres relacionados con el cambio climático y sus posibles acciones que disminuyan su ocurrencia o los efectos adversos, dentro del contexto de Ciudades Inteligentes.

En la figura siguiente se demuestra algunos factores que se deben considerar en el esquema propuesto:



Sobre la base de la premisa anterior, se definen los siguientes objetivos específicos.

Diagnóstico

- Identificar y catalogar los tipos de incidentes relacionados con el cambio climático más probables.
- Mapear las interdependencias entre sistemas urbanos críticos para identificar puntos vulnerables y potenciales efectos dominó.
- Evaluar el nivel de preparación actual de la ciudad mediante revisiones de infraestructura física, capacidades tecnológicas, protocolos de respuesta existentes y recursos humanos capacitados.
- Identificar poblaciones vulnerables (asentamientos periféricos, adultos mayores, personas con discapacidad, comunidades con bajos ingresos) y evaluar su nivel de exposición a riesgos climáticos.

Análisis de Riesgos e Impactos

- Desarrollar matrices de riesgo que cuantifiquen la probabilidad e impacto potencial de cada tipo de incidente identificado, Tener en cuenta datos históricos.
- Modelar escenarios de efectos dominó mediante simulaciones que representen cómo un incidente inicial puede desencadenar eventos secundarios y terciarios en sistemas interconectados.
- Analizar ocurrencias históricas de desastres en ciudades comparables (incluso no siendo Ciudades Inteligentes) para poder tener ejemplos de patrones de cascada, tiempos de respuesta efectivos y estrategias de mitigación exitosas que se hayan realizado.
- Cuantificar impactos potenciales en términos de vidas humanas, infraestructura crítica, actividad económica, servicios esenciales y medio ambiente.

Diseño de Sistemas de Respuesta

- Diseñar una arquitectura tecnológica integrada que podría incluir por ejemplo: redes de sensores IoT, sistemas de alerta temprana, plataformas de datos centralizadas, tablero de control en tiempo real y canales de comunicación redundantes.
- Desarrollar protocolos de respuesta específicos para cada tipo de incidente individual y protocolos coordinados para múltiples incidentes simultáneos.
- Establecer un órgano de revisión multidisciplinario con representación de gobierno, sector privado y residentes.
- Crear marcos de gobernanza que faciliten la coordinación, analicen los factores de falla y responsabilidades y autoridades en emergencias.

Desarrollo de Capacidades

- Implementar programas de educación ciudadana sobre preparación ante desastres, sistemas de alerta temprana, rutas de evacuación, puntos de reunión y comportamientos de autoprotección, con especial énfasis en poblaciones vulnerables.
- Capacitar a personal de respuesta de emergencia (bomberos, ambulancias, policía) en el uso de tecnologías, protocolos de coordinación entre organismos y gestión de efecto dominó.
- Desarrollar programas de capacitación para tomadores de decisiones gubernamentales en interpretación de datos en tiempo real,

activación de protocolos de emergencia y comunicación efectiva durante crisis.

- Establecer mecanismos de participación ciudadana que permitan a la población contribuir con información en tiempo real mediante aplicaciones móviles y plataformas Web y redes sociales.

Fase de Implementación y Pruebas

- Implementar infraestructura física necesaria (sensores, centros de datos, redes de comunicación redundantes) siguiendo estándares de interoperabilidad y seguridad informática.
- Realizar simulacros periódicos que prueben tanto incidentes individuales como escenarios de múltiples incidentes en cascada, evaluando tiempos de respuesta, efectividad de comunicación y coordinación entre agencias.
- Establecer alianzas público-privadas que aseguren financiamiento sostenible, acceso a tecnologías de vanguardia y capacidad de escalamiento del sistema.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Se ha realizado una descripción pormenorizada de los objetivos específicos y los resultados esperados se encuentran en línea con la obtención de los productos o resultados mencionados. Los marcos internacionales de la ONU [7] [8], particularmente el ODS 11, ODS 13 y el Marco de Sendai, proporcionan directrices claras y metas medibles que este proyecto puede adoptar y contribuir a alcanzar. La implementación exitosa de los objetivos específicos propuestos puede generar resultados sustanciales en reducción de pérdidas humanas y materiales, mejora de tiempos de respuesta, fortalecimiento de resiliencia comunitaria y avances en conocimiento científico.

La implementación exitosa de este proyecto puede generar los siguientes resultados:

- Reducción de impacto en la vida de los residentes, y sus económicos para la Ciudad
- Protección de infraestructura crítica
- Mejora en tiempos de respuesta y coordinación
- Fortalecimiento de resiliencia comunitaria y participación ciudadana
- Desarrollo de modelos predictivos precisos
- Sostenibilidad y escalabilidad a otras Ciudades Inteligentes
- Impacto Ambiental Positivo

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se encuentra formado por tres investigadores formados, un investigador en formación y una alumna, de las carrera Licenciatura en Sistemas y Licenciatura en Ciencias del Ambiente. Se desarrollará un plan de trabajo de la carrera de Doctorado, un plan de trabajo para una Maestría y un Trabajo Final de Carrera.

5. BIBLIOGRAFÍA

[1]. Seong, K., & Jiao, J. (2024). Is a Smart City Framework the Key to Disaster Resilience? A Systematic Review. Journal of the American Planning Association. https://doi.org/10.1177/08854122231199462

[2]. Goonetilleke, A., et al. (2024). Bridging Community Engagement and Technological Innovation for Creating Smart and Resilient Cities: A Systematic Literature Review. Smart Cities, 7(6), 3823-3852. https://doi.org/10.3390/smartcities7060147

[3]. UN-Habitat (2024). World Cities Report 2024: Cities and Climate Action. United Nations Human Settlements Programme.

[4] State-of-the-art review of resilient smart cities: Progress and challenges (2025). ScienceDirect. https://doi.org/10.1016/j.scs.2025.102704

[5] Singh, A. (2023). Smart Cities and Technological Innovations Towards Disaster Resilience. In: International Handbook of Disaster Research. Springer, Singapore. https://doi.org/10.1007/978-981-19-8388-7_222

[6] Sensors (2023). Management of Climate Resilience: Exploring the Potential of Digital Twin Technology, 3D City Modelling, and Early Warning Systems. 23(5), 2659. https://doi.org/10.3390/s23052659

[7] UN DESA (2024). Goal 11: Sustainable Cities and Communities. Extended Report on SDG 11. https://unstats.un.org/sdgs/report/2024/extended-report/

[8] UN DESA & UNFCCC (2024). Harnessing Climate and SDGs Synergies: Global Report on Synergy between Climate and the SDGs. High-Level Political Forum on Sustainable Development.

Sistemas inteligentes para la gestión sostenible del recurso hídrico

Javier Díaz, Agustín Candia, Jorge Bellavita, Laura Fava, Vanessa Aybar, Emanuel Borda, Matías Pagano
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata, calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528
{jdiaz, lfava, acandia, jbellavita, mpagano, vaybar}@linti.unlp.edu.ar

Resumen

La gestión sostenible de acuíferos costeros constituye un desafío creciente en regiones sometidas a explotación intensiva, donde la intrusión salina y la sobreexplotación comprometen la disponibilidad futura del recurso. En muchos casos, los prestadores locales carecen de información en tiempo real que permita una toma de decisiones basada en evidencia.

En este contexto, el presente trabajo propone un enfoque basado en sistemas inteligentes distribuidos que integran tecnologías de Internet de las Cosas (IoT), redes de comunicación de bajo consumo, procesamiento en el borde (Edge Computing) y modelos de analítica avanzada.

Se presentan tres líneas de investigación y desarrollo complementarias: (i) monitoreo inteligente de acuíferos y sistemas de extracción mediante sensores distribuidos y plataformas de análisis en tiempo real, (ii) digitalización automática de instrumentos de medición y monitoreo predictivo de infraestructura mediante técnicas de visión por computadora e inteligencia artificial, y (iii) trazabilidad criptográfica de la producción y el consumo de agua mediante mecanismos basados en blockchain para fortalecer la integridad y confiabilidad del dato.

En conjunto, estas líneas conforman una arquitectura tecnológica orientada a mejorar la eficiencia operativa, la transparencia y la sostenibilidad en la gestión del recurso hídrico.

Palabras claves: Internet de las Cosas (IoT), LoRaWAN, smart-water, Inteligencia Artificial, Edge-computing, Blockchain, recurso hídrico.

Contexto

El Laboratorio de Investigación de Nuevas Tecnologías Informáticas LINTI de la Facultad de Informática, desarrolla desde hace más de una década proyectos relacionados con Internet de las Cosas en Smart Cities.

Estas actividades se encuentran enmarcadas en el proyecto de I+D *Tecnologías Digitales para la Inclusión, la Equidad y la Sostenibilidad*, acreditado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

En estos últimos años se han presentado resultados vinculados principalmente al desarrollo e implementación de infraestructuras IoT aplicadas al monitoreo de calidad de aire [1], sistemas de balizamiento en el Puerto La Plata [2], soluciones de suministro inteligente de agua [3] y predicción de enfermedades foliares [4], entre otros. Estos antecedentes consolidan la experiencia del equipo en el despliegue de redes distribuidas y plataformas de monitoreo en entornos reales, constituyendo la base tecnológica sobre la cual se sustentan las líneas actuales.

En el presente artículo se presentan los avances específicos de las líneas orientadas a la gestión del recurso hídrico, incluyendo desarrollos en monitoreo IoT, analítica avanzada y trazabilidad criptográfica.

1. Introducción

Los proyectos antes mencionados, dan cuenta del trabajo que se viene desarrollando con IoT en diferentes escenarios.

El equipo de trabajo interdisciplinario que lleva a cabo este proyecto vinculado con el agua, está formado por especialistas del área de hidrogeología, personal técnico de las cooperativas de agua[1] y nuestro equipo de informáticos y especialistas IoT. Como se mencionó, existen actualmente tres líneas de trabajo:

La primera de las líneas mencionadas en este artículo, aborda el diseño e implementación de infraestructuras IoT para el monitoreo en tiempo real de acuíferos costeros sometidos a explotación intensiva. El objetivo principal es incrementar la eficiencia operativa de los prestadores de agua usando tecnologías emergentes para dotar a las comunas de indicadores sobre el servicio de agua y otras prestaciones que les permita realizar autoevaluaciones periódicas con el fin de lograr los objetivos de desarrollo sostenible. Las áreas específicas que se abordaron inicialmente incluyen la supervisión en tiempo real del nivel de los pozos, la medición de la salinidad, la telegestión de los motores responsables de extraer agua subterránea, el control de nivel en cisternas y tanques de distribución, así como la medición del consumo en puntos representativos de grandes consumidores [5] [6].

Como continuidad de esta experiencia, se incorporaron macromedidores de caudal en puntos estratégicos del sistema, específicamente para medir la producción total de agua en la planta potabilizadora y el caudal a la salida del acueducto principal que abastece a la ciudad. Esta ampliación permitió comenzar a registrar métricas de producción y distribución a escala real, habilitando análisis agregados en distintos horizontes temporales (semanal, mensual, estacional y comparativo interanual). La disponibilidad simultánea de ambas mediciones posibilita, además, la conciliación entre el volumen producido y el efectivamente distribuido a los usuarios finales, dentro de márgenes de error aceptables propios de la instrumentación. Asimismo, esta información resulta clave para la detección temprana de pérdidas o fugas en el acueducto, el cual posee una extensión aproximada de 3 km y conecta la planta potabilizadora con el casco urbano, aportando una nueva capa de información para la gestión operativa y la mejora de la eficiencia del sistema de distribución.



Fig 1.: Macromedidor en Planta

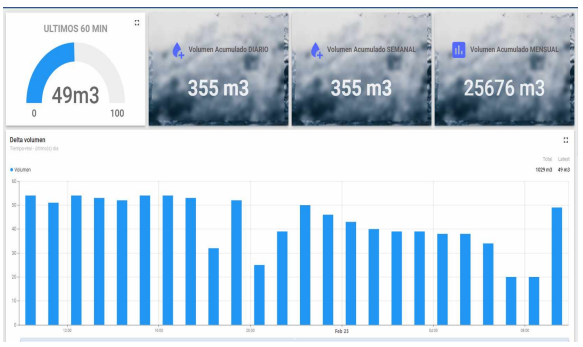


Fig 2.: Plataforma de monitoreo

La segunda línea de trabajo está

[1] Se comenzó con la Cooperativa de San Clemente (COS) como escenario piloto.

relacionada con el desarrollo de modelos avanzados de analítica e inteligencia artificial aplicados tanto a la digitalización de instrumentos de medición como al monitoreo predictivo de infraestructuras asociadas al recurso hídrico.

Para la digitalización del consumo o caudal, se ha comenzado a integrar técnicas de visión por computadora y reconocimiento automático de patrones para la extracción automatizada de información desde dispositivos analógicos tradicionales. Se han analizado y probado tecnologías como PyTorch y TensorFlow mediante arquitecturas YOLO y CNN, optimizando modelos a través de TensorFlow Lite para permitir detección de objetos en tiempo real directamente en microcontroladores [7][8]. Este enfoque permitiría dotar de conectividad a medidores existentes sin necesidad de sustituir el parque instalado.

Complementariamente, para el monitoreo predictivo y la detección temprana de anomalías en equipos electromecánicos, se ha iniciado la fase de experimentación con librerías especializadas en procesamiento de señales y Machine Learning. El análisis se enfoca en variables críticas como vibraciones, acústica e indicadores operativos, procesados mediante arquitecturas de computación distribuida en el borde (Edge Computing). Esta capacidad permite identificar patrones de falla antes de que ocurra una interrupción del servicio, optimizando los ciclos de mantenimiento. [9].

La última línea presentada se enfoca en la aplicación de blockchain para la trazabilidad de la producción y consumo del recurso hídrico. Se propone un modelo de confianza desde el origen del dato (producción o consumo), integrando dispositivos de medición con registros inmutables que reduzcan la dependencia de intermediarios y minimicen riesgos de alteración. Este enfoque se inscribe en modelos de gobernanza descentralizada donde la confianza se fundamenta en mecanismos criptográficos verificables más que en autoridades centralizadas [10].

2. Líneas de Investigación, Desarrollo e Innovación

Las líneas de investigación, desarrollo e innovación presentadas en este trabajo se orientan al diseño e implementación de soluciones tecnológicas para la gestión inteligente del recurso hídrico, considerando tres verticales principales: (i) monitoreo inteligente de acuíferos y sistemas de extracción de agua potable, (ii) digitalización inteligente de medición mediante y monitoreo predictivo de infraestructura mediante técnicas de inteligencia artificial, y (iii) trazabilidad criptográfica de la producción y consumo de agua potable mediante blockchain.

A partir de la integración de tecnologías IoT, infraestructuras de comunicación de bajo consumo y técnicas de Inteligencia Artificial y de trazabilidad criptográfica, se propone avanzar en los siguientes ejes de I+D+i:

- Análisis de nuevas tecnologías para la actualización del despliegue de las redes IoT distribuidas para el monitoreo en tiempo real.
- Profundizar la comparación de tecnologías de monitoreo hidrogeológico entre data loggers convencionales y sensores de bajo costo con transmisión remota. Análisis de correlación entre distintos dispositivos de medición.
- Análisis de datos hidrogeológicos para detección temprana de procesos de salinización. Evaluación de patrones asociados a sobreexplotación.
- Desarrollo e integración de modelos de visión por computadora para la digitalización automática de medidores mecánicos existentes y para detección de fallas.
- Diseñar e implementar módulos criptográficos embebidos en medidores

inteligentes capaces de firmar digitalmente cada registro de consumo en el momento exacto de su generación con el objetivo de garantizar la autenticidad, integridad e inmutabilidad del consumo de agua desde su generación.

3. Resultados Obtenidos/Esperados

Los resultados obtenidos hasta el momento se vinculan con la primera línea de investigación, orientada al desarrollo y validación de la infraestructura IoT para la adquisición confiable de datos en campo. Esta fase permitió consolidar la arquitectura de captura, transmisión, almacenamiento de la información y análisis estadístico convencional de los datos, estableciendo la base tecnológica necesaria para las etapas subsiguientes.

En el marco de la validación de la infraestructura IoT desarrollada, se implementó un esquema de doble monitoreo en perforaciones seleccionadas, con el objetivo de comparar el desempeño de sensores de bajo costo con comunicación inalámbrica frente a dataloggers convencionales utilizados en estudios hidrogeológicos de campo [11]. Ambos sistemas fueron instalados en paralelo en las mismas perforaciones, permitiendo registrar simultáneamente niveles freáticos y conductividad eléctrica.

El trabajo de campo incluyó la instalación coordinada de los dispositivos, la realización de mediciones manuales de control y el seguimiento continuo de las variables durante un período de operación intensiva del campo de bombeo. Esta metodología permitió analizar la respuesta del acuífero ante los ciclos de encendido y apagado de las bombas, así como evaluar la consistencia temporal de los registros obtenidos.

Los resultados evidencian una alta concordancia entre ambos sistemas de registro en la evolución de los niveles y la conductividad, reflejando de manera consistente los efectos del régimen de bombeo. Desde el punto de vista operativo, la disponibilidad de datos en tiempo real constituye un aporte significativo para la gestión del recurso, ya que permite identificar situaciones de riesgo y adoptar acciones correctivas de manera oportuna.

En forma complementaria al monitoreo instrumental, se llevó a cabo una transferencia tecnológica orientada al fortalecimiento de las capacidades operativas locales. En este marco, se realizaron calibraciones de todas las sondas de conductividad eléctrica en conjunto con el personal operativo de la planta, integrando estas tareas a la rutina de trabajo habitual. Asimismo, se elaboró y entregó un manual de procedimientos que documenta paso a paso la metodología de calibración y verificación de los sensores, con el propósito de que el propio personal pueda continuar de manera autónoma con esta operativa en el tiempo. Esta instancia permitirá no solo asegurar la calidad de los datos, sino también consolidar la apropiación de la tecnología por parte de los usuarios finales.

En relación con los resultados esperados, se plantea:

- Consolidar la integración de tecnologías IoT y redes LoRaPWAN en entornos operativos reales.
- Realizar junto a especialistas del tema hidrogeológico análisis de datos para detección de patrones y situaciones anómalas en la planta potabilizadora [10].
- Desarrollo de modelos predictivos y sistemas de alerta temprana basados en Inteligencia Artificial.
- Desarrollo de modelos de visión por computadora optimizados para ejecución en dispositivos edge orientados a la lectura automática de medidores de agua.
- Medidas de producción y consumo de agua firmadas criptográficamente para fortalecer la integridad y confiabilidad

del dato.

4. Formación de Recursos Humanos y proyectos conjuntos

El equipo de trabajo de las líneas de I+D+i presentada en este artículo se encuentra formado por docentes investigadores categorizados del LINTI y alumnos avanzados de la Licenciatura en Informática, Licenciatura en Sistemas e Ingeniería en Computación perteneciente a la Facultad de Informática y a la Facultad de Ingeniería. Asimismo se está trabajando con investigadores de la Facultad de Ciencias Naturales y Museo (UNLP) y del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

En relación a las tesinas de grado vinculadas con estas líneas de investigación, se está dirigiendo una tesina de grado donde participan dos estudiantes y dos PPS.

5. Referencias

[1] Porta, A. Marrone, L. *Solutions for SmartCities: proposal of a monitoring* system of air quality based on a LoRaWAN network with low-cost sensors. ISBN: 978-1-5386-5447-7, 2018.

[2] Díaz, J., Candia, A., Bellavita, J. Fava, L. (2024). *Sistema de balizamiento en Puertos usando LoRaWAN*. In: Pesado, P., Thomas, P., Chichizola, (eds) Computer Science - CACIC 2024. Communications in Computer and Information Science, Springer, Cham. https://doi.org/10.35537/10915/172755

[3] Díaz, J., Fava, L., Candia, A., Vilches, D., Pagano, M.(2024). *Soluciones IoT aplicadas al suministro de agua*, WICC 2024 - XXVI Workshop de Investigadores en Ciencias de la Computación, Chubut, Argentina.

[4] Castro, N., Queiruga, C., Pagano, A., Candia, A., *Una solución IoT de bajo costo para la predicción de enfermedades foliares en cultivos agroecológicos a escala agricultura familiar*. CAI 2025 - Congreso Argentino de Agroinformática, 54 Jornadas Argentinas de Informática, 2025.

[5] Kruse, E., Fabiano, M y otros, *Manual de procedimientos. Provisión de agua potable y sistemas de saneamiento. Con técnicas orientadas a Cooperativas de trabajo.* DOI 10.35537/10915/156766, Agosto de 2023.

[6] Díaz, J., Fava, L., Candia, A., Bellavita., J, Pagano, M. *Experiencias de aplicación de IoT a la producción sostenible de agua potable y en el Agro*, WICC 2025, XXVII Workshop de Investigadores en Ciencias de la Computación, Cuyo, Mendoza, Argentina.

[7] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. ArXiv, abs/1704.04861*.

[8] Reis, D., Kupec, J., Hong, J., & Daoudi, A. (2023). *Real-Time Flying Object Detection with YOLOv8*. *ArXiv, abs/2305.09972*.

[9] Xue, H., Huang, B., Qin, M., Zhou H., *Edge computing for Internet of Things: A survey*, in *Proc. 2020 Int. Conf. Internet Things IEEE Green Comput. Commun. IEEE Cybersecurity, Phys. Social Comput. IEEE Smart Data IEEE Congr. Cybermatics*, Rhodes, Greece, 2020, pp. 755–760.

[10] Reyna, A., Martín, C., Chen, J., Soler, E., Díaz, M. (2018). *On blockchain and its integration with IoT. Challenges and opportunities*. Future Generation Computer Systems. 88. 173-190. 10.1016/j.future.2018.05.046.

[11] Carretero, S., Rodriguez, L., Kruse, E. *Zonación química vertical del acuífero costero bonaerense: comportamiento de la interfase.* Revista de Geología Aplicada a la Ingeniería y al Ambiente. 10.59069/24225703e008, 2023.

TECNOLOGÍA Y DESARROLLO TURÍSTICO: APORTES PARA LA TRANSFORMACIÓN DIGITAL DE DESTINOS LOCALES

Adriana MARTIN, Viviana SALDAÑO, Gabriela GAETAN, Mariela COLOMBRES, Paula DIEZ, Silvia VILLAGRA, Claudia CARDOZO, Juan ARNEDO, Christian MONTAÑO TORRICO

Grupo de Investigación y Formación en Ingeniería de Software (GIFIS)
Instituto de Tecnología Aplicada (ITA)
Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia (UNPA-UACO)

adrianaelba.martin@gmail.com; vsaldanio@gmail.com; gabrielagaetan@uaco.unpa.edu.ar; mariela_colombres@hotmail.com, paulamabeldiez@gmail.com, svillagra@uaco.unpa.edu.ar; claudia_yoryi@yahoo.com.ar, juan.arnedo@unpa.edu.ar, christian.montano@unpa.edu.ar

RESUMEN

En este artículo se presentan avances y resultados del proyecto UNPA denominado "*Aportes para la Transformación Digital como Soporte para el Desarrollo del Turismo Local*" se centra en la identificación y aplicación de tecnologías para la Transformación Digital como Soporte para el Desarrollo del Turismo Local. Los objetivos específicos son: 1) Analizar tecnologías digitales aplicables al sector turístico. 2) Desarrollar productos de software orientados a satisfacer las necesidades de los usuarios del sector turístico local empleando las tecnologías digitales analizadas. 3) Evaluar el impacto de la incorporación de las tecnologías digitales al sector turístico local.

Palabras clave: *Turismo | Transformación Digital | Tecnologías para la Digitalización | Productos Digitales | Desarrollo Web y Móvil | Destinos Emergentes*

CONTEXTO

El proyecto actual es desarrollado por integrantes del Grupo de Investigación en Ingeniería de Software (GIFIS), perteneciente al Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia (UACO), de la Universidad Nacional de la Patagonia Austral (UNPA). El trabajo del equipo se ha centrado en el diseño, evaluación y desarrollo de soluciones digitales que incorporan principios de usabilidad, accesibilidad y experiencia de usuario, con el objetivo de generar productos tecnológicos de impacto social. A partir de esta experiencia, GIFIS ha desarrollado metodologías, guías de buenas prácticas e instrumentos de evaluación que permiten diagnosticar, rediseñar y optimizar plataformas digitales en distintos contextos institucionales.

En los últimos años, el grupo ha fortalecido su perfil de transferencia tecnológica, avanzando desde la producción académica hacia la implementación concreta de soluciones en entornos reales. Esto incluye procesos de asesoramiento técnico, rediseño de interfaces, desarrollo de sistemas web y móviles, así como acompañamiento en procesos de transformación digital en organismos públicos y organizaciones sociales.

Las actividades desarrolladas se han aplicado en diversos dominios, tales como: educación, salud, discapacidad, gobierno electrónico y patrimonio histórico-cultural, generando resultados. En particular, en el ámbito del turismo cultural —eje central del proyecto actual— se han obtenido avances significativos vinculados al diseño de soluciones digitales para la puesta en valor del patrimonio y la mejora de la experiencia del visitante [1][2][3][4].

Este proyecto se desarrolla en articulación con la Municipalidad de la ciudad de Caleta Olivia, a través de la Secretaría de Cultura, Deportes, Turismo y Juventud, en calidad de entidad adoptante. Esta vinculación institucional permite orientar el conocimiento científico-tecnológico hacia necesidades concretas del territorio, promoviendo la transferencia efectiva

de capacidades académicas al ámbito gubernamental. En este marco, el proyecto se concibe como una iniciativa de innovación aplicada destinada a fortalecer los procesos de transformación digital del sector turístico local, contribuyendo al desarrollo económico, cultural y social de la región.

1. INTRODUCCIÓN

La transformación digital se ha consolidado como un proceso estratégico para la modernización de sectores productivos y de servicios, generando cambios sustanciales en los modelos de gestión, en las cadenas de valor y en la experiencia de los usuarios [5]. En el ámbito del turismo, este proceso se manifiesta en la adopción de tecnologías como plataformas web y móviles, servicios de geolocalización, inteligencia artificial, análisis de datos masivos, Internet de las Cosas (IoT), realidad aumentada y virtual, y soluciones basadas en blockchain, entre otras. Estas herramientas posibilitan la personalización de servicios, la generación de información en tiempo real y la optimización en la gestión de flujos turísticos, contribuyendo a la evolución de destinos tradicionales hacia entornos más conectados e inteligentes [6].

No obstante, la integración de tecnologías digitales en la actividad turística no puede concebirse de manera aislada ni estandarizada, ya que su efectividad depende de múltiples factores contextuales. El desarrollo turístico constituye un sistema complejo que involucra variables como la localización geográfica, la trayectoria histórica, el clima, el nivel de desarrollo económico local, la calidad de los recursos naturales y paisajísticos, las políticas públicas y la infraestructura disponible [7].

De manera complementaria, los enfoques sobre Destinos Turísticos Inteligentes (DTI) subrayan la necesidad de alinear la adopción tecnológica con ejes estratégicos claramente definidos, con la capacidad institucional para gestionarlos y con la participación activa de múltiples actores locales [8][9].

En este entramado de variables, los destinos emergentes enfrentan el desafío de incorporar tecnologías digitales que acompañen su crecimiento sin comprometer su identidad ni su sostenibilidad. En particular, Caleta Olivia constituye una localidad de escala intermedia con una identidad cultural propia, lo que configura un entorno propicio para innovar en propuestas turísticas diferenciadas. En este sentido, los organismos gubernamentales y los actores territoriales impulsan estrategias orientadas a promover el turismo, fortalecer la identidad local y articular acciones con los ámbitos educativo y cultural, a fin de ampliar y consolidar su atractivo turístico.

En este marco, el presente proyecto se orienta a contribuir a la transformación digital del turismo en Caleta Olivia y en la región mediante el desarrollo de productos digitales estratégicos destinados a fortalecer la gestión del destino y a enriquecer la experiencia del visitante, en articulación con las capacidades institucionales y las particularidades del contexto local.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En este proyecto, el propósito del grupo GIFIS se centra en la identificación y aplicación de tecnologías para la Transformación Digital como soporte para el desarrollo del Turismo Local.

Para alcanzar este propósito se trabaja en las siguientes líneas de Investigación y Desarrollo:

- **Diseño Centrado en el Usuario para Turismo Digital.**
  Esta línea se orienta a la identificación, especificación y validación de requisitos funcionales y no funcionales para productos digitales turísticos, considerando principios de usabilidad, accesibilidad y experiencia de usuario (UX).
- **Desarrollo de Prototipos y Productos Digitales para la Gestión del Destino.**
  Esta línea se enfoca en el diseño, implementación y validación de soluciones tecnológicas orientadas a fortalecer la gestión del destino turístico y mejorar la experiencia del visitante.
- **Analítica de Datos y Apoyo a la Toma de Decisiones.**
  Esta línea propone la utilización de técnicas de análisis de datos para apoyar la planificación y gestión turística local.

- **Análisis de Tecnologías aplicadas al Turismo.** Esta línea se orienta a la identificación, análisis y evaluación de tecnologías emergentes con potencial impacto en el desarrollo turístico local y regional, con el propósito de anticipar oportunidades de innovación y orientar decisiones de posible inversión tecnológica.
- **Madurez Digital del Sector Público en Turismo.**
  Esta línea se orienta al diseño, validación y transferencia de un Modelo de Madurez Digital específico para organismos públicos con competencia en turismo, con el propósito de medir su nivel de desarrollo digital, identificar brechas y definir estrategias de mejora progresiva.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Este proyecto de investigación y desarrollo se enfoca en aportar a la transformación digital, proveyendo soporte estratégico al turismo local, a partir del desarrollo de productos software. Como resultado esperado se planteó el desarrollo de productos digitales cuyo objetivo sume valor a las propuestas ofrecidas por el sector turístico local, incorporando herramientas digitalizadas enriquecedoras y posibles de ser implementadas.

Como resultados obtenidos surge la familia de aplicaciones móviles "Caleta en un Click", orientada a brindar información y guía a los turistas. Hasta el momento se concretaron dos desarrollos:

1) **Caleta en un Click – Circuito Histórico**, funciona como guía turística digital, facilitando la exploración de lugares de valor histórico-cultural. Está disponible para dispositivos Android en Google Play Store, de forma gratuita[1] (Figura 1).
2) **Caleta en un Click – Circuito Flora**, actualmente en desarrollo, aborda un circuito de senderismo fuera del casco urbano (Figura 2). Como el recorrido carece de cartelería, la aplicación busca ofrecer material de apoyo a los guías turísticos, con imágenes y datos de las especies nativas y/o permitir que los visitantes recorran el circuito de manera autónoma. Está disponible para dispositivos Android en Google Play Store, de forma gratuita[2].



Figura 1. Pantalla inicial de la App "Caleta en un Click – Circuito Histórico"

[1] https://play.google.com/store/apps/details?id=com.caleta.turismo&hl=es_AR

[2] ttps://play.google.com/store/apps/details?id=com.kotlinnativo



Figura 2. Pantalla inicial de la App "Caleta en un Click – Circuito Flora"

Por otra parte, el grupo GIFIS ha conformado, en articulación con investigadores de otros proyectos, el Programa de Investigación UNPA denominado "*Propuestas de Innovación Tecnológica en Desarrollo de Herramientas Digitales Comunicacionales para la Puesta en Valor de Territorios Patrimoniales*". En el marco de este programa se desarrollaron y validaron diversos prototipos de productos digitales orientados a la valorización y difusión del patrimonio territorial, entre los que se destacan:

- Aplicación móvil para el reconocimiento de plantas autóctonas mediante técnicas de Inteligencia Artificial, destinada a fortalecer la educación ambiental y la interpretación del entorno natural.
- Aplicación móvil para recorridos turísticos guiados basada en servicios de geolocalización, orientada a mejorar la experiencia del visitante a través de información contextual en tiempo real.
- Aplicación móvil con estrategias de gamificación para la promoción y el aprendizaje interactivo en circuitos turísticos, incentivando la participación activa y el compromiso del usuario.

Estas experiencias consolidan la trayectoria del grupo en el diseño, desarrollo y evaluación de soluciones tecnológicas aplicadas a la puesta en valor del patrimonio cultural y natural, reforzando su capacidad para transferir conocimiento y productos digitales a actores territoriales.

4. FORMACIÓN DE RECURSOS HUMANOS

En la actualidad, GIFIS tiene nueve integrantes: seis docentes-investigadoras UNPA, una graduada de postgrado, cuatro alumnos de grado UNPA, un alumno becario de postgrado UNPA. Durante la ejecución de este proyecto un alumno se graduó como Ingeniero en Sistemas, siendo dirigido en el desarrollo de su Trabajo Final por docentes integrantes de GIFIS, y otro alumno se graduó como Analista de Sistemas. También se continúa con la capacitación y formación de postgrado de las integrantes docentes - investigadoras.

5. BIBLIOGRAFÍA

[1] Saldaño, V., Martin, A., Gaetán, G., Cardozo, C., Colombres, M., Arnedo, J., Montaño Torrico, C., (2025) Transformación Digital al Servicio del Turismo. Una experiencia de articulación Universidad-Gobierno local. VIII Jornadas Académicas de la REDVITEC: Transformación Digital y Tecnologías Emergentes.

[2] Montaño Torrico, C., Saldaño, V., Martin, A. (2023). Integrando Herramientas en una Aplicación Turística para el Aprendizaje de Realidad Aumentada y Geolocalización. ICT-UNPA, Vol. 14(3), pp. 142-169.
[3] Martin, A., Saldaño, V., Gaetán, G., Cardozo, C., Colombres, M., Sosa, H. (2022). Revalorizando el Patrimonio Histórico-Cultural de Caleta Olivia: Una Aplicación Móvil para Circuitos Turísticos y Educativos. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral.
[4] Cardozo, C., Martin, A., Saldaño, V., Gaetán, G. (2022). Aplicaciones de Turismo Cultural: Definiendo un Enfoque para el Desarrollo. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral.
[5] Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems Volume 28, Issue 2, Pages 118-144
[6] UNWTO. (2024). Digital Transformation | UN Tourism. Organización Mundial del Turismo.
[7] Cronjé, D. F., & du Plessis, E. (2020). A review on tourism destination competitiveness. Journal of Hospitality and Tourism Management, 45, 256-265.
[8] Gretzel, U. (2021). The Smart DMO: A new step in the digital transformation of destination management organizations. European Journal of Tourism Research.
[9] Wu, W., Xu, C., Zhao, M., Li, X., & Law, R. (2024). Digital tourism and smart development: State-of-the-art review. Sustainability, 16(23), 10382.

Procesamiento de Señales y Sistemas de Tiempo Real

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

CARACTERIZACIÓN DE TÉCNICAS DE SINCRONIZACIÓN HORARIA EN SISTEMAS DISTRIBUIDOS MEDIANTE GNSS Y RTC DE BAJO COSTO PARA TELEMETRÍA A ESCALA GLOBAL

María Macarena Repetto[1], Federico Ares[123], Sebastián Lucas[2], Nicolás Scivetti[12], Andrés Bilmes[12]

mrepetto@ing.unp.edu.ar {fares, slucas, nscivetti, abilmes}@cenpat-conicet.gob.ar

[1]Universidad Nacional de la Patagonia San Juan Bosco
[2]Instituto Patagónico de Geología y Paleontología
[3]Universidad Tecnológica Nacional – Facultad Regional Chubut

RESUMEN

La sincronización horaria precisa es esencial en sistemas distribuidos de telemetría y monitoreo remoto, donde se requiere precisión, bajo costo y bajo consumo energético. Este trabajo analiza un prototipo híbrido GNSS–RTC disciplinado por señal 1PPS, basado en una placa de desarrollo ESP32, que integra modelos dinámicos de reloj para estimar desplazamientos y variaciones de frecuencia respecto de la referencia GNSS. Se implementan estrategias de disciplinamiento que optimizan la estabilidad temporal y reducen el error de sincronización, incluso ante pérdida momentánea de referencia. Ensayos comparativos entre dispositivos independientes permiten cuantificar la incertidumbre relativa y demostrar la viabilidad de soluciones económicas para aplicaciones de telemetría distribuida y sistemas dispersos.

Palabras clave: tiempo global, modelos dinámicos, filtros de kalman, arquitectura embebida.

CONTEXTO

La presente línea de investigación se enmarca en las actividades del Grupo de Geomática Aplicada del Instituto Patagónico de Geología y Paleontología (IPGP- CCT CONICET-CENPAT) y la Facultad de Ingeniería de la Universidad Nacional de la Patagonia San Juan Bosco, sede Puerto Madryn.

1 INTRODUCCIÓN

Los relojes de tiempo real (RTC) ofrecen una solución de bajo costo y bajo consumo para mantener la hora en sistemas embebidos, aunque su precisión y estabilidad están condicionadas por la calidad del cristal de cuarzo interno, siendo susceptibles a deriva, envejecimiento y variaciones de temperatura [1].

La sincronización horaria constituye un componente crítico en sistemas distribuidos modernos, ya que permite correlacionar mediciones, ordenar eventos y garantizar coherencia temporal en procesos de adquisición y análisis de datos. Su relevancia se evidencia en aplicaciones como telemetría en ubicaciones remotas, monitoreo ambiental, instrumentación científica, redes eléctricas inteligentes y automatización industrial, donde

la correcta alineación temporal entre nodos asegura la integridad y utilidad de la información recolectada. Las soluciones basadas en NTP, si bien son ampliamente utilizadas a nivel global, requieren de conectividad IP estable, lo que no siempre es posible en escenarios rurales, marítimos o de difícil acceso. Además, al depender de la estimación estadística del retardo de ida y vuelta (RTT), su precisión puede degradarse significativamente ante variaciones en la latencia, congestión o asimetrías en los enlaces de comunicación, alcanzando errores del orden de milisegundos o incluso mayores en condiciones adversas [2][10].

Por otro lado, los Sistemas Globales de Navegación por Satélite (GNSS) proporcionan una referencia temporal de alta precisión basada en relojes atómicos a bordo de satélites, permitiendo reconstruir UTC con alta exactitud. La señal de pulso por segundo (1PPS) generada por receptores GNSS es ampliamente utilizada para aplicaciones que requieren precisión temporal, gracias a su baja deriva y estabilidad a largo plazo [3]. Sin embargo, su disponibilidad depende de la visibilidad satelital, y los receptores de bajo costo presentan variabilidad en forma de jitter de la señal 1PPS, así como retardos en interfaces serie como NMEA [4][5].

En este contexto, el desarrollo de esquemas híbridos GNSS–RTC se ha consolidado como una solución práctica, donde un RTC local mantiene la continuidad temporal mientras es disciplinado periódicamente por la señal 1PPS. Este enfoque permite un balance entre precisión, robustez y bajo costo, ideal para sistemas embebidos en entornos remotos, con escasa o nula cobertura de red. Investigaciones recientes aplican filtros avanzados, como Kalman o control adaptativo, para modelar el comportamiento no estacionario del reloj, corrigiendo offset y skew como procesos estocásticos [6][7]. El disciplinamiento puede realizarse a nivel de hardware —ajustando la frecuencia del oscilador interno [8][9]—, o a nivel de software mediante algoritmos de clock steering, e incluso mediante técnicas híbridas de ajuste digital en TCXO [11].

El estudio y desarrollo de estas técnicas forma parte de la investigación aplicada en sincronización horaria, orientada a optimizar precisión, estabilidad y eficiencia en sistemas distribuidos de adquisición de datos, abordando desafíos prácticos de filtrado, corrección de deriva y resolución temporal en plataformas embebidas de recursos limitados.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La investigación se centra en el estudio y desarrollo de soluciones de sincronización horaria para sistemas distribuidos, abordando desde el relevamiento del estado del arte en técnicas, arquitecturas y protocolos, hasta la definición de criterios, métricas y metodologías de evaluación de desempeño y robustez. Se contempla el diseño e implementación de un prototipo de hardware basado en ESP32 que integre un módulo GNSS y un RTC, junto con modelos dinámicos del reloj para estimar derivas y desplazamientos respecto de la señal 1PPS. Sobre esta base, se desarrollan estrategias de disciplinamiento del RTC y se realizan experimentos comparativos entre equipos independientes, con el fin de caracterizar su precisión, estabilidad y la incertidumbre relativa alcanzada.

3 RESULTADOS OBTENIDOS/ ESPERADOS

Al día de hoy se ha completado exitosamente la etapa de armado de hardware y el desarrollo de software para el manejo de comunicación entre el ESP32 y los módulos periféricos RTC y GNSS. Actualmente, el equipo se encuentra trabajando en la implementación del modelo de reloj mediante filtros de Kalman.

Se espera obtener una evaluación cuantitativa de la solución de sincronización horaria de bajo costo para dispositivos distribuidos basada en un receptor GNSS y un RTC. El resultado principal será un prototipo funcional capaz de mantener la hora con una precisión acotada y estable en el tiempo, utilizando hardware accesible y fácilmente replicable.

4 FORMACIÓN DE RECURSOS HUMANOS

El presente equipo de trabajo está conformado por un conjunto multidisciplinario que combina estudiantes, tutores y especialistas en investigación. La alumna de grado María Macarena Repetto, cursando la Licenciatura en Informática en la Facultad de Ingeniería de la UNPSJB, sede Puerto Madryn, lidera la ejecución del proyecto en el marco de su tesis de grado "Caracterización de técnicas de sincronización horaria en sistemas distribuidos mediante GNSS y RTC de bajo costo para telemetría a escala global" presentada formalmente en febrero de 2026, actualmente en proceso de evaluación. La supervisión y guía académica está a cargo del Ing. Federico Ares y del Lic. Lucas, quienes actúan como tutores, brindando apoyo técnico y metodológico. Además, el equipo cuenta con la colaboración de investigadores y docentes del CONICET y de la UNPSJB, Dr. Nicolás Scivetti y Dr. Andrés Bilmes, quienes aportan experiencia en investigación aplicada, y desarrollo experimental.

5 BIBLIOGRAFÍA

[1] Tirado-Andrés F., Araujo A., Performance of clock sources and their influence on time synchronization in wireless sensor networks, Int. J. Distrib. Sensor Networks, 2019.

[2] Novick, A. & Lombardi, M., Practical Limitations of NTP Time Transfer, IEEE FCS, 2015.

[3] Niu, K. et al., Quality evaluation of the pulse per second (PPS) signals from commercial GNSS receivers, GPS Solutions, 2015.

[4] u-blox AG, GPS-based Timing: Considerations with u-blox 6 GPS receivers, Application Note GPS.G6-X-11007, 2011.

[5] Veripos, Expected latency from PPS to NMEA ZDA telegram, 2025.

[6] Hamilton, R. et al., ACES: Adaptive Clock Estimation and Synchronization Using Kalman Filtering, MobiCom 2008.

[7] Li, Q. et al., An enhanced time synchronization method for a network based on Kalman filtering, Sci Rep, 2024.

[8] Riley, W. J., Handbook of frequency stability analysis, NIST SP 1065, 2008.

[9] Lombardi, M. A., GPS disciplined oscillators as primary frequency standards, NCSL Workshop, 2008.

[10] Mills, D. L., Internet time synchronization: The Network Time Protocol, IEEE Trans. Commun., 1991.

[11] Analog Devices, DS3231 Extremely Accurate RTC/TCXO, 2015.

Detección de anomalías en telemetría satelital utilizando Memoria Temporal Jerárquica

Pablo Soligo*, Jorge Ierache†, Martin Becerra‡ y Diego Sanz§

{*psoligo, †jierache, ‡mabecerra, §dsanz}@unlam.edu.ar

Grupo de Investigación y Desarrollo de Software Aeroespacial.
Departamento de Ingeniería e Investigaciones Tecnológicas.
Universidad Nacional de La Matanza.

RESUMEN

Este trabajo presenta los resultados preliminares de la aplicación de Memoria Temporal Jerárquica, del inglés Hierarchical Temporal Memory (HTM) para la detección de anomalías en telemetría satelital. Se exponen los resultados obtenidos al aplicar HTM a un conjunto de datos de telemetría satelital provisto por la Agencia Espacial Europea, del Inglés European Space Agency (ESA). El articulo muestra la capacidad de HTM para identificar patrones anómalos utilizando en este caso multiples canales. Los resultados sugieren que HTM es una herramienta prometedora para la detección temprana de fallas en sistemas espaciales especialmente debido a su potencial de implementación directamente en vuelo. *Palabras clave:* Detección de anomalías, telemetría satelital, memoria temporal jerárquica, sistemas espaciales.

CONTEXTO

La Universidad Nacional de La Matanza (UNLaM) como socio académico de la Comisión Nacional de Actividades Espaciales (CONAE) dicta una Maestría en Desarrollos Informáticos de Aplicación Espacial (MDIAE). Las experiencias realizadas en la maestría, tanto de manera directa, operando unidades de software de varias agencias, como mediante investigación general de las soluciones implementadas en el área, propiciaron la creación del Grupo de Investigación y Desarrollo de Software Aeroespacial (GIDSA). Este grupo está dedicado a investigar e implementar prototipos de software alternativos de bajo costo basados en las soluciones normalmente utilizadas en la industria del software de propósito general. Recientemente se vinculo esta investigación en el marco del proyecto SyMAT 2 UNLaM 2026-2027.

1. INTRODUCCIÓN

La ciencia de datos en general, el aprendizaje de máquina y la inteligencia artificial han generado en tiempo reciente un nuevo interés. Las técnicas de aprendizaje automático han captado la atención y se han vuelvo una de las áreas de mayor dinamismo en las ciencias de la computación. Los últimos avances, desarrollos e implementaciones en conjunto con el poder de cómputo actual representan nuevas oportunidades de aplicación. La detección automática y en tiempo real de anomalías forma parte de las herramientas disponibles en el universo de las ciencias de datos. En la industria espacial esta es una tarea crucial en las operaciones [1] donde históricamente se ha usado el control de límites [2] como su estrategia más extendida. Esta estrategia, aunque simple y eficiente, posee limitaciones; no contempla el contexto, ni los datos históricos y es sensible a datos extremos puntuales por errores de transmisión.

Algunas investigaciones sugieren que se puede observar en la telemetría histórica síntomas de las fallas futuras [3], la cantidad de datos y su naturaleza vuelven poco práctico el análisis manual realizado por un operador [4]. A medida que las misiones aumentan su tamaño, ya sea por complejidad del satélite o por cantidad (Constelaciones), más compleja y confusa puede ser la tarea de operadores e ingenieros, incluso hasta puede volverse impracticable [5].

La figura 1 muestra el crecimiento de la cantidad de variables (*channels*) que es necesario administrar en diferentes misiones de la ESA.La detección automáti-

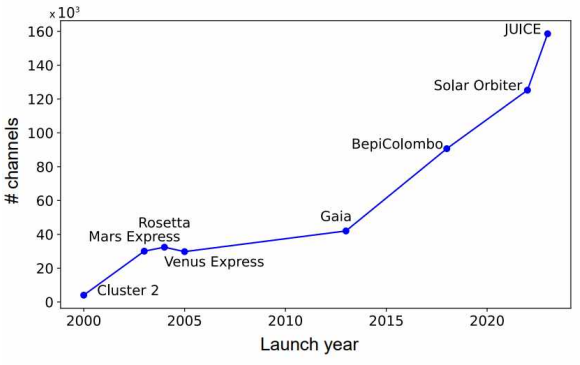


Figura 1: Evolución a traves de los años de la cantidad de variables para diferentes misiones de la ESA [2].

ca y en vuelo utilizando técnicas avanzadas ya es un requerimiento, explicitado por las agencias espaciales:

Moreover, potential solutions must be not only accurate but also fast enough to be run in real-time on computational resources accessible to mission control and, as the ultimate goal, on board satellites. [2]

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación han abarcado temas relacionados con el software terreno soporte y la incorporación de herramientas de uso extendido en la industria del software de propósito general incluyendo:

- La detección de fallas mediante técnicas de aprendizaje automático aplicadas a la telemetría satelital.
- La capa de visualización y operación basada en tecnologías web estándar [6].
- La recuperación y persistencia de datos utilizando bases de datos relacionales [7].

En este trabajo se ha profundizado en la primera línea de investigación, explorando técnicas de detección de anomalías en series temporales. Se ha realizado un primera aproximación utilizando HTM. HTM es una tecnología de inteligencia artificial inspirada en el funcionamiento del neocortex [8], es especialmente adecuada para la detección de anomalías debido a su capacidad para aprender patrones secuenciales. Además, puede operar en tiempo real con un costo computacional inferior a las estrategias de aprendizaje profundo.
La figura 2 muestra una representación conceptual de un modelo HTM simplificado. En la parte inferior se presenta en forma esquemática la entrada codificada que alimenta al modelo, en este caso un Codificador de Desplazamiento Aleatorio Distribuido, del inglés Random Distributed Scalar Encoder (RDSE). En la parte superior se representan el codificador espacial (*Spatial Pooler*) y la memoria temporal (*Temporal Memory*), que son los módulos principales de un modelo HTM. Las neuronas (*cells*) están organizadas en columnas (*Columns* o también descritas como *Minicolumns*), cada columna contiene múltiples neuronas. Las columnas pueden activarse en su totalidad (*Bursting*) o por el contrario si la fuerza de la conexiones generadas durante el proceso de aprendizaje lo determina solo la célula ganadora lo hará. El continuo ajuste de las fuerzas de la conexiones y los procesos de inhibición permiten que el modelo aprenda secuencias temporales de patrones de entrada.

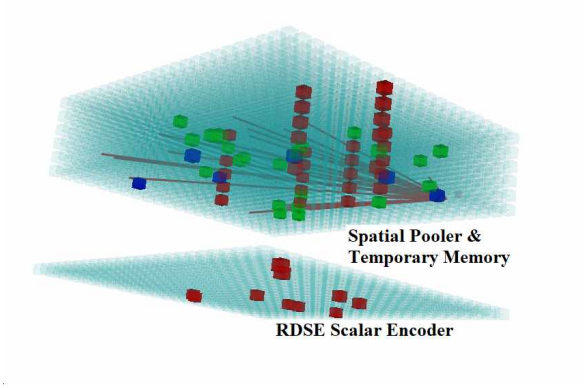


Figura 2: Representación conceptual del codificador, en este caso un codificador RDSE junto con un *SpatialPooler* y un *TemporaryMemory* en la sección superior. El color rojo indica una neurona activa, el azul una neurona predictiva y el verde una neurona ganadora. Se presenta una versión simplificada de alto nivel de un modelo HTM[9] con fines ilustrativos; la topología real, el total de neuronas y la cantidad de neuronas por columna puede variar respecto a esta esquematización. La figura es de generación propia utilizando la biblioteca pyvista[10].

En primer lugar se han desarrollado parcialemente módulos envoltorios que simplifiquen la exploración intentando respetar firmas de ambiente *sklearn*[11]. El objetivo final en este punto es simplificar una potencial integración.
Por otro lado se han desarrollado pruebas para evaluar la capacidad de detección de anomalías en series temporales, en una primera etapa usando señales sintéticas

con anomalías incrustadas artificialmente. La figura 3 muestra una señal senoidal y una evidente anomalía, se ha modificado el valor de la señal y se le ha agregado ruido gaussiano. El sector sombreado representa los puntos artificialmente modificados respecto a la señal original. Las barras muestran la propensión de anomalías detectadas por el modelo y los triángulos lugares donde se supera un umbral predefinido.

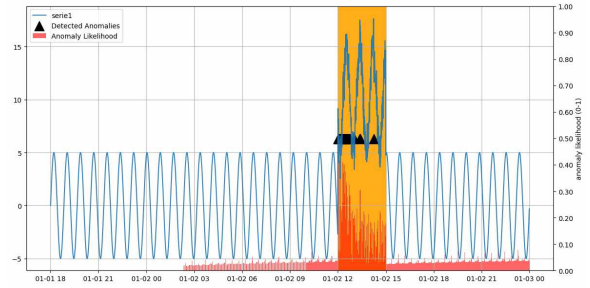


Figura 3: HTM[9] aplicado a una señal senoidal. Es sector sombreado representa las anomalías artificialmente generadas. Las barras muestran la propensión a una anomalías detectada por el modelo y los triangulos lugares donde se supera un umbral predefinido.

En segundo lugar y siguiendo las recomendaciones de la ESA se ha trabajado inicialmente sobre un conjunto de canales reducido (*lightweight subset of channels*)[2]. Cada canal ha sido estandarizado a números en el rango 0-1000 y se han codificado los canales utilizando el **ScalarEncoder** disponible en la biblioteca htm.core [9]. Se ha creado una única representación poco densa concatenando todos los canales aspirando a que la red reaccione a patrones en conjunto no vistos previamente. El código fuente se encuentra disponible en WICC2026_HTM_C44-C46.ipynb [12]. Las fechas no han sido codificadas por considerar que no tienen valor predictivo. La posibilidad de añadir información exógena, como periodos de eclipse o máximos solares ha sido descartada dado que las fechas han sido anonimizadas y no se conocen las características de la misión [2] [13].

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera que los algoritmos implementados o potenciales arquitecturas híbridas propuestas, logren un desempeño similar o superior a los benchmarks publicados por la ESA [2] a un costo computacional inferior que habilite su potencial implementación en vuelo. Hoy, variantes de Memoria a Largo Corto Plazo, del inglés Long Short-Term Memory (LSTM) como el Telemanom y Telemanom ESA Pruned ofrecen generalmente los mejores resultados aunque exigen datos etiquetados y pueden tener problemas para reconocer problemas no vistos previamente. Los algoritmos HTM han demostrado una buena capacidad para la detección de patrones sobre series de tiempo univariables y por tanto es razonable anticipar que su aplicación en la detección de fallas en sistemas espaciales será efectiva.

La figura 4 ilustra las pruebas exploratorias realizadas sobre los dataset de la ESA[13]. Sin explorar aún la optimización exhaustiva de hiperparámetros ni codificaciones de entradas alternativas, el modelo HTM muestra una capacidad prometedora para la detección de anomalías multicanal. En la figura 4 se puede observar una detección especialmente desafiante [2] utilizando los canales 44, 45 y 46 del dataset de la ESA [13]. El sector sombreado representa las anomalías etiquetadas por la ESA. Las barras muestran la propensión de anomalías detectadas por el modelo HTM y los triángulos lugares donde se supera un umbral predefinido mediante.

4. DISCUSIÓN

La métrica sugerida por la ESA para evaluar el desempeño de los modelos de detección de anomalías a nivel evento (*event-wise*) es una puntuación F con beta = 0.5 [2]. Aunque en estas pruebas el modelo HTM ha alcanzado una puntuación de 1.0, no se debe generar aquí una falsa ilusión de progreso. Los resultados son una primera aproximación que puede sufrir de sobre ajuste. La detección de la anomalía se realizó manteniendo el estado de aprendizaje durante toda la serie. La densidad realista de anotaciones cercana al 1.8 % (Sean anomalías o eventos raros) sobre un total de 774.856.895 puntos dificulta a la fecha el análisis exhaustivo de una parametrización con grados de libertad ajustados y por consiguiente la evaluación objetiva del desempeño del modelo.



Figura 4: HTM[9] aplicado a los canales 44, 45 y 46 del dataset de la ESA [13] mostrando capacidad de predicción de anomalías multicanal. Es sector sombreado representa las anomalías etiquetadas por la ESA. Las barras rojas al pie muestran la propensión de anomalías detectadas por el modelo y los triángulos lugares donde se supera un umbral predefinido a partir de ensayos.

5. TRABAJO A FUTURO

Las pruebas realizadas muestran una degradación en los tiempos de ejecución a medida que las series se vuelven mas largas. Problemas de sobre entrenamiento del modelo por parametrizaciones demasiado permisivas en términos de complejidad pueden estar contribuyendo a este comportamiento. El control de los periodos de exploración y explotación adecuados manteniendo constante la complejidad son aspectos clave a investigar.

Por otro lado, enfoques alternativos como combinar resultados de detecciones de anomalías individuales para cada canal, técnicas de preprocesamiento y reducción de la dimensionalidad vía Análisis de Componentes Principales, del inglés Principal Component Analysis (PCA) o suavizado de datos representan alternativas que merecen ser exploradas.

Como se mencionó en la sección 2, las fechas no fueron codificadas por no tener en principio carácter predictivo. Sin embargo, el tiempo entre muestras y la perdida de regularidad si puede ser indicativo de anomalía que no ha sido tenido en cuenta y que, adicionalmente, puede ser un auxiliar para cumplir con el requerimiento R8 [2] *"Algorithm should natively handle irregular timestamps and varying sampling rates"*. Este requerimiento, si bien no es considerado mandatorio figura entre las capacidades deseables de un sistema de detección automatizada de anomalías.

RECURSOS HUMANOS

El equipo esta conformado por dos investigadores formados, dos investigadores en formación y se encuentra en proceso de radicación una futura tesis de doctorado vinculada a la investigación presentada en este trabajo.

AGRADECIMIENTOS

Agradecemos al departamento de ingeniería e investigaciones tecnológicas (DIIT) de la UNLaM por apoyar las investigaciones

ACRÓNIMOS

MDIAE Maestría en Desarrollos Informáticos de Aplicación Espacial 1

UNLaM Universidad Nacional de La Matanza . 1

GIDSA Grupo de Investigación y Desarrollo de Software Aeroespacial 1

CONAE Comisión Nacional de Actividades Espaciales 1

ESA Agencia Espacial Europea, del Inglés European Space Agency 1

HTM Memoria Temporal Jerárquica, del inglés Hierarchical Temporal Memory 1

PCA Análisis de Componentes Principales, del inglés Principal Component Analysis . . 4

RDSE Codificador de Desplazamiento Aleatorio Distribuido, del inglés Random Distributed Scalar Encoder 2

LSTM Memoria a Largo Corto Plazo, del inglés Long Short-Term Memory 3

Referencias

[1] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. Anomaly detection in time series: a comprehensive evaluation. *Proceedings of the VLDB Endowment*, 15:1779–1797, 2022.

[2] Krzysztof Kotowski, Christoph Haskamp, Jacek Andrzejewski, Bogdan Ruszczak, Jakub Nalepa, Daniel Lakey, Peter Collins, Aybike Kolmas, Mauro Bartesaghi, Jose Martinez-Heras, et al. European space agency benchmark for anomaly detection in satellite telemetry. *arXiv preprint arXiv:2406.17826*, 2024.

[3] Kyle Hundman, Valentino Constantinou, Christopher Laporte, Ian Colwell, and Tom Soderstrom. Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 387–395, 2018.

[4] Takehisa Yairi, Minoru Nakatsugawa, Koichi Hori, Shinichi Nakasuka, Kazuo Machida, and Naoki Ishihama. Adaptive limit checking for spacecraft telemetry data using regression tree learning. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, volume 6, pages 5130–5135, 2004.

[5] L.Amoruso C.Brighenti, A.Brighenti D.Evans, M.Ricatto D.Moretto, and M.Carbone F.Ferrari. Advances in context aware spacecraft telemetry checking. In *2018 IAC*, 2018.

[6] Soligo Pablo and Ierache Jorge Salvador. Experiences and lessons learned developing a next-generation ground segment prototype. *2nd IAA Latin American Symposium on Small Satellites*.

[7] Soligo Pablo, Merkel German, and Ierache Jorge Salvador. Telemetría de altas prestaciones sobre base de datos de serie de tiempos. *REDDI 2021*, 2021.

[8] J. Hawkins, S. Ahmad, S. Purdy, and A. Lavin. Biological and machine intelligence (bami). Initial online release 0.4, 2016.

[9] M. Otahal, D. Keeney, D. McDougall, et al. HTM.core implementation of Hierarchical Temporal Memory. `https://github.com/htm-community/htm.core/`, 2019.

[10] Bane Sullivan and Alexander Kaszynski. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *Journal of Open Source Software*, 4(37):1450, May 2019.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[12] Soligo Pablo. Anomaly detection in satellite telemetry. `https://github.com/PabloSoligo2014/TlmyAnomaly_public.git`, 2025. Repositorio en GitHub.

[13] Gabriele De Canio, Krzysztof Kotowski, and Christoph Haskamp. Esa anomaly dataset, 6 2024.

EVALUACIÓN COMPARATIVA DE JSON Y CBOR PARA SISTEMAS INDUSTRIALES DE IIOT EN TIEMPO REAL SOBRE DISPOSITIVOS EDGE ESP32

Néstor Adrián Balich[1], Franco Adrián Balich[1],
Madiyar Zhunussov[2], Fabian Steinmetz[2]

[1] Laboratorio de Robótica Física e Inteligencia Artificial (LRFIA)
Centro de Altos Estudios en Tecnología Informática (CAETI)
Universidad Abierta Interamericana. Informática (UAI)
[2] Minerva University. 14 Mint Plaza Suite 300. San Francisco. United States
Montes de Oca 745. Ciudad Autónoma de Buenos Aires, Argentina.
{ nestor.balich, francoadrian.balich } @uai.edu.ar
{ zhunussov, fabian } @ uni.minerva.edu

RESUMEN

En entornos de Internet Industrial de las Cosas (IIoT), la eficiencia en la comunicación impacta directamente en el consumo energético, el determinismo, la escalabilidad y el desempeño en tiempo real. Aunque JSON continúa siendo el estándar de facto para el intercambio de datos debido a su legibilidad e interoperabilidad, su naturaleza textual impone una sobrecarga computacional significativa en sistemas embebidos con recursos limitados.

Este trabajo presenta una evaluación experimental comparativa entre JSON y CBOR (Concise Binary Object Representation) sobre un nodo edge basado en ESP32 operando a 240 MHz. Se realizaron cincuenta pruebas de benchmark utilizando el protocolo MQTT, midiendo tamaño de carga útil, ciclos de CPU y latencia de parseo.

Los resultados muestran que CBOR reduce el tamaño del payload en un 28,2 %, disminuye los ciclos de CPU aproximadamente un 99,8 % y mejora la latencia de parseo en un factor de 650× respecto de JSON. Los hallazgos demuestran que CBOR incrementa significativamente la viabilidad en tiempo real, la eficiencia energética y la escalabilidad en aplicaciones de IIoT (Internet industrial de las cosas).

Palabras clave: IIoT, CBOR, JSON, Computación en el Borde, Sistemas en Tiempo Real.

1. CONTEXTO Y MOTIVACIÓN

Los sistemas de IIoT industrial dependen cada vez más de nodos distribuidos en el borde (edge computing) [10], responsables de la adquisición de datos, el preprocesamiento y la actuación bajo estrictas restricciones temporales. Aplicaciones como mantenimiento predictivo, control robótico, automatización industrial y sistemas autónomos requieren baja latencia y comunicación determinística.

JSON, definido por la IETF en el RFC 8259 [7], se ha convertido en el formato de serialización dominante debido a su

legibilidad humana y amplia adopción en arquitecturas web [1]. Sin embargo, en entornos industriales embebidos, su codificación basada en caracteres introduce complejidad de parseo, fragmentación de memoria y mayor utilización de CPU [3].

CBOR (Concise Binary Object Representation), estandarizado en el RFC 8949 por la IETF [2], fue diseñado específicamente para entornos con recursos restringidos. Proporciona una codificación binaria sin esquema optimizada para compacidad y parseo rápido por máquina.

Este estudio evalúa si CBOR ofrece ventajas medibles sobre JSON en condiciones reales de edge industrial utilizando una plataforma de microcontrolador ampliamente adoptada: el ESP32 [9]. Diversos estudios han analizado formatos binarios de serialización para IIoT restringido [3], [4], [5], pero existen pocas evaluaciones experimentales enfocadas específicamente en escenarios industriales en tiempo real.

Este trabajo se enmarca dentro de la líneas de investigación de IoT, IIoT, Robótica y procesamiento en tiempo real.

1.1 Contribución del Trabajo

- Evaluación empírica en hardware real industrial (ESP32 240 MHz)
- Medición directa de ciclos de CPU (no simulación)
- Análisis de impacto en lazo de control de 100 Hz
- Discusión sistémica (determinismo + energía + escalabilidad)

2. METODOLOGÍA EXPERIMENTAL

2.1 Arquitectura de prueba

- Nodo emisor: Aplicación basada en Python
- Protocolo: MQTT [8]
- Nodo receptor: ESP32 @ 240 MHz
- Red: Wi-Fi
- Cantidad de pruebas: 50 por formato de serialización.
- Carga útil: Estructura de mensaje de detección de objetos.

2.2 Métricas evaluadas

- Sobrecarga de red (tamaño del payload en bytes)
- Ciclos de CPU requeridos para deserialización.
- Tiempo de ejecución (microsegundos)

Los ciclos de CPU fueron medidos directamente en el ESP32. La latencia se obtuvo utilizando temporizadores con resolución de microsegundos.

3. RESULTADOS EXPERIMENTALES

3.1 Tamaño de payload

La Tabla 1 resume el tamaño promedio de carga útil obtenido para cada formato.

Formato	Tamaño promedio
JSON	**~709.8 bytes**
CBOR	**~509.9 bytes**

Tabla 1. Comparación entre JSON y CBOR.

Como se muestra en la figura 1, CBOR logra una reducción significativa en los bytes transmitidos en comparación con JSON.

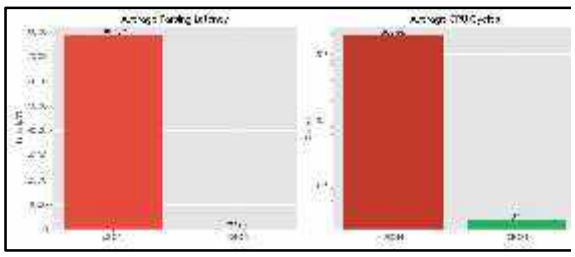


Fig. 1. Comparación del tamaño medio de la carga útil entre JSON y CBOR..

CBOR logró una reducción del 28,2 % en el tamaño de la carga útil. La mejora relativa se ilustra en la figura 2.

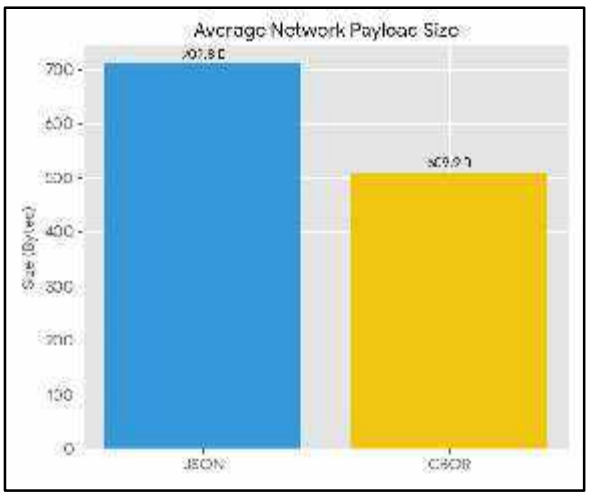


Fig. 2. Reducción relativa del tamaño de la carga útil lograda por CBOR en la transmisión IIoT.

Efecto en la industria:

La reducción del tiempo de transmisión disminuye directamente la duración de la radio, lo que reduce el consumo de energía y mejora el rendimiento de la red en implementaciones densas de IIoT. Como se muestra en la figura 3, CBOR reduce drásticamente la sobrecarga computacional.

3.2 Sobrecarga computacional

La Tabla 2 presenta los ciclos promedio de CPU medidos durante la deserialización.

Formato	Ciclos promedio de CPU
JSON	~1,870,000 ciclos
CBOR	~3,100 ciclos

Tabla 2. Promedio de ciclos de CPU necesarios para la deserialización.

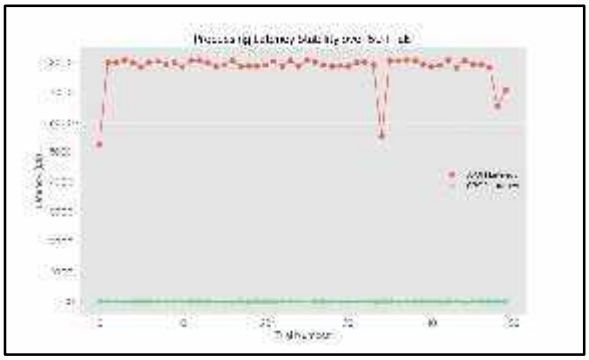


Fig. 3. Comparación del ciclo de CPU para la deserialización de JSON y CBOR en ESP32.

CBOR demostró una reducción del 99,8 % en ciclos de CPU.

El parseo de JSON requiere tokenización carácter por carácter, asignación dinámica de memoria y reconstrucción jerárquica de estructuras. En contraste, CBOR utiliza identificadores binarios de tipo que permiten una interpretación directa a nivel de bytes, evitando el procesamiento de cadenas.

Implicaciones industriales:

En dispositivos periféricos que funcionan de forma continua, la reducción del uso de la CPU permite:

- Mayor autonomía de batería.
- Menor estrés térmico.
- Mayor capacidad de cómputo disponible para análisis local.
- Mejor escalabilidad del sistema

3.3 Latencia y viabilidad en tiempo real

La Tabla 3 presenta la latencia promedio de parseo.

Formato	Tiempo promedio de parseo
JSON	**~7.8 ms**
CBOR	**~12 µs**

Tabla 3. Latencia media de análisis medida en ESP32

CBOR mostró una mejora de rendimiento de aproximadamente 650 en latencia de deserialización.

En un lazo de control de 100 Hz (presupuesto temporal de 10 ms):

- JSON consume aproximadamente el 78 % del ciclo.
- CBOR consume aproximadamente el 0,12 % del ciclo.

Conclusión:
JSON introduce una latencia no despreciable que puede comprometer el control determinístico en sistemas industriales. CBOR permite un desempeño cercano al tiempo real, compatible con lazos de control robótico o de automatización.

4. ANÁLISIS E IMPLICANCIAS INDUSTRIALES

Los resultados experimentales revelan que la limitación dominante en la comunicación industrial periférica no es el ancho de banda de transmisión, sino la sobrecarga computacional asociada con la deserialización. Aunque a menudo se hace hincapié en la optimización de la red en las arquitecturas IIoT, este estudio demuestra que el coste de procesamiento en el nodo periférico tiene un impacto significativamente mayor en el determinismo y la viabilidad en tiempo real.

En microcontroladores con recursos limitados, como el ESP32, el análisis JSON requiere tokenización a nivel de caracteres, asignación dinámica de memoria y reconstrucción de la estructura jerárquica [3]. Estas operaciones introducen fragmentación del paquete, mayor utilización de la CPU y tiempo de ejecución no determinista. En escenarios de automatización industrial en los que los bucles de control funcionan a 50-200 Hz, dicha variabilidad puede comprometer la estabilidad del sistema y las garantías temporales.

Por el contrario, CBOR emplea un esquema de codificación binaria con identificadores de tipo predefinidos [2], lo que permite una interpretación directa a nivel de bytes sin procesamiento de cadenas. Esta simplicidad estructural reduce drásticamente la complejidad del análisis y minimiza la variabilidad del tiempo de ejecución. El tiempo de ejecución casi constante observado en los experimentos sugiere una mayor previsibilidad, lo cual es fundamental para los sistemas industriales deterministas.

Desde el punto de vista energético, la deserialización intensiva en CPU aumenta directamente el tiempo de procesamiento activo y el consumo de energía [6]. En implementaciones industriales alimentadas por baterías o sensibles a la energía, como nodos de detección distribuidos o plataformas robóticas móviles, la reducción de la carga computacional se traduce en mejoras cuantificables en la autonomía

operativa y la gestión térmica.

La escalabilidad es otra dimensión crítica en las arquitecturas de IoT industrial [6]. En entornos IIoT densos con cientos o miles de nodos periféricos que publican telemetría simultáneamente, la reducción de los ciclos de CPU por mensaje permite un mayor rendimiento agregado y una menor fluctuación del sistema. Por lo tanto, la reducción de la sobrecarga de procesamiento observada con CBOR contribuye no solo a la eficiencia de los nodos individuales, sino también a la optimización del rendimiento a nivel de red.

Estos hallazgos indican que la selección del formato de serialización debe tratarse como una decisión arquitectónica a nivel de sistemas en el diseño del IoT industrial, en lugar de una mera elección de interoperabilidad. Si bien JSON sigue siendo ventajoso para la depuración, el registro y las interfaces legibles por humanos, CBOR demuestra una clara superioridad técnica en aplicaciones industriales periféricas de alta frecuencia, en tiempo real y con restricciones energéticas.

El experimento revela que el cuello de botella dominante en la comunicación periférica no es el ancho de banda de transmisión, sino la sobrecarga de deserialización.

En escenarios de IoT industrial, la sincronización determinista y la eficiencia energética son fundamentales [6].

El exceso de ciclos de CPU afecta directamente a:

- Consumo de potencia
- Fluctuación del sistema
- Límites de rendimiento de mensajes Previsibilidad de los sistemas de control

CBOR mejora significativamente:

Eficiencia energética

- Escalabilidad de mensajes (rendimiento potencial entre 30 y 50 veces superior).
- Comportamiento determinista.
- Idoneidad en tiempo real.

5. FORMACIÓN DE RECURSOS HUMANOS

El presente trabajo se enmarca en una estrategia institucional orientada al fortalecimiento de la formación en tecnologías de IoT industrial y robótica aplicada, promoviendo la integración entre docencia, investigación y cooperación académica internacional.

El equipo estuvo conformado por dos docentes-investigadores especialistas en IoT y Robótica pertenecientes al Laboratorio de Robótica Física e Inteligencia Artificial (LRFIA) de la Universidad Abierta Interamericana, junto con dos estudiantes pasantes de Minerva University, en el marco de un programa de intercambio académico y colaboración interuniversitaria.

Durante el desarrollo del proyecto, los estudiantes participaron activamente en:

- Programación de microcontroladores ESP32 en entornos de sistemas embebidos.

- Implementación de protocolos de comunicación orientados a IIoT (MQTT).
- Evaluación comparativa de formatos de serialización en escenarios industriales.
- Diseño experimental y medición de desempeño en tiempo real.
- Análisis crítico de resultados y elaboración de producción científica conjunta.

La experiencia permitió consolidar competencias técnicas en arquitectura de sistemas IoT industriales, al tiempo que fortaleció habilidades en metodología de investigación, análisis cuantitativo y redacción académica.

Asimismo, el proyecto fomentó el intercambio internacional y la producción científica colaborativa entre la Universidad Abierta Interamericana y Minerva University, promoviendo la participación estudiantil en proyectos de investigación aplicada con impacto tecnológico.

Este trabajo constituye, por lo tanto, no solo un aporte técnico en el área de Arquitectura, Redes y Sistemas Operativos, IIoT y Robótica, sino también una instancia concreta de formación de recursos humanos orientada a la investigación y la internacionalización académica.

6. CONCLUSIONES

Esta evaluación experimental demuestra que CBOR ofrece ventajas sustanciales de rendimiento con respecto a JSON en sistemas IoT industriales basados en ESP32.

Conclusiones principales:

- 28.2% menos de carga útil
- 99.8% menos de uso de CPU
- 650 veces más rápido en el análisis sintáctico

Para la automatización industrial IoT y robótica, los sistemas autónomos y las aplicaciones de telemetría de alta frecuencia, CBOR representa una opción de serialización técnicamente superior.

JSON sigue siendo adecuado para arquitecturas web legibles por humanos y no críticas; sin embargo, en entornos industriales periféricos con recursos limitados, CBOR permite sistemas escalables, energéticamente eficientes y con capacidad en tiempo real. Este trabajo aporta evidencia empírica cuantitativa sobre el impacto real del formato de serialización en el determinismo temporal de microcontroladores industriales, demostrando que la elección del formato debe considerarse una decisión arquitectónica crítica en el diseño de sistemas IIoT en tiempo real.

7. BIBLIOGRAFÍA

[1] Kaur, A., Ayyagari, S., Mishra, M., & Thukral, R. (2020). A literature review on device-to-device data exchange formats for iot applications. *Journal of Intelligent Systems and Computing*, *1*(1), 1-10. link

[2] Bormann, C., & Hoffman, P. (2020). RFC 8949 concise binary object representation (CBOR). *Internet Engineering Task Force (IETF)*. link

[3] Viotti, J. C., & Kinderkhedia, M. (2022). A benchmark of JSON-compatible binary serialization specifications. *arXiv preprint arXiv:2201.03051*. link

[4] Lenders, M. S., Bormann, C., Schmidt, T. C., & Wählisch, M. (2025). A Leaner and Faster Web: How CBOR Can Improve Dynamic Content Encoding in JSON and DNS over HTTPS. *arXiv preprint arXiv:2512.12067*. link

[5] Luis, Á., Casares, P., Cuadrado-Gallego, J. J., & Patricio, M. A. (2021). PSON: A serialization format for IoT sensor networks. *Sensors*, *21*(13), 4559. Link

[6] Ashish M. Jadhav, "Communication Protocols for IoT Applications," IJRASET, 2024.
General analysis of IoT protocols y comunicaciones.
https://www.ijraset.com/research-paper/communication-protocols-for-iot-applications

[7] T. Bray, "The JavaScript Object Notation (JSON) Data Interchange Format," RFC 8259, IETF, Dec. 2017. [Online]. Available: https://www.rfc-editor.org/rfc/rfc8259

[8] OASIS, "MQTT Version 3.1.1," OASIS Standard, Oct. 2014. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html

[9] Espressif Systems, "ESP32 Technical Reference Manual," 2023. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32_technical_reference_manual_en.pdf

[10] M. Satyanarayanan, "The Emergence of Edge Computing," Computer, vol. 50, no. 1, pp. 30–39, Jan. 2017. doi: 10.1109/MC.2017.9.

IMPLEMENTACIÓN DE INTERFACES CEREBRO-COMPUTADORA INTEGRANDO BIOFEEDBACK EEG EN VIDEOJUEGOS

Berenice Lourdes Balich, Carla Romina Galizio, Néstor Adrián Balich

Laboratorio de Robótica Física e Inteligencia Artificial (LRFIA), Centro de Altos Estudios en Tecnología Informática (CAETI), Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{Nestor.Balich,BereniceLourdes.Balich}@uai.edu.ar
CarlaRomina.Galizio@alumnos.uai.edu.ar

RESUMEN

El presente trabajo describe el diseño e implementación de una arquitectura modular para la integración de una Interfaz Cerebro-Computadora (BCI) basada en señales EEG con el motor de videojuegos Godot, orientada al desarrollo de entornos interactivos controlados mediante biofeedback en tiempo real.

La propuesta contempla la adquisición de señales neurofisiológicas mediante una vincha EEG comercial (Orbit/MindWave), su procesamiento a través de un middleware desarrollado en Python y su transmisión mediante protocolos de red de baja latencia hacia el motor gráfico. En el entorno virtual, las métricas cognitivas son mapeadas a parámetros dinámicos que modifican iluminación, ambientación y mecánicas de interacción.

Diversos estudios han demostrado la viabilidad del uso de BCI en videojuegos y sistemas interactivos [1][2][3][11], así como el potencial del procesamiento en tiempo real para experiencias adaptativas [5][8]. El proyecto tiene carácter experimental y no persigue fines clínicos. Se orienta a la exploración de nuevas modalidades de interacción humano-computadora en videojuegos neurocontrolados.

***Palabra clave**: BCI, EEG, videojuegos interactivos, biofeedback, interacción humano-máquina.*

CONTEXTO

El presente proyecto se desarrolla en el marco de las líneas de investigación en procesamiento de señales + sistemas de tiempo real , video juegos y tecnología aplicada.

Desde el punto de vista disciplinar, el trabajo se encuadra principalmente dentro de la línea de investigación PSSTR – Procesamiento de Señales y Sistemas de Tiempo Real, dado que aborda la adquisición, tratamiento y transmisión de señales electroencefalográficas (EEG) bajo restricciones temporales estrictas. La arquitectura propuesta contempla procesamiento digital de señal, normalización de métricas cognitivas, reducción de ruido y sincronización en tiempo real con un sistema interactivo.

El sistema implementado requiere garantizar baja latencia, estabilidad en la transmisión y coherencia temporal entre la señal

neurofisiológica y la respuesta visual del entorno virtual, aspectos centrales en sistemas de tiempo real y ampliamente documentados en la literatura sobre BCI [6][7][9].

Asimismo, el proyecto articula conceptos de interacción humano-computadora y videojuegos adaptativos [1][3][10] [11], pero su eje técnico principal radica en el procesamiento de señales biométricas y su integración en arquitecturas distribuidas de ejecución interactiva.

El trabajo tiene carácter experimental y no persigue fines clínicos ni diagnósticos. Se orienta a la exploración de nuevas modalidades de interacción basadas en biofeedback emocional, promoviendo desarrollos tecnológicos innovadores en el campo del neurogaming y los sistemas interactivos adaptativos.

1. Introducción

En este contexto se desarrolló “Feel or Fade”, un prototipo de videojuego experimental de género narrative-cozy que incorpora tecnología de biofeedback emocional para transformar el entorno virtual en tiempo real. El jugador explora una casa cálida y acogedora que presenta cambios dinámicos en función de sus estados cognitivos medidos mediante métricas EEG obtenidas con la vincha MindWave MW003.

La experiencia propone representar el equilibrio entre calma, ansiedad y concentración, evidenciando cómo el estado emocional interno puede influir en la percepción del entorno externo. El diseño del juego se basa en que jugador recorre una vivienda que simboliza metafóricamente su mente. En cada habitación se encuentran fragmentos de recuerdos materializados en objetos cotidianos (como por ejemplo una taza, una carta, una fotografía o un cuaderno) que, al ser activados, permiten reconstruir progresivamente el equilibrio emocional del personaje.

El diseño busca abordar la variabilidad emocional desde una perspectiva artística y no clínica, mostrando cómo estados como la ansiedad o la falta de concentración forman parte de la experiencia humana cotidiana, en concordancia con investigaciones que exploran el uso de videojuegos como herramientas de regulación emocional y apoyo en contextos de salud mental [14].
La casa funciona como un espejo emocional: sus luces, colores, objetos y ambientación sonora se modifican dinámicamente en función de las métricas neurofisiológicas recibidas, implementando un sistema de biofeedback interactivo en línea con investigaciones previas sobre integración entre videojuegos y señales EEG [3] [11].

Desde el punto de vista técnico, estas transformaciones se encuentran sincronizadas con el procesamiento en tiempo real de la señal, garantizando coherencia perceptual entre estado cognitivo y respuesta gráfica [8].

2.Línea de investigación y desarrollo

2.1 Conexión vía Bluetooth hacia el sistema de host

La vincha EEG (Figura 1) establece una conexión inalámbrica Bluetooth con el sistema host, permitiendo la transmisión continua de señales neurofisiológicas en tiempo real. Esta comunicación constituye la capa primaria de adquisición dentro de la arquitectura propuesta.
El enlace Bluetooth garantiza movilidad del usuario y estabilidad en la transmisión de

datos, permitiendo la recepción de métricas cognitivas con latencia reducida. La arquitectura contempla mecanismos de validación básica de integridad de paquetes, asegurando consistencia en el flujo de datos antes de su procesamiento.

Desde la perspectiva de sistemas de tiempo real, la etapa de adquisición resulta crítica, ya que cualquier pérdida o retraso en la señal impacta directamente en la coherencia perceptual del biofeedback [6]. Por ello, la comunicación se configura como una capa de transporte optimizada para disponibilidad inmediata de la señal, habilitando su posterior filtrado, normalización e interpretación dentro del sistema interactivo.

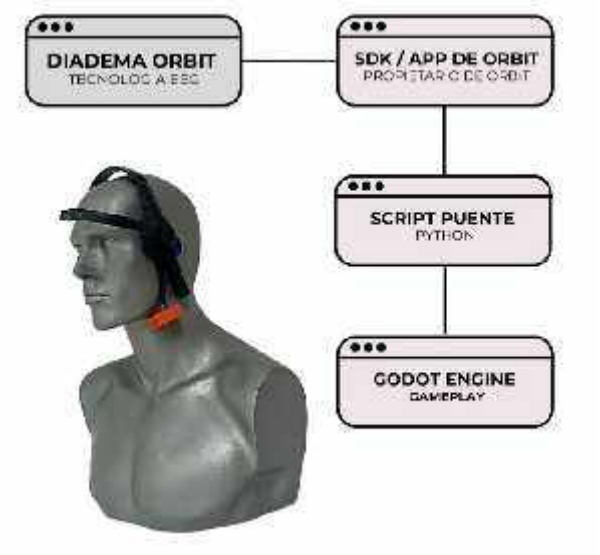


Figura 1. Visualización mapa conceptual de la arquitectura del sistema.

2.2 Desarrollo de la herramienta Script intermediario en Python

Se implementó un módulo intermediario en Python que actúa como middleware entre la capa de adquisición EEG y el motor de videojuegos. Este componente constituye el núcleo lógico del sistema, ya que desacopla el hardware de captura del entorno de aplicación, favoreciendo una arquitectura modular, mantenible y escalable.

El middleware recibe en tiempo real los datos generados por el SDK del dispositivo y ejecuta procesos iniciales de preprocesamiento orientados a:

- Estabilización de la señal.
- Reducción básica de ruido.
- Normalización de métricas cognitivas.
- Validación de integridad de datos.

Estos procedimientos se alinean con enfoques generales de procesamiento EEG descritos en la literatura [4][9], adaptados a un entorno interactivo no clínico.

Posteriormente, los datos neurofisiológicos son transformados en estructuras digitales estandarizadas (formato JSON), preparadas para su transmisión mediante protocolos de baja latencia (WebSocket/UDP). El middleware implementa control de frecuencia de actualización y encapsulamiento eficiente de paquetes para garantizar sincronización con el motor gráfico.

Desde el punto de vista arquitectónico, esta capa actúa como abstracción lógica entre adquisición y aplicación, evitando dependencias directas entre el motor de juego y el hardware EEG. Esta separación:

- Mejora la reutilización del sistema.
- Simplifica tareas de mantenimiento.
- Permite la incorporación futura de nuevos dispositivos o algoritmos de clasificación sin modificar la lógica del videojuego.

El diseño prioriza eficiencia temporal, estabilidad del flujo de datos y consistencia en la experiencia interactiva bajo restricciones de tiempo real.

2.3 Implementación de la Interfaz de Usuario en Godot

La interfaz de usuario fue implementada como la capa final de interacción entre el sistema neurocontrolado y el jugador.

El motor Godot fue seleccionado por su naturaleza open-source, su arquitectura basada en nodos y escenas, su bajo consumo de recursos y su facilidad para la integración mediante comunicación en red (WebSocket/UDP). Estas características resultan especialmente adecuadas para aplicaciones interactivas en tiempo real que requieren sincronización eficiente con sistemas externos de adquisición de datos, como dispositivos EEG [12][13].

Esta capa traduce los datos EEG procesados por el middleware en respuestas visuales y mecánicas dentro del entorno virtual.

La arquitectura de integración se basa en un modelo orientado a eventos [10], donde las métricas cognitivas recibidas en tiempo real son mapeadas a variables internas del sistema, enfoque que ha demostrado flexibilidad en entornos de simulación e investigación interactiva desarrollados con Godot [12].

Dichas variables modifican dinámicamente:

- Iluminación y temperatura cromática del entorno.
- Intensidad y tipo de ambientación sonora.
- Aparición o desaparición de objetos interactivos.
- Efectos visuales de desenfoque o suavizado.

Cuando se detectan estados asociados a relajación, el entorno se vuelve más luminoso y armónico; ante métricas vinculadas a ansiedad, se producen variaciones ambientales controladas; en casos de baja concentración, algunos elementos interactivos se difuminan temporalmente, incentivando el reenfoque del jugador.

La interfaz no cumple únicamente una función estética, sino que implementa un sistema de retroalimentación de biofeedback, permitiendo al usuario percibir la relación entre su actividad cognitiva y la respuesta del entorno virtual, en concordancia con estudios sobre videojuegos y EEG [3].

A nivel técnico, la implementación aprovecha el sistema de escenas, nodos y señales internas del motor Godot, garantizando bajo costo computacional y respuesta inmediata. Los valores neurofisiológicos son convertidos en variables de estado que se actualizan dentro del ciclo principal del motor, asegurando sincronización temporal entre señal y representación gráfica.

Para reducir el impacto del ruido fisiológico en la experiencia visual, se aplican mecanismos de interpolación y suavizado, lo que mejora la estabilidad perceptual y evita cambios abruptos no deseados. Este enfoque asegura:

- Baja latencia.
- Coherencia interactiva.
- Modularidad estructural.
- Escalabilidad futura del sistema.

3. Resultados

3.1 Resultados Obtenidos

Los resultados presentados corresponden a una etapa inicial de validación arquitectónica y funcional del sistema, orientada a comprobar la viabilidad técnica de la integración EEG–Videojuego. La implementación de estas primeras etapas permitió verificar el correcto funcionamiento de la arquitectura

propuesta para la integración de señales EEG en un entorno interactivo en tiempo real. Se obtuvieron los siguientes resultados:

- Integración funcional del sistema EEG–Godot, logrando comunicación estable entre la vincha Orbit, el middleware desarrollado en Python y el motor de videojuegos Godot, validando la arquitectura modular planteada [5].
- Transmisión continua de datos neurofisiológicos en tiempo real, permitiendo la recepción sostenida de métricas cognitivas sin pérdidas significativas de información durante la ejecución del sistema.
- Sincronización temporal adecuada entre adquisición, procesamiento y representación gráfica, garantizando coherencia perceptual en la experiencia de biofeedback.
- Respuesta interactiva consistente, donde las variaciones en métricas de atención y relajación se reflejaron de manera visible y controlada en el entorno virtual.
- Funcionamiento efectivo de la capa de biofeedback, confirmando que el usuario puede percibir la relación entre su actividad cerebral y la respuesta dinámica del sistema, aspecto central en aplicaciones BCI interactivas en tiempo real [5][8].

Los resultados obtenidos demuestran la viabilidad técnica de integrar dispositivos EEG comerciales en arquitecturas de procesamiento de señales bajo restricciones de tiempo real.

3.2 Resultados Esperados

En función del estado actual del desarrollo, se proyectan los siguientes resultados:

- Consolidación de la integración estable EEG–Videojuego, garantizando comunicación continua entre la vincha, el middleware en Python y el motor Godot, con mejoras en sincronización y tolerancia a fluctuaciones de señal.
- Optimización del procesamiento de señales [9], incorporando mejoras en filtrado y normalización que permitan reducir el impacto del ruido y aumentar la estabilidad de las métricas neurofisiológicas.
- Evaluación cuantitativa de desempeño, incluyendo medición de latencia extremo a extremo, estabilidad de transmisión y frecuencia de actualización.
- Refinamiento del sistema de biofeedback, validando experimentalmente la percepción del usuario respecto a la influencia de su actividad cerebral en el entorno interactivo.
- Escalabilidad de la arquitectura, posibilitando la incorporación futura de nuevos dispositivos EEG, algoritmos de clasificación emocional o mecánicas adicionales sin necesidad de rediseñar el sistema base.
- Implementación de arte y mecánicas para versión preliminar del juego.

Estas proyecciones consolidan la línea de investigación dentro del ámbito de Procesamiento de Señales y Sistemas de Tiempo Real, fortaleciendo la aplicabilidad del modelo propuesto en futuros desarrollos experimentales.

4. Formación de Recursos Humanos

El equipo de trabajo está conformado por:

- Dos docentes investigadores especialistas en IoT, robótica e impresión 3D, responsables de la supervisión técnica, metodológica y científica del proyecto, así como del diseño de la arquitectura del sistema y la validación de resultados.
- Un estudiante de Ingeniería en Sistemas y un estudiante de la Licenciatura en Producción de Simuladores y Videojuegos, quienes participaron activamente en el desarrollo e implementación del prototipo experimental.

Durante el desarrollo del proyecto se promovió una formación práctica interdisciplinaria, incorporando contenidos y competencias en:

- Desarrollo del prototipo en el motor Godot.
- Comunicación IoT y protocolos de transmisión de datos en tiempo real.
- Integración hardware-software.
- Lectura e interpretación de sensores biológicos (EEG).
- Procesamiento básico de señales neurofisiológicas.
- Diseño de arquitecturas modulares con middleware.
- Programación en Python para sistemas distribuidos.
- Implementación de sistemas interactivos adaptativos.

El proyecto implicó la articulación de diversas disciplinas, entre ellas:

- Procesamiento de señales.
- Sistemas de tiempo real.
- Arquitectura de software.
- Interacción humano-computadora.
- Computación gráfica aplicada a videojuegos.
- Sistemas embebidos e IoT.

Asimismo, se desarrollaron capacidades técnicas y profesionales tales como:

- Diseño e implementación de sistemas distribuidos.
- Análisis de latencia y sincronización.
- Trabajo colaborativo interdisciplinario.
- Metodología de investigación aplicada.
- Resolución de problemas en entornos experimentales.
- Integración de tecnologías emergentes en proyectos académicos.

La experiencia permitió fortalecer competencias tanto técnicas como investigativas, consolidando la formación en tecnologías biométricas aplicadas a entornos interactivos en tiempo real.

Bibliografía:

[1] Introducción / Estado del arte sobre EEG y videojuegos. Webb, N. (2023). *An Introduction to BCI and Its Use in Video Games: A Review.* DOI: https://doi.org/10.47611/jsrhs.v12i1.3924

[2] EEG para investigación de experiencia en juegos / análisis. Wehbe, R., & Nacke, L. (2014). *An Introduction to EEG Analysis Techniques and Brain-Computer Interfaces for Games User Researchers.* DOI: https://doi.org/10.26503/dl.v2013i1.641

[3] Revisión tecnológica de EEG y videojuegos (serious games / integración). Patiño, J., et al. (2025). *Integration Between Serious Games and EEG Signals: A Systematic Review.* Applied Sciences. DOI: 10.3390/app15041946

[4] BCI-BASED EMOTION RECOGNITION. Torres Proaño, Edgar Porfirio. (2022). *Bci-based emotion recognition (Reconocimiento de emociones basado en interfaces BCI).* Tesis de Licenciatura. Quito: EPN,Ecuador.Disponible en: https://bibdigital.epn.edu.ec/bitstream/15000/23420/1/CD%2012848.pdf

[5] Integración entre juegos serios y señales EEG: una revisión sistemática. García, A., et al. (2024). *Advances in Real-time EEG Processing for Interactive Systems.* Journal of Computer Behavior. DOI: 10.1016/j.chb.2024.12345

[6] Wolpaw, J. R. (2013). Brain–computer interfaces. In *Handbook of clinical neurology* (Vol. 110, pp. 67-74). Elsevier.

https://www.sciencedirect.com/science/article/pii/B978044452901500006X
[7] Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on biomedical engineering*, *51*(6), 1034-1043. https://ieeexplore.ieee.org/abstract/document/1300799
[8] Zander, T. O., & Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, *8*(2), 025005. https://iopscience.iop.org/article/10.1088/1741-2560/8/2/025005/meta
[9] Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21. https://www.sciencedirect.com/science/article/abs/pii/S0165027003003479
[10] Plass-Oude Bos, D., Reuderink, B., Van De Laar, B., Gürkök, H., Mühl, C., Poel, M., ... & Heylen, D. (2010). Brain-computer interfacing and games. In *Brain-computer interfaces: applying our minds to human-computer interaction* (pp. 149-178). London: Springer London. https://link.springer.com/chapter/10.1007/978-1-84996-272-8_10
[11] Dávila, J. A. V., Macías, J. V., & Lamas, M. V. (2017). Videojuegos basados en BCI (Interface cerebro computadora): Revisión sistemática literaria. *Programación matemática y software*, *9*(2), 10-23.
[12] Ranaweera & Mahmoud (2024) Ranaweera, M., & Mahmoud, Q. H. (2024). Deep reinforcement learning with Godot game engine. *Electronics, 13*(5), 985. https://www.mdpi.com/2079-9292/13/5/985
[13] Salmela (2022) Salmela, T. (2022). *Game development using the open-source Godot game engine* (Bachelor's thesis). Theseus Repository. https://www.theseus.fi/handle/10024/746943
[14] Fleming et al. (2017) Fleming, T. M., Bavin, L., Stasiak, K., Hermansson-Webb, E., Merry, S. N., Cheek, C., Lucassen, M., Lau, H. M., Pollmuller, B., & Hetrick, S. (2017). Serious games and gamification for mental health: Current status and promising directions. *Frontiers in Psychiatry, 7*, 215. https://doi.org/10.3389/fpsyt.2016.00215

Procesamiento computacional de series temporales para el estudio del caos y la aleatoriedad en sistemas físicos complejos

Juliana Bourdieu[1,2], Myriam Nonaka[2], Mónica Agüero[1,2], Marcelo Kovalsky[1]
[1] Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF)/ Departamento de Investigaciones en Láseres y sus aplicaciones (DEILAP-CONICET)/ Laboratorio de Láseres Sólidos (LLS). {jbourdieu,mnonaka,maguero,mkovalsky}@citedef.gob.ar
[2] Facultad de Ciencias Exactas y Naturales (FCEyN)/ Universidad de Buenos Aires (UBA).

RESUMEN

El Laboratorio de Láseres Sólidos (LLS) del DEILAP–(CITEDEF–CONICET) realiza investigación y desarrollo en sistemas físicos complejos basados en láseres de estado sólido, integrando experimentación, modelado y cómputo. El grupo promueve continuamente la formación de recursos humanos mediante la participación de estudiantes en trabajos de laboratorio, tesis de licenciatura y doctorado, consolidando un entorno de formación científica y tecnológica de alto nivel. Los sistemas experimentales implementados permiten revelar propiedades de señales tanto clásicas como cuánticas. En particular, el estudio de láseres con modulación de pérdidas permite, a partir de una serie temporal de alguna de las variables de salida del sistema, reconstruir el atractor subyacente. Este enfoque aporta una comprensión profunda de la dinámica no lineal y resulta especialmente relevante para el análisis de comportamientos caóticos y aleatorios. Asimismo, se desarrollan experimentos de generación de pares de fotones entrelazados en polarización (Experimento de Bell). Las series temporales obtenidas permiten cuantificar la aleatoriedad de secuencias, con implicancias en los fundamentos de la mecánica cuántica y en aplicaciones como la generación de números al azar.
Palabras clave: Láseres de estado sólido, series temporales, procesamiento.

CONTEXTO

En el Departamento de Investigaciones en Láseres y sus aplicaciones (DEILAP) se desempeñan más de 50 profesionales, entre doctores en Física, ingenieros, técnicos y becarios, conformando un equipo interdisciplinario dedicado a la investigación y desarrollo tecnológico. Las líneas de trabajo, tanto finalizadas como actualmente vigentes, se han sustentado en programas de investigación y desarrollo para la defensa (PIDDEF), subsidios del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y financiamiento internacional, lo que ha permitido sostener una actividad científica continua y de alto impacto. En todos los proyectos, el denominador común es la construcción, caracterización y aplicación de sistemas láser.

El Laboratorio de Láseres Sólidos (LLS), perteneciente al DEILAP, inició sus actividades en 1978 y se especializa en el estudio, desarrollo y aplicación de láseres de estado sólido. Desde 1992 es pionero en Argentina en la tecnología de láseres totalmente sólidos (all-solid-state lasers). Entre sus hitos de relevancia se destacan la primera medición de las desigualdades de Bell y la realización del primer láser de Nd:YAG con mode-locking por efecto Kerr y los estudios fundamentales sobre el régimen caótico en láseres de Ti:Zafiro de femtosegundos en Argentina, contribuyendo significativamente al avance del conocimiento en dinámica no lineal y óptica ultrarrápida.

1. INTRODUCCIÓN

Objetivo de estudio

El LLS realiza investigación y desarrollo a partir de diversos arreglos experimentales basados en cavidades láser de construcción propia, íntegramente diseñadas, montadas y optimizadas en el laboratorio. A partir de estos sistemas físicos complejos, se realiza el estudio de la dinámica no lineal y de sus propiedades. El grupo, impulsa desarrollos tecnológicos basados en láseres de estado sólido. En el ámbito académico, los estudios abarcan además implicancias fun-

damentales de la mecánica cuántica, estudios de aleatoriedad y caos en sistemas físicos complejos experimentales. Este enfoque integral no solo contribuye al avance del conocimiento básico y aplicado, sino que también promueve una sólida formación académica de recursos humanos altamente especializados, integrando estudiantes de grado y posgrado en actividades experimentales, desarrollo y procesamiento computacional avanzado.

Dinámica no lineal en láseres de estado sólido: análisis y desafíos tecnológicos

Los láseres de estado sólido [1,2] constituyen una de las tecnologías más ampliamente utilizadas tanto en aplicaciones industriales como en investigación científica, abarcando ámbitos civiles y militares. Su principal ventaja radica en su diseño compacto y robusto, lo que permite reducir dimensiones constructivas y favoreciendo sistemas más eficientes.
La incorporación de diodos láser de alta potencia marcó un avance significativo al reemplazar el tradicional bombeo mediante lámparas flash, mejorando la eficiencia, estabilidad y vida útil de estos dispositivos. No obstante, en el caso de láseres pulsados que emplean técnicas pasivas —como absorbentes saturables "rápidos" en el Kerr-lens mode-locking o "lentos" en el Q-switch pasivo— surgen inestabilidades características de los sistemas no lineales [3,4,5,6].
Estas dinámicas pueden manifestarse en fluctuaciones de la intensidad de los pulsos y en variaciones en la separación temporal entre pulsos sucesivos. Si bien dichas técnicas resultan atractivas por su simplicidad y bajo costo, pueden limitar su aplicabilidad en determinados contextos.
El estudio profundo de esta dinámica no lineal no solo representa un desafío científico de gran interés, sino que posee implicancias tecnológicas relevantes. La comprensión y eventual control de estas inestabilidades intrínsecas permiten el desarrollo de láseres con mayor alcance de aplicación.

Sistemas físicos complejos: abordaje experimental clásico y cuántico

A partir de distintos arreglos experimentales, se aborda el estudio del comportamiento de sistemas físicos complejos mediante el análisis de señales tanto cuánticas como clásicas.
Las señales cuánticas se obtienen a partir de un arreglo experimental diseñado para la generación y detección de pares de fotones entrelazados [7], recurso fundamental en experimentos de mecánica cuántica y en aplicaciones como comunicaciones cuánticas, distribución cuántica de claves (QKD) y generación de números aleatorios. El entrelazamiento constituye una propiedad exclusivamente cuántica, identificable mediante técnicas de adquisición por time stamping, que registra los instantes de detección de eventos individuales. La detección se realiza mediante fotodiodos de avalancha operando en modo Geiger (SPCM, Single Photon Counting Module).
Por otra parte, las señales clásicas se obtienen de sistemas complejos representativos de fenómenos de la naturaleza, como los láseres sólidos con modulación de pérdidas. En estos sistemas pueden observarse dinámicas caóticas [8], estudiadas a partir de series temporales (valores ordenados cronológicamente) adquiridos mediante fotodiodos. El análisis de estas señales permite determinar propiedades fundamentales de la dinámica no lineal del sistema.
El estudio comparativo de ambos tipos de señales resulta de gran relevancia tanto en investigación básica como en aplicaciones tecnológicas.

Adquisición y análisis computacional de datos experimentales

Las desigualdades de Bell establecen un límite a las correlaciones observables, clasicamente, entre mediciones realizadas sobre fotones en estaciones remotas (Alice y Bob). La evaluación experimental de dichas correlaciones requiere determinar el número de coincidencias entre eventos detectados. Dos detecciones se consideran coincidentes únicamente cuando su diferencia temporal es menor que una ventana de coincidencias Tw. En consecuencia, el valor de Tw influye directamente en el número de coincidencias registradas y, por lo tanto, en la magnitud de la correlación medida. En la mayoría de los experimen-

tos tradicionales, Tw es fijo y está determinado por el diseño de la electrónica de conteo [9].
Para el estudio detallado de procesos dinámicos y correlaciones con alta resolución temporal, resulta indispensable conocer con precisión el instante de detección de cada fotón. Esta información se adquiere mediante registros (time stamping) de tiempos de llegada utilizando un TDC (Time-to-Digital Converter) con resolución del orden de 10 ps. Los datos obtenidos son procesados mediante software de desarrollo propio, implementado para el análisis en las distintas estaciones remotas.
Por otro lado, para el estudio de dinámica no lineal en señales clásicas, el análisis se basa en la extracción de información a partir de series temporales generadas por el sistema. Estas incluyen series de amplitud —correspondientes a los máximos consecutivos de intensidad del láser— y series de diferencias temporales entre picos sucesivos, permitiendo caracterizar el comportamiento dinámico del sistema. La adquisición de estas señales se realiza mediante un osciloscopio digital PicoScope, que permite registrar la evolución temporal de la señal de salida del láser. Posteriormente, los datos son procesados utilizando herramientas computacionales como MATLAB y Python, donde se implementan algoritmos específicos para el filtrado, análisis y caracterización.

Caos y aleatoriedad en señales cuánticas y clásicas

A partir de los datos experimentales obtenidos, es posible realizar un procesamiento computacional orientado a la detección y caracterización de comportamientos caóticos y aleatorios.
En el caso de las señales cuánticas, los listados de time stamping —que contienen los tiempos de detección de eventos individuales— permiten construir secuencias binarias asociadas a resultados experimentales (por ejemplo, detecciones correlacionadas o no correlacionadas). Estas secuencias pueden analizarse desde el punto de vista estadístico y algorítmico para evaluar su grado de aleatoriedad, lo cual resulta de especial interés en aplicaciones como QKD.
En cuanto a las señales clásicas, a partir de las mediciones experimentales se construyen series temporales de amplitud o de intervalos entre eventos, que permiten estudiar la dinámica del sistema. Las señales caóticas se sitúan entre las totalmente predecibles, regulares, periódicas o cuasi periódicas y las totalmente irregulares, estocásticas, conocidas como ruido y que son completamente impredecibles. Aunque presentan evolución irregular, poseen una estructura determinista que puede evidenciarse mediante la reconstrucción del atractor en el espacio de fases [10,11], a partir de la serie temporal experimental y en el marco del teorema de embedding de Takens [12].
Para la detección de aleatoriedad, se emplean tanto pruebas estadísticas estandarizadas (como los tests del NIST) como enfoques algorítmicos basados en la complejidad de Kolmogorov, que cuantifica el grado de compresibilidad de una secuencia. De este modo, se combinan herramientas de la teoría de sistemas dinámicos y de la teoría de la información para distinguir entre caos y aleatoriedad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Actualmente, las actividades de investigación y desarrollo se estructuran en líneas estratégicas que incluyen el estudio de la dinámica no lineal en láseres de estado sólido [13,14], el análisis de aleatoriedad en sistemas físicos complejos, la generación de pares de fotones en estados entrelazados, los estudios fundamentales de Mecánica Cuántica [15] y la Distribución Cuántica de Claves (QKD) para criptografía.
Durante la última década, el grupo ha desarrollado investigación y desarrollo en sistemas físicos complejos, financiada mediante subsidios nacionales e internacionales y programas estratégicos. Estos proyectos han permitido consolidar infraestructura experimental, fortalecer colaboraciones internacionales y profundizar líneas científicas en óptica, dinámica no lineal y física cuántica.
Entre los principales proyectos financiados en los últimos diez años se destacan:
International Research Project (IRP) -

CNRS (Francia), titulado ”Complex Quantum Systems (CoQSys)” (2023-2027); PIP 00484 - CONICET, “Estudio experimental de aleatoriedad en sistemas cuánticos y complejos” (2023-2026); Laboratoire International Associé (LIA) - CNRS, “LICOQ - Complex Quantum Systems” (2019-2022); el proyecto financiado por Office of Naval Research Global (ONR), “Search for vulnerabilities of fundamental origin in Quantum Key Distribution” (2017-2021); PIP 027 - CONICET, “Desarrollo de láseres de estado totalmente sólido y de algunas aplicaciones” (2018-2021); el programa ECOS-Sud (CNRS-MinCyT), “Eventos Extremos en Sistemas no Lineales” (2015-2017); y el PIDDEF 01/14 - Ministerio de Defensa, “Láseres Sólidos para la Defensa II - LASODEF” (2016-2019); PIDDEF 18/14 - Ministerio de Defensa, “ Sistema de distribución de entrelazamiento a grandes distancias para criptografía” (2016-2019).

Estas líneas integran investigación básica y aplicada, combinando desarrollos experimentales propios, procesamiento computacional avanzado y colaboración internacional, posicionando al grupo en áreas estratégicas de óptica no lineal y tecnologías cuánticas.

3. RESULTADOS OBTENIDOS / ESPERADOS

En relación con la automatización, el control instrumental y el procesamiento computacional de series temporales experimentales, se realizaron aportes sustanciales que permitieron integrar y optimizar sistemas físicos complejos en los ámbitos cuántico y clásico. Se logró el diseño, automatización y control de dos plataformas experimentales [16] : una orientada al estudio de señales cuánticas y otra al análisis de señales clásicas provenientes de láseres de estado sólido. En el área cuántica, se implementó un sistema sincronizado de adquisición en dos estaciones remotas (Alice y Bob), separadas aproximadamente 60 m, incorporando el control automatizado para experimentos de desigualdades de Bell. Este desarrollo dio lugar a un arreglo experimental automatizado único a nivel nacional, con adquisición mediante TDCs de alta resolución y procesamiento propio de datos de time stamping, lo que permitió calcular correlaciones, determinar el parámetro de entrelazamiento S y generar secuencias binarias para el análisis de aleatoriedad. En el área clásica, se alcanzó la automatización en tiempo real de la adquisición de señales provenientes de un láser sólido con modulación de pérdidas, junto con el procesamiento y ajuste de parámetros como la frecuencia de modulación. A partir del análisis de series temporales experimentales, se identificaron regímenes caóticos determinísticos y se sentaron bases para futuras estrategias de investigación. Estos resultados dieron lugar a publicaciones en revistas internacionales de alto impacto y contribuyeron a consolidar la proyección científica del grupo. Asimismo, se prevé continuar y profundizar las investigaciones en las líneas actuales vinculadas a dinámica no lineal, aleatoriedad en sistemas complejos, generación de pares de fotones entrelazados y aplicaciones en tecnologías cuánticas.

4. FORMACIÓN DE RECURSOS HUMANOS

El LLS desarrolla una intensa actividad de formación académica. El grupo de trabajo está integrado por docentes universitarios, investigadores eméritos de amplia y reconocida trayectoria —cuyo aporte resulta clave en la consolidación de líneas estratégicas y en la formación de recursos humanos—, investigadores con experiencia posdoctoral en el exterior, así como investigadores dedicados principalmente a tareas científicas y tecnológicas, además de estudiantes de grado y doctorado en formación, conformando un esquema de trabajo colaborativo que integra investigación básica, desarrollo experimental y transferencia tecnológica.

Se desarrollaron en la última década 10 Laboratorios 6 y 7 , así como 2 tesis de licenciatura, y 2 tesis doctorales (1 más en curso) en Física y 1 en Ingeniería, en articulación con universidades como la Universidad de Buenos Aires (UBA) y la Universidad Tecnológica Nacional (UTN). Esta interacción fortalece la vinculación entre el sistema científico-tecnológico y el ámbito universitario, promoviendo una formación integral basada

en la investigación experimental y el desarrollo tecnológico, reflejando la continuidad y consolidación de la formación de recursos humanos altamente calificados.

5. BIBLIOGRAFÍA

1. W. Koechner, "Solid State Laser Engineering", Springer Science-Business Media, 1ra ed., EE.UU., 2006.

2. A. Siegman, "Lasers", University Science Books, 1ra ed., EE.UU., 1986.

3. A.A. Hnilo, M.G. Kovalsky, M.B. Agüero, J.R. Tredicce, "Characteristics of the extreme events observed in the Kerr-lens mode-locked Ti:sapphire laser", Physical Review A, 91, 013836, 2015. DOI: 10.1103/PhysRevA.91.013836

4. M.G. Kovalsky, A.A. Hnilo, "Chaos in the pulse spacing of passive Q–switched all-solid-state lasers", Optics Letters, Vol. 30, No. 20, pp. 3498–3500, 2010. DOI: 10.1364/OL.30.003498

5. M. Kovalsky, A. Hnilo, "Bistable behavior of a Kerr lens mode locked Ti:Sapphire laser", International Journal of Bifurcation and Chaos, Vol. 18, No. 6, p. 1719, 2008. DOI: 10.1142/S0218127408021349

6. M.G. Kovalsky, "On control of chaos in the Kerr lens mode locked Ti:Sapphire laser", Optics Communications, 260, p. 265, 2006. DOI: 10.1016/j.optcom.2005.11.082

7. M. Nonaka, M. Kovalsky, M. Agüero, A. Hnilo, "Testing how different levels of entanglement affect predictability in practical setups", Quantum Information Processing, 20, 165, 2021. DOI: 10.1007/s11128-021-03110-3

8. M. Nonaka, M. Agüero, M. Kovalsky, "Machine learning algorithms predict experimental output of chaotic lasers", Optics Letters, 2023. DOI: 10.1364/OL.483662

9. M. Agüero, "Fotones en estados entrelazados: mediciones con resolución temporal", Tesis Doctoral, Biblioteca Digital FCEN-UBA, 2013.

10. H. Abarbanel, "Analysis of Observed Chaotic Data", Springer Science-Business Media, 1ra ed., EE.UU., 1996.

11. R. Vilaseca, C.O. Weiss, "Dynamics of lasers", VCH Publishers, 1ra ed., EE.UU., 1991.

12. F. Takens, "Detecting strange attractors in turbulence", Dynamical Systems and Turbulence, Lecture Notes in Mathematics, Vol. 898, Springer-Verlag, pp. 366–381, 1981. DOI: 10.1007/BFb0091924

13. J.N. Bourdieu, M.B. Agüero, M.G. Kovalsky, "Experimental evidence of a strange non-chaotic attractor in a Nd:YVO_4 laser with saturable absorber", Physical Review A, 112, 043529, 2025. DOI: 10.1103/sxh7-fmlj

14. M. Nonaka, M. Agüero, M. Kovalsky, "Machine learning algorithms predict experimental output of chaotic lasers", Optics Letters, 2023. DOI: 10.1364/OL.483662

15. M. Agüero, A. Hnilo, M. Kovalsky, M. Nonaka, "Testing transient deviations from quantum mechanics' predictions in spatially spread optical entangled states", Journal of the Optical Society of America B, 2025. DOI: 10.1364/josab.547836

16. M. Nonaka, "Procesamiento en tiempo real de series temporales de láseres de estado sólido y su aplicación a la optimización de las señales de salida del láser", Tesis de Doctorado en Ingeniería, Biblioteca Digital, UTN – FRBA, Noviembre de 2021.

PRÓTESIS ROBÓTICA BASADA EN VISIÓN COMPUTACIONAL Y ARQUITECTURA DISTRIBUIDA

Santiago Andres Passerini, Francisco Tabelione, Pedro Daniel Lopez

Centro de Altos Estudios en Tecnología Informática (CAETI) Universidad Abierta Interamericana (UAI), Rosario, Argentina
{Pedro.Lopez}@uai.edu.ar
{SantiagoAndres.Passerini, Francisco.Tabellione}@alumnos.uai.edu.ar

RESUMEN

El presente artículo describe el diseño, modelado e implementación de un sistema robótico de mano funcional, fabricado mediante impresión 3D y controlado a través de una arquitectura distribuida. La solución integra un procesador de alto rendimiento para el procesamiento de visión artificial y un módulo ESP32 NodeMCU como unidad de control de actuadores. El sistema permite el reconocimiento y replicación de gestos en tiempo real con una latencia de 4.2 ms. Si bien el proyecto surge para necesidades médicas, su diseño modular permite aplicaciones en teleoperación industrial. Se detalla el proceso de ingeniería siguiendo el modelo de aprendizaje basado en competencias del CAETI-UAI.

Palabras clave: Prótesis Robótica, Visión Artificial, ESP32 NodeMCU, Edge Computing, Ingeniería de Software, CAETI.

CONTEXTO

Esta investigación se desarrolla en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana. El proyecto se alinea con la línea de "Sistemas Inteligentes y Robótica", vinculando las cátedras de Sistemas de Tiempo Real e Ingeniería de Software. El desarrollo busca la formación de recursos humanos capaces de proyectar soluciones tecnológicas de alto impacto social y versatilidad industrial [2].

1. INTRODUCCIÓN

La evolución de los sistemas ciberfísicos ha permitido que dispositivos de laboratorio hoy sean accesibles mediante fabricación aditiva. Este proyecto nace con un objetivo médico: reducir la brecha de accesibilidad para personas con amputaciones mediante una prótesis de bajo costo [1]. La versatilidad de la arquitectura propuesta permite que el dispositivo no se limite al área de la salud, abriendo posibilidades en la manipulación de materiales peligrosos y teleoperación industrial mediante el procesamiento de gestos a distancia.

2. METODOLOGÍA Y DISEÑO DEL SISTEMA

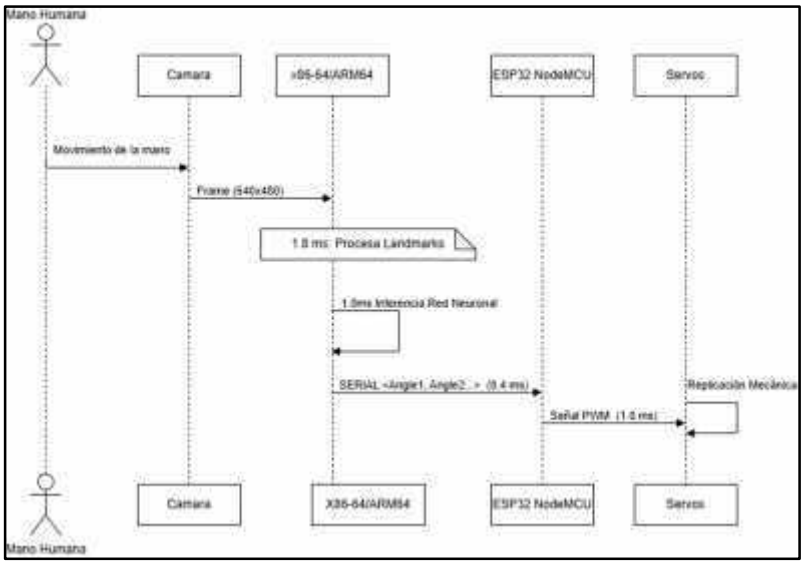
Para asegurar la calidad de la ingeniería de software, se utilizó la herramienta UAI Case para el modelado de la arquitectura [6]. Se definieron requisitos de alta disponibilidad y baja latencia, esenciales para la interacción hombre-máquina.

2.1 ARQUITECTURA DE HARDWARE

Se optó por una arquitectura de procesamiento distribuido para optimizar los tiempos de respuesta:

- Unidad de Procesamiento Central (x86-64 / ARM64): Encargada de la captura de video y la ejecución del modelo de inteligencia artificial. Al utilizar procesadores de alto desempeño, se logra una inferencia veloz de los puntos clave de la mano.
- Puente de Actuación (ESP32 NodeMCU): Recibe la información procesada vía comunicación serial. Su función principal es la generación de señales PWM (Pulse Width Modulation) para los servomotores [5]. Esta descarga de tareas asegura que los movimientos sean fluidos y no dependan de la carga de procesamiento de la imagen.

Figura 1. Diagrama de secuencia del flujo de datos. Se observa la distribución de carga entre la unidad de alto desempeño (PC) y el controlador de periféricos (ESP32), logrando un tiempo de respuesta end-to-end de 4.2 ms.



2.2 MODELADO MECÁNICO

La mano fue diseñada en piezas modulares utilizando material PLA. El sistema de transmisión de fuerza emplea tendones sintéticos de alta resistencia, logrando una estructura ligera y fácilmente reparable.

3. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

3.1. RECONOCIMIENTO DE GESTOS MEDIANTE VISIÓN ARTIFICIAL

El sistema emplea una cámara periférica conectada a la unidad central (PC/SBC). Se implementó una red neuronal convolucional optimizada para detectar 21 puntos clave (*landmarks*) de la anatomía de la mano [4].

La lógica de control traduce las distancias euclidianas entre estos puntos en comandos de movimiento. Una vez procesados, la unidad central envía una señal serial al ESP32 NodeMCU, el cual ejecuta la acción mecánica inmediata.

3.2 IMPLEMENTACIÓN DE TELEMETRÍA E INTERFAZ

Utilizando las capacidades de red del NodeMCU, se desarrolló una interfaz que permite observar en tiempo real el estado de los motores y la carga de la batería, permitiendo un control remoto seguro en aplicaciones industriales.

4. RESULTADOS OBTENIDOS Y EVALUACIÓN

Las pruebas de laboratorio dieron métricas de rendimiento superiores a los prototipos convencionales de su clase:

- Eficacia de Detección: El sistema fue capaz de detectar e identificar correctamente la mano en el 90% de las pruebas, incluso en condiciones de iluminación variable y fondos complejos.
- Velocidad de Respuesta: El tiempo transcurrido desde que la cámara capta el movimiento hasta que los dedos de la prótesis replican el gesto se estabilizó en 4.2 ms. Este valor sitúa al dispositivo dentro de los estándares de sistemas de alta performance [5], permitiendo una interacción fluida y natural.

Para validar el rendimiento del sistema y respaldar los datos anteriormente brindados, se realizaron 100 pruebas de ejecución en un entorno controlado. Los resultados demuestran que la división de tareas entre el ESP32 NodeMCU y el procesador permite procesar la información visual sin degradar la respuesta mecánica.

Como se detalla en la Tabla 1, el tiempo total de respuesta desde la captura de la imagen hasta el movimiento del servomotor se estabilizó en un promedio de 4.2 ms.

Tabla 1. Desglose de latencias y tiempos de procesamiento del sistema.

Etapa del Proceso	Latencia Promedio (ms)	Componente Responsable
Captura y preprocesamiento	1.8	X86-64/ARM64
Inferencia de red neuronal	1	X86-64/ARM64
Comunicación Serial (UART)	0.4	Bus de datos
Respuesta mecánica (PWM)	1	ESP32 NodeMCU
Tiempo Total (End-to-End)	4.2	Sistema Integral

Como se observa en el flujo temporal de la **Figura 1**, la segregación de la inferencia en la unidad de alto rendimiento permite que el sistema alcance los tiempos de respuesta detallados en la **Tabla 1**.

5. FORMACIÓN DEL RECURSO HUMANO

El proyecto es llevado adelante por alumnos de 4to año de la carrera de Ingeniería en Sistemas, integrándose dentro del cuadro de investigaciones institucionales de la Facultad de Tecnología Informática. Esta participación temprana en

proyectos del CAETI permite a los estudiantes aplicar conceptos avanzados de ingeniería de software y robótica en un entorno de investigación real, fortaleciendo sus competencias profesionales antes del proceso de tesis final.

6. BIBLIOGRAFÍA

[1] OMS. (2023). Informe mundial sobre tecnologías de asistencia.

[2] Haderne, M. F. (2025). Libro de Actas XXVII WICC. ISBN 978-987-575-267-2.

[3] Espressif Systems. (2024). ESP32 Series Datasheet.

[4] Rosebrock, A. (2021). Raspberry Pi for Computer Vision.

[5] Margolis, M. (2021). Arduino Cookbook. 3ra Ed.

[6] UAI Case Documentation. Facultad de Tecnología Informática - UAI.

Sistemas de Tiempo Real, Sistemas Embebidos y Técnicas de Simulación

Fernando Romero[1], Diego Encinas[1], Armando De Giusti[1,2], Lucas Maccallini[1], Alan Castelli[1], Horacio Villagarcía[1,3], Fernando G. Tinetti[1,3]

[1]Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática – Universidad Nacional de La Plata – Centro Asociado CIC
[2] CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas.
[3] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{fromero, dencinas, degiusti, hvw, fernando}@lidi.info.unlp.edu.ar,
{lucas.maccallini, alanfabcast}@gmail.com

Resumen

Esta línea de investigación se dedica al desarrollo y análisis de Sistemas de Tiempo Real. Dentro de esta temática se trabajó sobre dos sublíneas: 1) Modelado y Simulación de Sistemas de Tiempo Real. 2) Sistemas de robótica y comunicaciones. Las simulaciones estudiadas comprenden circulación de personas en ambientes hospitalarios, transmisión de enfermedades y de virus, circulación de pasajeros en trenes y aeropuertos; y vehículos en autopistas. Los sistemas de robótica y comunicaciones vistos este año comprenden un robot equilibrado en 2 ruedas, un planificador de trayectorias de un robot, y un robot rastreador.

Contexto

Esta línea de Investigación forma parte del proyecto F032, que comprende Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, Fog-Edge-Cloud, Big Data, Robótica, y Tiempo Real del Instituto de Investigación en Informática LIDI acreditado por la UNLP.

Palabras Claves: Tiempo Real, Simulación, Comunicaciones, Robótica, Microcontroladores, Cloud Computing, Inteligencia Artificial, Reconocimiento de patrones.

1. Introducción

Para que un sistema sea considerado de Tiempo Real (STR) debe incluir restricciones de tiempo estrictas en los plazos en que, ante una determinada entrada, deben producir una salida. Cuentan como otra característica interactuar con el mundo físico, a través de entradas conectadas a sensores y detectores y salida a actuadores [1] [2] [3] [6] [8] [11] [13] [14] [17][24]. De acuerdo con el sistema a controlar serán las restricciones temporales. Se requiere sincronizar el tiempo de procesamiento con el tiempo físico. Para el procesamiento de estas entradas y salidas se utilizan placas basadas en microcontroladores como, Arduino, NodeMCU, CIAA [10] [18] y Computadoras de Placa Simple como, Raspberry Pi, Raspberry Pi Zero utilizando diferentes SOTR (Linux RT-

Preempt, FreeRTOS, MQX, OSEK-OS, Zephyr, Raspbian, etc.) [4] [19].

En el campo del modelado y simulación [8] [9] [12] [14] [15] [22] [23] [25] [26] [27], se trata de predecir el comportamiento y la eficiencia de distintos sistemas ante diferentes escenarios, siendo calibrados con datos reales.

2. Resultados y Objetivos

Se han desarrollado tareas sobre los temas antes expuestos tales como:

- Péndulo invertido: se desarrolló un sistema de balanceo autónomo de un vehículo de dos ruedas utilizando una plataforma ESP32. El sistema logra mantener el equilibrio de forma autónoma gracias a un algoritmo de control PID
- Planificador de trayectoria: se diseñó un algoritmo eficiente de planificación de trayectorias capaz de operar en tiempo real en el ESP32 y se implementó en un robot [12].
- Q-learning: Se construyó un Robot Rastreador que utiliza el algoritmo de aprendizaje Q-Learning para aprender a moverse mediante un brazo articulado con dos grados de libertad.
- Tráfico (Autopista): se desarrolló un modelo, basado en otro trabajo anterior de la Cátedra, sobre el tráfico en la Autopista 25 de Mayo de CABA [17].
- Trenes (La Plata): Se modeló el movimiento de trenes y pasajeros en el ferrocarril de La Plata [16][28].
- Transporte (Aviones): se modeló el tráfico aéreo nacional utilizando el software ARENA y basado en el trabajo previo de los compañeros del año pasado [5].
- Dengue: simulación diseñada para estudiar la propagación del virus del dengue y evaluar el impacto de la vacunación en una población [10][20][21][29].

3. Formación de Recursos Humanos

Se desarrollan trabajos de alumnos en la Convocatoria a Proyectos de Desarrollo e Innovación de la Facultad de Informática de la UNLP y de promoción de la Cátedra Sistemas de Tiempo Real de Ingeniería en Computación.

Se dirigen Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Computación y una tesina de la Licenciatura en Sistemas.

De postgrado, investigadores del grupo están desarrollando un trabajo final de especialización y una tesis de Doctorado.

4. Referencias

[1] Integration of Sensor Networks with Cloud ComputingS. Medina, F. Romero, and F. G. Tinetti, Short papers of the 8th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2020), ISBN: 978-950-34-1927-4, págs. 2-5, 2020.

[2] Buttazzo, G. C., "Hard RealTime Computing Systems", Third edition, Springer, 2011.

[3] Burns, A, A, Wellings. "Real-Time Systems and Programming Languages: Ada, Real-Time Java and C/Real-Time POSIX", Addison-Wesley Educational Publishers Inc., 2009.

[4] FreeRTOS - market leading RTOS (real time operating system) for embedded systems supporting 34 microcontroller architectures. http:// www.freertos.org/.

[5] https://www.rockwellautomation. com/esmx/products/software/arena-simulation.html MinisteriodeTransporte. Agencia Nacional de AviaciónCivil. Recuperado de https://www.argentina. gob.ar/anac (octubre,2023).

[6] Kopetz. H., "Real-Time Systems, Design Principles for Distributed

Embedded Applications". Second Edition. Springer. 2011.
[7] Liu, J. W. S. Liu, "Real Time Systems", Integre Technical Publishing Co., Inc., 2000
[8] C. Macal, M. North, Tutorial on agent-based modeling and simulation part 2: how to model with agents, in: Proceedings of the Winter Simulation Conference, 2006.
[9] PHILLIP A. LAPLANTE, SEPPO J. OVASKA. REAL-TIME SYSTEMS DESIGN AND ANALYSIS Tools for the PractitionerFourth Edition. A JOHN WILEY & SONS, INC., PUBLICATION. IEEE PRESS. 2012.
[10] Chin,Pamela&Ehrbar,Dylan & Micieli,María&Fonseca,Dina&Ciota, Alexander& Kramer,Laura.(2018). "DifferentialEffectsofTemperatureandMosquitoGenetics DetermineTransmissibilityofArboviruses byAedesAegyptiinArgentina". The American JournalofTropicalMedicine andHygiene .99.10.4269/ajtmh.18-0097.
[11] R.S. Rosli, M.H. Habaebi, R. Islam On the analysis of received signal strength indicator from ESP8266, Bulletin of Electrical Engineering and Informatics · September 2019
[12] Análisis de una plataforma de simulación para Cloud Computing. Un caso de estudio. Tomás Rosales, Julián Spinelli, Marcos Di Nardo, Román Bond, Daniel Rosatto, Diego Encinas, Fernando Romero XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020) La Matanza
[13] F. G. Tinetti and O. C. Valderrama Riveros, "Unmanned Vehicles: Towards Heterogeneous Hardware Approaches," 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2018, pp. 919-924.
[14] Fernando G. Tinetti, Oscar C. Valderrama Riveros, Fernando L. Romero, "Unmanned Vehicles: Real Time Problems in Drone Receivers", Conf. on Computational Science & Computational Intelligence (CSCI'19), Las Vegas, Nevada, USA, 2019, pp. 1081-1085.
[15] D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
[16] Ministerio de Transporte Argentino, Cantidad de transacciones por fecha de la tarjeta SUBE (2020-2023) https:// datos. gob.ar/ dataset/transporte- sube — cantidad- transacciones- usos-por- fecha
[17] Buenos Aires Data, Flujo Vehicular por Unidades de Peaje AUSA, URL: https://data.buenosaires.gob.ar/dataset/fluj o-vehicular-por-unidades-de-peaje-ausa].
[18] NodeMcu http://www.nodemcu.com/ 2017
[19] https://www.ros.org/.
[20] Maccallini, L., Encinas, D. O., & Romero, F.. (2021). An Approach to the Modeling and Simulation of Intra-Hospital Diseases. Journal of Computer Science and Technology, 21(2), e14. https://doi.org/10.24215/16666038.21.e14
[21] Encinas, D. O., Matto, J., Stella, J., Maccallini, L., & Romero, F. (2024). Dengue Dynamics: Modelling Spread and Environmental Interactions. Journal of Computer Science and Technology, 24(2), e14.
https://doi.org/10.24215/16666038.24.e14
[22] Uri Wilensky (29 de septiembre de 2022). NetLogo User Manual version 6.3.0. Disponible en Internet en: https://ccl.northwestern.edu/netlogo/docs/
[23] M. Paradiso, L. Maccallini, A. Vericat, F. Romero and D. Encinas, "An approach to parallelization of respiratory disease spread simulations in emergency rooms," in 2022 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2022 pp. 1359-1365. doi: 10.1109/CSCI58124.2022.00244.
[24] Mauro Santos, Diego Encinas. Análisis de plataformas de Computación en la Nube para implementación de

protocolo de comunicaciones con una aplicación móvil 3D. XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022) La Rioja
[25] D. Encinas, S. Mendez, M. Naiouf, A. De Giusti, D. Rexachs del Rosario, and E. Luque, “An Agent-Based Model for Analyzing the HPC Input/Output System” International journal on advances in systems and measurements vol. 13, num. 3 & 4, págs. 192-202, 2020.
[26] D. Encinas, M. Naiouf, A. De Giusti, S. Mendez, D. Rexachs del Rosario, and E. Luque, “On the calibration, verification and validation of an agent-based model of the hpc input/output system” In Proceedings from The Eleventh International Conference on Advances in System Simulation (SIMUL 2019), 2019.
[27] D. Encinas et al., «Agent Based Simulation as an Efficient Way for HPC I/O System Tuning», en Computational Science – ICCS 2024, L. Franco, C. de Mulatier, M. Paszynski, V. V. Krzhizhanovskaya, J. J. Dongarra, y P. M. A. Sloot, Eds., Cham: Springer Nature Switzerland, 2024, pp. 129-136. doi: 10.1007/978-3-031-63759-9_16.
[28] M. Gonzalez; L. Gobbi; L. Maccallini; F. Romero; D. Encinas; M. Naiouf. Simulación de Ampliación del Tren Universitario de La Plata. Argentina. Viedma, Río Negro. 2025. Libro. Artículo Breve. Congreso. XXXI Congreso Argentino de Ciencias de la Computación.
[29] I. Di Loreto; R. Bond; M. Morales; D. Encinas. Estrategias de control del Dengue utilizando Modelado y Simulación basado en Agentes. Argentina. Viedma, Río Negro. 2025. Libro. Artículo Completo. Congreso. XXXI Congreso Argentino de Ciencias de la Computación.

ORCID autores:
Fernando Romero: 0000-0002-1498-3752
Diego Encinas: 0000-0002-6948-9786
A. De Giusti: 0000-0002-6459-3592
Santiago Medina: 0000-0001-6852-7165
Horacio Villagarcía:
Fernando G. Tinetti:
César Estrebou: 0000-0001-5926-8827

Procesamiento Distribuido y Paralelo

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

Adaptabilidad en la demanda para garantizar un uso eficiente de los recursos

Karina Cenci, José Moyano, Andrés Salamanca, Sebastián Farías

Laboratorio de Investigación en Sistemas Distribuidos (LISiDi)
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Bahía Blanca, Argentina
e-mails: {kmc, jose.moyano, cas,sebastian.farias} @cs.uns.edu.ar

Resumen

Los servicios digitales ofrecen a los usuarios (personas, empresas y agencias) la posibilidad de realizar tareas diferentes, como abonar un impuesto, consultar la humedad de una parcela o controlar el número de vehículos en un estacionamiento, entre otras, sin necesidad de gastar tiempo ni dinero en desplazamientos. La carga de las solicitudes para acceder a los servicios digitales puede variar en periodos de tiempo periódicos o aperiódicos. Para garantizar la atención equitativa de los requerimientos, los recursos utilizados deben adaptarse a las modificaciones de las cargas.

El uso eficiente de la infraestructura repercute en los costes de utilización, mantenimiento y consumo energético. Por otra parte, la inmersión de los ordenadores en el mundo requiere que consideremos el impacto que tiene en este. Uno de los aspectos en los que es importante trabajar es la sostenibilidad, que no solo involucra al medioambiente, sino también a un conjunto de dimensiones.

La elasticidad en la configuración del despliegue de la infraestructura y de los servicios tiene un impacto en otros aspectos, como la escalabilidad, la disponibilidad y la tolerancia a fallos. Además, el incremento de los recursos físicos requiere un aumento del consumo energético y del espacio para poder funcionar, así como de los desperdicios que producen los repuestos y las actualizaciones tecnológicas. Una de las posibles formas de abordar esta problemática es mejorar la sostenibilidad del hardware.

Palabras clave: Sistemas Distribuidos, Contenedores, Eficiencia, Sostenible

Contexto

Esta línea de investigación y desarrollo se realiza en el ámbito del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur (UNS). En particular, como parte de las tareas que se realizan en el Laboratorio de Investigación en Sistemas Distribuidos (LISIDI). El proyecto se financia parcialmente con fondos asignados por la UNS a proyectos de investigación.

1. Introducción

Los servicios digitales brindan a los usuarios – personas, empresas, agencias- la posibilidad de realizar diferentes tareas, como abonar un impuesto, consultar la humedad de una parcela, controlar la cantidad de autos en un estacionamiento, entre otras, sin la necesidad de consumir costos en tiempo y movilidad. La carga en las solicitudes para acceder a los servicios digitales puede variar en periodos de tiempo periódicos o aperiódicos. El garantizar la atención de los requerimientos en forma ecuánime requiere que los recursos utilizados se adapten a las modificaciones de las cargas.

El uso eficiente de la infraestructura impacta en los costos de utilización, de mantenimiento y de consumo energético. Por otra parte, la inmersión de las computadoras en el mundo requiere que consideremos el impacto que tiene en este. Uno de los aspectos en los cuales es importante trabajar es en la sostenibilidad, que no sólo involucra al medioambiente sino

integralmente a un conjunto de dimensiones. La sostenibilidad se define como un concepto sistémico y debe entenderse en un conjunto de dimensiones, que incluyen social, individual, ambiental, económica y técnica. La tecnología es parte del dilema para el desarrollo de soluciones sostenibles para mejorar el estado actual y futuro de nuestro mundo.

La elasticidad en la configuración del despliegue de la infraestructura y de los servicios es una calidad que impacta en otros aspectos como la escalabilidad, disponibilidad y tolerancia a fallos. Además, el incremento de los recursos físicos requiere crecimiento en el consumo energético y de espacio para que pueda funcionar, así como los desperdicios que producen los repuestos y las actualizaciones tecnológicas. Una de las posibles formas de abordar esta problemática se encuentra en mejorar la sustentabilidad de hardware. Esta mejora se propone a partir de la migración de las plataformas físicas convencionales hacia *System-on-Chip* (SoC) y *single-board computers*. Estas dos opciones mejoran significativamente el desempeño de la plataforma en relación con la cantidad de energía consumida, requieren mucho menos espacio para operar y reducen la cantidad de desperdicios. Sin embargo, este enfoque presenta nuevos desafíos técnicos que deben ser abordados.

El resto de este artículo se organiza de la siguiente manera: la sección 2 explica los objetivos de investigación, la sección 3 los resultados esperados/obtenidos y, por último, la sección 4 discute la formación de recursos humanos.

2. Líneas de Investigación y Desarrollo

Computación en la nube es un paradigma que se construye sobre un sistema distribuido entre pares para facilitar la movilidad de los usuarios por medio del uso de dispositivos móviles que se comunican a través de mensajes. Este proyecto está organizado en dos ejes temáticos de investigación: (a) Adaptabilidad en el uso de los recursos en una nube privada sobre un cluster con dispositivos de pequeña escala y (b) Métricas para la evaluación de los procesos de automatización para la recuperación y tolerancia a fallas en un cluster con dispositivos de pequeña escala.

Adaptabilidad en el uso de los recursos en una nube privada sobre un cluster con dispositivos de pequeña escala.

- El cúmulo de solicitudes de requerimientos varía en intervalos de tiempo; estos intervalos pueden estar predeterminados o depender del contexto. El estudio, análisis y despliegue de un modelo que permita utilizar en forma eficiente los recursos, reduciendo los costes de comunicación y consumo energético.

Métricas para la evaluación de los procesos de automatización para la recuperación y tolerancia a fallas en un cluster con dispositivos de pequeña escala.

- La identificación de los factores clave de rendimiento facilita y promueve la aplicación de métricas para la evaluación de los procesos.
- Comparación de alcance y desempeño de un clúster de dispositivos de pequeña escala con un clúster tradicional. Es necesario establecer cuáles funciones del clúster híbrido son candidatas para delegar en dispositivos de pequeña escala.

3. Resultados Obtenidos/Esperados

El diseño de la plataforma, en nuestro caso de un cluster de bajo costo, requirió un proceso de construcción compuesto por varias etapas: inicio, funcionamiento y automatización del estado de los nodos [8] Como resultado, se logró desplegar un cluster de contenedores Docker con un consumo energético de una fracción del que hubiera sido necesario para soportar un clúster tradicional con procesadores x86/amd64, a un costo muy

inferior y que puede operar en un espacio reducido.
Una plataforma con estas características presenta restricciones en la selección de herramientas de software para desplegar los servicios en un sistema con requerimientos que se expanden y contraen. Además, otro de los desafíos es la fragilidad y la limitada robustez de los dispositivos, que en algunos casos están asociados al bajo costo. Sin embargo, el equipo de trabajo validó imágenes de arquitectura ARM que permitieron el despliegue elástico de microservicios a pesar de las limitaciones de los nodos.

Los objetivos de una arquitectura eficiente requieren la evaluación de disponibilidad, escalabilidad, utilización de los recursos, costo de comunicación, seguridad.
Considerando las limitaciones de la plataforma, un primer paso es la clasificación de las metas para el desarrollo de experimentos para la evaluación (básica y avanzada).

La disponibilidad y utilización de recursos son aspectos indispensables en la evaluación de una arquitectura eficiente, por esto, son considerados básicos. El costo de comunicación, escalabilidad y seguridad son propiedades relevantes para la eficiencia de una solución; para este proyecto son consideradas avanzadas. Las preguntas directrices son:

- Cuáles indicadores de rendimiento son factibles de aplicar para cada una de las metas
- ¿Es posible que en algunos casos sea apropiada una solución híbrida, en la cual se replican los servicios o microservicios en los dispositivos que se encuentran en el borde (*edge computing*)?
- Implementar un modelo de evaluación diferencial para el hardware de pequeña escala, es decir, comparar el rendimiento en relación a las limitaciones y virtudes de los sistemas on chip (SoC). Ejemplos podrían ser la eficiencia por vatios de los nodos, la disponibilidad del clúster en relación a su robustez individual, la sostenibilidad económica (TCO, del inglés *Total Cost of Ownership*), o la adaptabilidad de grano fino (se pueden encender o apagar nodos que incorporan 4 o 8 GB de RAM de forma incremental y sin desperdicio).

4. Formación de Recursos Humanos

En relación con la formación de los recursos humanos, el trabajo a desarrollar permite una capacitación práctica adecuada de los profesionales involucrados, plasmando con logros concretos las investigaciones realizadas. Además, esta línea de investigación permitirá la dirección de tesis de licenciatura y trabajos finales de ingeniería afines. Asimismo, servirán para generar *courseware* para materias optativas a dictarse en el Departamento de Ciencias e Ingeniería de la Computación de la UNS.

5. Referencias

[1.] Leila Abdollahi Vayghan, Mohamed Aymen Saied, Maria Toeroe, Ferhat Khendek, A Kubernetes controller for managing the availability of elastic microservice-based stateful applications, Journal of Systems and Software, Volume 175, 2021, 110924, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2021.110924.

[2.] An Antonio Jose Alves Neto, Jose Aprigio Carneiro Neto, and Edward David Moreno Ordonez. 2022. Low-cost clusters on big data - A systematic study. In 11 Euro American Conference on Telematics and Information Systems (EATIS2022), June 1–3, 2022, Aveiro, Portugal. ACM, New York, NY, USA 7 Pages. https://doi.org/10.1145/3544538.3544635.

[3.] Bibal Benifa J.V., Dejey Dharma, HAS: Hybrid auto-scaler for resource scaling in cloud environment, Journal of Parallel and Distributed Computing, Volume 120, 2018, Pages 1-15, ISSN 0743-7315, https://doi.org/10.1016/j.jpdc.2018.04.016.

[4.] Dipto Biswas, Sohely Jahan, Sajeeb Saha, Md. Samsuddoha, A succinct state-of-the-art survey on green cloud computing: Challenges, strategies, and future directions, Sustainable Computing: Informatics and Systems, Volume 44, 2024,

101036, ISSN 2210-5379, https://doi.org/10.1016/j.suscom.2024.101036.
[5.] R. Dhaya, R Kanthavel, IoE based private multi-data center cloud architecture framework, Computers and Electrical Engineering, Volume 100, 2022, 107933, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2022.107933.
[6.] Tien Van Do, N.H. Do, H.T. Nguyen, Csaba Rotter, Attila Hegyi, Peter Hegyi, Comparison of scheduling algorithms for multiple mobile computing edge clouds, Simulation Modelling Practice and Theory, Volume 93, 2018.
[7.] Jing Huang, Yan Liu, Renfa Li, Keqin Li, JiYao An, Yang Bai, Fan Yang, Guoqi Xie, Optimal power allocation and load balancing for non-dedicated heterogeneous distributed embedded computing systems, Journal of Parallel and Distributed Computing, Volume 130, 2019, Pages 24-36.
[8.] J. Moyano, C. A. Salamanca, K. Cenci, J. Ardenghi. Evaluación de Docker en un Cluster de Bajo Costo y Alta Disponibilidad, en 10mo Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI). AJEA, (17), 1–2021.
[9.] Moyano, J., & Cenci, K. (2024). Repositorio Común para Mediciones de Estaciones Meteorológicas Automatizadas. *Memorias De Las JAIIO*, *10*(10), 81-92. https://publicaciones.sadio.org.ar/index.php/JAIIO/article/view/946
[10.] K. Cenci, J. Moyano y A. Salamanca. Utilización Eficiente y Sostenible de Recursos en Clúster de Bajo Costo. XXV Workshop de Investigadores en Ciencias de la Computación, WICC 2023, Universidad Nacional del Noroeste de Buenos Aires, Junín, Argentina.
[11.] Arul Rajagopalan, Dhivya Swaminathan, Mohit Bajaj, Issam Damaj, Rajkumar Singh Rathore, Arvind R. Singh, Vojtech Blazek, Lukas Prokop, Empowering power distribution: Unleashing the synergy of IoT and cloud computing for sustainable and efficient energy systems, Results in Engineering, Volume 21, 2024, 101949, ISSN 2590-1230, https://doi.org/10.1016/j.rineng.2024.101949.
[12.] Amani K. Samha, Strategies for efficient resource management in federated cloud environments supporting Infrastructure as a Service (IaaS), Journal of Engineering Research, Volume 12, Issue 2, 2024, Pages 101-114, ISSN 2307-1877, https://doi.org/10.1016/j.jer.2023.10.031.
[13.] Amritpal Singh, Gagangeet Singh Aujla, Rasmeet Singh Bali, Container-based load balancing for energy efficiency in software-defined edge computing environment, Sustainable Computing: Informatics and Systems, Volume 30, 2021, 100463, ISSN 2210-5379, https://doi.org/10.1016/j.suscom.2020.100463.
[14.] Kapil N. Vhatkar, Girish P. Bhole, Optimal container resource allocation in cloud architecture: A new hybrid model, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 5, 2022, Pages 1906-1918, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2019.10.009
[15.] Shangguang Wang, Yali Zhao, Jinlinag Xu, Jie Yuan, Ching-Hsien Hsu, Edge server placement in mobile edge computing, Journal of Parallel and Distributed Computing, Volume 127, Pages 160-168, 2019.
[16.] Xiyuan Xu, Shaobo Zang, Muhammad Bilal, Xiaolong Xu, Wanchun Dou, Intelligent architecture and platforms for private edge cloud systems: A review, Future Generation Computer Systems, Volume 160, 2024, Pages 457-471, ISSN 0167-739X, https://doi.org/10.1016/j.future.2024.06.024

Análisis Basado en Simulación de Sistemas de Salud e Infraestructuras de HPC.

Diego Encinas[1,2], Jimena Jara[1], Román Bond[1], Daniel Rosatto[1], Gustavo García Krahn[1], Ignacio Di Loreto[1], Eduardo Zarate[1], Verónica Scholz[1], Adriana Gaudiani[3], Martin Morales[1,4]

[1]Proyecto de Investigación SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - UNAJ

[2]Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

[3]Área de Computación - Instituto de Ciencias - UNGS

[4]Centro CodApli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, elchejime@gmail.com, rbond@unaj.edu.ar, drosatto@unaj.edu.ar, diegoalvarez3219@gmail.com, krahn_gustavo@hotmail.com, diloretoignacio@gmail.com, emzarate@exactas.unca.edu.ar, vero.scholz@gmail.com, agaudiani@campus.ungs.edu.ar, martin.morales@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de diferentes técnicas de modelado y simulación para entornos de Cómputo en Altas Prestaciones (HPC, High Performance Computing). El enfoque del estudio es la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios y reconfigurar el sistema físico. Además, se analizan los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o reconfigurarse.

Por otra parte, se ha llevado a cabo el desarrollo de un simulador para entornos en el área de salud, en el sector de emergencias hospitalarias y de propagación de enfermedades.

Palabras clave: *Arquitecturas Multiprocesador. Simulación. Sistema E/S paralela. Modelado y Simulación basado en agentes (Agent-Based Modeling and Simulation, ABMS). Cloud Computing. CloudSim. Simulación y Salud.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación "Simulación, Computación de Altas Prestaciones (HPC) y optimización de aplicaciones sociales – SimHPC" de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 183/21. Además, el proyecto aporta al Programa "Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social" – TICAPPS de la UNAJ.

En el tema existe un convenio de colaboración en actividades de Investigación y Postgrado con el Instituto de Investigación en Informática – LIDI de la Universidad Nacional de La Plata y el Área de Computación del Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

Introducción

El crecimiento sostenido en la demanda del poder de cómputo remarca la necesidad de sistemas con enfoques de paralelización masiva y cómputo de alta performance (HPC, High Performance Computing) [1]. Los clusters se han convertido en uno de los enfoques principales para lograr paralelismo a bajo costo. Una noción extendida lo constituye la utilización de grid computing y más recientemente cloud computing. Independientemente de la solución, estos sistemas constan de un gran número de componentes incluyendo nodos de procesamiento, bancos de memoria, discos, entre otros.

En cuanto a las herramientas de simulación para Cloud Computing, CloudSim es un framework desarrollado en Java que provee las APIs necesarias para que el usuario genere una simulación de un Data

Center funcionando como servidor de nube capaz de simular la ejecución de CloudLets.

Por otra parte, los servicios de urgencias hospitalarias son considerados como una de las unidades del sistema sanitario de mayor complejidad y fluidez, lo que unido a la variabilidad de su actividad da lugar a que su gestión operativa sea una tarea muy complicada [2]. Es por ello que resultaría de mucha utilidad para sus responsables disponer de un sistema de ayuda a la toma de decisiones (Decision Support System-DSS) tan flexible como un simulador, que permitiría tomar medidas disponiendo de información suficiente sobre las alternativas posibles.

Sistemas de E/S Paralela

Las exigencias en los sistemas de E/S paralelos se han incrementado debido al aumento en número, velocidad y potencia de las unidades de procesamiento en los clusters. También las aplicaciones científicas que utilizan cómputo de altas prestaciones acrecientan estos requerimientos.

En muchos casos, el cuello de botella de los sistemas paralelos es la E/S a causa de las exigencias que debe afrontar [3]. La E/S Paralela es esencial para emparejar el avance de las arquitecturas de los procesadores y el rápido crecimiento de la capacidad computacional. Aunque la arquitectura jerárquica de memoria multinivel puede evitar grandes pérdidas de prestaciones debido a los retardos de acceso a disco, la capacidad de memoria es limitada. Además, como la capacidad computacional aumentará, la disponibilidad de memoria por core decrecerá, especialmente si la escala de los sistemas de HPC se proyecta a millones de cores o más. Varias simulaciones científicas y de ingeniería de áreas críticas de investigación, tales como la nanotecnología, astrofísica, clima y energía física están convirtiéndose en aplicaciones intensivas de datos. Para poder disminuir la brecha entre CPUs-E/S se deben identificar los factores que influyen en las prestaciones y proponer nuevas soluciones [4] [5].

En el área de tolerancia a fallas en sistemas de cómputo de alta prestaciones se puede notar la importancia de la unidad de E/S en las arquitecturas paralelas como un punto a mejorar para lograr cubrir las exigencias de las aplicaciones que utilizan HPC.

Se puede disminuir la complejidad y la probabilidad de errores en la generación de sistemas híbridos desarrollando una simulación específica de éstos utilizando diferentes frameworks [6] [7] [8].

Las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos. Contar con una herramienta que permita predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores de aplicaciones paralelas como para administradores de centros de cómputo. Por otro lado, ABMS ha sido utilizado para modelar problemas y sistemas complejos en diversas áreas de la ciencia.

Evaluar las prestaciones del subsistema de E/S con diferentes configuraciones y la misma aplicación, permite adaptar la configuración de E/S teniendo en cuenta el patrón de acceso de la aplicación. Pero también puede ser una gran ventaja analizar las necesidades de las aplicaciones antes de configurar el sistema físico. Una manera de predecir el comportamiento de las aplicaciones en el sistema de cómputo, ante distintas configuraciones, es utilizando técnicas de modelado y simulación.

Se está desarrollando modelos e implementando una simulación de la arquitectura de E/S paralela, por medio de técnicas de simulación basadas en agentes o Sistemas Multi-Agente, (MAS-MultiAgentSystems), para evaluar el efecto de dimensionar el sistema de E/S o cambiar componentes como la red de almacenamiento, dispositivos de E/S, entre otros [9].

Simulación de arquitecturas de Cloud Computing

CloudSim [10] [11] es un Framework de simulación generalizado y extensible que permite el modelado y la simulación de diferentes infraestructuras y servicios de aplicaciones de Cloud Computing. Un ejemplo de utilización es la simulación de muchos centros de datos.

Su arquitectura consiste en entidades específicas que se representan como clases Java que pueden ser heredadas o instanciadas. Estas clases representan centros de datos, hosts físicos, máquinas virtuales, servicios a ejecutar en los centros de datos y servicios en la nube de usuarios [12] [13]. Además, CloudSim soporta la inserción dinámica de los elementos de simulación y proporciona aplicaciones de paso de mensajes y la topología de la red del centro de datos.

La versatilidad de CloudSim es la principal ventaja del sistema. La integración de nuevos parámetros y conceptos de la simulación es implementada desde abstracciones preestablecidas convenientemente por

los autores. Las abstracciones principales son SimEvent [14], SimEntity [15], DataCenterCharacteristics y Vm.

El aporte de un desarrollo de nuevos actores al componente Vm (Virtual Machine) posibilita que por medio de simulación se obtengan métricas de entrada/salida. Las estadísticas que aporta la nueva implementación de Vm dan soporte a un espacio de memoria ram en tiempos de simulación.

Simulación, Salud y Ciencia de Datos

Al analizar las necesidades presentes en los servicios de salud, se obtiene que dentro de una sala de urgencias se encuentre una gran diversidad de escenarios posibles donde cada uno de estos puede afectar a resultados sensibles como, por ejemplo, la tasa de mortalidad de personas. Para solucionar este problema se desarrolló un simulador con el objetivo de ser una herramienta capaz de recrear una gran cantidad de escenarios y así poder tomar decisiones rápidas.

El simulador desarrollado se centra en la prevalencia puntual de infecciones intrahospitalarias [16] en una sala de urgencias y cómo la afectan distintos factores relacionados con la gestión hospitalaria.

Para llevar a cabo el modelado del simulador se utilizó el paradigma de Modelado y Simulación basado en Agentes (ABMS). El framework utilizado es Repast Simphony [17], una herramienta especializada en ABMS, la cual provee una serie de ventajas respecto al modelado e implementación de agentes, así como también la coordinación concurrente de los agentes.

De esta manera, se clasificaron distintos agentes intervinientes en el ambiente de una sala de urgencias, como los pacientes y médicos, entre otros. Dichos agentes definen su comportamiento mediante máquinas de estado, las cuales determinan las acciones correspondientes tanto a la atención hospitalaria como también al estado de salud respecto de una enfermedad intrahospitalaria específica.

La propagación de la infección intrahospitalaria estará modelada mediante interacciones entre estos agentes, por ejemplo, uno de los focos de infección más importante de las salas de urgencias son las salas de espera. La forma más efectiva para calibrar el simulador se da mediante el grado de interacción de agentes, ajustando así la tasa de transmisión de la enfermedad en cuestión.

Una vez calibrado el simulador, se pueden obtener resultados. El trabajador de la salud dispone de diversos parámetros para configurar la simulación, por ejemplo, la cantidad de pacientes, la cantidad de camas disponibles, etc. Se pueden obtener resultados analizando distintos posibles cuellos de botella, configurando la cantidad de médicos clínicos disponibles, la cantidad de recepcionistas encargados de la admisión, el triage, entre otros. De esta manera, es posible la toma de decisiones respecto a la asignación de recursos y personal para agilizar la estancia hospitalaria de los pacientes y evitar los contagios producidos por sus interacciones.

Además, se analiza y trabaja la optimización y paralelización del simulador con los ambientes Repast HPC para clusters y Flame para ejecución sobre GPU, obteniendo mejor performance en términos de tiempos.

Relacionado a Ciencia de Datos, se trabaja en conjunto con el Hospital San Roque de La Plata en el análisis de datos. Se analizaron los datos obtenidos y se realizaron modificaciones o agregaciones de campos previamente calculados partiendo de la información básica, como segmentación de edad en base a la fecha de nacimiento comparado con el día del registro de ingreso. Además, es posible discriminar por por grupo etáreo y sexo, por ejemplo.

También se realizaron estudios de ingresos diarios, mensuales y anuales, para entender la periodicidad de una persona promedio en asistencias. Como el flujo promedio de personas por turnos de atención, para lo cual se podría definir la cantidad de personal por turnos y tipos de urgencias y de enfermedades relevadas.

El resultado final es un panel de control y un conjunto de datos limpio que es utilizado como entrada para el simulador previamente mencionado.

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.

- Arquitectura de E/S paralela considerando el software, hardware, comunicaciones entre módulos y dispositivos de almacenamiento.
- Nuevos aportes de desarrollos que mejoren los modelos de simulaciones con CloudSim para el análisis de la performance en sistemas de arquitecturas de software de Cloud Computing.
- Modelado y simulación para la administración de sistemas de salud.

Resultados y Objetivos

Investigación experimental

- Diseño y desarrollo de modelos mediante técnicas de Modelado y simulación basada en agentes (ABMS) para analizar el comportamiento de las distintas capas de la pila de software de E/S.
- Análisis, modelado e implementación de las operaciones típicas de E/S: read, write, open, close, flush. Se ha utilizado Amazon Web Services para creación de cluster virtuales y obtener métricas de la pila de software de E/S.
- Incorporación de tiempos de entrenamiento, obtenidos en AWS, en el simulador. Con esto se logra una salida más detallada y un método para validar tiempos y métricas del simulador con AWS.
- Implementación de comandos para ejecutar desde command center en NetLogo. Con esto se logró sintetizar el benchmark IOR correspondiente a la capa de aplicación de la pila de E/S, logrando introducir nuevos parámetros como tamaño de archivo y cantidad de nodos de E/S (metadata server y data server). De ejecutar este comando, se obtiene una nueva salida similar a la del benchmark IOR [18].
- Obtención de un método de desarrollo de nuevos actores genéricos CloudSim que mejoran el modelado y la producción de estadísticas virtuales. Implementación de la entidad InitEntity que procesa instrucciones en el espacio de memoria de las máquinas virtuales. Se vinculan exitosamente los tiempos de procesamiento de un cloudlet y las instrucciones en las máquinas virtuales [19].
- Diseño de prototipos de pruebas para estudiar entrada/salida como en el área de salud, utilizando microcontroladores [20] [21].
- En el área de tiempo real se trabajaron con microcontroladores en el reconocimiento de audio y en el control de escenarios 3D con Unity [22] [23].
- Modelado y simulación de procesos industriales, en específico sistemas de banda clasificadora [24].
- Contraste de una simulación de un cluster en la nube y uno idéntico desplegado en un sistema de cloud computing público [25].
- Modelado de infraestructuras de sistemas de salud. Implementación de un simulador para analizar el contagio de enfermedades intrahospitalarias [26] [27].
- Implementación de un panel de control con la herramienta PowerBI [28] para la toma de decisiones de un hospital a partir de un conjunto de datos real.
- Implementación de un sistema de reconocimiento emocional facial con Python y OpenCV [39].
- Modelado de sistema de línea de producción y banda en la industria [30].
- Mejora de un modelo de simulación del dengue. A partir del simulador original desarrollado en Netlogo se incorporaron nuevas funcionalidades: la incorporación del modelo SIR (Susceptible, Infectado, Recuperado), la integración de un agente climático que simula las condiciones ambientales y la diferenciación entre mosquitos macho y hembra.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería en Informática de la UNAJ. También aportan trabajos de alumnos de las materias Redes de Computadoras 2 y Programación en Tiempo Real. Por otro lado, algunos integrantes participan en el dictado de la Diplomatura y Maestría en Ciencia de Datos de la UNAJ.

Durante 2025 se han realizado publicaciones nacionales [32][33][34][35]. En esta línea de I/D existe cooperación a nivel nacional e internacional. Hay dos investigadores realizando estudios de postgrado, dos becarios de Iniciación a la

Investigación UNAJ y dos alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Grama A, Gupta A, Karypis G, Kumar V. “Introduction to parallel computing”. SecondEdition. Pearson Addison Wesley, 2003.
2. R. Galeano, C. Villalba, D. Rexachs, E. Luque. Agent-Based Model to Simulate Outpatient’s Consultations at the "Hospital de Clínicas". The Eighth International Conference on Advances in System Simulation (SIMUL 2016). 1:46-51
3. H Hennessy, J. L., Patterson, and D. A., Computer Architecture, Fourth Edition: A Quantitative Approach. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2006.
4. J. M. May, Parallel I/O for high performance computing. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001.
5. V. Balaji, Earth system modelling – Volume 4. IO and Postprocessing. Springer, 2013.
6. D. Encinas, Utilización de un reloj global para el modelado de un ambiente simulado distribuido. XVIII Congreso Argentino de Ciencias de la Computación. 2012
7. D. Encinas, Simulación de una red CAN para dimensionar las comunicaciones de una IMU. VII Congreso Argentino de Tecnología Espacial. 2013.
8. D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
9. D. Encinas et al.,Modeling I/O System in HPC: An ABMS Approach. The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015.
10. R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose and R. Buyya “CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms” Published online 24 August 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/spe.995.
11. http://www.cloudbus.org/cloudsim 2018.
12. Hamza Ouarnoughi, Jalil Boukhobza, Frank Singhoff, Stephane Rubini, Erwann Kassis. “Considering I/O Processing in CloudSim for Performance and Energy Evaluation”. OpenStack Cloud Software: Open source software for building private and public clouds. © Springer International Publishing AG 2016 M. Taufer et al. (Eds.): ISC High Performance Workshops 2016, LNCS 9945, pp. 591–603, 2016. DOI: 10.1007/978-3-319-46079-6 40.
13. Parikh, Hawanna, Haleema. P.K, Jayasubalakshmi.R and N.Ch.S.N.Iyengar. School of Computing Science and engineering Vellore Institute of Technology, Tamil Nadu,“Virtual Machine Allocation Policy in Cloud Computing Using CloudSim in Java.” 2015.
14. www.icsa.inf.ed.ac.uk/research/groups/hase/simjava 2018.
15. F. Howell, R Mc Nab. A discrete event simulation library for java. International Conference on Web-Based Modeling and Simulation. 1998.
16. Prevalencia puntual de infección nosocomial, disponible en internet: https://www.researchgate.net/publication/242363760_Prevalencia_puntual_de_infeccion_nosocomial. Fecha: 29/03/2020
17. Repast Simphony Frequently Asked Questions, disponible en internet: https://repast.github.io/docs/RepastFAQ/RepastFAQ.html. Fecha: 29/03/2020
18. D. Encinas, M. Naiouf, A. D. Giusti, R. Bond, S. Mendez, D. Rexachs, and E. Luque. Agent Based Simulation as an Efficient Way for HPC I/O System Tuning. En International Conference on Computational Science. Cham: Springer Nature Switzerland, 2024. p. 129-136.
19. D. Rosatto, R. Bond, M. Belizán, M. Morales, D. Encinas. Modelado y simulación de arquitecturas de Cloud Computing con CloudSim: comunicación entre entidades. XXIII Congreso Argentino de Ciencias de la Computación. ISBN 978-950-34-1539-9. 2017
20. L. Barrera, M. Rodriguez, R. Bond, M. Morales, D. Encinas Diseño de un oxímetro de pulso. Prototipo de pruebas. XXVIII Congreso Argentino de Ciencias de la Computación. ISBN 978-987-1364-31-2. 2022
21. M. López, R. Bond, D. Encinas. Diseño de un prototipo de pruebas de Entrada/Salida con Arduino. 10º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2022). Entre Ríos. 2022
22. Rodrigo Lago; Verónica Scholz; Roman Bond; Martin Morales; Diego Encinas. Experiencia de desarrollo de una aplicación de reconocimiento de canciones mediante la técnica de huella de audi". En XXIX Congreso Argentino de Ciencias de la Computación CACiC, Argentina. 2023.
23. David Verón; Roman Bond; Martin Morales; Diego Encinas. Control de iluminación en un entorno 3D con Unity. En 11° Congreso Nacional de Ingeniería Informática y Sistemas de la Información, Argentina. 2023.
24. Alan Dell Orto Marengo, Emanuel Julio, Diego Rolón, Román Bond, Martin Morales, Diego Encinas.Simulación de sistema de reciclaje mediante banda clasificadora. En 11° Congreso Nacional de Ingeniería Informática y Sistemas de la Información, Argentina. 2023.
25. T. Rosales, J. Spinelli, M. Di Nardo, R. Bond, D. Rosatto, D. Encinas, F. Romero. Análisis de una plataforma de simulación para Cloud Computing. Un caso de estudio. XXVI Congreso Argentino de Ciencias de la Computación. 2020
26. L. Maccallini, Encinas, and F. Romero. “An Approach to the Modeling and Simulation of Intra-Hospital Diseases”. Journal of computer science and technology (ISSN 1666-6038), vol. 21, num. 2, págs. 157-169, doi. 10.24215/16666038.21.e14, 2021.
27. J. Baez, A. Barreto, B. Galarza, M. Morales, D. Encinas. Simulación para estimar propagación de enfermedades. 8º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba. 2020
28. Becker & Gould (2019) Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data, Serials Review, 45:3, 184-188.
29. Albornoz, G.; Bond, R.; Morales, M.; Encinas, D.

"Implementación de un Sistema de Reconocimiento Emocional Facial con Python y OpenCV". 12º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2024). Catamarca.

30. Emiliano Camer; Martin Páez; Nicolás Vitek; Román Bond; Daniel Martín Morales; Diego Encinas. "Estudio de simulación de procesos aplicado a la optimización de sistemas de producción". En XXX Congreso Argentino de Ciencias de la Computación CACiC, Argentina. 2024.
31. Yamila Ferrufino; Román Bond; Daniel Martín Morales; Diego Encinas. "Evaluación de la Eficiencia Energética y el Impacto de las Aplicaciones de Redes Sociales". En XXX Congreso Argentino de Ciencias de la Computación CACiC, Argentina. 2024.
32. Ignacio DiLoreto; Román Bond; Martin Morales; Diego Encinas. "Estrategias de control del Dengue utilizando Modelado y Simulación basado en Agentes". En XXXI Congreso Argentino de Ciencias de la Computación (CACIC), Argentina. 2025.
33. Sabrina Russell; Franco Paiz; Román Bond; Martin Morales; Diego Encinas. "Diseño e implementación de un robot a tracción diferencial". En Actas 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2025), Argentina. 2025.
34. Cristian Iglesias, Jorge Osio, Diego Encinas. "Apache Flink para monitoreo energético en tiempo real". En Actas 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2025), Argentina. 2025.
35. Zarate E.M.; Encinas D.; Bond R. "Modelado de Fallas Vehiculares con Enfoque Predictivo Basado en Algoritmos Supervisados de Machine Learning". Jornadas Universitarias de Ciencias Exactas y Naturales (JUCEN), San Fernando del Valle de Catamarca, Argentina.

Análisis de Arquitecturas Distribuidas: Cloud, Fog y Edge Computing

Armando De Giusti [1][2], Marcelo Naiouf[1], Santiago Medina[1], Diego Montezanti[1][4], Laura De Giusti[1][3], Fernando G. Tinetti[1][3], Franco Chichizola[1], Enzo Rucci[1][3], Adrián Pousa[1], Victoria Sanz[1][3], Diego Encinas[1][4], Ismael Rodríguez[1], Sebastián Rodríguez Eguren[1], Leandro Libutti[1], Manuel Costanzo[1], Octavio Pérez Balcedo[1], Matías Manzin[1], Leonel Guerrico[1], Federico Walas Mateo[1][4]

[1]Instituto de Investigación en Informática LIDI (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
[2]CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
[3]CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
[4]Universidad Nacional Arturo Jauretche

{degiusti, mnaiouf, smedina, dmontezanti, ldgiusti, fernando, francoch, erucci, apousa, vsanz, dencinas, ismael, seguren, llibutti, mcostanzo, operezbalcedo, mmanzin, lguerrico}@lidi.info.unlp.edu.ar, fedewalas@gmail.com

Resumen

El eje principal de esta línea de I+D es el análisis de arquitecturas distribuidas para integrarlas con los paradigmas de Cloud, Fog y Edge Computing. Se estudia el despliegue de aplicaciones distribuidas que permitan almacenar, procesar y visualizar los datos generados por nodos sensores.

Los temas centrales de investigación son:

- El despliegue y configuración de servicios específicos para el control de datos.
- El análisis y uso de plataformas de internet de las cosas.
- La implementación de estrategias de resiliencia para procesamiento y almacenamiento distribuido en los diferentes niveles de la arquitectura.
- El estudio de la integridad de las comunicaciones en función a la distribución de los nodos.
- La comparación de los tiempos de respuesta de cada nivel de la arquitectura para la optimización del desempeño general del sistema.
- La optimización del consumo energético mediante la distribución de las tareas de procesamiento en cada una de las capas.
- La evaluación de las estrategias de resiliencia y optimización de consumo para el despliegue.
- El desarrollo de nodos sensores para la recolección de datos en diferentes entornos específicos.
- El modelado de sistemas complejos de tiempo real y el desarrollo de software inteligente para la toma de decisiones.
- El diseño y desarrollo de gemelos digitales para aplicaciones de tiempo real (industria de procesos/gestión sanitaria/área agrícola), en la búsqueda de respuestas predictivas inteligentes.

Palabras clave: *Arquitecturas Distribuidas. Cloud Computing. Fog Computing. Edge Computing. IoT. Aplicaciones de Tiempo Real. Eficiencia Energética. Resiliencia. Gemelos Digitales.*

Contexto

Esta línea de investigación se desarrolla dentro del III-LIDI de la Facultad de Informática de la UNLP y forma parte del proyecto “Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, *Fog-Edge-Cloud*, Big Data, Robótica, y Tiempo Real”.

Dentro de los temas relacionados a esta línea de investigación, existen cooperaciones con diferentes universidades de Argentina, de América Latina y Europa, en proyectos financiados por ERASMUS, CyTED y la OEI (Organización de Estados Iberoamericanos).

En particular, se realiza una colaboración con el proyecto “Computación de Altas Prestaciones Eficiente y Segura para

Aplicaciones de Servicios de Salud Inteligentes", de la Universidad Autónoma de Barcelona (UAB).

Se cuenta con financiamiento de diferentes empresas de Argentina, en lo referente a la formación de recursos humanos en las temáticas de *Cloud*, *Fog* y *Edge Computing*.

Por otro lado, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del Ministerio de Educación, Cultura, Ciencia y Tecnología de la Nación.

Además, desde la línea de investigación y el proyecto principal se participa en una iniciativa denominada "Consorcio de I+D+I en *Cloud Computing*, *Big Data* y *Emergent Topics*", conformado por diversas universidades de Argentina y España. Asimismo, se está trabajando en cooperación con la UAB (España), la UNNOBA y la UNAJ integrando proyectos específicos relacionados con las áreas de aplicación de Gemelos Digitales.

Introducción

El despliegue y utilización de nodos sensores heterogéneos en distintos entornos físicos conlleva la generación de grandes volúmenes de datos. Para gestionarlos, se requiere la implementación de arquitecturas distribuidas capaces de manejar las cargas de procesamiento y almacenamiento, de manera de cumplir con los requerimientos de las aplicaciones de tiempo real en cuanto a tiempos de respuesta.

El paradigma de arquitectura distribuida denominado "*Edge-Fog-Cloud Computing*" [1] permite integrar diversos niveles de procesamiento con diferentes servicios, buscando el beneficio de las aplicaciones en los siguientes aspectos:

- Reducir las comunicaciones hacia y desde el *Cloud*, mediante la realización de parte del procesamiento en los propios nodos sensores o en una capa intermedia (*Edge* o *Fog*).
- Permitir la flexibilidad del sistema mediante el traslado de la mayor cantidad de cómputo a los nodos sensores en función de la demanda y el contexto.
- Incorporar estrategias de resiliencia frente a fallos de comunicación o caídas de conexión a partir de la existencia de componentes redundantes en los distintos niveles.

Este modelo de arquitectura trae nuevos desafíos, entre los que se encuentran [2-5]:

- La administración del almacenamiento de datos heterogéneos en los distintos niveles, manteniendo la integridad y seguridad del sistema.
- La administración de los diferentes niveles de procesamiento para permitir la gestión inteligente de los recursos.
- La definición de protocolos y mecanismos de interoperabilidad entre las capas de la arquitectura distribuida y los diferentes servicios.
- El análisis e implementación de estrategias para tolerar diversos tipos de fallos.
- La reducción de los tiempos de respuesta realizando una mejor distribución de tareas entre los distintos niveles.
- La implementación de estrategias para disminuir el consumo energético, a partir del análisis sobre la conveniencia de procesar en los distintos niveles [30].
- La evaluación y utilización de diversas plataformas, aplicaciones y servicios en cada capa.
- El análisis de la escalabilidad cuando aumenta el número de nodos sensores y dispositivos finales.

Cloud Computing

El *Cloud Computing* (o cómputo en la nube), proporciona un acceso simplificado a grandes conjuntos de recursos físicos y lógicos (como infraestructura, plataformas de desarrollo, almacenamiento y/o aplicaciones), por medio de una interfaz de administración web, con un modelo de arquitectura "*virtualizada*" [6] [7]. Estos recursos se brindan en forma de servicios, pudiendo ser reconfigurados dinámicamente para adaptarse

a la variabilidad de la carga de trabajo. Así es posible lograr una utilización más flexible y elástica [8].

Cuando se integran distintos niveles, desde nodos sensores hasta cómputo en la nube, en una arquitectura distribuida, el nivel superior del Cloud se utiliza para procesar algoritmos complejos, que manejan grandes volúmenes de datos, de forma centralizada. Sin embargo, parte de estos datos pueden haber sido pre-procesados en las capas anteriores de *Edge* o *Fog* [9] [10].

Fog Computing

El paradigma de *Fog Computing* surge como respuesta al crecimiento sostenido de los desarrollos e implementaciones en el ámbito de IoT. Si bien estas soluciones requieren procesamiento en la nube, ciertas características propias de los entornos IoT dificultan su implementación cuando se emplean exclusivamente tecnologías de *Cloud* [10–12]. En este contexto, se introduce una capa de *Fog*, conformada por una plataforma intermedia que provee capacidades de cómputo, almacenamiento y comunicación entre los dispositivos *Edge* —conectados a los dispositivos encargados de la adquisición de datos— y la nube. Las características de *Fog Computing* [13-15] se orientan al manejo de:

- Aplicaciones distribuidas con requerimientos de baja latencia.
- Gran cantidad de nodos heterogéneos, que pueden escalar dinámicamente.
- Nodos con ubicaciones variables, que utilizan comunicaciones mayormente inalámbricas.

Edge Computing

La necesidad de utilizar un modelo de cómputo en el borde surge del crecimiento exponencial del número de nodos sensores disponibles con inteligencia local. Debido al aumento del tráfico generado por la IoT, es conveniente y necesario realizar tareas de preprocesamiento cerca de los sensores, de forma de poder integrar datos locales que se encuentran geográficamente dispersos, para luego enviarlos a los niveles superiores (*Fog* o *Cloud*) [16-19] generando una menor sobrecarga debida a las comunicaciones. Las principales ventajas del paradigma *Edge* son [10] [20] [21]:

- Proporciona respuesta al usuario en tiempo real, utilizando las capacidades de procesamiento local.
- Disminuye el volumen de datos transmitidos a capas superiores.
- Conexión privada con servidores locales, que brinda mayor seguridad.
- Posibilita la reducción del consumo energético, debido a la distribución del procesamiento a las características de los componentes individuales.
- Mejora la eficiencia global a partir de la utilización de una arquitectura distribuida, débilmente acoplada.

Aplicaciones de tiempo real.

El modelo de Edge–Fog–Cloud Computing resulta especialmente adecuado para su aplicación en problemas de tiempo real —como *cloud robotics*, vehículos autónomos, emergencias sanitarias y monitoreo de salud personalizado—, ya que permite que las capas más cercanas a los sensores y a los usuarios ejecuten parte del procesamiento con menor latencia y menor volumen de comunicaciones. De este modo, es posible brindar respuestas casi inmediatas, mientras que la capa de Cloud se reserva para el procesamiento masivo de datos en modalidad *offline* (por ejemplo, mediante algoritmos sofisticados de entrenamiento de modelos de comportamiento, cuyos resultados se traducen luego en mejoras en las tareas ejecutadas en las capas *Edge* y *Fog* [22][23]).

En particular, resulta de interés el modelado de sistemas complejos, tales como los propios de la industria de procesos, la agroindustria y la gestión de emergencias sanitarias, en los cuales es necesario recolectar datos de sensores en tiempo real y ejecutar software inteligente orientado a la toma de decisiones.

Eficiencia energética

Actualmente, la mejora de la eficiencia energética constituye un tema central en el ámbito de la informática, especialmente a partir de la aparición de plataformas con un

gran número de procesadores orientadas al alto desempeño. En este contexto, numerosos esfuerzos se concentran en incrementar la eficiencia energética y en desarrollar metodologías de medición que permitan obtener métricas de evaluación adecuadas [23].

El modelo *Edge–Fog–Cloud* presenta mejoras en el consumo energético principalmente debido a dos factores:

- Los procesadores utilizados en la capa *Edge* suelen ser de bajo consumo y, en muchos casos, cuentan con modos de suspensión (*sleep*)
- La reducción de las comunicaciones implica una disminución del consumo energético asociado al tráfico de red.

Líneas de Investigación, Desarrollo e Innovación

- Análisis de arquitecturas distribuidas para la aplicación de los paradigmas de *Edge* y *Fog Computing*.
- Integración y gestión de servicios en las capas de *Edge* y *Fog*.
- Evaluación de la seguridad e integridad de las comunicaciones entre las capas.
- Utilización de plataformas y servicios específicos de IoT para integrar de los niveles *Edge* y *Fog* con el *Cloud*.
- Análisis de estrategias de distribución del procesamiento y de los datos entre las distintas capas (por ejemplo, trasladando parte de la inteligencia al nivel *Edge* y a los nodos sensores), con el objetivo de mejorar los tiempos de respuesta y reducir el consumo energético [30].
- Análisis de la escalabilidad de las aplicaciones sobre la arquitectura distribuida.
- Obtención de métricas de eficiencia considerando tiempos de respuesta/costo de comunicaciones/consumo energético.
- Evaluación y aplicación de estrategias de resiliencia y control del consumo energético sobre plataformas basadas en *Edge*, *Fog* y *Cloud* [18] [24].
- Desarrollo de diferentes dispositivos para aplicaciones específicas: nodos sensores, robots, drones, etc.
- Utilización de tecnologías de *Cloud, Fog* y *Edge Computing* para controlar procesos de Transformación Digital en la industria y las organizaciones, mediante el desarrollo de aplicaciones de tiempo real [25] [26].
- Investigación de herramientas de desarrollo de Gemelos Digitales para aplicaciones industriales y de sistemas inteligentes de tiempo real.

Objetivos y Resultados Esperados

Investigación experimental por realizar

- Despliegue de arquitecturas distribuidas en donde se integren *Edge*, *Fog* y *Cloud Computing*.
- Integración de plataformas IoT, servicios específicos y servicios provistos por *Clouds* públicos.
- Evaluación de diversos protocolos de comunicación para la gestión de nodos de IoT, respecto de la integridad y los tiempos de respuesta.
- Análisis e implementación de estrategias de resiliencia [18] y control de consumo energético en aplicaciones que utilizan los niveles de *Edge, Fog* y *Cloud*.
- Definición de estrategias de balance de carga en aplicaciones que utilizan la arquitectura distribuida.
- Aplicación: red de nodos sensores inteligentes para monitoreo y control de consumo energético y parámetros ambientales en entornos institucionales.
- Aplicación: exploración y mapeo de entornos mediante robots, monitoreados y controlados desde la capa *Edge*.
- Aplicación: evaluación y utilización de modelos y heurísticas de procesamiento de imágenes en nodos robots.
- Aplicación: configuración y despliegue de ROS para el control de robots navegadores.

Resultados obtenidos

- Se han desplegado y evaluado diversas plataformas IoT (como *Thingsboard* o

Home Assistant) para la gestión de nodos sensores heterogéneos [27] [31].

- Se han utilizado servicios específicos para la gestión de nodos IoT.
- Se configuraron, probaron y compararon modelos de inteligencia artificial para el procesamiento de imágenes en vehículos robots.
- Se ha ampliado la capacidad de gestión del nodo *Edge* (basado en *Raspberry Pi*) para controlar los movimientos de múltiples robots heterogéneos incorporando resiliencia [32][33].
- Se ha desarrollado un nodo inteligente para la medición del consumo energético, basado en sensores de potencia, capaz de comunicarse con el sistema distribuido [35].
- Se han aplicado diversos protocolos de comunicación para estudiar los volúmenes de tráfico y los tiempos de respuesta en aplicaciones distribuidas en tiempo real.
- Se ha desarrollado una plataforma *ad-hoc* para la administración de nodos que controlan el consumo energético y parámetros ambientales en diversos entornos.
- Se ha desarrollado un vehículo robot con cámara y distintos sensores, conectado a un nodo *Edge* para su uso en diversas aplicaciones distribuidas.
- Se ha estudiado el impacto de aplicar diversas estrategias de resiliencia sobre una arquitectura distribuida basada en microservicios [28].
- Se ha realizado una estancia postdoctoral en el III-LIDI, vinculada con Gemelos Digitales y su aplicación en la industria de procesos [29].

Organización de Eventos

Desde la línea de investigación, se ha participado en 2025 de la organización de las XIII Jornadas de *Cloud Computing*, *Big Data & Emerging Topics* (JCC-BD&ET 2025), con participación de especialistas académicos del país y el exterior, además de empresas con experiencia en los temas de interés. Además, se ha participado de la 11va Expo Ciencia y Tecnología en la Facultad de Informática con el proyecto “Integración de robots y nodos sensores a servicios y plataformas”, habiendo contado con el trabajo de 2 alumnos de grado de la carrera de Ingeniería en Computación.

Formación de Recursos Humanos

Dentro de la línea de I/D están en curso 1 Tesis de Doctorado, 1 Tesis de Maestría y 1 Trabajo Final Integrador de Especialización.

Además, se están llevando adelante 3 trabajos finales de grado y 2 prácticas profesionales supervisadas.

Se finalizaron 2 trabajos de grado y 2 prácticas profesionales supervisadas.

Se presentó el proyecto CyTED "Gemelos Digitales en la gestión predictiva e inteligente de servicios, productos, procesos o sistemas" junto a otras instituciones académicas de Iberoamérica. (Argentina, Uruguay, Chile, Colombia, Brasil, Ecuador y España).

Varios integrantes de la línea de investigación participan en el dictado de las carreras de Doctorado en Ciencias Informáticas, Maestría y Especialización en Cómputo de Altas Prestaciones y Maestría en Gestión y Tecnología de Ciudades Inteligentes de la Facultad de Informática de la UNLP (acreditadas por CONEAU), por lo que potencialmente pueden generarse tesis de doctorado y maestría, además de trabajos finales de especialización.

Se dictaron tres cursos de postgrado: “Conceptos de *Cloud*, *Fog* y *Edge Computing*”, “Gobernanza y Gestión de Tecnologías de la Información” y “Tópicos de Fallos y Resiliencia en Arquitecturas Distribuidas”.

Docentes de la línea de investigación dictaron el curso *“Introducción a las redes de sensores, Edge Computing y plataformas IoT”* en la XXIX Escuela Internacional de Informática del XXXI Congreso Argentino de Ciencias de la Computación en la Sede Atlántica de la Universidad Nacional de Río Negro, además de participar en la coordinación de *workshops* del mismo evento.

Se dictó el primer curso relacionado a *Cloud*, *Edge* y *Fog* Computing, perteneciente a la Maestría en Transformación Digital. Por

otra parte, se dictaron diversos módulos de cursos en el contexto del Consorcio UNI UEAR (Alianza Universitaria Argentina Europea para la Transformación Digital). Por último, algunos integrantes de la línea de investigación participan en el dictado de dos asignaturas directamente relacionadas con los temas de ésta: "*Cloud Computing* y *Cloud Robotics*" y "Conceptos y Aplicaciones en *Big Data*", dentro de las carreras de grado.

Existe cooperación con grupos de otras universidades del país y del exterior, con posibilidades de realizar tesis en colaboración.

Referencias

[1] Mohan N., Kangasharju J. "Edge-Fog Cloud: A Distributed Cloud for Internet of Things Computations", 2016 Cloudification of the Internet of Things (CIoT), Paris, France, 2016, pp. 1-6.

[2] P. Garcia Lopez et al. "Edge-centric computing: Vision and challenges," *SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 37–42, Sep. 2015.

[3] M. Yannuzzi et al. "Key ingredients in an iot recipe: Fog computing, cloud computing, and more fog computing," in *IEEE CAMAD*, 2014.

[4] K. Hong et al. "Mobile fog: A programming model for large-scale applications on the internet of things," in *ACM SIGCOMM Workshop on Mobile Cloud Computing*,2013.

[5] A. Chandra, J. Weissman, and B. Heintz, "Decentralized edge clouds," *Internet Computing, IEEE*, vol. 17, no. 5, pp. 70–73, Sept 2013.

[6] E. Rucci, M. Naiouf, F. Chichizola, and L. De Giusti "Cloud Computing, Big Data & Emerging Topics. 8th Conference, JCC-BD&ET 2020, La Plata, Argentina, September 8-10, 2020, Proceedings". Springer CCIS, ISBN: 978-3-030-61218-4, 2020.

[7] Xing, Y., Zhan, Y.: "Virtualization and Cloud Computing". In: Proceedings pp.305-312, Springer Link. ISBN 978-3-642-27323-0. (2012). Morgan Kaufmann. 2013.

[8] Velte, A.T., Velte, T.J., Elsenpeter, R.: "Cloud Computing: A Practical Approach". McGraw Hill Professional. 2009.

[9] Ashkan Yousefpour, Caleb Fung, Tam Nguyen, Krishna Kadiyala, Fatemeh Jalali, Amirreza Niakanlahiji, Jian Kong, Jason P. Jue. "All one needs to know about fog computing and related edge computing paradigms: A complete survey", Journal of Systems Architecture, Volume 98, 2019, Pages 289-330, ISSN 1383-7621,

[10] P. Lea, "IoT and Edge Computing for Architects: Implementing edge and IoT systems from sensors to clouds with communication systems, analytics, and security". Packt Publishing, 2020.

[11] Keyan Cao, Yefan Liu, Gongjie Meng, Quimeng Sun "An Overview on Edge Computing Research" in IEEE Access, vol. 8, pp. 85714-85728, 2020.

[12] W. S. Shi, X. Z. Zhang, and Y. F.Wang, "Edge computing: State-of-the-art and future directions," J. Comput. Res. Develop., vol. 56, no. 1, pp. 1_21, 2019.

[13] X. Hong and Y. Wang, "Edge computing technology: Development and countermeasures," *Chin. J. Eng. Sci.*, vol. 20, no. 2, p. 20, 2018.

[14] D. Evans. "The Internet of Things How the Next Evolution of the Internet is Changing Everything". Available:
https://www.researchgate.net/publication/30612290

[15] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30_39, Jan. 2017.

[16] Y. Q. Gao, H. Bguan, and Z. W. Qi, "Service level agreement-based energy-Efficient resource man agreement in cloud data centers," *Comput. Elect. Eng.*, vol. 40, no. 5, pp. 1621_1633, 2014.

[17] W. Shi, H. Sun, J. Cao, Q. Zhang, and W. Liu, "Edge computing-an emerging computing model for the Internet of everything era," *J. Comput. Res. Develop.*, vol. 54, no. 5, pp. 907_924, May 2017.

[18] Medina, S., Montezanti, D., Gómez D'Orazio, L., Garay, F., De Giusti, A., & Naiouf, M. (2023, June). Distributed Architectures Based on Edge Computing, Fog Computing and End Devices: A Conceptual Review Incorporating Resilience Aspects. In Conference on Cloud Computing, Big Data & Emerging Topics (pp. 31-44). Cham: Springer Nature Switzerland.

[19] F. Bonomi, R. Milito, J. Zhu and S. Addepalli, "Fog Computing and Its Role in the Internet of Things", in *MCC '12: Proceedings of the first edition of the MCC workshop on Mobile cloud computing.*
Association for Computing Machinery, New York,NY, United States, 2012.

[20] M. Asemani, F. Jabbari, F. Abdollahei and P. Bellavista, "A Comprehensive Fog-enabled Architecture for IoT Platforms", *High-Performance Computing and Big Data Analysis. TopHPC 2019. Communications in Computer and Information Science, vol 891. Springer, Cham,* 2019.

[21] W. S. Aung and S. Aung Nyein Oo, "Monitoring and Controlling Device for Smart Greenhouse by using Thinger.io IoT Server". *International Journal of Trend in Scientific Research and Development*, 2019.

[22] Dr. S. K. Selvaperumal, W. Al-Gumaei, R. Abdulla and V. Thiruchelvam, "Integrated Wireless Monitoring System Using LoRa and Node-Red for University Building". *Journal of Computational and Theoretical Nanoscience, Volume 16, Number 8, American Scientific Publishers*, 2019.

[23] M. Pi Puig, J. M. Paniego, S. Medina, S. Rodriguez Eguren, L. Libutti, J. Lanciotti, J. de Antueno, C. Estrebou, F. Chichizola and L. De Giusti, "Intelligent Distributed System for Energy Efficient Control", *Cloud Computing and Big Data. JCC&BD*

2019. Communications in Computer and Information Science, vol 1050. Springer, Cham, 2019.
[24] Medina, S., Montezanti, D. M., Gomez D'Orazio, L., Compagnucci, E., De Giusti, A. E., & Naiouf, M. (2022). Incorporating Resilience to Platforms based on Edge and Fog Computing. In X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022).
[25] F. Walas, A.Redchuk, J.Tornillo. "Gestión de datos de la operación en la industria a través de plataformas IIoT. Experiencia en una empresa del sector alimentos en Argentina"
[26] Borek, A., Prill, N. "Driving Digital Transformation through Data and AI: A Practical Guide to Delivering Data Science and Machine Learning Products." Kogan Press, 2023.
[27] Gomez D´Orazio, L., Medina, S., Montezanti, D. "Integración de una red de sensores con una plataforma IoT para control inteligente de aulas". Proceedings del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022) – EUDELAR, pp. 1002 – 1007. ISBN 978-987-1364-31-2 (2022). http://sedici.unlp.edu.ar/handle/10915/149102.
[28] Suarez, Sergio Leonel. "Análisis de patrones de resiliencia en una arquitectura basada en microservicios". Tesina de Licenciatura en Sistemas. Universidad Nacional de La Plata. 2022. http://sedici.unlp.edu.ar/handle/10915/149187
[29] Federico Walas (2023) Tesis Doctoral. "Nuevos modelos de negocio en el paradigma Industria 5.0. Inteligencia Artificial y Aprendizaje Automático para optimizar procesos industriales"
[30] De Giusti, A., Montezanti, D., Medina, S. y Libutti, L. (2024). Edge Computing vs Cloud Computing: Optimizando el Consumo Energético. In XXX Congreso Argentino de Ciencias de la Computación (CACIC)(La Plata, 7 al 11 de octubre de 2024).
[31] Dell´Arciprete, D., Medina, S., & Montezanti, D. M. (2024). Despliegue y análisis de la plataforma Home Assistant para la integración de nodos sensores heterogéneos. In XXX Congreso Argentino de Ciencias de la Computación (CACIC)(La Plata, 7 al 11 de octubre de 2024).
[32] Pérez Balcedo, O., Medina, S., & Montezanti, D. M. (2024). Estrategia para el control resiliente de vehículos robot utilizando Edge Computing para mapeo de entornos. In XXX Congreso Argentino de Ciencias de la Computación (CACIC)(La Plata, 7 al 11 de octubre de 2024).
[33] Pérez Balcedo, O., Medina, S., & Montezanti, D. M. (2024). Adaptación de un sistema distribuido para el manejo de robots heterogéneos orientados al mapeo de entornos. In XXXI Congreso Argentino de Ciencias de la Computación (CACIC)(Viedma, 6 al 10 de octubre de 2025).
[34] Guerrico, L., Medina, S., & Montezanti, D. M. (2024). Análisis comparativo de modelos de Inteligencia Artificial para el Procesamiento de Imágenes en Vehículos Robots. In XXXI Congreso Argentino de Ciencias de la Computación (CACIC)(Viedma, 6 al 10 de octubre de 2025).
[35] Manzin, M., Medina, S., Libutti, L., Montezanti, D., & Navarria, L. (2025, June). Implementation of a smart electricity meter. In Conference on Cloud Computing, Big Data & Emerging Topics (pp. 3-14). Cham: Springer Nature Switzerland.

Aplicaciones Distribuidas en Arquitectura de Microservicios. Desafíos técnicos y mejores prácticas

Nelson Rodríguez[1], María Murazzo[1], Marcelo Moreno[1], Diego Medel[1], Fabiana Píccoli[2], Miguel Méndez Garabetti[3], Daniel Arias Figueroa[4], Melina Daroni Perini [5], Paulina Lenzo[5], Maximiliano Alcalá[5], Valentín Luján Vila[5]

[1]Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
[2] Departamento de Informática - F.C.F.M. y N – UNSL
[3] Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat
[4] Consejo de Investigación, FCE, Universidad Nacional De Salta
[5] Alumno Avanzado Licenciatura en Sistemas de Información y Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.
Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

nelson@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar, mpmoren@gmail.com, mpiccoli@unsl.edu.ar, mendez-garabettimiguel@uch.edu.ar, daaf@cidia.unsa.edu.ar, melinadaroni@gmail.com, paulilenzo@gmail.com, alcalamaximiliano03@gmail.com, valenlujan111@gmail.com

Resumen

Las aplicaciones distribuidas han revolucionado las prácticas de desarrollo de software al descentralizar los componentes, facilitar la escalabilidad y permitir la agilidad en el diseño y despliegue de sistemas. Si bien muchos problemas presentes al principio se resolvieron, otros como la pérdida de control, la seguridad, la comunicación eficiente entre componentes y la tolerancia a falla aún siguen sin resolverse completamente, lo cual la presentan como un área de investigación desafiante. Además, los entornos distribuidos pueden ocasionar otros problemas adicionales, como sincronización, escalabilidad, consistencia, optimización del rendimiento, entre otros. Los microservicios son arquitecturas de software esenciales para aplicaciones distribuidas modernas, pero requieren un enfoque disciplinado en diseño, seguridad, observabilidad y automatización. Adoptar buenas prácticas permite superar los retos técnicos y construir sistemas más robustos y sostenibles. El presente proyecto tiene por objetivo analizar y definir un conjunto de mejores prácticas para el diseño, implementación y mantenimiento de aplicaciones distribuidas basadas en arquitectura de microservicios, considerando los principales desafíos técnicos actuales.

Palabras clave: *Distributed Applications. Microservices Architecture. Distributed Systems. Technical challenges. Best Practices*

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es la línea marco del proyecto de investigación Aplicaciones Distribuidas en Arquitectura de Microservicios. Desafíos técnicos y mejores prácticas, cuya propuesta se presentó en convocatoria para el período 2026-2027 y se encuentra en evaluación. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil, distribuida y de alta performance, desde hace más de 25 años, al que se han incorporado estudiantes de informática. Como en proyectos anteriores, se continúa el trabajo con

investigadores de otras universidades, lo cual favorece a las instituciones participantes.

Introducción

Los sistemas distribuidos han surgido como herramientas indispensables para abordar las crecientes demandas de potencia de cálculo, escalabilidad y utilización eficiente de los recursos, que resulta imposible de obtener mediante un modelo centralizado. Además recientemente, el rápido crecimiento de las cargas de trabajo de inteligencia artificial, ha impulsado la necesidad de sistemas de computación capaces de procesar conjuntos de datos que superen escalas de petabytes, como los necesarios para entrenar grandes modelos de lenguaje, que implican cientos de miles de millones de parámetros [1].

Los microservicios son un desarrollo que ofrecen muchas ventajas en comparación con las antiguas arquitecturas monolíticas. Por eso, muchas de las grandes empresas tecnológicas han logrado con éxito realizar la transición a microservicios [2].

El tamaño del mercado de microservicios en la nube ha alcanzado los 2.270 millones de dólares en 2025. Se espera que crezca hasta 5.820 millones en 2030 con una tasa de crecimiento anual compuesta del 20,7% [3].

Este crecimiento está impulsado por empresas que trasladan aplicaciones de aplicaciones monolíticas a arquitecturas modulares nativas de la nube. Alrededor del 62,3% de las organizaciones informaron del uso de microservicios y tecnologías de contenedores para apoyar una codificación más flexible y escalar el desarrollo de aplicaciones y el despliegue de software más rápido en entornos distribuidos [4].

La investigación sobre aplicaciones distribuidas en arquitectura de microservicios es de gran relevancia debido a su impacto en distintos ámbitos: científico, tecnológico, social y estratégico. Desde una perspectiva científica, este estudio permite profundizar en los principios de diseño modular, comunicación distribuida y gestión eficiente de datos en entornos complejos. Analizar los desafíos técnicos y las mejores prácticas proporciona evidencia empírica que contribuye al conocimiento sistemático sobre patrones de diseño, escalabilidad, tolerancia a fallos y eficiencia en sistemas distribuidos, fortaleciendo el corpus académico en ingeniería de software y tecnologías cloud-native [5].

En el ámbito tecnológico, la investigación es crucial para optimizar el desarrollo de aplicaciones modernas que deben responder a demandas de alta disponibilidad, escalabilidad y resiliencia. La adopción de microservicios permite reducir tiempos de despliegue, facilitar la integración continua y mejorar la eficiencia operativa, lo que impacta directamente en la competitividad de organizaciones y empresas tecnológicas. Identificar estrategias óptimas de comunicación, escalabilidad y resiliencia permitirá implementar soluciones de software más robustas, rápidas y confiables [6].

Desde una perspectiva social, los sistemas basados en microservicios sustentan plataformas que afectan directamente la vida cotidiana de los usuarios, tales como servicios financieros digitales, comercio electrónico, educación virtual y salud en línea. Mejorar la disponibilidad, eficiencia y seguridad de estas aplicaciones contribuye a la calidad del servicio, la confianza de los usuarios y la reducción de interrupciones en actividades críticas, generando un beneficio tangible para la sociedad [7].

En términos estratégicos, el proyecto permite a las organizaciones optimizar sus recursos tecnológicos y mantener ventaja competitiva. La identificación de desafíos técnicos y la sistematización de buenas prácticas facilitan la planificación de infraestructuras distribuidas, la implementación de patrones de resiliencia y

el diseño de políticas de desarrollo sostenibles. Los resultados del estudio aportan insumos valiosos para guías, estándares y recomendaciones técnicas que favorecen la toma de decisiones informada y la gestión eficiente de sistemas complejos [5].

En síntesis, el estudio de aplicaciones distribuidas basadas en microservicios es pertinente y necesario, ya que genera conocimiento científico, impulsa la innovación tecnológica, mejora la experiencia de los usuarios y proporciona ventajas estratégicas a las organizaciones. Los aportes esperados incluyen la identificación de mejores prácticas, la comprensión de los desafíos técnicos y la propuesta de soluciones que optimicen la escalabilidad, resiliencia y eficiencia de sistemas distribuidos modernos [9].

Aplicaciones Distribuidas y Microservicios

Las aplicaciones distribuidas son sistemas de software cuyos componentes se ejecutan en nodos múltiples interconectados, trabajando de manera coordinada para cumplir objetivos específicos. Este enfoque permite mejorar la escalabilidad, la tolerancia a fallos y la flexibilidad de las aplicaciones, al distribuir la carga de procesamiento y almacenamiento entre distintos servidores o servicios. La coordinación entre nodos se basa en modelos de comunicación sincrónica o asincrónica, mientras que la replicación y la sincronización de datos garantizan consistencia y disponibilidad [7].

En el desarrollo moderno de software, la creciente complejidad de las aplicaciones ha impulsado la adopción de arquitecturas distribuidas que ofrecen mayor modularidad y escalabilidad. En este contexto, los microservicios representan un paradigma clave. La arquitectura de microservicios consiste en descomponer una aplicación en servicios pequeños, autónomos y especializados, cada uno encargado de funciones concretas dentro del dominio de negocio. Esta aproximación permite que cada servicio sea desarrollado, desplegado y escalado de manera independiente, mejorando la flexibilidad, resiliencia y mantenibilidad de los sistemas.

Los microservicios se integran naturalmente como aplicaciones distribuidas, que se caracterizan por ejecutar componentes del software en nodos múltiples, interconectados por red y coordinados para cumplir un objetivo común. Esta distribución permite escalabilidad horizontal, tolerancia a fallos y transparencia en la localización de recursos. Cada microservicio actúa como un nodo autónomo, interactuando con otros mediante protocolos de comunicación sincrónica, como REST o gRPC, o asincrónica, a través de sistemas de mensajería y eventos, lo que optimiza el rendimiento y la resiliencia del sistema [5].

En la actualidad las organizaciones están comenzando a migrar sus sistemas monolíticos, a este tipo de entornos, ya sea dejando la lógica del negocio on premise o llevarla paulatinamente al cloud. Esta transición no solo implica un cambio tecnológico, sino también conceptual y organizacional. Seedat et al. [9], señalan que la migración exitosa requiere identificar límites de dominio claros, definir servicios autónomos y establecer mecanismos eficientes de coordinación y comunicación. Investigaciones recientes también han explorado la integración de microservicios con edge computing, inteligencia artificial para automatizar la partición de servicios y herramientas avanzadas de observabilidad que permiten monitorear sistemas distribuidos en tiempo real, anticipar fallos y optimizar la performance [10]; [7].

Por lo tanto, es claro que las aplicaciones distribuidas en arquitectura de microservicios representan una convergencia de principios de

sistemas distribuidos, modularidad y servicios desacoplados, ofreciendo escalabilidad, resiliencia y agilidad en el desarrollo de software.

Los microservicios presentan algunas limitaciones debido a que las app distribuidas son más difíciles de programar, las llamadas remotas son lentas, se enfrentan a fallas y se debe realizar el mantenimiento de código en diversos lenguajes y frameworks. Cada servicio requiere pruebas y monitoreo individualizados. No es sencillo lograr la coherencia de los datos, porque cada proveedor tiene su propia base de datos y sistema de gestión de transacciones [11]. El mecanismo de comunicación entre servicios es uno de los aspectos más importantes que se deben considerar [12].

Los microservicios introducen varios desafíos debido a que se incrementa la complejidad operacional, la administración del sistema distribuido y lo concerniente a la consistencia de los datos. Obstáculos clave incluyen: administración de la comunicación intra-servicios, asegurar la tolerancia a fallas, descubrimiento de servicios, depuración distribuida y riesgos de seguridad altos, al presentar una superficie de ataque expandida.

Existen además aspectos muy importantes para analizar en este tipo de arquitectura, como son: la observabilidad, dado que estrategias robustas aseguran alta confiabilidad de sistema y performance. Si bien los microservicios habilitan la escalabilidad y flexibilidad, introducen complejidad tal como varios puntos de falla, bloqueo de la performance y administración de recursos. [13]

Las buenas prácticas claves incluyen diseñar los límites del servicio basándose en las capacidades del negocio, asegurar un acoplamiento débil y utilizar protocolos de comunicación estandarizados como APIs RESTful o colas de mensajes para promover interacciones fluidas.

Además, el uso de tecnologías de contenedores como Docker y plataformas de orquestación como Kubernetes es esencial para garantizar la escalabilidad, la tolerancia a fallos y la facilidad de despliegue en distintos entornos [14]. Se deben fomentar además buenas prácticas para mejorar la seguridad utilizando técnicas como OAuth, autenticación JWT y control de acceso basado en roles para mitigar posibles vulnerabilidades, entre otras.

Adoptar un enfoque de gestión de datos descentralizado, mejora la consistencia y reduce la dependencia entre servicios. La monitorización y el registro efectivo, junto con los marcos de pruebas automatizadas, son fundamentales para mantener la salud y el rendimiento de los microservicios en todas las industrias.

Líneas de la Investigación

Objetivos

Los objetivos del grupo de investigación en esta línea de conocimiento son los siguientes:

Objetivo general:

- Analizar y proponer un conjunto de mejores prácticas para el diseño, implementación y mantenimiento de aplicaciones distribuidas basadas en arquitectura de microservicios, considerando los principales desafíos técnicos actuales.

Objetivos específicos:

- Analizar beneficios y limitaciones de la arquitectura de microservicios para el desarrollo de aplicaciones distribuidas.
- Investigar los principales desafíos técnicos que enfrentan las aplicaciones distribuidas.
- Documentar las mejores prácticas en áreas clave: diseño de servicios, comunicación entre componentes, gestión de datos, seguridad, despliegue y monitoreo.

Resultados Obtenidos

Si bien, la presentación del presente proyecto es reciente, en proyectos anteriores se han

desarrollado trabajos finales en el área y publicaciones científicas. Específicamente en el marco del último proyecto realizado por el grupo, cuya denominación fue: ‘‘Soporte Serverless para aplicaciones móviles de nueva generación’’ (2022-2025), se llevaron a cabo variadas publicaciones en congresos internacionales y además se concretaron trabajos finales de grado. Como parte de las tareas de investigación, se inició el estudio de la arquitectura de microservicios, generando publicaciones [16] y trabajos finales de alumnos de grado en arquitecturas distribuidas y en particular en microservicios.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los investigadores que figuran en este trabajo de la UNSJ, UNSL, UNSA, y UCH de Mendoza y ocho alumnos de grado, de la UNSJ..

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis de maestría en áreas afines y tres tesinas de grado en el área de Serverless computing, Virtualización y Computación distribuida.

Se espera aumentar el número de publicaciones, y además también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.

Referencias

[1] Achiam, J.; Adler, S.; et al. (2023) Gpt-4 technical report. arXiv:2303.08774.

[2] Gao, W.; Ren, S.; et al. (2025) Lattice-Based Group Signature with VLR for Anonymous Medical Service Evaluation System. Electronics 2025, 14, 680. https://doi.org/10.3390/electronics14040680

[3] Cloud Microservices Market Report 2026. ttps://www.thebusinessresearchcompany.com/report/cloud-microservices-global-market-report

[4] Cloud Microservices Market Size 2025 – 2034. https://www.gminsights.com/industry-analysis/cloud-microservices-market.

[5] Velepucha, V., & Flores, P. (2023). Survey on Microservices Architecture: Principles, Patterns and Migration Challenges. IEEE Access.

[6] Söylemez, M., Tekinerdogan, B., & Kolukısa Tarhan, A. (2022). Feature-Driven Characterization of Microservice Architectures. Applied Sciences, 12(9), 4424.

[7] Castro, L. F. S. de, & Rigo, S. (2023). Relating Edge Computing and Microservices: Systematic Literature Review. arXiv.

[8] Domakonda, D. (2025). Secure and Scalable Microservices Architecture: Principles, Benefits, and Challenges. IJSCSEIT.

[9] Seedat, M., Abbas, Q., & Ahmad, N. (2023). Systematic Mapping of Monolithic Applications to Microservices Architecture.

[10] Narváez, D., Battaglia, N., Fernández, A., & Rossi, G. (2025). Designing Microservices Using AI: Systematic Literature Review.

[11] Chitrak Vimalbhai D., Abhishek P., Utkarsh K. (2021) Microservices Software Architecture: A Review. In: International Journal for Research in Applied Science & Engineering Technology ISSN: 2321-9653; Vol 9 Issue XI - Available at www.ijraset.com.

[12] Almog, D.; Chassidim, H., Shlomo. M. (2021) Testability and Testing of Microservices-complex challenge. In: International Journal of Computers, vol. 6.

[13] U. Faseeha et al. (2025): Observability in Microservices: An In-Depth Exploration of Frameworks. IEEE Access, Vol. 13.

[14] Srinivasan Jayaraman & Anand Singh International Journal of Computer Science and Engineering (IJCSE) ISSN (P): 2278–9960; ISSN (E): 2278–9979. Vol. 13, Issue 2,

[15] Creswell, J. W. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.). Sage Publications.

[16] Atencio, Rodríguez, Murazzo (2024). Evaluación de parámetros asociados a la comunicación síncrona entre microservicios. XXX CACIC

Arquitectura IoT multiprotocolo para series temporales en el continuo borde–nube

Claudio Omar Biale [1], Selva Nieves Ivaniszyn [1], Rubén Luis María Castaño [1], Ariel Lutenberg [2]

[1] Departamento de Informática, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones.

[2] CONICET-GICSAFe, Laboratorio de Sistemas Embebidos, Facultad de Ingeniería, Universidad de Buenos Aires

claudiobiale@fceqyn.unam.edu.ar, selvai@fceqyn.unam.edu.ar, rcastano@fceqyn.unam.edu.ar, alutenb@fi.uba.ar

RESUMEN

Las arquitecturas IoT centralizadas en la nube suelen presentar alta latencia, consumo elevado de ancho de banda y vulnerabilidad ante conectividad intermitente, especialmente en despliegues geográficamente distribuidos. Este trabajo presenta SensorWave, una plataforma IoT distribuida para el continuo borde-nube que integra interoperabilidad multiprotocolo, almacenamiento local con compresión y motor de reglas, un servicio despachador sin estado para coordinar consultas globales y un lenguaje de consulta declarativo tipo SQL para operar sobre series temporales. Como validación, se desarrolló Huerta-IoT, una aplicación de monitoreo agrícola construida sobre SensorWave. Los resultados muestran que la propuesta reduce el tráfico ascendente y mejora los tiempos de respuesta, manteniendo la continuidad operativa y la disponibilidad de datos en escenarios con conectividad limitada.

Palabras clave: IoT, almacenamiento distribuido, continuo borde-nube, middleware, interoperabilidad, series temporales.

CONTEXTO

Este trabajo de investigación se desarrolla en el marco del proyecto "16/Q2875-TI" de la Universidad Nacional de Misiones denominado “Gestión distribuida de datos para aplicaciones de IoT”.

1. INTRODUCCIÓN

La expansión del Internet de las Cosas (IoT) ha impulsado la generación de volúmenes masivos de datos provenientes de sensores y actuadores heterogéneos, desplegados en dominios tan diversos como ciudades inteligentes, monitoreo de la salud, agricultura de precisión e Industria 4.0 [1, 2]. La naturaleza distribuida, dinámica y de gran escala de estos flujos de información excede las capacidades de arquitecturas centralizadas basadas exclusivamente en la computación en la nube, dando lugar a congestión en la red, incrementos significativos en la latencia de respuesta y elevados costos asociados con la transmisión continua de datos [1, 2].

Como respuesta a estas limitaciones, ha surgido el paradigma del continuo borde–nube como un modelo arquitectónico distribuido orientado a optimizar el procesamiento y la gestión de datos en

entornos IoT [2-4]. Este enfoque integra recursos de computación descentralizados y físicamente próximos a las fuentes de generación de datos con infraestructuras centralizadas en la nube, posibilitando una provisión de servicios coordinada, escalable y eficiente [1, 3, 4]. Al trasladar una parte significativa de las tareas de procesamiento hacia el borde de la red, se contribuye a mitigar la congestión del tráfico, reducir los tiempos de respuesta y optimizar el uso del ancho de banda, superando así las limitaciones inherentes a los modelos estrictamente centralizados [1, 2, 5].

Para materializar estos beneficios en despliegues de IoT a gran escala, se emplean estrategias de reducción de datos que incluyen métodos de compresión (tanto con pérdida como sin pérdida), esquemas de agregación para consolidar múltiples lecturas en representaciones sintéticas y enfoques de predicción capaces de anticipar patrones futuros, de modo que únicamente se transmitan las actualizaciones estrictamente necesarias [2, 5]. En conjunto, estas técnicas buscan reducir el volumen de datos transferidos y almacenados, y extender la autonomía de dispositivos con recursos restringidos sin degradar la utilidad de la información [2, 4, 5]. Asimismo, el modelo propuesto busca reforzar la seguridad y la soberanía de los datos mediante el procesamiento local de información sensible, evitando su exposición durante la transmisión hacia centros de datos remotos [1-3].

Desde la perspectiva de los servicios modernos de gestión de datos en la nube, el desacoplamiento entre los recursos de cómputo y almacenamiento se ha consolidado como un principio fundamental de diseño, ejemplificado por sistemas de almacenamiento de objetos como Amazon S3 [6-8]. Esta separación permite escalar de manera independiente las capacidades de procesamiento y persistencia, facilitando la elasticidad frente a cargas de trabajo variables e impredecibles [7, 8]. Al delegar la durabilidad de los datos en servicios especializados de almacenamiento, los recursos computacionales pueden dedicarse de forma prioritaria a la ejecución de consultas y tareas analíticas, evitando los costos derivados del acoplamiento estrecho característico de las arquitecturas de tipo *shared-nothing* y promoviendo una mayor flexibilidad operativa [7, 8].

No obstante, la gestión de datos a lo largo del continuo borde–nube enfrenta desafíos significativos asociados con la masividad, distribución y dinamismo del Big Data, así como a la heterogeneidad de dispositivos, protocolos y plataformas involucradas [1-5]. Entre estos retos se incluyen la dificultad de garantizar la consistencia de los datos entre las distintas capas de la arquitectura, la complejidad de orquestar recursos de manera eficiente en entornos altamente dinámicos y la necesidad de diseñar mecanismos agnósticos que aseguren la interoperabilidad en infraestructuras de gran escala [2-4].

En particular, los entornos IoT heterogéneos introducen problemas de interoperabilidad derivados de la coexistencia de múltiples protocolos, formatos de datos y estándares [9-11]. Para mitigar esta complejidad, los middlewares desempeñan un rol central al proporcionar servicios comunes de abstracción, traducción de protocolos y coordinación entre plataformas dispares, facilitando la integración fluida de dispositivos y sistemas heterogéneos [9-11].

Asimismo, consultar y analizar datos en entornos IoT distribuidos exige abstraer la complejidad derivada de su distribución física

y de su localización dinámica a lo largo del continuo borde–nube [12-14]. En este sentido, los lenguajes declarativos ofrecen una interfaz unificada para conjuntos de datos masivos y heterogéneos, ocultando diferencias de almacenamiento, esquemas y protocolos de comunicación [12, 14, 15]. Así, las consultas de alto nivel pueden formularse sin depender de la ubicación de los datos y el sistema asume la planificación, optimización y ejecución distribuida, ampliando el acceso a los datos IoT a usuarios no especializados en sistemas distribuidos [12-15].

Finalmente, la aplicabilidad de estas arquitecturas resulta especialmente relevante en dominios productivos como la agricultura de precisión, donde los sistemas IoT deben operar en entornos geográficamente distribuidos, con conectividad intermitente y fuertes restricciones de recursos. Estos escenarios constituyen un banco de pruebas adecuado para evaluar soluciones IoT integrales orientadas a la eficiencia, la escalabilidad y la autonomía operativa.

En este marco, el presente trabajo se alinea con los desafíos actuales de la gestión de datos IoT en el continuo borde-nube, particularmente en lo relativo a la interoperabilidad en entornos heterogéneos, la eficiencia en el uso de recursos y la explotación de datos distribuidos. Se espera contribuir con lineamientos, métodos y componentes arquitectónicos que favorezcan el diseño de soluciones escalables y sostenibles para la captura, gestión y acceso a datos IoT, facilitando su adopción en aplicaciones reales de alto impacto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El trabajo de investigación aborda los siguientes ejes principales:

- Análisis y diseño de arquitecturas de middleware multiprotocolo orientadas a habilitar mecanismos de interoperabilidad transparente y escalable en entornos de Internet de las Cosas (IoT) heterogéneos, considerando la integración de dispositivos, servicios y plataformas con distintos estándares de comunicación, modelos de datos y requisitos de calidad de servicio.
- Gestión e integración de recursos de almacenamiento distribuido entre nodos de computación en el borde y servicios en la nube, mediante la definición y aplicación de políticas avanzadas de retención, replicación y migración de datos.
- Diseño y desarrollo de estrategias de compresión diferencial adaptativa aplicadas a series temporales, orientadas a la optimización del uso del ancho de banda y de los recursos de almacenamiento local.
- Articulación de los niveles de cómputo en el borde y en la nube mediante servicios escalables capaces de coordinar consultas globales distribuidas.
- Diseño de interfaces de administración unificadas que habiliten la operación integrada y coherente de sistemas desplegados a lo largo del continuo borde-nube, mediante el uso de lenguajes de consulta formales para la gestión, monitorización y orquestación de recursos distribuidos.
- Implementación de aplicaciones que integren y optimicen el uso coordinado de tecnologías de computación en el borde y en la nube para el monitoreo, control y automatización de procesos en sistemas de producción agropecuaria.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Durante el desarrollo del presente trabajo se diseñó e implementó un sistema distribuido, denominado SensorWave, orientado a la captura, el almacenamiento y la consulta de datos de forma transparente para el usuario final. La información se gestiona mediante un enfoque híbrido, alojándose tanto en el borde de la red como en un servicio de almacenamiento de objetos en la nube.

A continuación, se presentan los resultados obtenidos [16]:

- Implementación de un middleware con soporte para los protocolos HTTP, CoAP y MQTT, que proporciona niveles de garantía de entrega de mensajes entre emisor y receptor de "como máximo una vez" y "al menos una vez", mantenidos de forma consistente entre protocolos, con capacidad de traducción bidireccional de mensajes y compatibilidad con el uso de comodines (wildcards) en todos los protocolos soportados.
- Clientes para Go, ESP-IDF (C) y MicroPython con una API unificada, soporte de reconexión automática y formato de mensajes uniforme.
- Sistema de almacenamiento en el borde basado en PebbleDB, que incorpora compresión configurable por serie temporal tanto a nivel de muestras (XOR, Delta-Delta, RLE, diccionario, bit-packing) como a nivel de bloque (LZ4, ZSTD, Snappy, Gzip). El sistema permite definir parámetros específicos por serie, tales como el tamaño de bloque y el tiempo de elegibilidad para migración, y dispone de migración automática por bloques hacia un sistema de almacenamiento en la nube compatible con la interfaz S3.
- Motor de inferencia de reglas en el borde con soporte para la evaluación de condiciones basadas en agregaciones temporales y para la definición de acciones de publicación personalizables.
- Módulo despachador sin estado, escalable horizontalmente, diseñado para la gestión de consultas globales en el continuo borde-nube.
- Lenguaje de consulta declarativo, con una sintaxis similar a SQL, diseñado para operar tanto en entornos de computación en el borde, proporcionando soporte para operaciones completas sobre series temporales, como en un modo de despacho orientado a la emisión de consultas globales de solo lectura a lo largo del continuo borde–nube. El lenguaje se expone mediante una interfaz de tipo REPL (*Read–Eval–Print Loop*), que habilita la exploración interactiva y reiterativa de los datos.
- Huerta-IoT es una aplicación web orientada al ámbito agrícola para el monitoreo de cultivos hortícolas, desarrollada sobre la plataforma SensorWave. Esta solución tecnológica permite a los productores gestionar y supervisar nodos sensores distribuidos espacialmente en el campo, facilitando la recolección, transmisión y análisis de datos agroambientales

A continuación, se detallan las principales líneas de acción para dar continuidad a este trabajo:

- Ampliación del conjunto de

protocolos compatibles con el middleware y fortalecimiento de los mecanismos de seguridad y la implementación de esquemas avanzados de autenticación.

- Implementación de esquemas de clustering de nodos en el borde orientados a garantizar alta disponibilidad.
- Desarrollo de innovaciones tecnológicas y aplicaciones especializadas orientadas a la optimización y fortalecimiento de la producción agrícola en ámbitos locales y regionales.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por docentes e investigadores que pertenecen a las diferentes instituciones universitarias. En el marco del presente trabajo de investigación, se encuentra en desarrollo la tesis de Maestría en Internet de las Cosas del Esp. Lic. Claudio Omar Biale bajo la dirección académica del Dr. Ing. Ariel Lutenberg.

Asimismo, se dará continuidad al desarrollo de actividades de extensión en el marco del ciclo de conferencias titulado "AgriTech: soluciones IoT y tecnológicas para la agricultura", que estarán a cargo de integrantes del equipo de trabajo.

5. BIBLIOGRAFÍA

[1] Andriulo, F. C., Fiore, M., Mongiello, M., Traversa, E., & Zizzo, V. (2024). Edge computing and cloud computing for internet of things: A review. In *Informatics* (Vol. 11, No. 4, p. 71). MDPI.

[2] Kreković, D., Krivić, P., Žarko, I. P., Kušek, M., & Le-Phuoc, D. (2025). Reducing communication overhead in the IoT–edge–cloud continuum: A survey on protocols and data reduction strategies. *Internet of things*, *31*, 101553.

[3] Belcastro, L., Marozzo, F., Orsino, A., Talia, D., & Trunfio, P. (2025). Navigating the edge-cloud continuum: A state-of-practice survey. *arXiv preprint arXiv:2506.02003*.

[4] Gkonis, P., Giannopoulos, A., Trakadas, P., Masip-Bruin, X., & D'Andria, F. (2023). A survey on IoT-edge-cloud continuum systems: Status, challenges, use cases, and open issues. *Future Internet 15, 12*.

[5] Cheikh, I., Roy, S., Sabir, E., & Aouami, R. (2026). Energy, scalability, data and security in massive IoT: Current landscape and future directions. *IEEE Internet of Things Journal*.

[6] Brantner, M., Florescu, D., Graf, D., Kossmann, D., & Kraska, T. (2008). Building a database on S3. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 251-264).

[7] Dageville, B., Cruanes, T., Zukowski, M., Antonov, V., Avanes, A., Bock, J., Claybaugh, J., Engovatov, D., Hentschel, M., Huang, J., Lee, A. W., Motivala, A., Munir, A. Q., Pelley, S., Povinec, P., Rahn, G., Triantafyllis, S., & Unterbrunner, P. (2016). The snowflake elastic data warehouse. In *Proceedings of the 2016 International Conference on Management of Data* (pp. 215-226).

[8] Durner, D., Leis, V., & Neumann, T. (2023). Exploiting cloud object storage for high-performance analytics. *Proceedings of the VLDB Endowment*, *16*(11), 2769-2782.

[9] Aly, M., Khomh, F., Guéhéneuc, Y. G., Washizaki, H., & Yacout, S. (2018). Is fragmentation a threat to the success of the internet of things?. *IEEE Internet of Things Journal*, *6*(1), 472-487.

[10] Bouguettaya, A., Sheng, Q. Z., Benatallah, B., Neiat, A. G., Mistry, S., Ghose, A., Nepal, S., & Yao, L. (2021). An internet of things service roadmap. *Communications of the ACM*, *64*(9), 86-95.

[11] Noura, M., Atiquzzaman, M., & Gaedke, M. (2019). Interoperability in internet of things: Taxonomies and open challenges. *Mobile networks and applications*, *24*(3), 796-809.

[12] Ramadhan, H., Indikawati, F. I., Kwon, J., & Koo, B. (2020). MusQ: a Multi-store query system for IoT data using a datalog-like language. *IEEE Access*, *8*, 58032-58056.

[13] Sneps-Sneppe, M., & Namiot, D. (2015). On web-based domain-specific language for internet of things. In *2015 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* (pp. 287-292). IEEE.

[14] Wu, W., Bernstein, P. A., Raizman, A., & Pavlopoulou, C. (2022). Factor windows: Cost-based query rewriting for optimizing correlated window aggregates. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)* (pp. 2722-2734). IEEE.

[15] Kolev, B., Bondiombouy, C., Levchenko, O., Valduriez, P., Jimenez-Péris, R., Pau, R., & Pereira, J. (2016). Design and implementation of the CloudMdsQL multistore system. In *Workshop on Towards Convergence of Big Data, SQL, NoSQL, NewSQL, Data streaming/CEP, OLTP and OLAP* (Vol. 2, pp. 352-359). SciTePress.

[16] SensorWave. (s. f.). SensorWave [Organización de GitHub]. GitHub. Recuperado el 19 de febrero de 2026, de https://github.com/sensorwave-dev

Arquitecturas Multiprocesador en HPC: Portabilidad, Eficiencia Energética y Computación Cuántica

Armando De Giusti [(1)(2)], Marcelo Naiouf[(1)], Fernando G. Tinetti [(1)(3)], Horacio Villagarcía [(1)], Franco Chichizola[(1)], Laura De Giusti[(1)(3)], Enzo Rucci[(1)(3)], Adrián Pousa[(1)], Victoria Sanz [(1)(3)], Diego Montezanti [(1)], Diego Encinas [(1)], Ismael Rodríguez[(1)], Sebastián Rodríguez Eguren[(1)], Leandro Libutti[(1)], Manuel Costanzo [(1)], César Estrebou[(1)], Matías Manzin [(1)], Javier Balladini[(4)]

[1]Instituto de Investigación en Informática LIDI (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
[2]CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
[3]CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
[4]Universidad Nacional del Comahue

{degiusti,mnaiouf,fernando,hvw,francoch,ldgiusti,erucci,apousa,vsanz,dmontezanti,dencinas,ismael,seguren,llibutti,mcostanzo,cesarest,mmanzin}@lidi.info.unlp.edu.ar; javier.balladini@gmail.com

Resumen

El eje de esta línea de I/D lo constituye el estudio de las arquitecturas multiprocesador que integran sistemas distribuidos y paralelos. Incluye como temas centrales:

- Arquitecturas many-core (GPU, procesadores MIC, TPUs), FPGAs, híbridas (diferentes combinaciones de multicores y aceleradores), asimétricas, y cuánticas.
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas y su evaluación de rendimiento computacional y energético.
- Desarrollo y evaluación de algoritmos cuánticos. Comparación con algoritmos paralelos clásicos.
- Estudio y optimización de código heredado.
- Desarrollo y evaluación de estrategias de resiliencia.
- Modelado y simulación de E/S en HPC.

Palabras clave: *Sistemas Paralelos. Asimétricos. GPU, FPGA, TPU. Computación cuántica. Eficiencia energética. Resiliencia. Portabilidad. E/S paralela.*

Contexto

La línea de investigación que se presenta es parte del proyecto “Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, Fog-Edge-Cloud, Big Data, Robótica, y Tiempo Real.” del III-LIDI y de proyectos específicos apoyados por organismos nacionales e internacionales. También es parte del proyecto “Planificación eficiente de algoritmos de ML en contenedores utilizando multicores y GPUs” financiado por la Facultad de Informática de la UNLP.

En los temas relacionados existen cooperaciones con varias Universidades de Argentina, y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por ERASMUS, CyTED y la OEI (Organización de Estados Iberoamericanos). En particular, se colabora con el proyecto “Computación de Altas Prestaciones Eficiente y Segura para Aplicaciones de Servicios de Salud Inteligentes”, de la Universidad Autónoma de Barcelona.

Por otra parte, se cuenta con financiamiento de diferentes empresas de Argentina, en particular para la formación de recursos humanos en las temáticas de Cloud, Fog y Edge Computing.

Se participa en iniciativas como el Consorcio de I+D+I en Cloud Computing, Big Data y Emergent Topics, conformado por Universidades de Argentina y España.

Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) de la Secretaría de Innovación, Ciencia y Tecnología (SICYT) de la Nación.

Introducción

Una de las áreas de creciente interés lo constituye el cómputo de altas prestaciones, en el cual el rendimiento está relacionado con dos aspectos: por un lado, las arquitecturas de soporte, y por otro, los algoritmos que hacen uso de éstas.

A la aparición de arquitecturas *many-core* (como las GPU o los procesadores MIC), se ha sumado el uso de FPGAs, debido a su potencia de cómputo y rendimiento energético. Su combinación en sistemas HPC da lugar a plataformas híbridas con diferentes características [22]. Lógicamente, esto trae aparejada una revisión de los conceptos del diseño de algoritmos paralelos (incluyendo los mismos lenguajes de programación y el software de base), así como la evaluación de las soluciones que éstos implementan.

Si bien la heterogeneidad del hardware permite maximizar las ganancias de rendimiento y eficiencia energética, implica el uso de lenguajes, librerías y herramientas específicas para cada una de ellas, lo que aumenta el costo de desarrollo y mantenimiento del software [3].

GPUs y Cluster de GPUs

Las GPUs son el tipo de acelerador dominante en la comunidad de HPC hoy en día debido a su alto rendimiento y bajo costo de adquisición. En los últimos años, las GPUs han evolucionado hacia arquitecturas heterogéneas, que combinan diferentes clases de instrucciones. El principal desafío reside en cómo aprovechar esta vasta capacidad computacional de forma eficiente. En ese sentido, la tecnología Multi-Instance GPU (MIG) de NVIDIA representa un avance para los centros de datos [25].

Asimismo, la combinación de GPUs con otras plataformas paralelas, como clusters y multicores, brindan un vasto conjunto de posibilidades de investigación en arquitecturas híbridas. Los desafíos que se plantean aquí son múltiples, sobre todo en lo referido a distribución de datos y procesos en tales arquitecturas híbridas, a fin de optimizar el rendimiento de las soluciones.

FPGAs

Una FPGA (*Field Programmable Gate Array*) es una clase de acelerador basado en circuitos integrados reconfigurables. La capacidad de adaptar sus instrucciones de acuerdo con la aplicación objetivo le permite incrementar la productividad de un sistema y mejorar el rendimiento energético para ciertos tipos de aplicaciones. En los últimos años, se ha extendido su uso debido a la integración de estos dispositivos en arquitecturas híbridas [10][11]; al desarrollo de nuevas herramientas de programación, que ayudan a reducir los tradicionales tiempos y costos de programación [27][31]; y a la incorporación de FPGAs a los servicios de Cloud, lo que abre nuevas oportunidades para la explotación de esta clase de aceleradores.

TPUs

Las unidades de procesamiento tensorial (TPU) son una clase de Circuitos Integrados de Aplicación Específica (ASIC) desarrolladas por Google con el propósito de acelerar las cargas de trabajo de aprendizaje automático que requieren las aplicaciones desarrolladas en su framework TensorFlow [29]. Su uso provee una alternativa a otras arquitecturas ya conocidas como CPUs, GPUs y MICs. En ese sentido, interesa analizar las tasas de aceleración y eficiencia energética provistas por esta nueva arquitectura, en comparación con el resto.

Computación cuántica

La Computación Cuántica (QC) surge como un nuevo paradigma que aprovecha propiedades de la mecánica cuántica en computación. La QC posee características inherentemente paralelas y se espera que resuelva algunos problemas de forma más rápida que la computación clásica. Varios autores coinciden que la QC deberá integrarse al cómputo clásico HPC y esto abre nuevos desafíos [37][38].

Eficiencia energética

La mejora de la eficiencia energética es una de las principales preocupaciones en la informática actual, principalmente a partir de

las plataformas con gran cantidad de procesadores. Muchos esfuerzos están orientados a tratar la eficiencia energética y a las metodologías para medirla como ejes de I/D, como una métrica de evaluación relevante. Entre los puntos de interés se pueden mencionar la evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas, y su posterior optimización a través del estudio y aplicación de técnicas que reducen el consumo energético a nivel de aplicación, middleware o sistema.

Código heredado

La mayoría de los programas de simulación numérica que se emplean hoy en día fueron desarrolladas cuando las arquitecturas paralelas no existían. Por ello, este conjunto de aplicaciones abre oportunidades para desarrollar técnicas y herramientas que permitan optimizar el código, tanto desde el punto de vista computacional como desde la ingeniería de software [28].

Resiliencia

En la actualidad, lograr sistemas resilientes resulta un verdadero desafío considerando el creciente número de componentes, la cercanía a los límites físicos en las tecnologías de fabricación y la complejidad incremental del software. La corrección de las aplicaciones y la eficiencia en su ejecución se torna más importante en HPC debido a los extensos tiempos de ejecución. En ese sentido, resulta relevante desarrollar estrategias de detección y recuperación de fallos, especialmente a través de librerías de software [15]. Por otro lado, resulta de interés la aplicación de mecanismos de resiliencia en aplicaciones distribuidas [21].

Simulación y Entrada/Salida paralela

A pesar de los avances tecnológicos, las operaciones de E/S en los centros de supercómputo siguen siendo un cuello de botella para determinadas aplicaciones HPC. El rendimiento de un sistema depende de la carga de trabajo (patrones de E/S de las aplicaciones) y de su configuración (hardware y software) [19]. Contar con herramientas que permitan modelar y predecir el comportamiento de este tipo de aplicaciones en HPC resulta fundamental para mejorar su rendimiento [5].

El análisis y diseño de modelos de simulación basados en la arquitectura de E/S paralela, permite disminuir la complejidad y cubrir las exigencias de las aplicaciones en HPC, al poder identificar y evaluar los factores que influyen en las prestaciones [6]. Además. considerando la relevancia que ha cobrado la simulación de transmisión de enfermedades intra-hospitalarias, resulta de interés utilizar la experiencia desarrollada para el modelado de estos sistemas [13][32].

Dispositivos de bajo costo con capacidades para cómputo paralelo

En la actualidad se comercializan placas de bajo costo como Raspberry PI [20] u Odroid [33], que poseen múltiples núcleos simples. Asimismo, existen diversos dispositivos móviles con capacidades similares. Es de interés estudiar la manera de explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones [1].

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas many-core (CPU, GPU y TPU) y FPGA. Análisis de este tipo de máquinas y de técnicas para desarrollar código optimizado.
- Arquitecturas híbridas (diferentes combinaciones de clusters, multicores, manycores y FPGAs). Diseño de algoritmos paralelos sobre las mismas. Técnicas de resiliencia.
- Exploración de nuevos lenguajes, modelos y estándares de programación para HPC.
- Consumo energético en las diferentes arquitecturas de alto desempeño, en particular en relación con los algoritmos paralelos y la configuración de la arquitectura. Modelado y estimación del consumo de potencia de arquitecturas HPC y distribuidas con capas de Cloud, Fog y Edge Computing.

- Análisis y desarrollo de modelos e implementación de simuladores de la pila de software de E/S en HPC.

Resultados y Objetivos

Investigación experimental por realizar

- Desarrollar y evaluar algoritmos paralelos sobre nuevas arquitecturas paralelas. Analizar rendimiento, eficiencia energética, costo de programación y portabilidad.
- Desarrollar y evaluar algoritmos en máquinas cuánticas. Comparar con algoritmos paralelos en máquinas tradicionales.
- Analizar las capacidades de lenguajes no convencionales para procesamiento paralelo, considerando rendimiento, costo de programación y portabilidad.
- Realizar el desarrollo de nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [24][25].
- Analizar proceso de migración de algoritmos entre arquitecturas diferentes, considerando portabilidad, rendimiento, esfuerzo de programación y portabilidad.
- Evaluar el consumo energético debido al procesamiento en los distintos niveles de la arquitectura distribuida y a las comunicaciones entre ellos
- Desarrollar técnicas de tolerancia a fallas que permitan aumentar la resiliencia de sistemas paralelos y distribuidos.
- Calibrar y sintonizar el modelado y simulación de E/S en HPC conseguido para distintos escenarios e infraestructuras.

Resultados obtenidos

- Se llevó a cabo una breve comparación del rendimiento proporcionado por diferentes compiladores C al aplicar autovectorización [23].
- Se realizaron y analizaron modificaciones al framework TensorFlow para permitir la maleabilidad de hilos en CPU [9][12].
- En el ámbito de la bioinformática, se propuso un método para acelerar la compresión de datos genómicos [41] y se trabajó en la migración de una librería popular para comparación de secuencias biológicas [39].
- Se realizó un estudio de migración de códigos CUDA a SYCL usando el ecosistema oneAPI para aplicaciones del ámbito biológico y se evaluó el esfuerzo de programación requerido, la portabilidad funcional y de rendimiento en arquitecturas GPU, CPU y CPU-GPU[2][3].
- Se analizaron los desafíos en la integración de unidades de procesamiento cuántico a sistemas HPC tradicionales [37][38].
- Se diseñó la herramienta SEDAR y evaluó su desempeño para detección y recuperación de fallos transitorios [15].
- Se implementaron y evaluaron diferentes estrategias de resiliencia en una arquitectura basada en microservicios de gran escala [21][26].
- Se desarrolló un simulador que permite predecir cuál es la mejor estrategia para ahorrar energía ante una falla permanente en un sistema HPC [16].
- Se realizó un primer análisis comparativo del consumo energético cloud vs edge computing [40].
- Se analizó la resiliencia de un servicio de emergencias hospitalarias ante situaciones extremas mediante el uso de simulación [14][35][36].
- Se analizó el comportamiento del Dengue en la ciudad de La Plata mediante simulación buscando estimar su impacto de propagación [13].
- Se desarrollaron técnicas de modelado y simulación de E/S en HPC que permiten predecir cómo los cambios realizados en los diferentes componentes afectan a la funcionalidad y al rendimiento del sistema, para posteriormente ser capaces de optimizar [5][6].
- Se analizó el impacto de la E/S en aplicaciones de aprendizaje automático profundo [19].
- Se exploró el uso de distintos microcontroladores para aplicaciones de aprendizaje automático [7] [8].
- Se desarrolló una metodología similar a la conocida como CI/CD (integración continua y

despliegue continuo), que incluye varias estrategias de prueba para mejorar el software heredado de computación científica [30].

Organización de Eventos

En el año 2025 se han organizado las XIII Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2025) en Argentina, con participación de especialistas académicos del país y el exterior, además de empresas con experiencia en Cloud Computing [4][17]. En junio de 2026 se tiene prevista la organización de las XIV JCC-BD&ET.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyó 1 tesis de doctorado [34]. Al mismo tiempo se encuentran en curso 1 tesis de Doctorado en Ciencias Informáticas y 2 tesis de Maestría.

Además, se participa en el dictado de las carreras de Doctorado en Ciencias Informáticas, y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática de la UNLP (acreditadas por la CONEAU con categoría A, B y A, respectivamente), por lo que potencialmente pueden generarse nuevas Tesis de Doctorado y Maestría, además de Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior, y hay tesistas de diferentes Universidades realizando su Tesis con el equipo del proyecto.

Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación tres materias directamente relacionadas con los temas de ésta: “Taller de Programación sobre GPUs”, “Cloud Computing y Cloud Robotics” y “Conceptos y Aplicaciones en Big Data”.

Referencias

[1] P. S. Rodríguez Eguren, F. Chichizola, and E. Rucci, “Análisis del uso de un cluster de Raspberry Pi para cómputo de alto rendimiento”. Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018), ISBN: 978-950-658-472-6, págs. 134-144, 2018.

[2] Manuel Costanzo, Enzo Rucci, Carlos García-Sánchez, Marcelo Naiouf, Manuel Prieto-Matías. “Assessing opportunities of SYCL for biological sequence alignment on GPU-based systems.” J Supercomput (2024). https://doi.org/10.1007/s11227-024-05907-2

[3] Manuel Costanzo, Enzo Rucci, Carlos García-Sánchez, Marcelo Naiouf, Manuel Prieto-Matías, “Analyzing the performance portability of SYCL across CPUs, GPUs, and hybrid systems with SW sequence alignment”, Future Generation Computer Systems, Volume 170, 2025, ISSN 0167-739X, https://doi.org/10.1016/j.future.2025.107838

[4] Marcelo Naiouf, Franco Chichizola, Laura De Giusti, Leandro Libutti (Editores) “13th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2025)”. ISBN: 978-950-34-2583-1. Facultad de Informática. Agosto 2025. https://doi.org/10.35537/10915/182427.

[5] D. Encinas, S. Mendez, M. Naiouf, A. De Giusti, D. Rexachs del Rosario, and E. Luque, “An Agent-Based Model for Analyzing the HPC Input/Output System” International journal on advances in systems and measurements vol. 13, num. 3 & 4, págs. 192-202, 2020.

[6] D. Encinas et al., «Agent Based Simulation as an Efficient Way for HPC I/O System Tuning», en Computational Science – ICCS 2024, L. Franco, C. de Mulatier, M. Paszynski, V. V. Krzhizhanovskaya, J. J. Dongarra, y P. M. A. Sloot, Eds., Cham: Springer Nature Switzerland, 2024, pp. 129-136. doi: 10.1007/978-3-031-63759-9_16.

[7] César A. Estrebou, Martín Fleming, Marcos D. Saavedra, Federico Adra, Armando E. De Giusti. “Lightweight Convolutional Neural Networks Framework for Really Small TinyML Devices.”. Third International Conference on Smart technologies, Systems and Applications (SmartTech-IC), ISBN 978-3-030-99170-8, págs. 3-16, Cham, doi: 10.1007/978-3-030-99170-8_1, marzo de 2022.

[8] César Armando Estrebou, Marcos David Saavedra, Federico Adra, Martín Fleming. “TinyML for Small Microcontrollers.” X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022), ISBN 978-950-34-2126-0, págs. 42-46, julio de 2022.

[9] Leandro Libutti. «Optimizing Resource Allocation for Elastic Tensorflow Container Scheduling on Multi-core Servers». CLEI Electronic Journal (ISSN 0717-5000), vol. 28, núm. 6, págs. 2:1-2:14, doi: 10.19153/cleiej.28.6.2, diciembre de 2025

[10] IBM. “IBM and Xilinx Announce Strategic Collaboration to Accelerate Data Center Applications”. Disponible en https://www-03.ibm.com/press/us/en/pressrelease/48074.wss

[11] Intel. “Intel Acquisition of Altera”. Disponible en http://intelacquiresaltera.transactionannouncement.com

[12] Leandro Ariel Libutti. “Thesis overview: Scheduling Elastic Machine Learning Process Through Containers”. Journal of Computer Science and Technology (ISSN 1666-6038), vol. 23, núm. 2, págs.

e17-e17, doi: 10.24215/16666038.23.e17, octubre de 2023
[13] D. O. Encinas, J. Matto, J. Stella, L. Maccallini, y F. Romero, «Dengue Dynamics: Modelling Spread and Environmental Interactions», Journal of Computer Science and Technology, vol. 24, n.o 2, Art. n.o 2, oct. 2024, doi: 10.24215/16666038.24.e14.
[14] Mercedes Planas, Eva Bruballa, Emilio Luque, Dolores Rexachs, Francisco Epelde, Armando De Giusti. "Evaluación del tiempo crítico para realizar acciones preventivas en el Servicio de Urgencias Hospitalarias". XXIX Congreso Argentino de Ciencias de la Computación – CACIC 2024. Octubre de 2024. En prensa.
[15] D. Montezanti, E. Rucci, A. D. De Giusti, M. Naiouf, D. Rexachs, and E. Luque, "Soft errors detection and automatic recovery based on replication combined with different levels of checkpointing". Future generation computer systems (ISSN 0167-739X), vol. 113, págs. 240-254, doi. https://doi.org/10.1016/j.future.2020.07.003, 2020.
[16] Marina Morán, Javier Balladini, Dolores Rexachs, Enzo Rucci. "Exploring Energy Saving Opportunities in Fault Tolerant HPC Systems" Journal of Parallel and Distributed Computing (ISSN 0743-7315), págs. 104797, doi: 10.1016/j.jpdc.2023.104797, noviembre de 2023
[17] Marcelo Naiouf, Laura De Giusti, Franco Chichizola, Leandro Libutti (editores). "Cloud Computing, Big Data and Emerging Topics. 13th Conference, JCC-BD&ET 2025, La Plata, Argentina, June 24–26, 2025, Revised Selected Papers" ISBN: 978-3-032-06336-6. Springer Cham. Octubre 2025. https://doi.org/10.1007/978-3-032-06336-6
[18] S. Calderón, E. Rucci, y F. Chichizola, «Enhanced OpenMP Algorithm to Compute All-Pairs Shortest Path on X86 Architectures», en Computer Science – CACIC 2023, P. Pesado, W. Panessi, y J. M. Fernández, Eds., Cham: Springer Nature Switzerland, 2024, pp. 46-61. doi: 10.1007/978-3-031-62245-8_4.
[19] E. Parraga, B. Leon, S. Mendez, D. Rexachs, R. Suppi, y E. Luque, «An Empirical Method for Processing I/O Traces to Analyze the Performance of DL Applications», en Cloud Computing, Big Data and Emerging Topics , Cham: Springer Nature Switzerland, 2025, pp. 74-90. doi: 10.1007/978-3-031-70807-7_6.
[20] Raspberry PI. https://www.raspberrypi.org/
[21] S. L. Suárez, E. Rucci, V. Betran, y D. M. Montezanti, «Implementando estrategias de resiliencia en una arquitectura basada en microservicios», en Actas del XXIX Congreso Argentino de Ciencias de la Computación (CACIC 2023), oct. 2023, pp. 467-488. Accedido: 24 de abril de 2024. [En línea]. Disponible en: http://sedici.unlp.edu.ar/handle/10915/164996
[22] Rucci, Enzo: "Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática". Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
[23] Matias Manzin, Enzo Rucci. "Breve Comparación de Rendimiento de Compiladores C para Autovectorización" XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025). Octubre 2025. En prensa.
[24] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias: "ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems". In: ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
[25] NVIDIA, "GPU NVIDIA de Múltiples Instancias. "https://www.nvidia.com/es-la/technologies/multi-instance-gpu/
[26] Sergio Leonel Suarez, Enzo Rucci, Victor Betran, Diego Montezanti. "Extendiendo las Estrategias de Resiliencia de una Arquitectura Basada en Microservicios" XXIX Congreso Argentino de Ciencias de la Computación – CACIC 2024. Octubre de 2024. En prensa.
[27] Sean Settle: "High-performance Dynamic Programming on FPGAs with OpenCL". In: IEEE High Performance Extreme Computing Conference. 2013.
[28] F. G. Tinetti, M. J. Perez, A. Fraidenraich, and A. E. Altenberg, "Legacy code and parallel computing: updating and parallelizing a numerical model". The journal of supercomputing (ISSN 1573-0484), doi. 10.1007/s11227-020-03172-7, 2020.
[29] Google Inc. "Cloud Tensor Processing Unit (TPU)" Disponible en https://cloud.google.com/tpu/docs/tpus?hl=es-419
[30] Fraidenraich, A., Santone, F., Altenberg, A.E., Tinetti, F.G. (2025). "Scientific Legacy Code Enhancement and Testing Strategies: Fortran Example". In: Arabnia, H.R., Deligiannidis, L., Shenavarmasouleh, F., Amirian, S., Ghareh Mohammadi, F. (eds) Computational Science and Computational Intelligence. CSCI 2024. Communications in Computer and Information Science, vol 2505. Springer, Cham. https://doi.org/10.1007/978-3-031-95127-5_23
[31] Xilinx Inc. "SDAccel Development Environment". [Online]. Disponible en http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html
[32] Maccallini, L., Encinas, D. O., & Romero, F. (2021). An Approach to the Modeling and Simulation of Intra-Hospital Diseases. Journal of Computer Science and Technology, 21(2), e14. https://doi.org/10.24215/16666038.21.e14
[33] Odroid http://www.hardkernel.com Accedido 21 de marzo de 2016.
[34] Marina Morán, "Modelización y gestión del consumo energético en un sistema de computación de altas prestaciones con tolerancia a fallos" Tesis de Doctorado en Ciencias Informáticas, UNLP. https://sedici.unlp.edu.ar/handle/10915/180599
[35] Mariela Rodriguez, Francesc Boixader , Francisco Epelde, Eva Bruballa, Armando De Giusti, Alvaro Wong, Dolores Rexachs, Emilio Luque. "Análisis de desempeño de Servicio de Urgencia Hospitalario con llegada creciente de pacientes". XXIX Congreso

Argentino de Ciencias de la Computación – CACIC 2024. Octubre de 2024. En prensa.
[36] M. Ansari Dogaheh, M. Taboada, F. Epelde, E. Luque, D. Rexachs, y A. Wong, «An ABMS COVID-19 Propagation Model for Hospital Emergency Departments», en Cloud Computing, Big Data and Emerging Topics, Cham: Springer Nature Switzerland, 2025, pp. 103-117. doi: 10.1007/978-3-031-70807-7_8.
[37] A. Pousa y V. Sanz, «Desafíos en la integración de QPUs a sistemas HPC clásicos», Memoria investig. ing. (Facultad Ing., Univ. Montev.), n.º 27, pp. 238–247, dic. 2024.
[38] Pousa, A., Sanz, V. M., Naiouf, M., & De Giusti, A. (2024). Trends in High Performance Computing and Quantum Computing. Journal of Computer Science and Technology, 24(2), e11. https://doi.org/10.24215/16666038.24.e11
[39] Juan Tomás Vicens, Enzo Rucci, Manuel Costanzo, Marcelo Naiouf. "Migración de Librería SSW para Alineamientos de Secuencias Biológicas a Tecnología AVX". XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025). Octubre 2025. En prensa.
[40] De Giusti, A., Montezanti, D., Medina, S. y Libutti, L. (2024). Edge Computing vs Cloud Computing: Optimizando el Consumo Energético. XXIX Congreso Argentino de Ciencias de la Computación – CACIC 2024. Octubre de 2024. En prensa.
[41] Victoria Sanz, Adrián Pousa, Marcelo Naiouf, Armando De Giusti. "A Fast and Scalable Genomic Data Compressor for Multicore Clusters". En: Computational Science – ICCS 2025. 25th International Conference, Singapore, Singapore, July 7–9, 2025, Proceedings, Part IV. ISBN: 978-3-031-97635-3. Págs. 180-188. Springer Cham. Julio 2025. https://doi.org/10.1007/978-3-031-97635-3

Ciencia computacional para la predicción y modelado en contextos de incertidumbre

Caymes-Scutari Paola[1,2], Bianchini Germán[2], Gómez Laura[1,3], Artaza Pablo[2]

[1]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
[2]Laboratorio de Investigación en Cómputo Paralelo/Distribuido (LICPaD, UTN-FRM)
[3]Instituto Argentino de Investigaciones de las Zonas Áridas (IADIZA, CCT-Mendoza)

paola.caymesscutari@frm.utn.edu.ar, german.bianchini@frm.utn.edu.ar,
lgomez@mendoza-conicet.gov.ar, pabloartaza927@gmail.com

RESUMEN

La predicción y el modelado de fenómenos, sistemas y parámetros constituyen ejemplos en los que la ciencia computacional, potenciada por el procesamiento paralelo/distribuido, brinda un soporte invaluable, tanto para representar el escenario de la situación, como para afrontar con cierto grado de precisión el efecto negativo que tienen las distintas fuentes de incertidumbre. La reducción de incertidumbre es un proceso para disminuir los efectos negativos de la falta de precisión en cualquier factor de un sistema de predicción. En los últimos años hemos desarrollado un conjunto de métodos para la reducción de incertidumbre, los DDM-MOS (Data Driven Methods with Multiple Overlapped Solutions), que combinan distintas estrategias algorítmicas (computación evolutiva y metaheurísticas, paralelismo, estadística, búsqueda por novedad, y sintonización) para potenciar la calidad (precisión, tiempo de respuesta, y/o uso de recursos) en la predicción de fenómenos de propagación. Para el presente proyecto se propone el estudio e incorporación de estrategias propias de la ciencia de datos y de la inteligencia artificial, como pueden ser las técnicas de aprendizaje de máquina. Dentro de los posibles casos de aplicación, en particular se propone abordar el modelado y predicción de parámetros ambientales relacionados con la hidrogeología de la zona árida irrigada de Mendoza.

***Palabras clave*:** propagación, hidrogeología, reducción de incertidumbre, oasis irrigados.

CONTEXTO

El presente proyecto ha sido propuesto para su desarrollo en el marco del LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido), de la UTN, Facultad Regional Mendoza, Departamento de Ingeniería en Sistemas de Información. El mismo cuenta con el aval de la Secretaría de Ciencia y Tecnología de la UTN, mediante disposición 45/2025, para su ejecución a partir de abril 2026. El mismo contempla la extensión de las líneas de investigación desarrolladas hasta el momento, a través del desarrollo de métodos en colaboración con miembros del IADIZA (Instituto Argentino de Investigaciones de las Zonas Áridas, CCT CONICET Mendoza) para el abordaje de la incertidumbre

hidrogeológica existente en las zonas áridas irrigadas con alta demanda, en pos de estudiar la evolución de las mismas, para su adecuado uso y protección.

1. INTRODUCCIÓN

La incertidumbre es la falta de conocimiento o exactitud causada por la dificultad para modelar todos los parámetros del sistema, la imposibilidad de censarlos de forma continua, los errores que introducen las estimaciones, las limitaciones en la implementación computacional, etc. (Bianchini 2010). La reducción de incertidumbre es un proceso para disminuir los efectos negativos de la falta de precisión en cualquier factor de un sistema de predicción (Griffin et al, 2023) (Bianchini 2010).

En el área informático-computacional, generalmente los procesos de predicción implementan modelos matemáticos, empíricos, o semiempíricos que describen el comportamiento del fenómeno o sistema en cuestión a través de un simulador, que a su vez requiere como dato de entrada un grupo de variables representando aquellos factores que condicionan el comportamiento del fenómeno o sistema. Constituye la forma clásica de realizar una predicción, ya que considera el estado inicial del fenómeno o sistema y lo utiliza como insumo para el programa de simulación, arrojando como resultado el estado que se predice para el instante siguiente. Este tipo de predicción clásica suele ser limitada en cuanto a la precisión de los resultados, ya que no contempla mecanismos para mitigar el efecto de la incertidumbre. Es por ello que existe una variedad de métodos computacionales que complementan y enriquecen las capacidades de la predicción o el modelado clásico, mediante la utilización y combinación de diferentes técnicas computacionales (como, por ejemplo, las que se mencionan en la sección 2).

La incertidumbre hidrogeológica se refiere a la incertidumbre atinente a un sistema natural como los acuíferos. Dicha incertidumbre existe, en particular, en las zonas áridas irrigadas de Mendoza con alta demanda, en las cuales la falta de datos espaciales y temporales tiene como consecuencia una gran dificultad para definir los procesos de recarga, descarga, y salinización de acuíferos, que permitan explorar posibles escenarios de evolución de los mismos. Es por ello que resulta de interés abordar el tratamiento de dicha incertidumbre mediante técnicas computacionales que permitan optimizar modelos para anticipar riesgos en acuíferos sin datos y con alta demanda en los oasis irrigados de Mendoza.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

En los últimos años hemos desarrollado un conjunto de métodos para la reducción de incertidumbre, los DDM-MOS (Data Driven Methods with Multiple Overlapped Solutions) (Bianchini et al, 2010) (Strappa et al, 2022). Los mismos han sido aplicados a la predicción de incendios forestales, pero tales métodos podrían utilizarse también para predecir otros tipos de fenómenos de propagación, como avalanchas, aludes, o inundaciones, por ejemplo.

Los DDM-MOS son métodos computacionales predictivos con reducción de incertidumbre, que combinan distintas estrategias estadísticas, de computación evolutiva, y

de paralelismo para potenciar la calidad en la predicción. La predicción del frente de propagación (por ejemplo una línea de fuego) en base a múltiples combinaciones de valores de parámetros, y a la agregación de las múltiples soluciones que arroja cada una de dichas combinaciones. Dado que la cantidad de combinaciones posibles de valores de parámetros constituye un espacio de búsqueda de tamaño considerable, la componente poblacional evolutiva de los métodos considera una población de individuos (también denominados escenarios, conformados por diferentes combinaciones de valores en los parámetros de entrada que describen, por ejemplo, el estado del incendio para el simulador, como velocidad y dirección del viento, humedad, etc.), los cuales constituyen una muestra de puntos en el espacio de búsqueda. Para cada uno de dichos escenarios, el simulador arroja un resultado generalmente diferente, y por lo tanto el sistema debe analizar los resultados de todos ellos y determinar la tendencia. Es por ello que, en los DDM-MOS, la funcionalidad del simulador se encuentra complementada y enriquecida por metaheurísticas, búsqueda por novedad, estadística y paralelismo. Las metaheurísticas (Talbi, 2009) se utilizan para orientar la búsqueda dentro del gran espacio de búsqueda hacia mejores soluciones, en base a la calidad de las múltiples simulaciones obtenidas. El cómputo paralelo brinda la potencia necesaria para ejecutar las múltiples simulaciones en paralelo y acelerar el procesamiento de la información. Y la Estadística se utiliza para obtener el patrón de comportamiento del frente de propagación en base a los resultados obtenidos en las múltiples simulaciones realizadas para los diferentes escenarios. Asimismo, algunos DDM-MOS incorporan el uso de técnicas de Sintonización (Naono et al, 2010) (Caymes-Scutari et al, 2016) a fin de mejorar su rendimiento frente a otros métodos, y en pos de brindar un resultado de mayor calidad y utilidad. Cada DDM-MOS se distingue por el grado de incorporación de los elementos mencionados, lo que permite mejoras en la calidad de predicción, el tiempo de ejecución, y el uso de recursos (Bianchini et al, 2010) (Bianchini et al, 2015) (Caymes-Scutari et al, 2016)(Méndez et al, 2016)(Tardivo et al, 2017a) (Tardivo et al, 2017b)(Strappa et al, 2022).

Para el presente proyecto proponemos el estudio e incorporación de elementos y/o estrategias propios de la ciencia de datos (Aws, 2025) como así también de la inteligencia artificial (Russell y Norvig, 2010). Si bien dentro de los DDM-MOS ya hemos explorado y experimentado con algunas técnicas de búsqueda y optimización provenientes de la inteligencia artificial (en particular con metaheurísticas poblacionales como algoritmos genéticos, algoritmos evolutivos, evolución diferencial, PSO, y más recientemente con el agregado de la búsqueda por novedad o novelty search (Talbi, 2009)), la presente propuesta está más orientada a explorar con técnicas de aprendizaje de máquina, ya sean supervisadas o no supervisadas (Russell y Norvig, 2010)(Ayyadevara, 2018)(Corea, 2018). Dado que se trata de algoritmos que requieren cierta cantidad de datos para ser entrenados y validados, se considera la posibilidad de dar tratamiento a diferentes casos de aplicación, según la disponibilidad y factibilidad de obtener datos suficientes. En particular, la propuesta está orientada al modelado y predicción de parámetros ambientales relacionados con la hidrogeología de los oasis irrigados de Mendoza (Gómez et al, 2014) a fin de

conocer la problemática existente, las necesidades particulares para realizar un aporte desde la informática a esta área de conocimiento, y la disponibilidad de datos de mediciones en la zona.

3. RESULTADOS ESPERADOS

Se espera desarrollar un método o herramienta capaz de modelar y predecir parámetros vinculados al comportamiento de acuíferos, como así también con capacidades para el diagnóstico y la toma de decisiones en situaciones de contaminación o desastres. Dada la limitada caracterización y el escaso conocimiento actual de tales sistemas, el uso de estas herramientas permitiría modelar e interpretar procesos de recarga, descarga y evolución de la salinización, en un contexto de riego intensivo en constante aumento, y al mismo tiempo, contribuir al conocimiento del principal reservorio de agua dulce del que dispone el planeta.

4. FORMACIÓN DE RECURSOS HUMANOS

El objetivo es promover y dar continuidad a la actividad científica y ponerla al servicio del desarrollo de la región y el país, promoviendo líneas orientadas al crecimiento social y ambiental, asegurando que la pertinencia y relevancia de esos conocimientos contribuyan al progreso colectivo, en un momento crítico en cuanto a las limitaciones que ofrecen las becas, frente a las competitivas ofertas laborales que reciben los estudiantes desde el sector privado, representando un gran obstáculo a la hora de incorporar recursos humanos de grado y postgrado, que perduren a lo largo del tiempo. El estudio del modelado hidrogeológico de acuíferos, contribuirá al conocimiento del principal reservorio de agua dulce del que dispone el planeta. Dicha temática permite continuar con la formación de los distintos integrantes del grupo de trabajo, de forma complementaria a la formación adquirida hasta el momento en el campo de los métodos predictivos y de reducción de incertidumbre. El grupo de trabajo cuenta con la participación de un estudiante de grado, que a partir de su participación en el proyecto podrá comenzar a familiarizarse con las diferentes dimensiones y temáticas que conlleva, e involucrarse en el desarrollo del mismo. Asimismo, el grupo de trabajo siempre está abierto a la incorporación de nuevos integrantes que deseen familiarizarse con las temáticas que aquí se describen.

5. BIBLIOGRAFÍA

AWS - ¿Qué es la ciencia de datos? https://aws.amazon.com/what-is/data-science/ (Acceso abril de 2025).

Ayyadevara, V.K. 2018. Pro Machine Learning Algorithms. Apress Berkeley, CA. https://doi.org/10.1007/978-1-4842-3564-5.

Bianchini, Germán, Denham, Mónica, Cortés, Ana, Margalef, Tomàs, Luque, Emilio. 2010. "Wildland Fire Growth Prediction Method Based on Multiple Overlapping Solution". Journal of Computational Science Vol 1 Issue 4, pp. 229-237. Elsevier.

Bianchini, G.; Caymes-Scutari, P.; Méndez Garabetti, M.2015. Evolutionary-

Statistical System: A Parallel Method for Improving Forest Fire Spread Prediction. J. Comput. Sci., 6, 58–66. https://doi.org/10.1016/j.jocs.2014.12.001

Caymes Scutari, Paola, Germán Bianchini, Anna Sikora, and Tomàs Margalef. 2016. Environment for Automatic Development and Tuning of Parallel Applications. 2016 International Conference on High Performance Computing & Simulation (HPCS), 743–50. Innsbruck, Austria: IEEE. https://doi.org/10.1109/HPCSim.2016.7568409.

Corea, F. 2018. AI Knowledge Map: How to Classify AI Technologies. Forbes. https://www.forbes.com/sites/cognitiveworld/2018/08/22/ai-knowledge-map-how-toclassify-ai-technologies/

Gómez, M. L., Aranibar, J., Wuilloud, R., Rubio, M., Martinez, D., Soria, N., Monasterio, R., Villagra, P., Goirán, S. 2014. Hydrogeology and hidrogeochemical modeling in phreatic aquifer of NE Mendoza, Argentina. Journal of Iberian Geology. http://dx.doi.org/10.5209/rev_JIGE.2014.v40.n3.43302. Springer.

Griffin E., Ledbetter A., Spark G. 2023. A first look at Communication Theory – 11th edition. Capítulo 9: Uncertainty Reduction Theory, de Charles Berger. (2023). McGraw Hill.

Mendez Garabetti, Miguel, Bianchini, Germán, Caymes Scutari, Paola, Tardivo María Laura. 2016. Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic. Fire Safety Journal, pp. 49-62. Elsevier.

Naono, Ken, Keita Teranishi, John Cavazos, and Reiji Suda, eds. 2010. Software Automatic Tuning. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-6935-4

Russell, S., Norvig, P. 2010. Artificial Intelligence: A Modern Approach. Prentice Hall.

Strappa J., Caymes Scutari P., and Bianchini G. 2022. Evolutionary Statistical System based on Novelty Search: a Parallel Metaheuristic for Uncertainty Reduction Applied to Wildfire Spread Prediction. (2022). Algorithms 2022, vol 15, issue 12, 478. MDPI.

Talbi, El Ghazali. 2009. Metaheuristics: From Design to Implementation. Metaheuristics: From Design to Implementation. https://doi.org/10.1002/9780470496916. John Willey & Sons.

Tardivo, María Laura, Paola Caymes Scutari, Germán Bianchini, and Miguel Méndez Garabetti. 2017. Hierarchical Parallel Model for Improving Performance on Differential Evolution: Hierarchical Parallel Model for Improving Performance on Differential Evolution. Concurrency and Computation: Practice and Experience 29 (10): e4087.

Tardivo, María Laura, Caymes-Scutari, Paola, Méndez-Garabetti, Miguel y Bianchini, Germán. 2017. Optimization for an Uncertainty Reduction Method Applied to Forest Fires Spread Prediction. Computer Science – Cacic 2017, pp. 13-23. Springer

Fundamentos, construcción y evaluación de rendimiento de algoritmos paralelos en HPC

Marcelo Naiouf[(1)], Armando De Giusti[(1)(2)], Laura De Giusti[(1)(3)], Franco Chichizola[(1)], Victoria Sanz[(1)(3)], Adrián Pousa[(1)], Enzo Rucci[(1)(3)], Mariano Sánchez[(1)], Manuel Costanzo [(1)], Emmanuel Frati[(1)(4)], Adriana Gaudiani[(5)]

[1]Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires
Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

[2] CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
[3]CIC – Comisión de Investigación Científica de la Provincia de Buenos Aires
[4] Universidad Nacional de Chilecito
[5] Universidad Nacional de General Sarmiento
{mnaiouf, degiusti, ldgiusti, francoch, vsanz, apousa, erucci, msanchez, mcostanzo, fefrati}@lidi.info.unlp.edu.ar, agaudi@ungs.edu.ar

RESUMEN

El eje central de la línea de I/D es investigar en temas de cómputo paralelo y distribuido de alto desempeño, tanto en lo referido a los fundamentos como a la construcción, evaluación y optimización de las aplicaciones en arquitecturas multiprocesador. Se aplican los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos con el fin de obtener soluciones de alto rendimiento.
También incluye la construcción de ambientes para la enseñanza de la programación concurrente y paralela.
En la dirección de tesis de postgrado existe colaboración con el grupo HPC4EAS (High Performance Computing for Efficient Applications and Simulation) de la Universidad Autónoma de Barcelona y el Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid, entre otros.
Palabras clave: Cómputo paralelo y distribuido de altas prestaciones. Algoritmos paralelos y distribuidos. Arquitecturas multiprocesador. Ambientes de enseñanza.

CONTEXTO

La línea de I/D que se presenta es parte del Proyecto "Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, Fog-Edge-Cloud, Big Data, Robótica, y Tiempo Real" del III-LIDI acreditado por el Ministerio de Educación, y de proyectos acreditados y subsidiados por la Facultad de Informática de la UNLP. Además, existe cooperación con Universidades de Argentina, Latinoamérica y Europa a través de proyectos acreditados. Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD).

1. INTRODUCCIÓN

El área de cómputo de altas prestaciones (HPC, High-Performance Computing) es clave dentro de las Ciencias de la Computación, debido al creciente interés por el desarrollo de soluciones a problemas con alta demanda computacional y de almacenamiento, produciendo transformaciones profundas en las líneas de I/D [1].
El rendimiento en este caso está relacionado con dos aspectos: las arquitecturas de soporte y los algoritmos que hacen uso de las mismas, y el desafío se centra en cómo aprovechar las prestaciones obtenidas a partir de la evolución de las arquitecturas físicas. En esta línea la mayor importancia está en los algoritmos paralelos y en los métodos utilizados para su construcción y análisis a fin de optimizarlos.
Uno de los cambios de mayor impacto ha sido el uso de manera masiva de procesadores con más de un núcleo (*multicore*), produciendo plataformas distribuidas híbridas (memoria compartida y distribuida) y generando la necesidad de desarrollar sistemas operativos, lenguajes y algoritmos que las usen adecuadamente. También creció la incorporación de placas aceleradoras a los sistemas multicore constituyendo plataformas paralelas de memoria compartida con paradigma de programación propio asociado como pueden ser las unidades de procesamiento gráfico (GPU, Graphic Processing Unit) de NVIDIA y AMD, los coprocesadores Xeon Phi de Intel [2] o los aceleradores basados en circuitos integrados reconfigurables (FPGAs, Field Programmable Gate Array) [3]. Asimismo, los entornos de computación cloud introducen un nuevo foco desde el punto de vista del HPC, brindando un soporte "a medida" sin la necesidad de adquirir el hardware.
En el último tiempo ha cobrado relevancia la investigación en el área de la computación cuántica [4][5]. La idea de una computadora cuántica surge de aplicar leyes de la física cuántica en computación. Existe mucha expectativa en este tipo de máquinas, con gran capacidad de paralelismo, ya que se espera que resuelvan algunos problemas más eficientemente que las computadoras clásicas.
La creación de algoritmos paralelos en arquitecturas multiprocesador no es un proceso directo [6]. El costo puede ser alto en términos del esfuerzo de programación y

el manejo de la concurrencia adquiere un rol central en el desarrollo. Si bien en las primeras etapas el diseñador de una aplicación paralela puede abstraerse de la máquina sobre la que ejecutará el algoritmo, para obtener buen rendimiento debe tenerse en cuenta la plataforma de destino. En las máquinas multiprocesador, se deben identificar las capacidades de procesamiento, interconexión, sincronización y escalabilidad. La caracterización y estudio de rendimiento del sistema de comunicaciones es de interés para la predicción y optimización de performance, así como la homogeneidad o heterogeneidad de los procesadores [7].

Muchos problemas algorítmicos se vieron impactados por los multicore y clusters de multicore. El desarrollo de algoritmos que aprovechen adecuadamente las arquitecturas motiva el estudio de performance en sistemas híbridos. Es necesario estudiar la utilización de lenguajes y bibliotecas ya que aún no se cuenta con un standard, aunque puede mencionarse el uso de los tradicionales MPI, OpenMP y Pthreads [8][9] o los más recientemente explorados UPC, Chapel y Titanium del modelo PGAS [10].

La combinación de arquitecturas de múltiples núcleos con aceleradores dio lugar a plataformas híbridas con diferentes características. Actualmente, se plantea la idea de incorporar computadoras cuánticas a estas plataformas híbridas. Varios autores suponen que las computadoras cuánticas no reemplazarán a las computadoras clásicas, sino que se incorporarán como aceleradores a sistemas paralelos existentes [11]. Más allá del acelerador utilizado, la programación de estas plataformas representa un desafío. Para lograr aplicaciones de alto rendimiento, los programadores enfrentan dificultades como: estudiar características específicas de cada arquitectura y aplicar técnicas de programación y optimización particulares de cada una, lograr un balance de carga adecuado entre los dispositivos de procesamiento y afrontar la ausencia de estándares para este tipo de sistemas.

Por otra parte, los avances en las tecnologías de virtualización han llevado a que Cloud Computing sea una alternativa a los tradicionales sistemas de cluster [12]. El uso de cloud para HPC presenta desafíos atractivos, brindando un entorno reconfigurable dinámicamente sin la necesidad de adquirir hardware, y es una excelente plataforma para testear escalabilidad de algoritmos.

Métricas de evaluación del rendimiento y balance de carga

La diversidad de opciones vuelve complejo el análisis de performance de los Sistemas Paralelos, ya que los ejes sobre los cuales pueden compararse dos sistemas son varios. Existe un gran número de métricas para evaluar el rendimiento, siendo las tradicionales: tiempo de ejecución, speedup, eficiencia. Por su parte, la *escalabilidad* permite capturar características de un algoritmo paralelo y la arquitectura en que se lo implementa. Posibilita testear la performance de un programa sobre pocos procesadores y predecirla en un número mayor, así como caracterizar la cantidad de paralelismo inherente en un algoritmo.

Un aspecto de interés que se ha sumado como métrica, a partir de las plataformas con gran cantidad de procesadores, es el del *consumo* y la *eficiencia energética* [13]. Muchos esfuerzos están orientados a tratar el consumo como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirlo.

El objetivo principal del cómputo paralelo es reducir el tiempo de ejecución haciendo uso eficiente de los recursos. El *balance de carga* es un aspecto central y consiste en, dado un conjunto de tareas que comprenden un algoritmo y un conjunto de procesadores, encontrar el mapeo (asignación) de tareas a procesadores tal que cada una tenga una cantidad de trabajo que demande aproximadamente el mismo tiempo, y esto es más complejo si hay heterogeneidad. Dado que el problema general de mapping es *NP*-completo, pueden usarse enfoques que dan soluciones subóptimas aceptables. Las técnicas de planificación a nivel micro (dentro de cada procesador) y macro (en un cluster) deben ser capaces de obtener buen balance de carga. Existen técnicas estáticas y dinámicas cuyo uso depende del conocimiento que se tenga sobre las tareas de la aplicación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Investigar en temas de cómputo paralelo y distribuido de alto desempeño, en lo referido a los fundamentos y a la construcción y evaluación de las aplicaciones. Esto incluye los problemas de software asociados con el uso de arquitecturas multiprocesador:
  - Lenguajes, modelos y paradigmas de programación paralela (puros e híbridos a distintos niveles).
  - Asignación de procesos a procesadores optimizando el balance de la carga de procesamiento.
  - Métricas de evaluación de complejidad y rendimiento: speedup, eficiencia, escalabilidad, consumo energético, costo de programación.
- Construir, evaluar y optimizar soluciones utilizando algoritmos concurrentes, paralelos y distribuidos sobre diferentes plataformas de software y arquitecturas con múltiples procesadores:
  - Arquitecturas de trabajo homogéneas, heterogéneas e híbridas: multicores, clusters, GPU, Xeon Phi, FPGA, entornos cloud y nuevas plataformas como las computadoras cuánticas.
  - Aplicar los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos (aplicaciones científicas, búsquedas, simulaciones, imágenes, realidad virtual y aumentada, bioinformática, big data, n-body).
- Analizar y desarrollar ambientes para la enseñanza de programación concurrente y paralela.

- Caracterizar diferentes modelos de arquitecturas paralelas.
- Representar distintos modelos de comunicación/sincronización.
- Definir métricas de evaluación de rendimiento y eficiencia energética.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollar y optimizar algoritmos paralelos sobre diferentes modelos de arquitectura. En particular, en aplicaciones numéricas y no numéricas de cómputo intensivo y tratamiento de grandes volúmenes de datos.
- Estudiar y comparar los lenguajes sobre las plataformas multiprocesador para diferentes modelos de interacción entre procesos.
- Investigar la paralelización en plataformas que combinan clusters, multicore y aceleradores. Comparar estrategias de distribución de trabajo teniendo en cuenta las diferencias en potencias de cómputo y comunicación, dependencia de datos y memoria requerida.
- Evaluar la performance (speedup, eficiencia, escalabilidad, consumo energético) de las soluciones propuestas. Analizar el rendimiento de soluciones paralelas a problemas con diferentes características (dependencia de datos, relación cómputo / comunicación, memoria requerida).
- Mejorar y adecuar las técnicas disponibles para el balance de carga (estático y dinámico) entre procesos a las arquitecturas consideradas.

En este marco, pueden mencionarse los siguientes resultados:

- Para la experimentación se han utilizado y analizado diferentes arquitecturas homogéneas o heterogéneas, incluyendo multicores, cluster de multicores (con 128 núcleos), GPU y cluster de GPU, Xeon Phi y FPGA.
- Se experimentó la paralelización en arquitecturas híbridas, con el objetivo de estudiar el impacto del mapeo de datos y procesos, así como de los lenguajes y librerías.
- Respecto de las aplicaciones y temas estudiados, se trabajó fundamentalmente con los siguientes problemas:

➢ **Aceleración de aplicaciones con cómputo colaborativo CPU-GPU.** Las computadoras comerciales actuales incluyen decenas de cores y al menos una GPU. El uso de ambas unidades de procesamiento de forma colaborativa puede mejorar significativamente el rendimiento de una aplicación. Sin embargo, esto supone un desafío para los programadores ya que dichas unidades difieren en arquitectura, modelo de programación y rendimiento. En [14] se propuso un modelo híbrido para estructurar código a ser ejecutado sobre un sistema heterogéneo con múltiples cores y 1 GPU (utilizando todos los recursos disponibles). Utilizando este modelo se desarrolló un algoritmo paralelo de pattern matching para sistemas heterogéneos CPU-GPU [15]. Los resultados revelaron que este algoritmo supera en rendimiento a trabajos previos, desarrollados para sistemas multicore y GPUs, para datos de tamaño considerable. En [16] se presentó una solución al problema de pattern matching que aprovecha toda la potencia computacional de los procesadores Intel Xeon Phi KNL 7230 mediante el uso de SIMD y paralelismo de hilos. Se mostró que el algoritmo propuesto alcanza aceleraciones significativas. En trabajos futuros interesa investigar sobre el cómputo colaborativo incluyendo este tipo de arquitecturas.

➢ **Bioinformática y biología computacional.** Estas áreas incluyen aplicaciones que se nutren de las capacidades de HPC para brindar soluciones en un tiempo de respuesta aceptable, siendo cada día más debido al crecimiento exponencial de datos genómicos en los últimos años. Es posible encontrar aplicaciones en alineamiento de secuencias, acoplamiento y dinámica molecular, predicción y búsqueda de estructuras moleculares, entre otras.

➢ **Alineamiento de secuencias.** Si bien el algoritmo de Smith-Waterman es considerado el método de alineamiento más preciso, este algoritmo resulta costoso debido a su complejidad computacional cuadrática. En el contexto de CPUs, se propuso la migración de una librería popular (SSW) a tecnologías actuales [17]. Además, dado el surgimiento del estándar SYCL y de implementaciones como Intel oneAPI, se migró una versión CUDA de este algoritmo a su equivalente en SYCL, analizando el esfuerzo de programación requerido y la relación costo-beneficio entre overhead y portabilidad para diferentes arquitecturas CPU y GPUs [18] [19]. A futuro, se propone completar la migración de SSW, además de extender el estudio realizado contemplando otras generaciones de GPUs y arquitecturas como FPGAs.

➢ **Identificación de biomarcadores:** entre todas las plataformas, servicios y paquetes de software disponibles para esta aplicación, se encuentra Multiomix [20], quien se destaca del resto por facilitar el proceso de integración y minería de datos oncogenómicos públicos y cargados por los usuarios, proporcionando una interfaz gráfica de usuario amigable para los usuarios no expertos. A pesar de ello, Mutiomix aun funciona de forma secuencial por ser un desarrollo reciente (una ejecución habitual en Multiomix puede tomar varias horas para datasets de pequeño porte). Se propone acelerar el procesamiento de Multiomix de manera de disminuir los tiempos de respuesta y así mejorar la productividad de sus usuarios.

➢ **Cálculo de los caminos mínimos**. Es uno de los problemas básicos y de mayor antigüedad de la teoría de grafos teniendo aplicación en el dominio de las

comunicaciones, del ruteo de tráfico, de la bioinformática, entre otros. El algoritmo de Floyd-Warshall (FW) permite computar la distancia mínima entre todos los pares de un grafo. Además de poseer una alta demanda de ancho de banda, FW resulta costoso computacionalmente al ser $O(n^3)$. Habiendo desarrollado una implementación optimizada para arquitecturas x86 [21], se diseñó y desarrollo una librería C que permite usar este código de manera sencilla y práctica en aplicaciones de terceros, dando soporte a diferentes tipos de datos y formatos de archivos de entrada y salida [22].

- **Problemas de optimización de simulación de sistemas dinámicos complejos mediante heurísticas**. En una primera etapa se trabajó con una metodología de sintonización de modelos físicos usando para las experimentaciones un modelo de cauce de ríos. Estos modelos requieren de técnicas que detecten la falta de calibración de sus parámetros debido a muchas clases de incertidumbre que impactan en la calidad de los resultados. Se buscó un conjunto ajustado de parámetros que puedan sintonizar el simulador y se propuso un metamodelo de optimización mediante simulación basado en una heurística combinada entre un método de búsqueda Monte Carlo y un algoritmo de agrupamiento KMeans. Se lograron resultados alentadores que mejoraron la calidad de la simulación con gran reducción de los recursos y tiempo de cómputo. Se actuó "por proximidad" entre las distintas configuraciones de parámetros, lo cual fue incluido en el método combinado. En una segunda etapa se actuó por "similitud" de los valores de los parámetros, aprovechando que los valores de magnitudes físicas en regiones contiguas del dominio tienen poca variación. Se agrega conocimiento del sistema simulado, lo que disminuyó considerablemente el espacio de búsqueda de los parámetros y el costo computacional. Actualmente, se aumentó este conocimiento del sistema aprovechando el conocimiento de eventos disruptivos del pasado y su sintonización construyendo una base histórica de eventos que pueden aprovecharse en eventos actuales, disminuyendo aún más el costo computacional de calibración del modelo [23][24].

- **Ambientes para la enseñanza de concurrencia.** Se desarrolló el entorno CMRE para la enseñanza de programación concurrente y paralela a partir de cursos iniciales en carreras de Informática. Incluye un entorno visual que representa una ciudad en la que pueden definirse varios robots que interactúan. Combina aspectos de memoria compartida y distribuida mediante instrucciones para bloquear y liberar esquinas de la ciudad y el concepto de pasaje de mensajes a través de primitivas de envío y recepción. Además, se incluyen los conceptos de heterogeneidad (diferentes velocidades de los robots) y consumo energético [25]. Se ha integrado con el uso de robots físicos (Lego Mindstorm 3.0) que ejecutan en tiempo real las mismas instrucciones que los robots virtuales y se comunican con el entorno mediante bluetooth [26]. Se ha ampliado para incorporar conceptos básicos de computación en la nube (Cloud Computing) [26]. Actualmente, se está desarrollando una nueva herramienta para la enseñanza de programación concurrente en cursos avanzados. Su objetivo principal es visualizar los conceptos de sincronización y comunicación entre procesos.

- **Cifrado de grandes volúmenes de datos.** Hoy en día, la cantidad de datos sensibles que se generan para ser almacenados y/o transmitidos a través de la red aumenta constantemente. Para proteger los datos confidenciales de amenazas potenciales, se utilizan estrategias de encriptación. Además, el tiempo involucrado en el cifrado de datos está directamente relacionado con la cantidad de datos que se cifrarán y puede ser significativo. Para reducir el tiempo de cifrado es natural recurrir a soluciones de encriptación paralelas, que explotan todo el poder computacional proporcionado por las arquitecturas emergentes. AES (Advanced Encryption Standard) es uno de los algoritmos de cifrado más utilizados y el gobierno de los Estados Unidos lo considera lo suficientemente seguro como para proteger la información nacional. Hay varias implementaciones de AES, tanto en hardware como en software. En [27] se presenta una comparación de rendimiento de una solución AES basada en hardware para CPU multinúcleo con el de otras dos soluciones AES basadas en software para CPU multinúcleo y GPU, respectivamente. El primero se implementa con las nuevas instrucciones de Intel AES y el segundo con la biblioteca OpenSSL. Los resultados revelan que utilizar cómputo paralelo para el proceso de cifrado reduce significativamente el tiempo de ejecución respecto a una solución secuencial. Los resultados también muestran que el mayor rendimiento se alcanza utilizando una solución multicore con AES basado en hardware. Sin embargo, la solución por software utilizando 2 GPUs puede ser una alternativa competitiva a la solución multicore por hardware cuando se tienen pocos núcleos o una CPU que no soporta AES. Asimismo, en [28] se muestra que una solución colaborativa, que utiliza al mismo tiempo múltiples cores y 2 GPUs para el cifrado de grandes volúmenes de datos, permite alcanzar mejoras en el rendimiento.

- **Compresión de grandes volúmenes de datos.** Grandes volúmenes de datos requieren grandes volúmenes de almacenamiento. Por esta razón, se hace necesario desarrollar algoritmos de compresión que ayuden a reducir el costo de almacenamiento sin consumir demasiado tiempo. Un ejemplo de esto es el almacenamiento de datos genómicos. La genómica ha ganado relevancia ya que permite prevenir, diagnosticar y tratar enfermedades de forma personalizada. Es a partir de esta mayor demanda que se requiere una gran

capacidad de almacenamiento de datos genómicos. En [29] se propuso una mejora a un algoritmo multihilo ya existente obteniendo mejoras en el tiempo de ejecución mientras se conserva un bajo ratio de compresión. En [30] se propuso una solución híbrida sobre un cluster multicore mejorando el tiempo de ejecución a partir de reducir la contención por recursos compartidos.

- **Computación cuántica.** Aunque las capacidades de las computadoras cuánticas actuales son muy limitadas, existen muchas líneas de investigación en el área. Estas abarcan el estudio de la fabricación de arquitecturas cuánticas eficientes, simuladores, control de errores por decoherencia, traspilación, el desarrollo de algoritmos cuánticos, la migración de algoritmos clásicos a cuánticos, integración de sistemas cuánticos y clásicos, entornos de desarrollo y lenguajes de programación cuánticos, entre otras.

Organización de Eventos

En 2025 se organizaron las XIII Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2025), con participación de especialistas académicos del país y el exterior, además de empresas con experiencia en Cloud Computing [31]. En junio de 2026 se tiene prevista la organización de las XIV JCC-BD&ET [32].

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D se concluyó 1 tesis doctoral y 2 Tesinas de Licenciatura. Se encuentran en curso en el marco del proyecto 2 tesis doctorales, 2 de maestría, y 2 Tesinas.

Se participa en el dictado de las carreras de Doctorado en Cs. Informáticas y Maestría y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática UNLP, por lo que potencialmente pueden generarse más Tesis y Trabajos Finales.

Asimismo, en grado, los integrantes de la línea de I/D dictan materias relacionadas como Programación Concurrente, Sistemas Paralelos y Taller de Programación sobre GPU.

Existe cooperación con grupos de otras Universidades del país y del exterior, y tesistas de diferentes Universidades realizan su trabajo con el equipo del proyecto.

5. BIBLIOGRAFÍA

[1]. Giles MB, Reguly I. “Trends in high-performance computing for engineering calculations”. Phil.Trans.R.Soc.A 372: 20130319. 2014. http://dx.doi.org/10.1098/rsta.2013.0319

[2]. Jeffers J., Reinders J. “Intel Xeon Phi Coprocessor High Performance Programming”. Morgan Kaufmann.

[3]. Sean Settle. “High-performance Dynamic Programming on FPGAs with OpenCL”. IEEE High Performance Extreme Computing Conf (HPEC ’13). 2013. https://doi.org/10.1016/j.physa.2017.04.159.

[4]. IBM Quantum Computing https://www.ibm.com/quantum

[5]. Google Quantum AI https://quantumai.google/quantumcomputer

[6]. McCool M. “Structured Parallel Programming: Patterns for Efficient Computation”. Morgan Kaufmann.

[7]. De Giusti L, Naiouf M., Chichizola F., Luque E., De Giusti A. “Dynamic Scheduling in Heterogeneous Multiprocessor Architectures. Efficiency Analysis”. Computer Science and Technology Series – XV Argentine Congress of Computer Science Selected Papers. La Plata (Buenos Aires): Editorial de la Universidad de La Plata (edulp). 2010. p85 - 95. isbn 978-950-34-0684-7.

[8]. Chapman B., Jost G., Van der Pas. “Using OpenMP – Portable Shared Memory Parallel Programming”. 2008. UK: MIT Press.

[9]. Hager G., Wellein G. “Introduction to HPC for Scientists and Engineers”. 2011. EEUU: CRC Press.

[10]. De Wael M., Marr S., De Fraine B., Van Cutsem T., De Meuter W. “Partitioned Global Addess Space Languages”. ACM Computing Surveys 47 (4), 2015.

[11]. Pousa, A., Sanz, V. M., Naiouf, M., & De Giusti, A. (2024). Trends in High Performance Computing and Quantum Computing. Journal of Computer Science and Technology, 24(2), e11. https://doi.org/10.24215/16666038.24.e11

[12]. Amazon Elastic Compute Cloud (Amazon EC2). http://aws.amazon.com/es/ec2/. Febrero 2013.

[13]. Balladini J., Rucci E., De Giusti A., Naiouf M., Suppi R., Rexachs D., Luque E. “Power Characterisation of Shared-Memory HPC Systems”. Computer Science & Technology Series – XVIII Argentine Congress of Computer Science Selected Papers. ISBN 978-987-1985-20-3. Pp. 53-65. EDULP, La Plata (Argentina), 2013

[14]. Sanz V., Pousa A., Naiouf M., De Giusti A. “Accelerating Pattern Matching with CPU-GPU Collaborative Computing”. Proceedings of the International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2018), ISBN: 978-3-030-05051-1, págs. 310-322, doi. https://doi.org/10.1007/978-3-030-05051-1_22, 2018.

[15]. Sanz V., Pousa A., Naiouf M., De Giusti A. “Efficient Pattern Matching on CPU-GPU Heterogeneous Systems”. In: Wen S., Zomaya A., Yang L. (eds) Algorithms and Architectures for Parallel Processing. ICA3PP 2019. Lecture Notes in Computer Science, vol 11944. Springer, Cham. 2020.

[16]. Sanz V., Pousa A., Naiouf M., De Giusti A. “Accelerating Pattern Matching on Intel Xeon Phi Processors”. In: Qiu M. (eds) Algorithms and Architectures for Parallel Processing. ICA3PP 2020. Lecture Notes in Computer Science, vol 12452. Springer, Cham. https://doi.org/10.1007/978-3-030-60245-1_18.

[17]. Juan Tomás Vicens (UNLP), Enzo Rucci (UNLP), Manuel Costanzo (UNLP), Marcelo Naiouf (UNLP). “Migración de Librería SSW para Alineamientos de Secuencias Biológicas a Tecnología AVX”. XXXI

Congreso Argentino de Ciencias de la Computación (CACIC 2025). Octubre 2025. En prensa.

[18]. Manuel Costanzo, Enzo Rucci, Carlos García-Sánchez, Marcelo Naiouf, Manuel Prieto-Matías. "Assessing opportunities of SYCL for biological sequence alignment on GPU-based systems." J Supercomput (2024). https://doi.org/10.1007/s11227-024-05907-2.

[19]. Manuel Costanzo, Enzo Rucci, Carlos García-Sánchez, Marcelo Naiouf, Manuel Prieto-Matías, "Analyzing the performance portability of SYCL across CPUs, GPUs, and hybrid systems with SW sequence alignment", Future Generation Computer Systems, Volume 170, 2025, ISSN 0167-739X, https://doi.org/10.1016/j.future.2025.107838

[20]. Genaro Camele, Sebastian Menazzi, Hernán Chanfreau, Agustin Marraco, Waldo Hasperué, Matias D Butti, Martin C Abba, Multiomix: a cloud-based platform to infer cancer genomic and epigenomic events associated with gene expression modulation, Bioinformatics, Volume 38, Issue 3, 1 February 2022, Pages 866–868, https://doi.org/10.1093/bioinformatics/btab678

[21]. S. Calderón, E. Rucci, y F. Chichizola, «Enhanced OpenMP Algorithm to Compute All-Pairs Shortest Path on X86 Architectures», en Computer Science – CACIC 2023, P. Pesado, W. Panessi, y J. M. Fernández, Eds., Cham: Springer Nature Switzerland, 2024, pp. 46-61. doi: 10.1007/978-3-031-62245-8_4.

[22]. Jerónimo Lambre, Enzo Rucci. "Librería C para Cómputo Paralelo de Caminos Mínimos en Grafos sobre Arquitecturas Multicore". XXIX Congreso Argentino de Ciencias de la Computación – CACIC 2024. Octubre de 2024. En prensa

[23]. Trigila, M., Gaudiani, A., Wong, A., Rexachs, D., Luque, E. (2023). Reduction of the Computational Cost of Tuning Methodology of a Simulator of a Physical System. In: Mikyška, J., de Mulatier, C., Paszynski, M., Krzhizhanovskaya, V.V., Dongarra, J.J., Sloot, P.M. (eds) Computational Science – ICCS 2023. ICCS 2023. Lecture Notes in Computer Science, vol 10475. Springer, Cham. https://doi.org/10.1007/978-3-031-36024-4_49.

[24]. Gaudiani, A., Wong, A., Luque, E. et al. A computational methodology applied to optimize the performance of a river model under uncertainty conditions. J Supercomput 79, 4737–4759 (2023). https://doi.org/10.1007/s11227-022-04816-6

[25]. Castro J., De Giusti L., Gorga G., Sánchez M., Naiouf M. "ECMRE: Extended Concurrent Multi Robot Environment". Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3, Springer, Cham, págs. 285-294. 2018.

[26]. De Giusti L., Chichizola F., Rodríguez Eguren S., Sánchez M., Paniego J. M., De Giusti A. "Introduciendo conceptos de Cloud Computing utilizando el entorno CMRE". Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) – Workshop de Innovación en Educación en Informática. 2016.

[27]. Sanz V., Pousa A., Naiouf M., De Giusti A. "Comparison of Hardware and Software Implementations of AES on Shared-Memory Architectures". In: Naiouf M., Rucci E., Chichizola F., De Giusti L. (eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2021. Communications in Computer and Information Science, vol 1444. Springer, Cham. 2021. https://doi.org/10.1007/978-3-030-84825-5_5.

[28]. Sanz, V., Pousa, A., Naiouf, M., De Giusti, A. "Performance Analysis of AES on CPU-GPU Heterogeneous Systems". In: Rucci, E., Naiouf, M., Chichizola, F., De Giusti, L., De Giusti, A. (eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2022. Communications in Computer and Information Science, vol 1634. Springer, Cham. 2022. https://doi.org/10.1007/978-3-031-14599-5_3

[29]. Sanz V., Pousa A., Naiouf M., De Giusti A. "Fast Genomic Data Compression on Multicore Machine" XII Jornadas de Cloud Computing, Big Data & Emerging Topics. DOIhttps://doi.org/10.1007/978-3-031-70807-7_1

[30]. Victoria Sanz, Adrián Pousa, Marcelo Naiouf, Armando De Giusti. Computational Science – ICCS 2025: 25th International Conference, Singapore, Singapore, July 7–9, 2025, Proceedings, Part IV Pages 180 – 188. https://doi.org/10.1007/978-3-031-97635-3_22

[31]. Marcelo Naiouf, Laura De Giusti, Franco Chichizola, Leandro Libutti "Cloud Computing, Big Data and Emerging Topics 13th Conference, JCC-BD&ET 2025, La Plata, Argentina, June 24–26, 2025, Revised Selected Papers". Communications in Computer and Information Science. Springer Cham. 2025. https://doi.org/10.1007/978-3-032-06336-6.

[32]. XIII Jornadas de Cloud Computing, Big Data & Emerging Topics. https://jcc.info.unlp.edu.ar/

HPC en la Nube para el procesamiento de datos: Simulación y Despliegue en Arquitecturas IaaS.

Diego Encinas[1,2], Brian Galarza[1], Román Bond[1], Gonzalo Zaccardi[1], Nicolás Benquerença Mendes[1], Rodrigo Lago[1], Cristian Iglesias[1], Rodrigo Iván Duarte[1], Jorge Osio[1], David Duarte[1], Martin Morales[1,3]

[1]Proyecto de Investigación SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - UNAJ

[2]Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

[3]Centro CodApli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, bgalarza@unaj.edu.ar, rbond@unaj.edu.ar, gzaccardi@unaj.edu.ar, nbenquerenca@unaj.edu.ar, rodrigazo98@gmail.com, roodrigo4@hotmail.com, josio@unaj.edu.ar, davito.duarte.22@gmail.com, martin.morales@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio del rendimiento de las arquitecturas tipo cloud a través del despliegue de IaaS y utilización de IaaS públicos, en particular en el área de cómputo paralelo de altas prestaciones (HPC). Enfocando en la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente, especialmente la entrada/salida y las comunicaciones.

Palabras clave: *Cloud Computing. OpenStack. Sistemas de Archivos en clústers. Redes definidas por Software.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación "Simulación, Computación de Altas Prestaciones (HPC) y optimización de aplicaciones sociales – SimHPC" de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 183/21. Además, el proyecto colabora con el Programa "Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social" – TICAPPS de la UNAJ.

Introducción

Cloud Computing es un paradigma que ha estado en constante crecimiento. Cada vez más compañías y grupos de investigación trabajan en conjunto con el fin de explotar las oportunidades ofrecidas por el mismo [1]. Dicho paradigma ofrece muchas ventajas, tales como el bajo costo de implementación, ya que no se necesitan computadoras de última tecnología debido a que éstas trabajan conjuntamente (Clustering) con la posibilidad de escalar horizontalmente de manera sencilla. Además, hay software Open Source disponible para los nodos en el clúster como las infraestructuras Eucalyptus, OpenNebula, CloudStack u OpenStack

integradas con GNU/Linux y compatibles, por ejemplo, con Amazon WebServices.

Despliegue de IaaS

Las comunicaciones en Cloud Computing son una parte fundamental del paradigma que consisten en utilizar distintos nodos y lograr hacerlos funcionar conjuntamente.

Para lograr una comunicación sincronizada entre estos nodos se propone utilizar OpenStack [2] como también OpenNebula [3].

OpenStack es una plataforma de tecnología open source que utiliza recursos virtuales agrupados para diseñar y gestionar nubes privadas y públicas a través de múltiples servicios que, de manera coordinada, cumplen diferentes propósitos para lograr el correcto funcionamiento de, por ejemplo, una ”Infraestructure as a service” (IaaS). Algunos de los servicios ofrecidos por OpenStack son: hypervisor (Nova), autentificación (Keystone), Imágenes (Glance), Dashboard (Horizon), Networking (Neutron) y block storage (Cinder). Según las necesidades se pueden requerir de ciertos servicios u otros. La Arquitectura básicamente consiste en dos tipos de nodos: “Compute Node” y “Controller Node”. Se llaman Compute Node a todos aquellos que se encargan del procesamiento de servicios específicos mientras que Controller Node es aquel que comunica a cada uno de los anteriores [4] [5] [6].

Fuel es una herramienta open source desarrollada por Mirantis en la cual se ejecuta un script que permite configurar, de manera más amigable respecto a OpenStack, los recursos que se desean otorgar a la infraestructura, como la cantidad de nodos, los núcleos de procesador, la memoria RAM, entre otros [7].

Fuel trabaja con un nodo master el cual es el encargado de controlar a los nodos slaves que contendrán la infraestructura OpenStack. Es decir, desde el nodo Fuel Master se indican qué paquetes se van a instalar en cada nodo slave (Glance, Nova-Compute, Keystone, etc.) para luego en los slaves tener armados los nodos compute y controller, sin necesidad de realizar configuraciones manuales en cada uno de los mismos.

OpenNebula es un software de código abierto que permite el despliegue de IaaS. Busca reducir la complejidad generada por OpenStack y ofrece soporte con hypervisores tales como KVM y VMware vCenter.

OpenNebula clasifica a los nodos en dos tipos, Front – end los cuales entran en contacto con los usuarios y a su vez se comunican con los nodos de la infraestructura en los cuales se lanzarán las instancias y los nodos virtualizados los que a su vez deben contar con los paquetes correspondientes de storage, autenticación y networking para poder funcionar correctamente.

La implementación de estas infraestructuras ofrece ventajas en las cuales los clústeres virtualizados trabajan en conjunto ofreciendo un buen rendimiento a bajos costos y con posibilidad de escalabilidad al poder agregar mayor cantidad de nodos para procesamiento de manera sencilla.

Sistemas de Archivos Paralelos en clústeres

Amazon Webservices [8], mediante el servicio EC2 (Amazon Elastic Compute Cloud) permite desplegar clústeres virtuales mediante instancias de VMs y almacenamiento para las mismas. Mediante este servicio y el correspondiente clúster conformado, se utiliza un sistema de archivos paralelo (PVFS2) que permite la

gestión de datos particionados y distribuidos en los distintos nodos, mediante múltiples tareas de una aplicación ejecutada sobre el clúster.

PVFS2 utiliza una estructura cliente-servidor. Dependiendo de su rol, existen tres tipos de nodos dentro de un clúster con este sistema de archivos: servidores de datos, servidores de metadatos y clientes, en donde cada uno de los nodos puede cumplir los tres roles.

PVFS2 [9] contiene una herramienta interna que permite volcar información en logs durante la ejecución de tareas dentro del sistema de archivos. Por cada uno de los nodos, el administrador del clúster puede obtener información de depuración (GOSSIP) conformado por registros de debug, de acceso, contadores de rendimiento y errores producidos durante la ejecución. Referido a los contadores de rendimiento, se obtienen mediante el software de monitoreo Atop. Esta herramienta permite obtener reportes de la actividad de los procesos y la utilización de los diferentes recursos del sistema (memoria, disco, CPU, red, etc) [10].

Asimismo, es posible especificar los datos a obtener en función de las distintas capas de PVFS2, del rol específico de cada nodo (cliente-servidor), así como también de otros factores relacionados con el funcionamiento interno y operaciones asociadas a la gestión de archivos.

Teniendo conocimiento del funcionamiento del código fuente correspondiente a PVFS2 y del modo en el que realiza la escritura de registros en los logs, resulta factible realizar modificaciones que permitan obtener otro tipo de parámetros de interés a partir de esta herramienta.

Finalmente se propone el análisis, uso y configuración de distintas herramientas no invasivas para determinar la performance del sistema de archivos en clústeres virtuales.

Evaluación del rendimiento en BigData

En los últimos años, la generación de datos ha crecido exponencialmente. A partir de esto, diferentes tecnologías de procesamiento surgieron con el fin de poder lidiar con el gran volumen de datos existente. Algunas de estas tecnologías son Apache Spark [11], Dask [12] y Apache Ray [13]. Estos frameworks permiten realizar cómputo distribuido utilizando lenguajes de alto nivel de abstracción como Python en entornos y arquitecturas HPC.

Por lo tanto evaluar la eficiencia en el procesamiento de cargas de trabajo de datos intensivas en un entorno de clúster controlado resulta un desafío para administradores y proveedores de servicios. Por tal motivo, se configura un clúster de seis nodos con un sistema de archivos distribuido HDFS (Hadoop) y el gestor de recursos YARN. Sobre esta infraestructura, se ejecutarán scripts de procesamiento de datos (por ejemplo: tareas de ETL, agregaciones o entrenamiento de modelos de machine learning) adaptados a cada framework. El rendimiento no solo se medirá en términos de tiempo total de ejecución, sino que se analizará a un nivel más profundo utilizando las herramientas de monitoreo de sistema collectl y atop. Estas herramientas permitirán capturar métricas detalladas de consumo de CPU, uso de memoria y tráfico de red.

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de

memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.
- Plataformas de software para implementar y administrar Clouds públicos, privados e híbridos.
- Sistemas de Archivos Paralelos.
- Procesamientos de Grandes Volúmenes de datos en entornos cloud con herramientas del ecosistema Apache Hadoop y afines.

Resultados y Objetivos

Investigación experimental

- Implementación de un IaaS encargado de realizar operaciones en procesamiento paralelo aumentando la eficiencia y reduciendo los costes generados.
- Implementación de OpenStack Dashboard y de un sistema desarrollado para poder controlar/administrar de manera visual (web) y más básica cada uno de los servicios.
- Implementación de OpenNebula en un sistema con las mismas características que el implementado por OpenStack con el fin de poder realizar pruebas en entornos similares.
- Utilización de Fuel para administrar OpenStack como sistema de administración de nube (Cloud Computing) a partir de la infraestructura de 2 nodos compute y e1 controller [14].
- Análisis del rendimiento de un Cloud privado en la ejecución de instancias personalizadas.
- Ejecutar diferentes benchmarks en la infraestructura desplegada sobre OpenStack y OpenNebula para así comparar los resultados obtenidos de ambas infraestructuras y poder realizar un análisis del rendimiento en cada caso.
- Introducción al estudio e integración de sensores físicos y los servicios en la nube [15].
- Medición de servicios en la nube enfocados a IaaS y PaaS [16].
- Análisis y configuración de clústeres virtuales. Análisis y configuración de herramientas no invasivas para la obtención de métricas en las distintas capas de software de los sistemas de archivos paralelos.
- Utilización de otros sistemas de archivos paralelos como Lustre [17] y Beegfs [18] para obtener métricas en Metadataservidores [19] [20] [21].
- En el área de tiempo real se trabajó en el reconocimiento de audio mediante la técnica huella de audio [22]. Por otro lado, se estudió mediante simulación el funcionamiento de una banda clasificadora utilizada en reciclaje [23]. En este mismo contexto se presentó un trabajo que explora mediante la simulación el funcionamiento de una línea de producción [24].
- Implementación de un sistema de reconocimiento emocional facial con Python y OpenCV [25].
- En el área de eficiencia energética, se trabajó en el impactó de las aplicaciones en dispositivo móvil y su relación con la emisión de gases de efecto invernadero (GEI) [26].

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería Informática de la UNAJ. También aportan trabajos de alumnos de las materias Sistemas Operativos 1, Redes de

Computadoras 2, Programación en Tiempo Real y Organización y Arquitecturas de Computadoras. Por otro lado, algunos integrantes participan en el dictado de la Maestría y Diplomatura en Ciencia de Datos de la UNAJ.

Durante 2025 se han realizado publicaciones nacionales [27][28][29[30][31]. En esta línea de I/D existe cooperación a nivel nacional. Hay 3 investigadores realizando carreras de postgrado y alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Kondo, D., Javadi, B., Malecot, P., Cappello, F., Anderson, D. P.: "Cost-benefit analysis of Cloud Computing versus desktop grids". In: IPDPS '09 Proceedings. IEEE International Symposium on Parallel and Distributed Processing. Washington, USA (2009).
2. OpenStack Cloud Software: Open source software for building private and public clouds. http://www.openstack.org. Febrero 2015.
3. OpenNebula. https://opennebula.org/. Febrero 2019
4. Galarza, B.; Tuamá, C.; Zaccardi, G.; Encinas, D.; Morales, M. "Implementaciones de Cloud Computing y aplicaciones en el ámbito universitario". 1° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2013). Ciudad de Córdoba, Argentina.
5. Zaccardi, G.; Galarza, B.; Encinas, D.; Morales, M. "Implementación de Cloud Computing utilizando OpenStack". 2° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2014). Ciudad de San Luis, Argentina.
6. Galarza, B.; Zaccardi, G.; Encinas, D.; Morales, M. "Análisis de despliegue de una IaaS utilizando Openstack". XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015). Ciudad de Junín, Argentina.
7. OpenStack Deployment Fuel. https://www.mirantis.com/products/mirantis-openstack-software/openstack-deployment-fuel/. Febrero 2016
8. Amazon Web Services (AWS)-Cloud Computing Services. https://aws.amazon.com Marzo 2019
9. T. PVFS2, "PVFS 2 File System Semantics Document," tech. rep., PVFS Development Team, 2015
10. Atop Tool. https://www.atoptool.nl/index.php
11. Salloum, S., Dautov, R., Chen, X., Peng, P. X., & Huang, J. Z. (2016). Big data analytics on Apache Spark. International Journal of Data Science and Analytics, 1(3), 145-164.
12. Peters, T. (2023). Parallel Python with Dask: Perform distributed computing, concurrent programming and manage large dataset. GitforGits.
13. Kato, K., Takefusa, A., Nakada, H., & Oguchi, M. (2018, December). A study of a scalable distributed stream processing infrastructure using Ray and Apache Kafka. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 5351-5353). IEEE.
14. Zaccardi, G.; Galarza, B.; Morales, M.;Encinas, D. "Despliegue y ejecución de un cloud privado". 4°Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2016). Ciudad de Salta, Argentina.
15. Armanno, M.; Navarro, G.; Morales, M.; Encinas, D. "Utilización de servicios de Cloud Computing y sensores". 8º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.
16. González, N.; Lescano, N.; Pinto, C.; Morales, M.; Encinas, D. "Análisis de rendimiento de IaaS y PaaS". 8º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.
17. Lustre Manual. https://doc.lustre.org/lustre_manual.xhtml#idm140436306123424. Febrero 2021.
18. Heichler, Jan. "An introduction to BeeGFS." (2014).
19. Benquerença Mendes, N.; Bond, R.; Morales, M.; Encinas, D. "Rendimiento de sistema de archivos en arquitecturas

distribuidas y paralelas**”.** 8º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.

20. E. Párraga, B. León, R. Bond, D. Encinas, A. Bezerra, S. Mendez, D. Rexachs, E. Luque. (2021) Analyzing the I/O Patterns of Deep Learning Applications. Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2021. Communications in Computer and Information Science, vol 1444. Springer, Cham. https://doi.org/10.1007/978-3-030-84825-5_1
21. Benquerença Mendes, N.; Bond, R.; Morales, M.; Encinas, D. “Sistema de Archivos Paralelos con Aplicaciones de Machine Learning.”. XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022). La Rioja, Argentina.
22. Rodrigo Lago; Verónica Scholz; Roman Bond; Martin Morales; Diego Encinas. Experiencia de desarrollo de una aplicación de reconocimiento de canciones mediante la técnica de huella de audio". En XXIX Congreso Argentino de Ciencias de la Computación CACiC, Argentina. 2023.
23. Alan Dell Orto Marengo, Emanuel Julio, Diego Rolón, Román Bond, Martin Morales, Diego Encinas. Simulación de sistema de reciclaje mediante banda clasificadora. En 11° Congreso Nacional de Ingeniería Informática y Sistemas de la Información, Argentina. 2023.
24. Emiliano Camer; Martin Páez; Nicolás Vitek; Román Bond; Daniel Martín Morales; Diego Encinas. “Estudio de simulación de procesos aplicado a la optimización de sistemas de producción”. En XXX Congreso Argentino de Ciencias de la Computación CACiC, Argentina. 2024.
25. Albornoz, G.; Bond, R.; Morales, M.; Encinas, D. “Implementación de un Sistema de Reconocimiento Emocional Facial con Python y OpenCV”. 12º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2024). Catamarca.
26. Yamila Ferrufino; Román Bond; Daniel Martín Morales; Diego Encinas. “Evaluación de la Eficiencia Energética y el Impacto de las Aplicaciones de Redes Sociales”. En XXX Congreso Argentino de Ciencias de la Computación CACiC, Argentina. 2024.
27. Ignacio DiLoreto; Román Bond; Martin Morales; Diego Encinas. "Estrategias de control del Dengue utilizando Modelado y Simulación basado en Agentes". En XXXI Congreso Argentino de Ciencias de la Computación (CACIC), Argentina. 2025.
28. Sabrina Russell; Franco Paiz; Román Bond; Martin Morales; Diego Encinas. "Diseño e implementación de un robot a tracción diferencial". En Actas 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2025), Argentina. 2025.
29. Guillermo Cruz; Jorge Gamez; Carlos Miers; Román Bond; Martin Morales; Diego Encinas. "Prototipo de Bebedero Automático para Mascotas". En Actas 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2025), Argentina. 2025.
30. David Baltazar; Juan Dominguez; Román Bond; Martin Morales; Diego Encinas. "Sistema automatizado de envasado y control de medicamentos en polvo con monitoreo ambiental". En Actas 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2025), Argentina. 2025.
31. Cristian Iglesias, Jorge Osio, Diego Encinas. “Apache Flink para monitoreo energético en tiempo real”. En Actas 13º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2025), Argentina. 2025.

Marco de Investigación para el Modelado, Procesamiento y Nuevas Tecnologías para la Gestión de Datos

Arroyuelo Jorge, Barrionuevo Mercedes, Labella Danilo, Loor Fabricio,
Palacios Cecilia, Palacio Gabriela, Printista A. Marcela, Sarubi Franco, Tissera Pablo Cristian
LIDIC, Dpto. de Informática, FCFMyN - Universidad Nacional de San Luis
{jarroyu; mercedes.barrionuevo; dglabella; loorfabricio; mcecipalacios; gdpalaci; marprinti;sarubbi.fe; ptissera}@gmail.com

Gil Costa G. Verónica
LIDIC, Dpto. de Minería, FCFMyN - Universidad Nacional de San Luis. CONICET
ggvcosta@gmail.com

Ganuza M. Luján
VyGLAB, Dpto. de Ciencias e Ingeniería en Computación, Universidad Nacional de Sur
mlg@sc.uns.edu.ar

RESUMEN

La presente línea de investigación se orienta al estudio, diseño e implementación de modelos, algoritmos y arquitecturas computacionales avanzadas para la gestión integral de ingentes cantidades de datos en entornos complejos. Se abordan problemáticas vinculadas al modelado estadístico y al aprendizaje automático, la simulación basada en agentes, la incorporación de tecnologías emergentes como blockchain y el desarrollo de herramientas de visualización interactiva orientadas a la toma de decisiones. Como una meta a desarrollar en los próximos años, nos proponemos que todas las soluciones que se desarrollen sean capaces de gestionar grandes volúmenes de datos de manera eficiente en entornos de Computación de Alto Rendimiento (HPC).

El enfoque adoptado integra investigación, desarrollo e innovación (I+D+i), promoviendo la transferencia tecnológica y la generación de soluciones aplicadas en ámbitos ambientales, educativos y organizacionales.

Palabras clave**:** *modelado de datos, análisis de datos, aprendizaje automático, simulación basada en agentes, visualización interactiva, procesamiento de alto rendimiento*.

1. CONTEXTO

La línea de investigación denominada "Modelado, Procesamiento y Nuevas Tecnologías para la Gestión de Datos" se enmarca en el Proyecto de Investigación "*Soluciones Computacionales para la Gestión de Datos*", desarrollado en la Universidad Nacional de San Luis (2026-2029). Este proyecto es una reformulación de un proyecto anterior denominado "*Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos*" que propone avanzar hacia un enfoque innovador capaz de abordar problemas concretos, con el fin de analizar los resultados en un contexto específico mediante la aplicación de estrategias de Ciencia de Datos.

Las actividades se llevarán a cabo en el Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC), en articulación con otros grupos académicos locales y externos, consolidando un ambiente interdisciplinario orientado al desarrollo de soluciones basadas en ciencia de datos y tecnologías emergentes. A

través de los años, investigadores de las diferentes áreas de Computación de la UNSL y de otras instituciones con las que se realiza colaboración, han asociado a este laboratorio iniciativas de investigación y de cooperación científica, en función de sus objetivos y temáticas de interés.

Esta sinergia favorecerá la generación de un entorno propicio para alcanzar los siguientes objetivos generales de investigación y desarrollo:

a) Promover la generación de resultados transferibles mediante la articulación sistemática de Investigación, Desarrollo e Innovación (I+D+i).
b) Consolidar la vinculación con el entorno socio-productivo, favoreciendo la producción de conocimiento aplicable que contribuya a la toma de decisiones estratégicas.
c) Integrar los desarrollos tecnológicos en un esquema de innovación alineado con procesos de transformación digital y sostenibilidad.

2. INTRODUCCIÓN

La creciente disponibilidad de datos heterogéneos y de gran volumen plantea desafíos sustanciales en términos de almacenamiento, procesamiento, modelado y análisis. En este contexto, esta línea de investigación propone el estudio, desarrollo e integración de metodologías algorítmicas y arquitecturas tecnológicas que permitan transformar datos en conocimiento accionable.
Se abordarán técnicas de estadística avanzada, aprendizaje automático supervisado y no supervisado, pensamiento computacional aplicado al análisis estructurado de problemas y simulación basada en agentes para el estudio de sistemas complejos. Asimismo, se investigará la incorporación de tecnologías como blockchain para garantizar la integridad, trazabilidad e inmutabilidad de la información, y plataformas de procesamiento distribuido para el manejo eficiente de grandes volúmenes de datos.
Dicha incorporación se basa en la existencia de datos corruptos o manipulados dando lugar a fallas en modelos predictivos. Por lo tanto, las tecnologías de blockchain pueden actuar como etapa de certificación de conjunto de datos.
Partiendo de una investigación que revele las brechas y oportunidades del estado actual de la tres áreas (Big Data, HPC y Ciencia de Datos), se han definido los siguientes objetivos:
a) Diseñar e implementar algoritmos de aprendizaje automático adaptados a diferentes dominios de aplicación:
b) Evaluar e integrar tecnologías emergentes -como big data y blockchain- en arquitecturas de gestión de datos, mediante el desarrollo de prototipos funcionales.
c) Desarrollar modelos de simulación basada en agentes orientados al análisis y optimización de flujos de información en entornos dinámicos.
d) Explorar la integración entre modelos de simulación, HPC, ciencia de datos y mecanismos de blockchain para mejorar transparencia y eficiencia en la gestión de datos.
e) Investigar técnicas avanzadas de visualización interactiva aplicadas tanto a datos reales como a datos generados por simulación.
f) Definir métricas y criterios de evaluación que permitan analizar desempeño, seguridad y robustez de las soluciones implementadas.

Iniciar investigaciones que permitan la convergencia entre HPC, ciencia de datos y blockchain permitirá contar con una arquitectura computacional de próxima generación en la cual el procesamiento masivo, la inteligencia analítica y la confianza distribuida, se integren para habilitar sistemas escalables, verificables y autónomos.

3. ÁREAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se presentan con mayor profundidad algunas de las áreas de investigación en las que estamos actualmente involucrados. En las secciones siguientes, se esbozan otras áreas de desarrollo que se explorarán progresivamente a medida que avance el nuevo proyecto, recientemente aprobado

Simulación basada en agentes y gemelos digitales: la función de un Gemelo Digital es replicar el comportamiento y las condiciones de un sistema real, lo que permite simular diferentes escenarios, predecir resultados y optimizar procesos en tiempo real.
El grupo de investigación ha desarrollado un Simulador basado en agentes (ABS), el cual tiene la capacidad de instanciar el modelo a diferentes situaciones complejas, como lo es, por ejemplo, la simulación de evacuaciones de personas, aunque no está restringida a esta área. Cada instancia permite describir el entorno físico, los diferentes actores (agentes), su interacción, flujos de información y proporciona un contexto para la toma de decisiones, basado en modernos algoritmos de comportamientos inteligentes. Los agentes en este Simulador están capacitados para responder a condiciones cambiantes que pueden surgir durante la simulación (siguiendo nuestro ejemplo, la aparición de obstáculos, variaciones en la visibilidad, comportamiento en pánico, etc.), por lo que el ABS ha resultado de gran utilidad cuando se quiere observar y estudiar el comportamiento de los agentes.
En base a los trabajos científicos e implementaciones realizadas, nuestros próximos objetivos son: a) Implementar herramientas de visualización y monitoreo para dotar al ABS con la capacidad de tener conexión con el estado del sistema en tiempo real y realizar ajustes sobre la marcha y b) Incluir análisis de datos históricos para predecir comportamientos futuros. Una vez incorporadas estas facilidades, y en base a una interacción bidireccional, el gemelo digital se concebirá como una plataforma inteligente implementada con tecnologías innovadoras y podrá proporcionar una herramienta poderosa para la planificación y gestión de diferentes sistemas reales.

Análisis visual de datos meteorológicos: un enfoque integrado aplicado a la provincia de San Luis:

El análisis de datos meteorológicos ha evolucionado dramáticamente en los últimos años, en gran medida debido al crecimiento exponencial de las redes de sensores, las capacidades computacionales y la acumulación de conjuntos de datos históricos. En esta área se ha pasado del procesamiento de datos a escala de gigabytes a abordar desafíos de escala de Terabytes y Petabytes [AFZAL19]. En la Provincia de San Luis, la Red de Estaciones Meteorológicas (REM) [REM25] compuesta por más de 59 estaciones meteorológicas recopila lecturas cada minuto, resultando en un archivo histórico que abarca 15 años. Estos datos de series temporales de gran escala y alta granularidad [Ouedraogo24] plantean desafíos significativos para los métodos tradicionales de procesamiento y visualización. El sistema propuesto permitirá detectar y caracterizar eventos climáticos extremos (por ejemplo, olas de calor, eventos de heladas, fenómenos relacionados a eventos El Niño/La Niña, entre otros), facilitando la comunicación clara y eficiente de dichos patrones a través representaciones gráficas intuitivas integradas en dashboards interactivos que apoyen la toma de decisiones en varios entornos pero en particular, en el ámbito ambiental y agronómico.

El objetivo de este sistema busca integrar: a) Modelos de aprendizaje automático para detección de anomalías en series temporales multivariadas, b) Procesamiento de alto rendimiento y

almacenamiento escalable, c) Visualización interactiva mediante dashboards dinámicos, d) Incorporación de curvas Intensidad–Duración–Frecuencia (IDF) [NOBOA25] para caracterizar eventos de precipitación extrema y f) aplicar técnicas de interpolación espacial [CUADRADO22], las cuales requieren gran procesamiento de datos para la completitud de datos faltantes, utilizando datos de localidades vecinas, para completar las series temporales.

4. RESULTADOS ESPERADOS

El Proyecto de Investigación "*Soluciones Computacionales para la Gestión de Datos*" al cual pertenece esta línea de trabajo, se ha iniciado en febrero de 2026. Paulatinamente, se prevé que esta investigación proporcione los siguientes resultados:

- Desarrollar prototipos funcionales validados en contextos reales.
- Definir métricas comparativas para evaluar desempeño, robustez y escalabilidad de las soluciones implementadas.
- Consolidar un marco metodológico replicable para la gestión avanzada de datos en dominios complejos.
- Publicar resultados en congresos y revistas especializadas.
- Generar herramientas transferibles al sector público y privado.
- Capacitación continua de los miembros de la línea de investigación.

Es importante mencionar que en todas las áreas presentadas se considera trabajar en el tratamiento de datos a gran escala, por lo que hay tres etapas que se mantendrán comunes a todas ellas. La primera etapa plantea investigar e implementar nuevas metodologías para el procesamiento eficiente y eficaz de grandes volúmenes de datos; mientras que la segunda etapa contempla investigar para utilizar tecnologías innovadoras para la obtención de información útil. Como en cada uno de estos procesos hay involucrados un gran número de problemas computacionalmente costosos, el uso de nuevas técnicas y arquitecturas puede contribuir a mejorar su rendimiento; es por ello que la tercer etapa implicará la búsqueda y selección de técnicas de computación de altas prestaciones (HPC) en área de trabajo, para poder alcanzar con eficiencia cada uno de sus objetivos.

5. FORMACIÓN DE RECURSOS HUMANOS

Desde la línea de investigación se contribuirá activamente a la formación de recursos humanos altamente especializados. Actualmente, el grupo de investigación cuenta con la supervisión de tres doctores en ciencias de la computación, quienes están asignados a las investigaciones de las áreas descritas en este artículo.

Además, se cuenta con la participación de dos investigadores que están finalizando su formación de posgrado en las siguientes carreras: *Maestría en Inteligencia de Datos Orientada a Big Data*, cuya tesis está centrada en técnicas avanzadas de visualización (tema de aplicación desarrollado en la sección 3 de este paper y con sus primeras publicaciones disponibles en congresos [CACIC25]) y *Maestría en Ciencia de Datos e Innovación Empresarial*, cuyo tema central de tesis es la búsqueda de modelos predictivos de Machine Learning en aplicaciones de interés al ámbito educativo-universitario tomando como bases distintos trabajos previos [HIMMEL18, ISTVAN19].

Por otra parte, dos docentes/investigadores están interesados en generar estrategias y recursos digitales para prevenir la deserción en la UNSL. En este caso, estarán enfocados a la implementación del pensamiento computacional en el ámbito de la ciencia de datos, centrándose en el desarrollo de modelos eficientes que permitan la detección de patrones. Estos

patrones son cruciales para facilitar intervenciones tempranas y prevenir el problema de deserción planteado.

Por último, se menciona la formación de dos estudiantes de grado, quienes se centran en la recolección de datos relacionados con variables ambientales. Estos datos serán capturados mediante sensores y alimentarán al pipeline de ciencia de datos, donde se llevarán a cabo los procesos de almacenamiento, procesamiento escalable y eficiente, análisis y posterior muestra de resultados en entornos de comunicación efectiva.

6. BIBLIOGRAFÍA

[AFZAL19] Afzal, S., Hittawe, M. M., Ghani, S., Jamil, T., Knio, O., Hadwiger, M., & Hoteit, I. (2019, June). The state of the art in visual analysis approaches for ocean and atmospheric datasets. In Computer graphics forum (Vol. 38, No. 3, pp.881-907). https://doi.org/10.1111/cgf.13731

[CACIC25] Plataforma de Visualización para el Monitoreo Meteorológico: J. Arroyuelo, M. Barrionuevo, V. Gil-Costa y C. Tissera. Universidad Nacional de San Luis (UNSL). XXXI Congreso Argentino de Ciencias de la Computación - CACIC 2025. pp. 466-475.

[CUADRADO22] J.Cuadrado, Erick Horna. ESPOCHInterpolación espacial y estructura espaciales. 2022-06-13. https://rpubs.com/JhonaCuad5/914016

[HIMMEL18] Himmel, E. Modelo de análisis de la deserción estudiantil en la educación superior. Calidad en la Educación, ISSN-e 0718-4565, No. 17, pag. 91 - 108. 2018

[ISTVAN19] Istvan, R. M. Sistema informático para la detección temprana de deserción estudiantil universitaria. Innovación y Desarrollo Tecnológico y Social (IDTS); vol. 1, no. 2, pag. 1-15. (2019).

[JONES20] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, Characterising the digital twin: A systematic literature review. CIRP journal of manufacturing science and technology, vol. 29, pp. 36–52, 2020.

[LEW20] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., Lewis, M., Yih, W.-t., Rockt¨aschel, T., et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in Neural Information Processing Systems 33, 9459–9474 (2020)

[NOBOA25] Noboa, J. P. B., Vaca, N. E. P., Coba, M. E. L., & Sarabia, D. A. P. (2025, July). Generación de Curvas IDF para la Quebrada Las Abras a partir de Regresión Lineal Múltiple y el Método de Talbot. In ASCE (Vol. 4, No. 3, pp. 685-710).

[Ouedraogo24] Ouedraogo, I. Collaborative Immersive Analytics for Meteorological and Climatological Data. Human-Computer Interaction. Doctoral dissertation, Université Paris-Saclay, 2024. English.https://theses.hal.science/tel-04764690/file/128017_OUEDRAOGO_2024_archivage.pdf

[REM25] Red de Estaciones Meteorológicas, San Luis (REM). www.clima.sanluis.gob.ar. Último acceso: Agosto de 2025.

[SUN24] Y. Sun, Q. Zhang, J. Bao, Y. Lu, and S. Liu, Empowering digital twins with large language models for global temporal feature learning. Journal of Manufacturing Systems, vol. 74,pp. 83–99, 2024.

[ZHENG22] X. Zheng, J. Lu, and D. Kiritsis, The emergence of cognitive digital twin: vision, challenges and opportunities. International Journal of Production Research, vol. 60, no. 24, pp. 7610–7632, 2022.

Marco de Investigación y Desarrollo a través de la integración de Simulación, Big Data y Ciencia de Datos

Arroyuelo Jorge, Barrionuevo Mercedes, Labella Danilo, Printista A. Marcela, Tissera Pablo Cristian
LIDIC, Dpto. de Informática, FCFMyN - Universidad Nacional de San Luis
{*jarroyu; mercedes.barrionuevo; dglabella; marprinti; ptissera}@gmail.com*

Gil Costa G. Verónica
LIDIC, Dpto. de Minería, FCFMyN - Universidad Nacional de San Luis. CONICET
ggvcosta@gmail.com

Debattista Alfredo
LEIS, Dpto. de Electrónica, FCFMyN - Universidad Nacional de San Luis
adebattista@gmail.com

RESUMEN

La investigación desarrollada en esta publicación se centra en la optimización de algoritmos de procesamiento de datos, el diseño de arquitecturas de sistemas escalables, la aplicación de técnicas avanzadas de análisis y visualización de datos como también la simulación mediante computadoras. Asimismo, se incorporan métodos de aprendizaje automático y minería de datos con el propósito de extraer conocimiento relevante a partir de grandes volúmenes de información heterogénea.
Este enfoque integral tiene como objetivo no solo mejorar la eficiencia del procesamiento de datos, sino también permitir la generación de modelos de simulación precisos y útiles que puedan ayudar en la toma de decisiones en una amplia gama de aplicaciones y dominios.

Palabras clave: *Procesamiento eficiente, análisis de datos, simulación, ciencia de datos, visualización.*

1. CONTEXTO

En este artículo se presenta un trabajo articulado entre las líneas de investigación "Modelado y simulación de sistemas a gran escala" y "Análisis, procesamiento y visualización eficiente de grandes volúmenes de datos", desarrolladas en el marco del Proyecto Consolidado (PROICO) de la Universidad Nacional de San Luis y de Incentivos: "Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos". El mismo tiene como objetivo central el diseño e investigación de nuevas metodologías y arquitecturas orientadas al procesamiento eficiente, escalable y efectivo de datos masivos, así como la generación de tecnologías que permitan transformar datos en información estratégica para diversos dominios de aplicación.
Asimismo, se propone la continuidad y el fortalecimiento de la colaboración con investigadores con quienes el grupo mantiene un contacto fluido, pertenecientes a diversas instituciones académicas y científicas. Entre estas, se destacan la Universidad Autónoma de Barcelona (UAB), el Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) de la UNS-CIC, y el Instituto de Investigación en Informática LIDI (III-LIDI-UNLP).
Finalmente, es relevante mencionar el soporte recibido para el desarrollo de una de las líneas presentadas en este artículo, del convenio Marco de colaboración establecido entre el Ministerio de Ciencia e Innovación del Gobierno de la Provincia de San Luis, la Red de Estaciones Meteorológicas (REM) y la Facultad de Ciencias Físico Matemáticas y Naturales de la UNSL. Este convenio tuvo por objetivo la cooperación técnica y la

coordinación del trabajo conjunto para desarrollar una solución para el tratamiento de datos climatológicos recopilados por la red de estaciones meteorológicas (REM) combinando tecnologías emergentes de ciencias de datos e inteligencia artificial (IA).

2. INVESTIGACIÓN Y DESARROLLO

En la era digital actual, la proliferación de datos ha dado lugar a la necesidad de herramientas avanzadas y metodologías eficaces para su análisis y procesamiento. El Big Data y la Ciencia de Datos emergen como campos esenciales que permiten la extracción de conocimientos y patrones a partir de grandes volúmenes de información. Por su parte, la simulación incorpora capacidad experimental a la Ciencia de Datos, permitiendo generar escenarios sintéticos mediante el uso de gemelos digitales, evaluar políticas bajo condiciones controladas y validar modelos mediante la experimentación computacional sobre sistemas complejos. Esto es de gran utilidad en contextos donde los datos observacionales resultan de poca calidad, insuficientes o incompletos.
Estos enfoques han transformado diversos sectores al ofrecer soluciones innovadoras y eficientes para la toma de decisiones y la gestión de situaciones críticas.
Por lo tanto, integrar las áreas mencionadas anteriormente no solo refuerza la capacidad para analizar y comprender datos complejos, sino que también facilita la identificación de patrones y tendencias relevantes que pueden influir en la acción social y la toma de decisiones informadas, como así también, permite ampliar el alcance analítico.

2.1. LÍNEAS DE TRABAJO

ARQUITECTURAS PARALELAS PARA EL PROCESAMIENTO DE DOCUMENTOS A GRAN ESCALA

La transformación digital de las empresas ha generado un volumen de documentos PDF que requieren procesamiento automatizado. Los bufetes de abogados analizan contratos para la debida diligencia, las instituciones sanitarias digitalizan los historiales clínicos de los pacientes, y las instituciones financieras extraen datos de los documentos regulatorios. El objetivo común es transformar imágenes estáticas de documentos en bases de conocimiento con capacidad de búsqueda, capaces de alimentar sistemas de Recuperación-Generación Aumentada (RAG) [LEW20] para la búsqueda semántica y la respuesta a preguntas impulsada por IA.
Por lo tanto, varias organizaciones utilizan servicios comerciales de OCR como Google Vision API, AWS Textract y Azure Cognitive Services. Sin embargo, dichos servicios presentan las siguientes desventajas: cobran por página, imponen límites de velocidad y generan dependencia del proveedor. Además, el procesamiento secuencial en hardware estándar requiere horas para corpus de tamaño moderado, y las API comerciales, a pesar de su coste, no pueden superar las limitaciones de velocidad que hacen que los proyectos de millones de documentos sean imprácticos.
Por lo tanto, esta línea de investigación aborda el procesamiento de OCR a gran escala, utilizando hardware disponible en las universidades y software de código abierto, eliminando así la dependencia de costosos servicios en la nube. Se propone diseñar e implementar arquitecturas paralelas para el procesamiento masivo de documentos PDF: (1) un pipeline distribuido basada en Ray con capacidades RAG integrada con tolerancia a fallos; (2) una arquitectura multiproceso local que utiliza ProcessPoolExecutor; y (3) un diseño híbrido que combina la orquestación de Ray con trabajadores locales optimizados.

GEMELOS DIGITALES PARA LA PREVENCIÓN DE INCENDIOS FORESTALES

En los últimos años, los gemelos digitales han emergido como herramientas clave para la representación virtual dinámica de sistemas

físicos complejos [JONES20]. Su valor radica en la capacidad de integrar datos en tiempo real, simulaciones avanzadas y análisis predictivo en un entorno digital que refleja, anticipa y retroalimenta el comportamiento del sistema real. En el contexto de esta línea, el gemelo digital se concibe como una plataforma inteligente que no sólo replica las condiciones y evolución de un territorio vulnerable a incendios forestales, sino que también permite simular escenarios, analizar respuestas institucionales y sociales, y generar recomendaciones estratégicas. La integración de simulación basada en agentes (ABS) y modelos de lenguaje de gran escala (LLM) en este tipo de sistema representa un enfoque novedoso [ZHENG22], con potencial transformador en la gestión preventiva de desastres.

A medida que avanza la ejecución de una simulación basada en agentes, la dinámica del sistema se vuelve cada vez más compleja e incluso impredecible, debido a las múltiples interacciones no lineales entre los agentes y su entorno. Comprender esta evolución en tiempo de ejecución es fundamental para identificar patrones emergentes, anticipar puntos críticos y reconocer los momentos clave en los que deben tomarse decisiones estratégicas. No obstante, interpretar dicha complejidad representa un desafío técnico y cognitivo significativo, especialmente cuando se requiere comunicar los hallazgos de manera accesible para actores de diversa formación.

La Figura 1 muestra los componentes de la arquitectura que se propone para diseñar un gemelo digital que integra un motor de simulación basado en agentes asistido por LLM, orientado a la prevención de desastres.

En este escenario, los LLM ofrecen una herramienta prometedora para asistir en la interpretación de la dinámica de las simulaciones [SUN24].

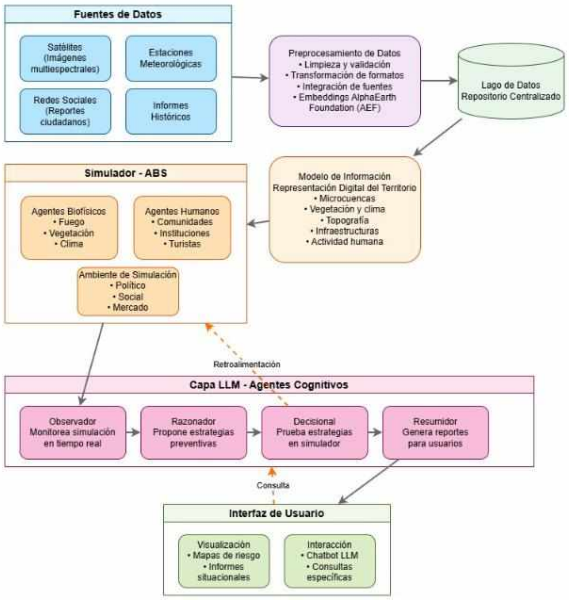


Figura 1: Arquitectura del gemelo digital.

SIMULACIÓN EFICIENTE PARA ATENCIÓN PRIMARIA DE SALUD

Los sistemas de Atención Primaria de Salud (APS) se enfrentan a una presión insostenible debido al envejecimiento poblacional y la carga de enfermedades crónicas, lo que genera ineficiencias críticas en el flujo de trabajo y afecta la atención al paciente. Este desafío afecta particularmente al seguimiento de pacientes crónicos, donde los gestores de casos suelen verse abrumados por la excesiva carga de trabajo y las complejidades de la coordinación de la atención.

En este contexto, la simulación de eventos discretos (DES) es esencial para evaluar el impacto de las nuevas políticas sanitarias antes de su implementación. Sin embargo, simular estos procesos a escala conlleva altos costos computacionales, lo que limita la viabilidad de los enfoques estándar de DES para explorar escenarios de sistemas de APS a gran escala [ADAM23], [HOU23].

Para abordar este desafío, proponemos utilizar un marco de Simulación Paralela de Eventos Discretos (PDES) diseñado específicamente para sistemas de APS. Nuestra propuesta implementa un protocolo aproximado de simulación paralela basado en el modelo Bulk

Synchronous Parallel (BSP), que mejora la eficiencia al simplificar la ordenación de eventos y eliminar la reversión.
El estudio se enfoca en la asignación de turnos y la gestión de tareas realizadas por los profesionales de la salud, quienes no sólo coordinan la atención de los pacientes, sino que también supervisan el seguimiento de sus tratamientos. Además, para abordar el problema del balance de carga en la simulación paralela, se implementan estrategias que optimicen la distribución del trabajo entre los procesadores.

SIMULACIÓN DE EVACUACIONES EN AMBIENTES CERRADOS

El cloud computing ha transformado la operación y el desarrollo de aplicaciones sociales, proporcionando a los usuarios acceso a servicios y plataformas desde cualquier ubicación y dispositivo con conexión a Internet. Esta tecnología no solo facilita el almacenamiento y el procesamiento de grandes volúmenes de datos, sino que también permite que las aplicaciones ofrezcan experiencias personalizadas y en tiempo real. La aplicación de servicios en la nube a la simulación de evacuaciones pedestres representa un caso de estudio significativo en este ámbito. Partiendo de un simulador de evacuaciones basados en agentes (ABS) y con el objetivo de optimizar el proceso, el equipo de investigación se ha enfocado en varios objetivos claves: a) Paralelización e implementación en la Nube: Investigar e implementar diversas metodologías de simulación en entornos cloud con el fin de reducir los costos computacionales asociados. b) Desarrollo de un Framework SaaS: Crear e implementar un nuevo marco de simulación tipo cloud SaaS, fundamentado en la interacción de autómatas celulares y agentes inteligentes, que permita instanciar múltiples realidades y escenarios para su análisis a través de experimentación simulada. c) Visualización de Datos: Investigar y aplicar técnicas y herramientas de visualización de datos que faciliten el análisis de la información obtenida a partir de las simulaciones. Este último objetivo es transversal a todas las líneas de trabajo del proyecto y es esencial para la identificación y análisis de patrones emergentes derivados de la simulación. De esta manera, se busca no solo elevar la eficacia de las simulaciones, sino también contribuir significativamente a la planificación y gestión de evacuaciones en ambientes cerrados.

CIENCIA DE DATOS PARA EL ANÁLISIS DE DATOS METEOROLÓGICOS

En la provincia de San Luis, la existencia de una extensa Red de Estaciones Meteorológicas (REM), que recopila datos en tiempo real sobre variables como temperatura, precipitación, viento y humedad, constituye un recurso fundamental para el análisis ambiental. Sin embargo, el volumen de datos recolectados (en un periodo de 15 años) requieren herramientas avanzadas para procesarlos, analizarlos, visualizarlos y obtener conocimiento de los mismos.



Figura 2: Plataforma de datos meteorológicos

Con el objetivo de aprovechar esta valiosa fuente de información, en la Figura 2 se muestra la plataforma tecnológica desarrollada la cual integra técnicas de procesamiento, limpieza y visualización de datos.
Esta plataforma permite analizar datos históricos, detectar patrones relevantes históricos, analizar tendencias, y servirá de base para proponer estrategias futuras de prevención y alertas ante eventos extremos como heladas, incendios, inundaciones u olas de calor.

3. RESULTADOS ESPERADOS

Las líneas de investigación y las técnicas descritas en la sección anterior involucran una serie de desarrollos individuales que en su conjunto logran obtener el objetivo planteado. Este objetivo contempla por un lado el modelado y diseño de sistemas complejos mediante diferentes herramientas, algunas de las cuales han sido desarrolladas por los integrantes del proyecto. Por otro lado, se propone utilizar técnicas de ciencias de datos para analizar los datos obtenidos y encontrar comportamientos o tendencias que puedan ser utilizadas en diferentes situaciones que requieran soporte a la toma de decisiones. Finalmente, el procesamiento de este gran volumen de datos debe realizarse en forma eficiente y en el menor tiempo posible. Esto implica investigar formas de despliegue de los algoritmos en diferentes plataformas y tecnologías que permitan lograr el objetivo deseado.

En el próximo año se espera poder avanzar en formular nuevas metodologías y técnicas para la adquisición, tratamiento y análisis de datos masivos, investigar el uso de simuladores eficientes y eficaces para evaluar el rendimiento de nuevos algoritmos para aplicaciones de sistemas de gran escala con datos masivos, profundizar en el uso de herramientas IA, abordar problemas reales que involucren datos reales, desarrollando soluciones eficientes para el análisis de grandes volúmenes de datos, continuar formando estudiantes de posgrado en la temática y capacitar profesionales en la temática mediante talleres y cursos de posgrado.

4. FORMACION DE RECURSOS HUMANOS

Actualmente, el Proyecto de Investigación cuenta con la supervisión de tres doctores en ciencias de la computación, quienes están asignados a las investigaciones de las líneas descritas en este artículo.

Además, tres investigadores están finalizando su formación de posgrado en las siguientes carreras: Maestría en Sistemas Embebidos, Maestría en Inteligencia de Datos Orientada a Big Data utilizando técnicas avanzadas de visualización y Maestría en Ciencia de Datos e Innovación Empresarial.

Por último, se destaca la formación de 4 estudiantes a nivel de grado y posgrado, realizando sus trabajos finales en las temáticas presentadas en este artículo.

5. BIBLIOGRAFÍA

[ADAM23] Adam C and Arduin H (2023) Agent-based epidemics simulation to compare and explain screening and vaccination prioritization strategies. SIMULATION 100(4): 335–355. DOI:10.1177/0037549723117833.

[HOW23] Hou X, Sharma J and Zhao F (2023) Silver Opportunity: Building Integrated Services for Older Adults around Primary Health Care. Washington, DC:World Bank.

[JONES20] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, “Characterising the digital twin: A systematic literature review,” CIRP journal of manufacturing science and technology, vol. 29, pp. 36–52, 2020.

[LEW20] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., Lewis, M., Yih, W.-t., Rockt¨aschel, T., et al.: Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in Neural Information Processing Systems 33, 9459–9474 (2020)

[SUN24] Y. Sun, Q. Zhang, J. Bao, Y. Lu, and S. Liu, “Empowering digital twins with large language models for global temporal feature learning,” Journal of Manufacturing Systems, vol. 74,pp. 83–99, 2024.

[ZHENG22] X. Zheng, J. Lu, and D. Kiritsis, “The emergence of cognitive digital twin: vision, challenges and opportunities,” International Journal of Production Research, vol. 60, no. 24, pp. 7610–7632, 2022.

RENDIMIENTO, RESILIENCIA Y SOLUCIONES INTELIGENTES EN APLICACIONES DE CÓMPUTO PARALELO Y DISTRIBUIDO

Javier Balladini[1], Claudio Zanellato[1], Remo Suppi[2], Marina Morán[1], Guillermo Torres[2], Rodrigo Cañibano[1], Conchillo Mariano[1], Héctor Mella[1], Julián Prandi[1], Aldana Vega[1], Candelaria Álvarez[2], Enzo Rucci[3], Armando De Giusti[3], Dolores Rexachs[2], Emilio Luque[2]

[1] Facultad de Informática, Universidad Nacional del Comahue

{javier.balladini, claudio.zanellato, marina, rcanibano, mariano.conchillo}@fi.uncoma.edu.ar

{julian.prandi, aldana.vega}@est.fi.uncoma.edu.ar

hectorjaviermella@gmail.com

[2] Universidad Autónoma de Barcelona

{remo.suppi, guillermo.torres, anacandelaria.alvarez, dolores.rexachs, emilio.luque}@uab.cat

[3] Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata

{erucci, degiusti}@lidi.info.unlp.edu.ar

RESUMEN

Esta línea de investigación se enfoca en el desarrollo de modelos, arquitecturas y herramientas para mejorar la eficiencia, la tolerancia a fallos y la adaptabilidad de sistemas de cómputo de altas prestaciones (HPC) y entornos *cloud–fog–edge*, con especial énfasis en aplicaciones críticas del sector salud.

Uno de los ejes aborda los sistemas inteligentes para la salud. Se desarrollan plataformas para detección temprana del deterioro clínico mediante Sistemas de Alerta Temprana, implementadas sobre arquitecturas distribuidas resilientes, priorizando baja latencia, continuidad operativa y eficiencia energética. Además, se estudian modelos trazables basados en datos para la progresión de enfermedad renal crónica, y técnicas de análisis automatizado de imágenes para diagnóstico temprano de cáncer de pulmón.

Otro eje aborda la ejecución eficiente y resiliente en sistemas paralelos y distribuidos. Se estudia el modelado del tiempo de ejecución y del consumo energético en HPC, la optimización de estrategias de *checkpoint/restart*, la migración y partición de servicios entre *cloud*, *fog* y *edge*, y el monitoreo inteligente en tiempo real. El enfoque integra modelado teórico, desarrollo de prototipos y validación experimental en entornos reales y controlados, promoviendo transferencia tecnológica y formación de recursos humanos de grado y posgrado.

Palabras claves: *cómputo paralelo y distribuido; HPC; sistemas resilientes; cloud–fog–edge; eficiencia energética; inteligencia artificial en salud; sistemas de alerta temprana; diagnóstico asistido por IA.*

CONTEXTO

La línea de I/D presentada se inserta en el proyecto “Rendimiento, resiliencia y soluciones inteligentes en aplicaciones de cómputo paralelo y distribuido” acreditado por la Universidad Nacional del Comahue (UNCo), a través de su Secretaría de Ciencia y Técnica, con inicio en 2026 y finalización en 2029.

El proyecto no dispone de financiamiento externo específico y se ejecuta con recursos institucionales, equipamiento informático disponible para actividades docentes y de investigación, y equipamiento adquirido en proyectos previos.

Se mantienen colaboraciones académicas con la Universidad Nacional de La Plata (UNLP) y la Universitat Autònoma de Barcelona (UAB), particularmente en temas de cómputo de altas prestaciones (HPC) y resiliencia en sistemas distribuidos. Estas colaboraciones permiten acceso a infraestructuras remotas de cómputo paralelo y fortalecen la formación de posgrado.

Se participa en el Consorcio de I+D+I en Cloud Computing, Big Data y Emergent Topics, conformado por Universidades de Argentina y España.

El proyecto tiene previsto transferir resultados a instituciones públicas de salud del Alto Valle de Río Negro y Neuquén, contribuyendo al sector socioeconómico regional.

1. INTRODUCCIÓN

El crecimiento exponencial de datos, la proliferación de dispositivos conectados y la demanda de servicios en tiempo real han consolidado al cómputo paralelo y distribuido como infraestructura esencial para aplicaciones críticas. En este contexto, la integración de técnicas avanzadas de modelado, optimización e Inteligencia Artificial (IA) ha dado lugar a lo que se denomina computación inteligente [1–3].

En el ámbito de la computación de altas prestaciones (HPC), el incremento sostenido en la escala de los sistemas ha puesto en primer plano dos desafíos estructurales: el consumo energético y la tolerancia a fallos [4, 5]. A medida que aumenta el número de nodos, crece la probabilidad de fallos y se incrementa la potencia total requerida. En este contexto, la predicción del tiempo de ejecución, el modelado del consumo energético y la optimización de estrategias de *checkpoint/restart* resultan líneas de investigación estratégicas [6].

En paralelo, los sistemas distribuidos han evolucionado hacia arquitecturas jerárquicas basadas en el paradigma *cloud–fog–edge* [7, 8]. Este modelo permite distribuir la carga computacional entre el nivel de *cloud*, *fog* y *edge*, reduciendo latencias y mejorando la eficiencia energética. Desde el punto de vista arquitectónico, la resiliencia en estos entornos implica no solo tolerancia a fallos, sino también capacidad de adaptación ante perturbaciones como cambios de carga, interrupciones de conectividad o ciberataques, manteniendo la funcionalidad básica del sistema [9, 10].

El sector salud puede beneficiarse ampliamente de la IA [11]. Este es un dominio de aplicación crítico, donde el rendimiento y la resiliencia no son solo atributos técnicos sino requisitos fundamentales. La detección temprana del deterioro clínico mediante Sistemas de Alerta Temprana (EWS) y la vigilancia continua, permite reducir eventos adversos graves y optimizar recursos hospitalarios [12]. Asimismo, el modelado de progresión de enfermedades y el diagnóstico asistido por IA ofrecen oportunidades para mejorar el pronóstico y la calidad de atención, exigiendo infraestructuras robustas y eficientes.

En este escenario, la línea de I/D presentada integra investigación en rendimiento y resiliencia computacional con el desarrollo de

soluciones inteligentes aplicadas a la salud. El enfoque combina modelado teórico, desarrollo de prototipos y validación experimental en entornos reales y controlados.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La organización de la línea de I/D se organiza en los siguientes dos ejes.

2.1. Sistemas inteligentes en salud

a. Detección temprana del deterioro clínico

Se desarrollan plataformas para salas de internación general y unidades críticas que permiten la vigilancia continua y el cálculo automático de puntuaciones de riesgo basadas en EWS, clasificando pacientes en categorías de riesgo. Mediante el uso de dispositivos móviles, es posible emitir alertas tempranas al personal de salud. A su vez, el almacenamiento de datos históricos hace posible la mejora continua de los modelos predictivos.

Estas soluciones se implementan bajo arquitecturas *cloud–fog–edge* resilientes, minimizando latencias y dependencia de conectividad.

b. Modelo basado en datos para la progresión de la enfermedad renal crónica

Los modelos basados en datos pueden simular la progresión de enfermedades y generar sistemas de apoyo a la decisión clínica. Aunque los modelos de aprendizaje automático son muy utilizados, su falta de trazabilidad presenta desafíos en contextos médicos donde la transparencia es crítica. En este contexto, resulta prometedora la construcción de un modelo trazable basado en datos que no emplee aprendizaje automático.

c. Diagnóstico temprano de cáncer de pulmón

Se exploran técnicas de análisis automatizado de imágenes de tomografía computarizada, utilizando IA como herramienta de apoyo al diagnóstico, con el objetivo de mejorar la detección temprana y reducir barreras de acceso.

2.2. Ejecución eficiente y resiliente

a. Eficiencia y resiliencia en arquitecturas *cloud–fog–edge* para aplicaciones críticas de tiempo real

Se estudian estrategias de migración y partición de servicios entre niveles cloud, fog y edge, replicación para tolerancia a fallos, monitoreo del flujo de datos y detección de anomalías, y optimización energética en dispositivos de borde. El objetivo es garantizar disponibilidad y cumplimiento de restricciones temporales estrictas.

b. Otros temas de interés

Ejecución eficiente y resiliencia en HPC (modelado del tiempo de ejecución y potencia, análisis de oportunidades de ahorro energético en sistemas tolerantes a fallos) y el monitoreo inteligente en Unidades de Cuidados Intensivos (procesamiento paralelo y distribuido para análisis en tiempo real de datos fisiológicos).

3. RESULTADOS OBTENIDOS Y ESPERADOS

3.1 Resultados obtenidos

- Se propuso una herramienta para mejorar el rendimiento de la entrega de servicios de cuidados intensivos integrados [13, 14].
- Se analizaron desafíos tecnológicos y se propusieron soluciones para un sistema de detección del deterioro temprano y automático de pacientes en unidades de cuidados intensivos. [15, 16]
- Se desarrolló un mecanismo de inicio de sesión automatizado y seguro para un sistema resiliente, con arquitectura cloud-fog-edge, basado en dispositivos móviles para entornos hospitalarios [17].
- Se desarrolló un prototipo de una aplicación para tratar pacientes, implementando un EWS para COVID-19 [18, 19].
- Se desarrolló un modelo inicial dirigido por datos para la progresión de enfermedad renal crónica [20].
- Se implementó un algoritmo para el procesamiento eficiente de señal de electrocardiograma, para determinación de frecuencia cardíaca, en tiempo real [21].
- Se exploraron oportunidades de ahorro energético en sistemas de HPC [22] y sus mecanismos de tolerancia de fallos [23, 24].
- Se propuso un enfoque radiómico inteligente para la detección temprana del cáncer de pulmón [25].

3.2 Resultados esperados

- Consolidación de plataforma resiliente cloud-fog-edge para el sistema de detección temprana del deterioro clínico en salas de internación común. Validación de prototipos funcionales en instituciones de salud.
- Desarrollo de una solución para el diagnóstico temprano de cáncer de pulmón mediante análisis automatizado de tomografías computarizadas usando inteligencia artificial.
- Consolidación del modelo basado en datos para la progresión de la enfermedad renal crónica.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D el último año se concluyó 1 tesis de doctorado [26], y 1 tesis de grado. Al mismo tiempo se encuentran en curso 1 tesis de Doctorado en Ciencias Informáticas, y 2 tesis de grado.

La línea promueve articulación con programas de posgrado en cooperación con la UNLP y la UAB, fortaleciendo la formación en HPC, resiliencia y aplicaciones inteligentes para la salud.

5. BIBLIOGRAFÍA

[1] Fakhrul Safitra, M.; Prastika Putri, D.; Lubis, M.; Fakhrurroja, H.; Saidi Lubis, F.; Satria, V.: Intelligent Computing in Communication Networks: Challenges, Solutions, and Research Prospects. In: IEEE International Conference on Computing, ICOCO 2023. DOI: 10.1109/ICOCO59262.2023.10397643. EID: 2-s2.0-85184852200. Part of ISBN: 9798350302684

[2] Gozalo-Brizuela, R. & Merchan, E. E. G. (2024). A Survey of Generative AI Applications. Journal of Computer Science, 20(8), 801-818. https://doi.org/10.3844/jcssp.2024.801.818

[3] S. Aldoihi, K. Alblalaihid, F. Alzemaia and A. Almoajel, "The Role of System Modeling on Artificial Intelligence: A Review of Emerging Trends," 2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), Milano, Italy, 2023, pp. 161-165, doi: 10.1109/MetroXRAINE58569.2023.10405725.

[4] Navaux, P. O. A., Lorenzon, A. F., & da Silva Serpa, M. (2023). Challenges in high-performance computing. *Journal of the Brazilian Computer Society*, *29*(1), 51-62.

[5] Reed, D., Gannon, D., & Dongarra, J. (2022). Reinventing high performance computing: challenges and opportunities. *arXiv preprint arXiv:2203.02544*.

[6] Marina Morán, Javier Balladini, Dolores Rexachs, and Enzo Rucci. 2024. Exploring energy saving opportunities in fault tolerant HPC systems. J. Parallel Distrib. Comput. 185, C (Mar 2024). https://doi.org/10.1016/j.jpdc.2023.104797

[7] Flavio Bonomi, Rodolfo Milito, Jiang Zhu, and Sateesh Addepalli. 2012. Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing (MCC '12). Association for Computing Machinery, New York, NY, USA, 13–16. https://doi.org/10.1145/2342509.2342513

[8] Prokhorenko, Victor & Ali Babar, Muhammad. (2020). Architectural Resilience in Cloud, Fog and Edge Systems: A Survey. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2971007.

[9] J. Lofstead, J. Dayal, I. Jimenez and C. Maltzahn, "Efficient, Failure Resilient Transactions for Parallel and Distributed Computing," 2014 International Workshop on Data Intensive Scalable Computing Systems, New Orleans, LA, USA, 2014, pp. 17-24, doi: 10.1109/DISCS.2014.13.

[10] Amiri, Z., Heidari, A., Navimipour, N. J., & Unal, M. (2023). Resilient and dependability management in distributed environments: A systematic and comprehensive literature review. *Cluster Computing*, *26*(2), 1565-1600.

[11] Al Kuwaiti, A., Nazer, K., Al-Reedy, A., Al-Shehri, S., Al-Muhanna, A., Subbarayalu, A. V., Al Muhanna, D., & Al-Muhanna, F. A. (2023). A Review of the Role of Artificial Intelligence in Healthcare. *Journal of Personalized Medicine*, *13*(6), 951. https://doi.org/10.3390/jpm13060951

[12] Lee, J. R., Kim, E. M., Kim, S. A., & Oh, E. G.: A systematic review of early warning systems' effects on nurses' clinical performance and adverse events among deteriorating ward patients. Journal of patient safety, 16(3), e104-e113. LWW (2020).

[13] Balladini, J., Bruno, P., Zurita, R., Orlandi, C., Suppi, R., Rexachs, D. and Luque, E., 2019. A tool for improving the delivery of integrated intensive health care performance. International Journal of Integrated Care, 19(4), p.222. DOI: http://doi.org/10.5334/ijic.s3222. E-ISSN: 1568-4156. Indexada por: Scimago. JCR mejor cuartil Q1, factor de Impacto de 2,753. [2.b]

[14] Cañibano, R., Rozas, C., Orlandi, C., Balladini, J. (2020). Data Management Optimization in a Real-Time Big Data Analysis System for Intensive Care. In: Rucci, E., Naiouf, M., Chichizola, F., De Giusti, L. (eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2020. Communications in Computer and Information Science, vol 1291. Springer, Cham. https://doi.org/10.1007/978-3-030-61218-4_7

[15] Balladini J., Bruno P., Zurita R., & Orlandi C. (2018). An Automatic and early detection of the deterioration of patients in Intensive and Intermediate Care Units: technological challenges and solutions. Journal of Computer Science and Technology, 18(03), e25. https://doi.org/10.24215/16666038.18.e25

[16] Balladini, Javier, C. Rozas, F. E. Frati, N. Vicente and Cristina Orlandi. "Big Data Analytics in Intensive Care Units: challenges and applicability in an Argentinian Hospital." Journal of Computer Science and Technology 15 (2015): 61-67.

[17] Claudio Zanellato, Rodrigo Cañibano, Javier Balladini. Automated and Secure Login in ALERTAR, a Resilient Cloud–Fog–Edge mHealth System for Hospital Environments. XII Jornadas de Cloud Computing, Big Data & Emerging Topics. 2025.

[18] Javier Balladini. COVINDEX: una app para tratar pacientes con COVID-19. https://www.editores-srl.com.ar/autor/javier_balladini/20200925_covindex_una_app_para_tratar_pacientes_con_covid19 (Accedido el 14 de Junio de 2025).

[19] Latorre, M., Conchillo, M., Cañibano, R., & Balladini, J. (2024). Implementación del núcleo de la aplicación móvil de ALERTAR, un sistema para vigilancia de pacientes hospitalizados y alertas tempranas de gravedad. In XXX Congreso Argentino de Ciencias de la Computación (CACIC)(La Plata, 7 al 11 de octubre de 2024).

[20] Alvarez, C., Suppi Boldrito, R., Ibeas, J., & Balladini, J. (2024). Data-driven model for chronic kidney disease progression: a work in progress. In 38th annual European Simulation and Modelling Conference (pp. 110-113). EUROSIS-ETI Publication.

[21] María Belén Casanova: Procesamiento eficiente de señal de electrocardiograma, para determinación de frecuencia cardíaca, en tiempo real. Tesis de Licenciatura en Ciencias de la Computación, Universidad Nacional del Comahue. Junio 2023.

[22] Paniego J.M., Gallo S., Pi Puig M., Chichizola F., De Giusti L., Balladini J. (2018) Analysis of RAPL Energy Prediction Accuracy in a Matrix Multiplication Application on Shared Memory. In: De Giusti A. (eds) Computer Science – CACIC 2017. CACIC 2017. Communications in Computer and Information Science, vol 790. Springer, Cham. https://doi.org/10.1007/978-3-319-75214-3_4

[23] Marina Morán, Javier Balladini, Dolores Rexachs, and Enzo Rucci. 2024. Exploring energy saving opportunities in fault tolerant HPC systems. J. Parallel Distrib. Comput. 185, C (Mar 2024). https://doi.org/10.1016/j.jpdc.2023.104797. Indexada por: Scimago. JCR mejor cuartil Q1, factor de Impacto de 4.

[24] M. Morán, J. Balladini, D. Rexachs and E. Luque, "Prediction of Energy Consumption by Checkpoint/Restart in HPC", in IEEE Access, vol. 7, pp. 71791-71803, 2019, doi: 10.1109/ACCESS.2019.2919970. E-ISSN: 2169-3536. Indexada por: Scimago. JCR mejor cuartil Q1, factor de Impacto de 3,745.

[25] Torres, G., Baeza, S., Sanchez, C., Guasch, I., Rosell, A., & Gil, D. (2022). An Intelligent Radiomic Approach for Lung Cancer Screening. Applied Sciences, 12(3), 1568. https://doi.org/10.3390/app12031568

[26] Marina Morán, "Modelización y gestión del consumo energético en un sistema de computación de altas prestaciones con tolerancia a fallos" Tesis de Doctorado en Ciencias Informáticas, UNLP. https://sedici.unlp.edu.ar/handle/10915/180599

Seguridad Informática

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

Análisis de Seguridad en Dispositivos Biométricos de Control de Acceso: Caso de Estudio ZMM220

Mg. Ing. Rodrigo Atilio Elgueta[a], Dr. Ing. Miguel Méndez-Garabetti[a], Ing. Diego Germán Gomez[a]
[a]Univerisdad de Champagnat, Facultad de Informática y Diseño, General Juan Manuel Belgrano 721,CP: 5501 Godoy Cruz, Mendoza, Argentina
elguetarodrigo@uch.edu.ar, mendez-garabettimiguel@uch.edu.ar, gomezdiego@uch.edu.ar

RESUMEN

La autenticación biométrica se ha convertido en un estándar para el control de accesos en organizaciones. Sin embargo, la implementación práctica en dispositivos comerciales como el ZMM220 de ZKTeco presenta vulnerabilidades críticas que comprometen su efectividad. Este trabajo propone una metodología reproducible para el análisis de seguridad en dispositivos biométricos, validada mediante un caso de estudio en este equipo (kernel Linux 3.0.8 para arquitectura MIPS). Utilizando herramientas estandarizadas de pruebas de penetración (Nmap, Wireshark, Hydra), se identificaron fallos como credenciales predeterminadas (CWE-798), uso del protocolo Telnet que transmite datos en texto plano (RFC 854), y bases de datos SQLite sin cifrado que almacenan información biométrica sensible. Los resultados evidencian violaciones a principios fundamentales de seguridad (defensa en profundidad, mínimo privilegio) y permitieron demostrar la modificación de registros de asistencia sin dejar trazas de auditoría. Como contribución, se presentan recomendaciones de mitigación basadas en estándares internacionales (NIST, ISO/IEC), incluyendo la migración a SSH, el cifrado de bases de datos y la actualización del kernel. Este enfoque metodológico busca establecer un precedente para la evaluación sistemática de otros sistemas biométricos en entornos organizacionales.

Palabras clave: Seguridad biométrica, pruebas de penetración, IoT, credenciales predeterminadas, CWE-798.

CONTEXTO

Esta línea de I/D se desarrolla en el marco de la cátedra de "Seguridad de Sistemas", radicada en la Facultad de Informática y Diseño de la Universidad de Champagnat (UCH). El proyecto cuenta con apoyo institucional mediante la Dirección de Investigación de la UCH y busca generar vínculos de transferencia con el sector productivo y gubernamental. Para la validación de los hallazgos, se establecieron acuerdos de colaboración con organizaciones que utilizan estos dispositivos, permitiendo la realización de pruebas controladas en entornos reales bajo estrictos acuerdos de confidencialidad. El financiamiento proviene de becas internas de investigación y recursos propios de la universidad.

1. INTRODUCCIÓN

La masificación de los sistemas biométricos para control de acceso físico y lógico ha transformado la seguridad perimetral en organizaciones de todo tipo. Dispositivos como el ZMM220 de ZKTeco prometen un alto nivel de seguridad al basarse en características biológicas únicas e indelegables [2]. Sin embargo, la experiencia práctica y reportes de seguridad [3] demuestran que estos dispositivos embebidos pueden convertirse en el eslabón más débil de la cadena de seguridad si no se consideran aspectos fundamentales de su arquitectura y configuración.

A diferencia de los estudios que se centran únicamente en listar vulnerabilidades, esta investigación propone y valida una metodología reproducible para el análisis de seguridad en dispositivos biométricos. El objetivo no es solo encontrar fallas en un modelo específico, sino establecer un procedimiento sistemático que permita a investigadores y administradores evaluar la postura de seguridad de estos sistemas. El caso de estudio en el dispositivo ZMM220 (con kernel Linux 3.0.8 para MIPS) sirve como validación de la metodología propuesta.

El trabajo se estructuró en torno a la aplicación de técnicas de pentesting en sistemas embebidos [4][5], evaluando el cumplimiento de principios de seguridad clásicos [9] y el alineamiento con guías de referencia para IoT [6]. Las preguntas que guiaron esta investigación son: ¿Es posible establecer un procedimiento estándar para auditar la seguridad de un dispositivo biométrico comercial? ¿Qué desviaciones respecto a las buenas prácticas de seguridad son más frecuentes y críticas en estos sistemas?

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Diseño de Metodologías de Evaluación para IoT Biométrico

Este eje se centró en el desarrollo de un ciclo de auditoría de seguridad específicamente adaptado a las características particulares de los dispositivos embebidos utilizados para control de acceso biométrico. Se observó que estos ejecutan sistemas operativos recortados y utilizan protocolos de comunicación que en muchos casos son propietarios, lo que requirió enfoques de evaluación diferenciados. La metodología propuesta comprendió fases secuenciales e iterativas que fueron refinadas a lo largo de la investigación. Se comenzó con una fase de reconocimiento inicial para identificación de activos y superficies de exposición, donde se relevó información pública sobre cada dispositivo, sus características técnicas y posibles vulnerabilidades conocidas. Luego se procedió al escaneo técnico para detección de puertos y servicios activos, para ello se utilizaron herramientas como Nmap que permitieron identificar no solo la conectividad sino también versiones de software y sistemas operativos subyacentes. La fase siguiente fue la explotación controlada de vulnerabilidades en entornos de prueba y en infraestructuras productivas con las debidas autorizaciones y acuerdos de confidencialidad. Posteriormente se realizó el análisis de los datos extraídos para comprender el alcance real del compromiso, identificando qué información sensible había quedado expuesta y qué niveles de acceso se habían podido alcanzar. Finalmente se procedió a la generación de informes técnicos y

ejecutivos, diferenciando el nivel de detalle según el destinatario. Se puso especial énfasis en la reproducibilidad del proceso y en el mínimo impacto en la disponibilidad del servicio, para lo cual se desarrollaron protocolos de intervención que incluyeron la definición de ventanas de prueba en horarios de bajo impacto, la implementación de mecanismos de rollback y el monitoreo continuo durante todo el proceso de evaluación.

Análisis Forense en Sistemas Embebidos

En éste eje se abordó la investigación de técnicas especializadas para la extracción y análisis de evidencia digital en dispositivos biométricos. Se observó que éstos, utilizan sistemas de archivos embebidos diseñados para memorias flash con características diferentes a los sistemas de archivos tradicionales. Adicionalmente, se comprobó que emplean bases de datos livianas optimizadas para entornos con recursos limitados, siendo SQLite el caso más frecuente [7]. La investigación se orientó a comprender los mecanismos de almacenamiento persistente, identificando las rutas específicas donde residen los datos de interés forense y los formatos de archivo utilizados. Particularmente se detectó que al carecer de mecanismos de autenticación nativos y almacenarse sin cifrado en la mayoría de los dispositivos analizados, las bases de datos exponen toda la información sensible a cualquier actor con acceso al sistema de archivos.

Evaluación de Cumplimiento de Estándares y Principios de Seguridad

Este eje se propuso realizar un análisis crítico de dispositivos biométricos comerciales contrastándolos con marcos normativos establecidos y con principios fundamentales de seguridad de la información. Se tomó como referencia principios clásicos de diseño seguro como la defensa mediante implementación de múltiples capas de controles de seguridad superpuestos de manera que si una capa falla, las siguientes continúen protegiendo el activo subyacente [10]. Otro principio fundamental considerado fue el de mínimo privilegio, según el cual cualquier usuario, proceso o sistema debe operar únicamente con los permisos estrictamente necesarios para cumplir su función específica. Estos se complementaron con el análisis de alineamiento a estándares internacionales como la guía NISTIR 8259, que establece actividades de ciberseguridad fundamentales para fabricantes de dispositivos IoT, y la norma ISO/IEC 27001 [1], que define los requisitos para sistemas de gestión de seguridad de la información. El objetivo fue identificar y caracterizar las brechas existentes entre lo que la teoría y los estándares recomiendan como buenas prácticas, y lo que efectivamente implementan los fabricantes. Los hallazgos de este eje permitieron construir un mapa de madurez de seguridad que puede servir tanto a fabricantes para mejorar sus productos como a organizaciones usuarias para tomar decisiones de compra más informadas.

Desarrollo de Contramedidas y Guías de Hardening

A partir de los hallazgos obtenidos en los ejes anteriores, esta línea de trabajo se orientó a la investigación y propuesta de contramedidas efectivas que permitieran mitigar los riesgos identificados en dispositivos ya desplegados. El enfoque adoptado fue eminentemente práctico y

buscó generar soluciones que pudieran ser aplicadas en el corto plazo, sin requerir necesariamente la intervención del fabricante. Se desarrollaron configuraciones seguras que en conjunto constituyen procesos de hardening, incluyendo la desactivación de servicios innecesarios que amplían la superficie de ataque. También se trabajó en la implementación de mecanismos de autenticación robustos que reemplazaran las débiles credenciales predeterminadas, mediante la configuración de contraseñas fuertes, autenticación de dos factores cuando el dispositivo lo soporta, o integración con servidores de autenticación centralizada. Especial atención recibió el reemplazo de protocolos inseguros, particularmente la migración de Telnet [8] a SSH con autenticación basada en claves criptográficas ed25519. También se abordó el cifrado de bases de datos embebidas mediante herramientas como SQLCipher, que permiten proteger la información sensible almacenada sin requerir modificaciones estructurales en el funcionamiento del dispositivo. El objetivo final fue democratizar el acceso a prácticas de seguridad efectivas, reduciendo la dependencia de las organizaciones respecto de las actualizaciones que puedan o no proveer los fabricantes.

3. RESULTADOS OBTENIDOS

Se llegó a las siguientes conclusiones:

En relación a las credenciales predeterminadas, este estudio confirmó su presencia en el ZMM220 mediante Hydra (root:password), mientras que The Hacker News (2024) lo reportó genéricamente. Como contribución, se aportaron comandos de explotación específicos como telnet y find para localizar bases de datos. La solución propuesta consiste en un hardening inmediato que incluya el cambio de todas las contraseñas por defecto en la primera configuración y la implementación de políticas de contraseñas robustas según NIST SP 800-63B.

En cuanto a las bases de datos SQLite sin cifrado, este estudio logró explotarlas con sqlite3 para modificar registros, mientras que The Hacker News solo lo mencionó sin detalles técnicos. La contribución principal fue una prueba de concepto con consultas SQL que permitió modificar registros de asistencia. Como solución se propone implementar SQLCipher o similar para cifrar la base de datos en reposo y aplicar control de acceso a nivel de archivos.

Respecto al kernel 3.0.8 para arquitectura MIPS, este estudio identificó su vulnerabilidad a DirtyCow (CVE-2016-5195), un aspecto no analizado por The Hacker News. La contribución radica en evidenciar este riesgo en arquitectura MIPS, no documentado previamente. La solución propuesta implica gestionar con el fabricante la actualización del kernel y firmware, así como aplicar parches de seguridad o mitigaciones como deshabilitar módulos vulnerables mediante modprobe.

Finalmente, sobre el uso de Telnet en texto plano, este estudio capturó tráfico con Wireshark utilizando filtros específicos que revelaron credenciales expuestas, mientras que The Hacker News lo citó como riesgo general. La contribución

consiste en aportar trazas de paquetes con credenciales expuestas como evidencia empírica. La solución propuesta es el reemplazo del protocolo deshabilitando Telnet e implementando SSH con autenticación basada en claves ed25519, configurado para la versión 2 del protocolo.

Actualmente, las organizaciones que pusieron a disposición de este estudio sus dispositivos, están siendo asistidas y acompañadas en el parcheo de las vulnerabilidades detectadas.

A mediano plazo, se espera completar el desarrollo de una guía de hardening para el ZMM220 que incluya la migración a SSH, el cifrado de la base de datos con SQLCipher y la actualización del firmware. y se espera aplicar la misma metodología a otros modelos de dispositivos biométricos para establecer un análisis comparativo de seguridad en el mercado local.

4. FORMACION DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada a la Dirección de Investigación de la UCH y desarrollada por el equipo de trabajo integrado por los siguientes investigadores: Dr. Ing. Miguel Méndez-Garabetti, Mg. Ing. Rodrigo Atilio Elgueta e Ing. Diego Germán Gomez.

5. BIBLIOGRAFÍA

[1] International Organization for Standardization. (2013). *ISO/IEC 27001:2013*.

[2] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition.

[3] Lakshmanan, R. (2024). ZKTeco Biometric System Found Vulnerable to Critical Flaws. thehackernews.com.

[4] Lyon, G. (2008). Nmap Network Scanning.

[5] MITRE. (2023). *CWE-798: Use of Hard-coded Credentials*.

[6] National Institute of Standards and Technology (NIST). (2020). NISTIR 8259.

[7] Owens, M., & Owens, M. (2006). The Definitive Guide to SQLite.

[8] Postel, J., & Reynolds, J. (1983). RFC 854: Telnet Protocol.

[9] Saltzer, J. H., & Schroeder, M. D. (1975). The protection of information in computer systems.

[10] Shostack, A. (2014). Threat Modeling: Designing for Security.

Análisis y detección de hábitos digitales orientados a la mitigación de fugas de datos personales

Parisi Germán Nicolás, Notreni Juliana María, Ruhl Analia Lorena, Auquer Marcelo Javier, Zea Cárdenas Milagros Ninfa, Bertola Federico Javier, Ciceri Leonardo Ramón, Barrionuevo Ileana Maricel, Sanchez Ignacio, Gibellini Fabián Alejandro.

GISSIC - Grupo de Investigación en Seguridad de Sistemas de Información y Ciberseguridad. / Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas / Universidad Tecnológica Nacional / Facultad Regional Córdoba
Cruz Roja S/N, 5016.

germannparisi@gmail.com, julinotreni@gmail.com, analialorenaruhl@gmail.com, marcelo.auquer@gmail.com, milyzc@gmail.com, fedebertola@gmail.com, leonardorciceri@gmail.com, ilebarrionuevo@gmail.com, ignaciojsb@gmail.com, fabiangibellini@gmail.com.

Resumen

La Protección de Información Personal Identificable (Personally Identifiable Information, PII) constituye uno de los principales desafíos actuales en materia de ciberseguridad y privacidad de datos [1]. Diversos incidentes de alto impacto, tanto a nivel internacional como nacional, evidencian que las filtraciones de datos continúan incrementándose en frecuencia, volumen y costo económico [2][3].

Estudios recientes señalan que el factor humano y los errores no intencionales de los usuarios representan una de las principales causas de estas brechas de seguridad [4][5]. En este contexto, el presente trabajo propone un enfoque de investigación orientado a la detección temprana de hábitos de uso de tecnologías digitales que atentan contra la ciberseguridad de los usuarios y facilitan la fuga de PII.

El objetivo principal es analizar, clasificar y ponderar dichos hábitos, así como estudiar métodos y técnicas que permitan identificarlos y advertir al usuario final en tiempo oportuno. Como resultado esperado, se plantea el desarrollo de un software como prueba de concepto que implemente mecanismos de detección y concientización, contribuyendo a la reducción del riesgo de exposición de datos sensibles y al fortalecimiento de la cultura de la seguridad de la información.

Palabras Clave

PII, ciberseguridad, protección de datos, factor humano, hábitos digitales, Argentina.

Contexto

En un escenario global caracterizado por una acelerada transformación digital, la ciberseguridad se ha consolidado como una de las principales preocupaciones económicas, sociales y tecnológicas, tanto para organizaciones públicas y privadas como para los usuarios finales. El crecimiento sostenido del uso de plataformas digitales, servicios en la nube, aplicaciones móviles y dispositivos conectados ha incrementado de manera significativa la superficie de exposición a incidentes de seguridad.

En este contexto, la PII se configura como un activo crítico. La exposición,

filtración o uso indebido de este tipo de información puede derivar en consecuencias legales, económicas y reputacionales de alto impacto, tales como fraudes, suplantación de identidad o vulneraciones a la privacidad de las personas [1][6].

A nivel normativo, distintos marcos regulatorios buscan mitigar estos riesgos mediante la protección de los datos personales. En Argentina, la Ley N.º 25.326 establece los principios fundamentales para la recolección, almacenamiento y uso de datos personales [7], mientras que en la Unión Europea el Reglamento General de Protección de Datos (RGPD) amplía el alcance de la protección incorporando criterios contextuales para la identificación de personas [6][8]. Legislaciones similares existen en otros países, como la Ley Federal N.º 152-FZ de la Federación Rusa, que regula de forma integral la protección de datos personales [9].

Sin embargo, la existencia de estos marcos regulatorios no ha sido suficiente para frenar el incremento de incidentes de seguridad. Casos recientes como la filtración de National Public Data, la brecha de seguridad de Ticketmaster o el incidente ocurrido en el Registro Nacional de las Personas (RENAPER) en Argentina evidencian la magnitud del problema y sus consecuencias sociales, legales y económicas [2][3][10]. Según el informe Cost of a Data Breach Report 2024 de IBM, el costo promedio de una filtración de datos alcanzó los 4,88 millones de dólares, representando el mayor incremento registrado en los últimos años [11].

Diversos estudios coinciden en que el factor humano constituye uno de los principales vectores de riesgo en materia de ciberseguridad. El comportamiento impredecible de los usuarios, sumado a hábitos digitales inseguros, limita la efectividad de los controles técnicos tradicionales y amplía la superficie de ataque [4][5][12][13]. Estas prácticas incluyen, entre otras, la reutilización de contraseñas, el uso de redes no seguras, la instalación de software no confiable y la exposición involuntaria de información sensible [14][15][16].

En este marco, resulta necesario abordar la protección de la PII desde una perspectiva integral que contemple no sólo aspectos tecnológicos y normativos, sino también el comportamiento humano. La identificación temprana de prácticas de riesgo y la generación de mecanismos de advertencia y concientización emergen como estrategias relevantes para reducir la probabilidad de fuga de datos y fortalecer la cultura de la seguridad de la información [17][18].

Introducción

La PII constituye un concepto central en la regulación de la privacidad de la información. De acuerdo con Schwartz y Solove, el alcance de las leyes de privacidad depende directamente de la existencia de PII involucrada; sin embargo, no existe una definición uniforme de este concepto a nivel legal o técnico [1].

Organismos internacionales como el National Institute of Standards and Technology (NIST) definen la PII como cualquier información que permita identificar, distinguir o rastrear la identidad de un individuo, de manera directa o indirecta [19]. No obstante, la delimitación entre PII y otros tipos de información no siempre resulta clara, ya que requiere una evaluación contextual del riesgo de identificación, tal como lo destaca la Administración de Servicios Generales de los Estados Unidos [20].

A pesar de los marcos regulatorios vigentes, las filtraciones de datos continúan en aumento. Los ataques dirigidos a la obtención de PII representan uno de los principales objetivos de los ciberdelincuentes, debido al valor económico y estratégico de esta información [2][3][10]. Este escenario evidencia la necesidad de complementar las medidas normativas con estrategias preventivas que aborden las causas subyacentes de estos incidentes.

Numerosas investigaciones identifican al factor humano como el eslabón más débil de la cadena de seguridad. El error humano explica un porcentaje significativo de las brechas de seguridad, superando incluso a las fallas técnicas [4][5]. Kobis señala que el comportamiento de los usuarios constituye uno de los principales factores que facilitan la infiltración de información confidencial, ya sea mediante la apertura de correos maliciosos, la instalación de software no verificado o la gestión inadecuada de credenciales [12][14].

Líneas de Investigación y Desarrollo

La presente investigación se sustenta en conocimientos actuales del campo de la ciberseguridad, con un enfoque particular en la protección de la PII y en el análisis del factor humano como componente crítico en la ocurrencia de incidentes de seguridad, formando parte de un proyecto denominado "Detector de hábitos que atentan a la ciberseguridad de usuarios para la minimización de fuga de datos", homologado por la Secretaría de Ciencia y Tecnología bajo el código COSIPP1000. La línea principal de investigación se inscribe en el área de seguridad informática, integrando de manera transversal aportes del análisis del comportamiento de los usuarios y la concientización en seguridad.

El enfoque metodológico adoptado es empírico, basado en el método científico, orientado a la observación, identificación y análisis de hábitos digitales que puedan derivar en la exposición o fuga de PII. El proyecto se desarrolla bajo los principios de Investigación, Desarrollo e Innovación (I+D+i), estructurándose en etapas que incluyen el relevamiento del estado del arte, la clasificación de hábitos de riesgo, el desarrollo de un software como prueba de concepto y la evaluación cualitativa y exploratoria de su impacto potencial.

Resultados y Objetivos

Objetivo general

Investigar, analizar y desarrollar mecanismos orientados a la detección temprana de hábitos de uso de tecnologías digitales que representen riesgos para la ciberseguridad de los usuarios y favorezcan la exposición de PII, desde una perspectiva centrada en el factor humano.

Objetivos específicos

- Identificar y caracterizar hábitos digitales que impliquen riesgos para la protección de la PII.
- Analizar enfoques y metodologías existentes para la detección de comportamientos inseguros.
- Definir criterios de evaluación del nivel de riesgo asociado a los hábitos detectados.
- Diseñar y desarrollar un software como prueba de concepto orientado a la detección y concientización.
- Evaluar de forma cualitativa y cuantitativa la eficacia del software desarrollado.

Resultados esperados

Se espera obtener una clasificación sistematizada de hábitos digitales inseguros vinculados a la exposición de PII, así como el desarrollo de un software funcional que permita detectar tempranamente prácticas de riesgo. Estos resultados permitirán validar enfoques preventivos centrados en el usuario y aportar conocimiento relevante al área de la ciberseguridad aplicada al factor humano.

Formación de Recursos Humanos

El proyecto involucra docentes e investigadores de la Universidad Tecnológica Nacional – Facultad Regional Córdoba, integrantes del Grupo de Investigación en Seguridad de Sistemas de Información y Ciberseguridad (GISSIC). Asimismo, participan alumnos becarios de la carrera de Ingeniería en Sistemas de Información, quienes inician su formación en investigación científica y tecnológica, profundizando en temáticas vinculadas a la seguridad de la información y al desarrollo de software.

Como parte de las actividades de difusión, se prevé la realización de charlas, talleres y conferencias a nivel provincial y nacional.

Referencias

[1] Schwartz, P., & Solove, D. (2011). The PII Problem: Privacy and a New Concept of Personally Identifiable Information. New York University Law Review, 86.

[2] Keepnet Labs. (2024). Top 10 data breaches of 2024. https://keepnetlabs.com/blog/top-10-data-breaches-of-2024-and-their-financial-impacts

[3] Framework Security. (2024). Ticketmaster breach. https://www.frameworksec.com/post/ticketmaster-breach-a-deep-dive-into-the-may-2024-cyberattack-and-the-history-of-the-alleged-hackers

[4] Enzoic. (2024). Why data breaches keep happening. https://www.enzoic.com/blog/why-data-breaches-keep-happening

[5] Acar, Y., et al. (2016). The impact of information sources on code security. IEEE Symposium on Security and Privacy.

[6] Redscan. (s.f.). What is personal data and why is it so important? https://www.redscan.com/news/personal-data-important-keep-safe/

[7] Argentina. Ley N.º 25.326 de Protección de Datos Personales. https://www.argentina.gob.ar/justicia/derechofacil/leysimple/datos-personales

[8] Unión Europea. Reglamento General de Protección de Datos (RGPD). https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_es.htm

[9] Federación Rusa. Ley Federal N.º 152-FZ sobre datos personales. https://ciberseguridad.com/normativa/rusia/#Ley_Federal_No_152-FZ_sobre_datos_personales

[10] Marina, R. (2024). Filtración de datos personales en el RENAPER. Chequeado. https://chequeado.com/el-explicador/filtracion-de-datos-personales-en-el-renaper-que-es-y-que-consecuencias-puede-tener/

[11] IBM. (2024). Cost of a Data Breach Report 2024. https://www.ibm.com/downloads/documents/us-en/107a02e94948f4ec

[12] Kobis, P. (2021). Human factor aspects in information security management. Operations Research and Decisions, 31(1).

[13] Maalem Lahcen, R. A., et al. (2020). Behavioral aspects of cybersecurity. Cybersecurity, 3(10).

[14] Richardson, M. D., et al. (2020). Planning for Cyber Security in Schools: The Human Factor. Educational Planning, 27(2).

[15] Check Point / World Economic Forum. (2023). Internet habits that compromise cybersecurity.

[16] Kedem, M. (2019). 11 bad habits that destroy cybersecurity efforts. SentinelOne.

[17] Majumdar, N., & Ramteke, V. (2022). Human elements impacting risky habits in cybersecurity. AIP Conference Proceedings.

[18] Ncubukezi, T. (2022). Human Errors: A Cybersecurity Concern. ICCWS.

[19] National Institute of Standards and Technology. PII – Glossary. https://csrc.nist.gov/glossary/term/PII

[20] U.S. General Services Administration. Protecting PII – Privacy Act. https://www.gsa.gov/reference/gsa-privacy-program/rules-and-policies-protecting-pii-privacy-act

ARQUITECTURA HÍBRIDA PARA GESTIÓN DINÁMICA DE FACTORES DE ACCESO: INTEGRACIÓN DE AGENTES DE IA CON WSO2 Y KEYCLOAK

Miguel Méndez-Garabetti[1,2,3], Eduardo Piray[3,4], Leonardo Requena[4], Ricardo González[3,4], Guillermo Romero Arregin[2,3], Federico Schwemler[4], Pablo Bernal[4], Yago González[4]

[1]*Facultad de Ciencias Sociales y Administrativas, Universidad del Aconcagua, Mendoza, Argentina.*

[2]*Free and Open Source Software/Hardware Research Laboratory (FOSSHLab), Argentina.*

[3]*Departamento de Sistemas, Universidad CAECE, Mar del Plata, Argentina.*

[4]*Universidad Siglo 21, Córdoba, Argentina.*

mmendezgarabetti@gmail.com, edupiray@gmail.com, lprequena@gmail.com, eliangonzalez045@gmail.com , gromeroarregin@gmail.com, schwemlerfederico@gmail.com, bpabloso12@gmail.com, yago06_2000@hotmail.com

RESUMEN

Vivimos en un mundo conectado. Las aplicaciones en la actualidad están permanentemente conectadas a redes de datos, y a la red Internet. Dichas redes pueden ser inseguras, por lo tanto, las aplicaciones y diversas soluciones de software están expuestas a ciberataques. Es clave efectuar una correcta Gestión de la Identidad y Acceso (Identity and Access Management IAM por su siglas en inglés), porque la identidad digital de los usuarios y los activos digitales deben estar protegidos, y no deben permitirse accesos no autorizados. En los últimos años la irrupción de la inteligencia artificial ha sido determinante en muchos ámbitos, la ciberseguridad y la gestión de identidad y acceso no son la excepción. Los avances en algoritmos de machine learning y modelos de lenguaje de gran tamaño (LLM) como GPT de OpenAI, Claude de Anthropic o DeepSeek-R1-Zero y DeepSeek-R1 de DeeepSeek, han impulsado el desarrollo de agentes de inteligencia artificial. Es vital poder demostrar cómo detectar dinámicamente amenazas y cómo optimizar la gestión de identidad y acceso mediante agentes. Tomando como base las limitaciones de los sistemas de gestión tradicionales que no utilizan inteligencia artificial, se estudiará cómo los agentes de inteligencia artificial pueden administrar la autenticación, la detección de anomalías y la autorización en tiempo real, en muchas ocasiones tomando sus propias decisiones, y en otras decidiendo en base a la supervisión humana. Este enfoque busca avanzar hacia un modelo IAM autónomo, seguro y escalable. Se pretende seguir una metodología de estudio basada en un análisis cuantitativo. Por medio de estudios empíricos en ambientes controlados, tanto locales como en producción que permitan evaluar la efectividad de los agentes de IAM generados.

Palabras claves: Gestión de Identidad y Acceso (IAM). Agentes inteligentes. Gestores de identidad open source. Ciberseguridad adaptativa.

CONTEXTO

La presente línea de investigación está integrada por tres proyectos de investigación llevados a cabo en universidades diferentes: 1) Universidad Siglo 21, en el marco de la carrera Licenciatura en Informática, 2) Universidad CAECE, en el marco de la Licenciatura en Sistemas, la Ingeniería en Sistemas y la Licenciatura en Gestión de Sistemas, 3) Universidad del Aconcagua. Cada proyecto está financiado individualmente por su respectiva universidad.

1. INTRODUCCIÓN

La IAM depende cada vez más de la inteligencia artificial lo que permite a las empresas adoptar una respuesta mucho más detallada y adaptable a la autenticación y el control de acceso [1] [2]. Además, se requiere inteligencia artificial para el análisis de comportamiento de los usuarios y entidades, que se utiliza para detectar actividades sospechosas. La autenticación y el control de acceso son necesarios para la seguridad de los datos [1].

La inteligencia artificial (IA) se ha convertido en una fuerza determinante en la configuración de los sistemas digitales. Las tecnologías de IA, en particular el aprendizaje automático (ML), el procesamiento del lenguaje natural y la visión artificial, se integran cada vez más en las infraestructuras de identidad para mejorar la detección de fraudes, agilizar la incorporación, habilitar el reconocimiento biométrico y automatizar el control de acceso. Estas aplicaciones prometen mayor eficiencia, escalabilidad y adaptabilidad, lo que las hace atractivas para los actores de los sectores público y privado que buscan proteger y personalizar los servicios de identidad. Sin embargo hay que tener presente que la convergencia de la IA y la identidad en línea introduce nuevas complejidades. Esto se debe a que, a medida que la IA se convierte en el intermediario en decisiones relacionadas con la identidad, como la verificación de una muestra biométrica, el análisis del comportamiento del usuario o la detección de anomalías en tiempo real, redefine la arquitectura de la confianza . La IA se está implementando en diferentes arquitecturas de identidad para mejorar la prestación de servicios: arquitecturas centralizadas, arquitecturas de identidad federada, arquitecturas descentralizadas y de identidad autosoberana (Self-Sovereign Identity SSI por sus siglas en inglés) [3].

La inteligencia artificial (IA) está evolucionando rápidamente desde sistemas aislados y deterministas a sistemas en constante aprendizaje, basados en objetivos y con una autonomía sin precedentes. La aparición de los Modelos de Lenguaje de Gran Tamaño (LLM) y los agentes de inteligencia artificial marca una revolución en los procesos empresariales, permitiendo una toma de decisiones sofisticada y la automatización del flujo de trabajo. Sin embargo, su potencial se amplifica aún más cuando se combinan con cadenas de herramientas y combinaciones de API, que permiten a los agentes de IA ejecutar tareas de varios pasos, como la organización de reuniones, el cálculo de nóminas o la ejecución de pipelines de DevOps. Además, los sistemas de IA descomponen los objetivos generales en subtareas factibles, aprenden de forma adaptativa de los resultados y corrigen el rumbo por sí solos con el tiempo. Estas tecnologías tienen un gran potencial, pero traen consigo nuevos desafíos de gestión y gobernanza de identidades que escapan a la capacidad de los sistemas IAM tradicionales [4].

Por otro lado, el uso de la inteligencia artificial por parte de los ciberdelincuentes para llevar a cabo ataques cada vez más automatizados y amplios, se está volviendo más común. La reacción obvia sería restringir el control de acceso en la mayor medida

posible, obligando a los usuarios a autenticarse varias veces durante el día mediante MultiFactor Authentication (MFA) cada vez que necesiten acceder a sus aplicaciones [5]. Si bien un sistema IAM puede permitir implementar tales reglas de seguridad, no todos pueden aumentar o reducir dinámicamente los requisitos de autenticación necesarios [1].

Finalmente es necesario resaltar que al aprovechar el poder de la inteligencia artificial, los sistemas de detección de amenazas pueden mejorar la precisión y la eficiencia de los sistemas de gestión de identidad y acceso. Con el apoyo de las tecnologías de inteligencia artificial, los sistemas IAM pueden analizar cantidades masivas de datos en tiempo real, lo que les permite detectar patrones o comportamientos anormales que pueden indicar una posible amenaza a la seguridad. Este enfoque proactivo para la detección de amenazas proporciona a las organizaciones la capacidad de responder rápidamente a los incidentes de seguridad y prevenir posibles infracciones antes de que ocurran [6].

Los antecedentes mencionados plantean distintas problemáticas de ciberseguridad que pueden ser analizadas y probadas empíricamente en el presente proyecto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación se centra en la ciberseguridad tomando como eje la protección de la identidad del usuario en arquitecturas de sistemas centralizados y descentralizados. A continuación, se describe brevemente la línea de investigación.

1. Por un lado se trabaja con el objetivo de demostrar cómo detectar dinámicamente amenazas y cómo optimizar la gestión de identidad y acceso mediante inteligencia artificial en Web 2.0 y Web 3.0 para proteger la identidad digital y los accesos no autorizados. Se tiene como objetivo diseñar agentes inteligentes de IAM basados en machine learning e inteligencia artificial para proteger la identidad digital y los accesos no autorizados en aplicaciones web. Se tiene planificado trabajar en Python. Además se busca comparar los agentes de IAM generados con sistemas de IAM tradicionales (que no utilizan inteligencia artificial) en aspectos de seguridad y experiencia de usuario para determinar los más apropiados de implementar en aplicaciones web reales. Se pretende desarrollar e integrar los agentes inteligentes en los gestores de gestión de identidad y acceso de código abierto Keycloak y WSO2. WSO2 ya ha evolucionado en esa dirección, confirmando esta tendencia en WSO2 Identity Server 7.2, el cual está diseñado para "Garantizar la seguridad de los agentes de IA" [7].
2. Otro aspecto en lo que se está trabajando es en una técnica que consiste en un modelo doble factor de autenticación basado en smart contracts desplegados en blockchain que no haga uso de proveedores de identidad de terceros, con un enfoque en la privacidad y la seguridad. Esta es una técnica que se podría integrar en un gestor tradicional basado en Web 2.0, como Keyclock o WSO2, para ofrecer una alternativa de gestión de identidad y acceso basada en blockchain. En lo que va del trabajo, se concluye que las herramientas de software vinculadas a la gestión de identidad y acceso, tanto en Web 2.0 y Web 3.0, podrían ser interoperables e integrarse para brindar más alternativas de seguridad a los usuarios de aplicaciones web. Sin poder llegar a afirmar de forma fehaciente que el uso de tecnologías

vinculadas a la Web3 mejora o no los niveles de seguridad en aplicaciones web respecto a tecnologías vinculadas a la Web2.

3. Otro eje con el cual se viene trabajando desde 2024 es el diseño y ejecución de pruebas de tipo read team a los sistemas de IAM insertos en aplicaciones web. Hasta el momento se han llevado a cabo pruebas en entorno en nube y en un Servidor Privado Virtual (VPS por sus siglas en inglés). El entorno en nube es con el servicio de Render, e implicó el despliegue de una versión de Keycloak y de un sample web en React. El entorno en la VPS implicó el despliegue de una versión de WSO2 Identity Server, y se lo integró con otra aplicación web tipo sample React desplegada en Cloudflare. El despliegue en la VPS y en Render se realizó mediante Docker [8]. Concretamente uno de los enfoques de las pruebas fue auditar el servicio WSO2 Identity Server en su versión 7.1.0. Para ello se seleccionaron vulnerabilidades reportadas en dicho gestor, teniendo como base el estándar de criticidad CVSS (Common Vulnerability Scoring System) y el programa CVE (Common Vulnerabilities and Exposures). Los test fueron de caja negra y caja gris de que validaron la mitigación de vulnerabilidades críticas tales como: path traversal (CVE-2022-29464), XSS almacenado (CVE-2024-3509) y denegación de servicio por atributos repetidos (WSO2-2023-2678). Otras vulnerabilidades analizadas y que se confirmó mitigación fueron CVE-2024-6914, CVE-2024-7097, CVE-2024-8122, CVE-2024-7487 y CVE-2024-3509. Por otro lado, también se realizaron pruebas de carácter más general, no directamente alineadas con los avisos propios de WSO2. Esto tuvo como objetivo dar indicios del grado de fortaleza o hardering del gestor de identidad, frente a posibles ataques o eventos no esperados que afecten la resistencia y resiliencia de WSO2. Estas pruebas fueron: configuración por defecto de superficies de ataque, tomando como base científica el CVE-2022-29464, el cual se refiere a la carga de archivos sin restricciones, lo que resulta en la ejecución remota de código; la segunda se basó en una prueba de resiliencia a DoS (Denegación de Servicios) [8].

   Los resultados más destacados es que el sistema WSO2 no bloquea cuentas tras múltiples intentos fallidos y permite ataques de fuerza bruta persistentes contra la consola administrativa. También se verificó que no existe bloqueo de cuenta ante este tipo de ataque, lo cual valida la ausencia de mecanismos anti fuerza bruta. Además que WSO2 Identity Server 7.1.0 es resiliente frente a ataques de DoS. Los tiempos de respuesta se mantuvieron estables y en términos razonables ($\Delta <= \pm 0{,}07$ s).

   Con la experiencia adquirida se continúa trabajando en este ámbito para lograr pruebas y resultados más significativos.

4. Otro aspecto que se está analizando desde un punto de vista de concientización, es el de la identidad autosoberana (SSI). Comprendiendo el contexto y la situación actual en Argentina, en Latinoamérica y en la Comunidad Europea.

3. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la presente línea de I+D+i cuenta con la dirección del Dr. Miguel

Mendez-Garabetti. Además, participan docentes investigadores de las diferentes universidades. Se cuenta además con la participación de cinco de grado. Uno de los estudiantes de grado se encuentra presentando su proyecto final basado en la línea de investigación.

4. RESULTADOS ESPERADOS

Existen diferentes resultados esperados relacionados en los diferentes proyectos, los cuales confluyen al resultado esperado de la línea de I+D+i general. El primer resultado esperado es generar, al menos, un agente de inteligencia artificial que se pueda integrar a los gestores de identidad Keycloak y WSO2 Identity Server. Que sea de carácter autónomo, seguro y escalable, y que permita administrar dinámicamente factores de autenticación en base a patrones de comportamiento del usuario y del entorno. Se harán estudios para auditar las decisiones del agente.

Otro de los resultados esperados es obtener mediciones y establecer una ponderación lo más exacta posible, acerca del rendimiento y niveles de seguridad del agente generado, con el objetivo de definir el grado de protección que ofrece a la identidad del usuario y cómo influyen en la experiencia de usuario.

Finalmente, otro resultado esperado es concluir si el uso de tecnologías descentralizadas Web3 mejoran los niveles de seguridad en aplicaciones web respecto a tecnología vinculadas a las Web2.

5. BIBLIOGRFÍA

[1] I. A. Mohammed, “A literature review on the application of AI to Identity Access Management”, vol. 5, núm. 10, 2018.

[2] D. Cole, “Artificial intelligence and personal identity”, Synthese, vol. 88, núm. 3, pp. 399–417, sep. 1991, doi: 10.1007/BF00413555.

[3] L. O. Nweke, Artificial Intelligence, Trust, and the Architecture of Online Identity: Security Implications in an Algorithmic Era. IntechOpen, 2025. doi: 10.5772/intechopen.1012880.

[4] B. Bhushan, “An Explainable Zero Trust Identity Framework for LLMs, AI Agents, and Agentic AI Systems”, International Journal of Computer Applications, vol. 187.

[5] B. Wang, D. Liu, y M. Ji, “Research on Management System of Mold Manufacturing Enterprise Based on RFID Technology”, MATEC Web of Conferences, vol. 95, p. 10002, feb. 2017, doi: 10.1051/matecconf/20179510002.

[6] Nikhil Ghadge, “Enhancing threat detection in Identity and Access Management (IAM) systems”, Int. J. Sci. Res. Arch., vol. 11, núm. 2, pp. 2050–2057, abr. 2024, doi: 10.30574/ijsra.2024.11.2.0761.

[6] S. Talluri, “AI In Identity And Access Management: Current Practices, Challenges, And Opportunities”, vol. 12, núm. 1, 2022.

[7] “Presentamos WSO2 Identity Server 7.2: Impulsando el futuro de los agentes CIAM y de IA segura | WSO2”. Consultado: el 19 de febrero de 2026. [En línea]. Disponible en: https://wso2.com/library/blogs/introducing-identity-server-7-2

[8] M. Mendez-Garabetti, E. Piray, L. Requena, R. González, G. Romero Arregin, F. Schwemler, P. Bernal, Y. González, Matías Moll y D. Fraser, “Estudio de vulnerabilidades y robustez en WSO2 para una integración segura entre Web2 y Web3”, nov. 2025.

Atacantes en redes OT: impacto funcional sobre PLC y procesos físicos

Gonzalo Heinen, Valentina Heinen, Jorge Kamlofsky
CAETI - Centro de Altos Estudios en Tecnología Informática.
Universidad Abierta Interamericana, Facultad de Tecnología Informática

Abstract

Se analiza el impacto que puede generar un atacante con acceso a la red de Tecnología Operacional (OT) sobre sistemas de control industrial, poniendo especial énfasis en la manipulación directa de controladores lógicos programables (PLC) mediante protocolos industriales sin mecanismos de autenticación. En este contexto, se describe un experimento práctico orientado a evaluar cómo la escritura maliciosa de variables internas del PLC puede alterar el comportamiento de un proceso físico real, sin modificar la lógica de control del sistema.

Para ello, se emplea un banco de pruebas portátil previamente desarrollado, que permite recrear una celda OT física, controlada y reproducible, integrada por un PLC, una estación de ingeniería, un servidor SCADA y un proceso físico representativo.

Los resultados obtenidos evidencian que es posible inducir estados de encendido, apagado e interferencia intermitente en una maquinaria simulada, demostrando que la ausencia de mecanismos de autenticación y control de acceso en entornos OT constituye un vector de riesgo significativo. De este modo, el trabajo busca aportar una aproximación experimental y académica al análisis del impacto funcional de ataques sobre sistemas de control industrial.

Palabras claves: *Ciberseguridad OT, Sistemas de Control Industrial, PLC, Modbus TCP, Infraestructuras Críticas, Redes OT, Ataques a PLC*

CONTEXTO

Esta investigación se desarrolla en el marco del proyecto “Ciberdefensa en redes OT”, llevado adelante en el Laboratorio CAETI de la Universidad Abierta Interamericana. En dicho proyecto se diseñó y documentó previamente un banco de pruebas portátil basado en PLC, orientado al estudio de vulnerabilidades en sistemas de control industrial dentro de entornos OT físicos y controlados.

1. INTRODUCCION

La ciberseguridad en infraestructuras críticas constituye un aspecto central para la operación segura de los sistemas industriales, ya que la presencia de vulnerabilidades puede derivar en fallos operativos y daños sobre equipamiento físico. Esta problemática se evidencia a nivel internacional a partir de incidentes como el ataque a las instalaciones nucleares iraníes en 2010, el cual demuestra que la manipulación maliciosa de sistemas de control industrial puede inducir fallas físicas deliberadas en infraestructuras críticas [1]. Este tipo de eventos pone de manifiesto que las vulnerabilidades presentes en entornos de Tecnología Operacional pueden trasladarse directamente al plano físico, generando consecuencias operativas reales.

Diversos estudios y relevamientos posteriores documentan incidentes adicionales que refuerzan esta problemática. En particular, se reportan múltiples casos de afectación a sistemas SCADA y controladores industriales asociados a accesos no autorizados, errores de configuración y exposición de servicios de control a redes externas, evidenciando la fragilidad de este tipo de infraestructuras cuando no se implementan mecanismos de protección adecuados [4]. Asimismo, se registran eventos en los que la conectividad Ethernet de PLC industriales permite la recepción de tráfico no previsto, provocando la pérdida de control del proceso y la interrupción de la operación normal del sistema [5]. Del mismo modo, el acceso remoto no autorizado a sistemas SCADA ha sido identificado como

un vector recurrente de ataque, habilitando la manipulación de variables de proceso y la degradación del funcionamiento de infraestructuras industriales críticas [6].

Los sistemas de control industrial incorporan de manera creciente tecnologías de red originalmente desarrolladas para entornos de Tecnología de la Información (IT), muchas veces sin que este proceso sea acompañado por mecanismos de seguridad adecuados. Como consecuencia, numerosos protocolos industriales continúan operando bajo supuestos de confianza implícita en la red, careciendo de autenticación o controles de acceso robustos, lo que expone a los controladores lógicos programables a ataques capaces de afectar directamente los procesos físicos que controlan [2]. En este escenario, y tomando como referencia los incidentes documentados, el presente trabajo emplea un banco de pruebas para analizar, mediante un experimento concreto y reproducible, el impacto de acciones maliciosas sobre un sistema OT físico en operación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1. ENTORNO EXPERIMENTAL Y SUPUESTOS DEL EXPERIMENTO

El experimento se desarrolla sobre un entorno de tecnología operacional físico implementado mediante un banco de pruebas portátil [3]. Dicho dispositivo integra una red Ethernet organizada como una celda OT aislada, en la cual el PLC, la estación de ingeniería y el servidor SCADA se interconectan a través de un switch Ethernet, conformando un único dominio de red.

Esta configuración habilita la comunicación directa entre los distintos componentes del sistema mediante el uso de protocolos industriales estándar, como Modbus TCP, sin depender de infraestructura externa. La topología adoptada replica de forma simplificada un entorno industrial real, lo que permite tanto la operación legítima del proceso desde la estación de ingeniería como la ejecución de pruebas controladas de ciberseguridad sobre el PLC.

El escenario experimental plantea un atacante con acceso a la red OT, donde se asume que el mismo ya tiene acceso a la red. No se consideran en este trabajo vectores de intrusión previos, escaladas de privilegios ni técnicas de explotación del sistema operativo del controlador. El foco del análisis se centra exclusivamente en las capacidades de un atacante una vez obtenido el acceso a la red, bajo un modelo de amenaza realista y común en entornos industriales.

2.2. ACTORES DE LA RED OT

En el dominio interactúan distintos actores con roles bien definidos dentro de la red interna del banco de pruebas portátil.

En primer lugar, el **controlador lógico programable** constituye el núcleo del proceso, siendo el encargado de ejecutar la lógica de control y de accionar las entradas y salidas físicas del sistema. El PLC expone variables internas a través de la red Ethernet para su supervisión y control desde otros componentes de la celda OT.

La **estación de ingeniería**, implementada sobre una PC tipo NUC, cumple el rol de operador legítimo del sistema. Desde esta estación se realiza la programación, descarga y monitoreo de la lógica Ladder del PLC mediante software propietario de ingeniería, así como la supervisión del estado del proceso durante la operación normal. Este actor representa el uso autorizado del sistema dentro del entorno OT.

El **servidor SCADA**, desplegado sobre una plataforma basada en Raspberry Pi, actúa como punto central de visualización y supervisión del proceso. A través de este componente se consumen las variables del PLC, se presentan interfaces HMI y se registran eventos relevantes del sistema, permitiendo observar el comportamiento del proceso en tiempo real dentro de la red local de la plataforma.

Finalmente, el entorno contempla la presencia de un **actor no autorizado (atacante)**, representado por un equipo conectado al mismo segmento de red OT. El atacante no interactúa con el PLC mediante software de ingeniería ni modifica la lógica de control, sino que se comunica directamente con el controlador a través de protocolos industriales estándar, aprovechando la ausencia de mecanismos de autenticación.

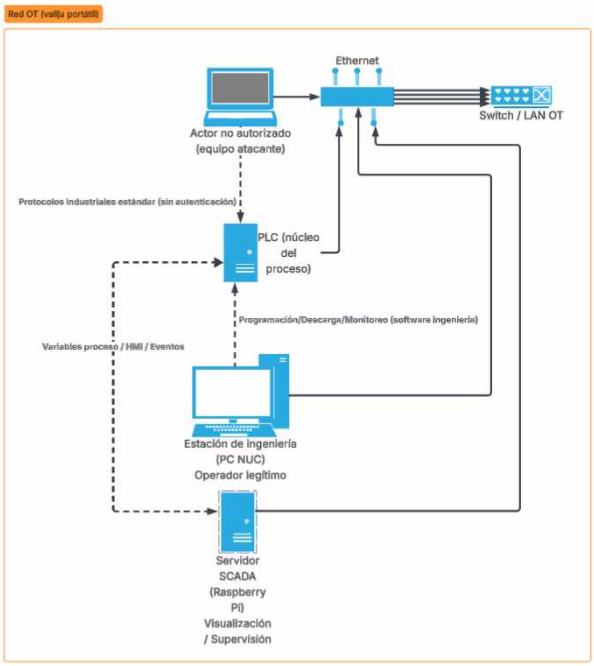


Figura 1. Topología de red OT del banco de pruebas portátil.

2.3. ACCIONES DEL ATACANTE Y EFECTOS ESPERADOS SOBRE EL PROCESO FÍSICO

El atacante emplea un script desarrollado en Python que implementa un cliente Modbus TCP. Dicho script permite interactuar directamente con el PLC mediante la escritura de bobinas y registros internos. De este modo, las acciones del atacante se limitan al uso legítimo del protocolo industrial, explotando la ausencia de mecanismos de autenticación.

El código implementa distintos modos de operación, que incluyen la escritura puntual de bobinas específicas y, en aquellos casos en los que no se conoce su direccionamiento, la posibilidad de realizar una sobreescritura masiva de salidas. Asimismo, permite la modificación de registros utilizados por la lógica de control y contempla un modo de interferencia cíclica que fuerza cambios de estado de manera continua. Estas acciones se ejecutan en tiempo de ejecución y actúan directamente sobre las variables internas del PLC, sin afectar el programa original ni requerir reinicios del sistema.

Los efectos de estas acciones se observan de forma inmediata sobre el proceso físico implementado en el banco de pruebas. El sistema cuenta con un conjunto de indicadores luminosos, un mecanismo de encendido por llave y una turbina que representa una maquinaria industrial en operación. A partir de la manipulación de las variables del PLC, es posible forzar distintos estados del proceso, tales como el encendido estable de la turbina, su detención completa o un funcionamiento intermitente caracterizado por variaciones cíclicas en la velocidad de rotación.

Estos últimos comportamientos resultan particularmente relevantes, ya que simula una condición de falla operativa en la que la

maquinaria alterna entre estados de arranque y parada sin intervención del operador legítimo. En todos los casos, la lógica Ladder continúa ejecutándose en el PLC, pero el control del proceso queda supeditado a los valores impuestos externamente por el atacante.

3. RESULTADOS OBTENIDOS

3.1 EJECUCIÓN DEL EXPERIMENTO

Durante la ejecución del experimento, el atacante inició la comunicación con el PLC mediante un script desarrollado en Python que utiliza la librería pymodbus, implementando un cliente Modbus TCP sobre el puerto estándar 502. Una vez establecida la conexión con el controlador, el programa presentó un menú de opciones que permitió seleccionar distintas acciones de escritura sobre las variables internas del PLC, incluyendo la manipulación puntual y masiva de bobinas, así como la modificación de registros.

Entre los modos evaluados, se observó que la ejecución de las funciones de escritura de bobinas (write coil) permitió modificar directamente el estado de las salidas asociadas al proceso físico. Al forzar el estado de bobinas determinadas, el PLC impactó de manera directa sobre la turbina del banco de pruebas, reflejando cambios funcionales equivalentes a los de una maquinaria industrial en operación. Estas acciones se realizaron sin modificar la lógica Ladder cargada en el controlador y sin intervención del operador legítimo.

Asimismo, el uso de operaciones masivas, implementadas mediante iteraciones sucesivas sobre el espacio de direcciones de bobinas, permitió cambiar estados globales del sistema, generando condiciones de encendido o apagado total. Este comportamiento se reflejó tanto en el conjunto de indicadores luminosos como en la turbina, que pasó a un estado de detención completa o funcionamiento continuo no previsto por la lógica de control original.

Un caso particularmente relevante fue la ejecución del modo de interferencia cíclica, implementado en el script mediante una secuencia repetitiva de escrituras con retardos temporales (sleep). Este modo provocó un comportamiento intermitente del proceso, caracterizado por variaciones cíclicas en la velocidad de rotación de la turbina, simulando una falla operativa en la maquinaria. El efecto observado reproduce un escenario de inestabilidad funcional, en el cual el proceso alterna entre estados de encendido y apagado sin una causa aparente desde la perspectiva del operador.

3.2 CONCLUSIÓN DE LOS RESULTADOS

En conclusión, los resultados obtenidos muestran que, bajo el supuesto de un atacante con acceso al segmento de red OT, es posible interferir de manera directa y efectiva en el funcionamiento de un proceso físico controlado por un PLC. Mediante el uso exclusivo de un protocolo industrial estándar y sin mecanismos de autenticación, el atacante logró modificar variables internas del controlador en tiempo de ejecución, provocando comportamientos anómalos en la maquinaria simulada.

La posibilidad de forzar estados de encendido, apagado o funcionamiento intermitente sin alterar la lógica de control ni utilizar herramientas de ingeniería resalta la fragilidad de los sistemas industriales cuando se asume implícitamente la confianza en la red.

En este sentido, el experimento valida que la ausencia de mecanismos de autenticación y control de acceso en protocolos industriales ampliamente utilizados, constituye un vector de riesgo significativo, capaz de traducirse en

fallas operativas visibles y potencialmente críticas, incluso en entornos controlados y de pequeña escala.

Para el corriente año, se espera profundizar esta línea de investigación mediante la ampliación de los escenarios experimentales y el análisis de nuevas variantes de interacción maliciosa sobre sistemas OT. En particular, se prevé evaluar el impacto de distintos mecanismos de protección a nivel de red y protocolo, así como incorporar métricas que permitan caracterizar con mayor precisión los efectos funcionales de este tipo de ataques. Asimismo, se considera la extensión del banco de pruebas para incluir configuraciones más complejas y representativas de entornos industriales reales, con el objetivo de fortalecer la comprensión de los riesgos asociados y contribuir al desarrollo de estrategias de mitigación más efectivas.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto se desarrolla bajo la dirección del Mg. Lic. Jorge Kamlofsky en el ámbito del Laboratorio CAETI de la Universidad Abierta Interamericana. En el marco de su ejecución, se prevé la participación de estudiantes de grado de la carrera de Ingeniería en Sistemas de Información, quienes intervienen en las distintas etapas del trabajo como parte de su formación académica. Asimismo, participan estudiantes de posgrado, quienes profundizan sus conocimientos en ciberseguridad y sistemas de control industrial, utilizando los resultados obtenidos como insumo para el desarrollo de sus trabajos de investigación y tesis doctorales.

Esta integración de recursos humanos permite fortalecer la formación técnica y académica de los participantes, al tiempo que contribuye al desarrollo, validación y transferencia de los resultados alcanzados en el proyecto.

5. BIBLIOGRAFÍA

[1] Denning, D. E. (2012). Stuxnet: What has changed? Future Internet, 4(3), 672–687. https://doi.org/10.3390/fi4030672

[2] Heinen, G., Milano, M. A., & Kamlofsky, J. (2025). Uso de técnicas de machine learning para la detección temprana de ransomware. ResearchGate. https://surl.lu/refnbc

[3] Heinen, G., & Heinen, V. (2025). Prototipado de una plataforma SCADA portátil para el análisis de vulnerabilidades en redes OT. ResearchGate. https://surl.li/qmdlmj

[4] Miller, B., & Rowe, D. (2012). A survey of SCADA and critical infrastructure incidents. Proceedings of the 1st Annual Conference on Research in Information Technology. https://surl.li/dajyzs

[5] Knapp, E. D., & Langill, J. T. (2015). Industrial network security: Securing critical infrastructure networks for smart grid, SCADA, and other industrial control systems (2nd ed.). https://surl.li/dtzpho

[6] Byres, E., Franz, M., & Miller, D. (2004). The use of attack trees in assessing vulnerabilities in SCADA systems. Proceedings of the International Infrastructure Survivability Workshop. https://surl.lt/usdumc

Ciberdefensa Desplegable

Jonatan Schmidt, Daniel Manrique, Jorge Kamlofsky
CAETI - Centro de Altos Estudios en Tecnología Informática.
Universidad Abierta Interamericana, Facultad de Tecnología Informática

Abstract

El término «ciberdefensa desplegable» se entiende como la capacidad de desplegar recursos y equipos de ciberseguridad con una capacidad expedicionaria/móvil y de operar in situ en lugares donde existe una necesidad inmediata. El documento se centra en la investigación del concepto y algunos de sus términos relacionados, como la ciberdefensa móvil y las operaciones cibernéticas expedicionarias.

A partir de aquí, la investigación se embarca en un examen de la evolución conceptual (incluidas las definiciones y los marcos doctrinales) y también de la evolución práctica que analiza las implementaciones y los casos de uso de grado militar.

El documento aclara los principales contextos de aplicación, que van desde la defensa nacional hasta las operaciones de coalición (OTAN, UE), y traza un mapa de las regiones geográficas que han sido testigo de debates activos en este campo. Por último, se evalúa la madurez del campo, se destacan las lagunas en la investigación y se determina si la ciberdefensa desplegable es un área emergente validada o si sigue estando en una fase inicial poco investigada.

Palabras claves*: ciberdefensa desplegable,*

ciberdefensa móvil, operaciones expedicionarias, ciberseguridad militar.

CONTEXTO

Esta es una línea de trabajo dentro del proyecto "Ciberseguridad en Redes Operacionales", radicado en el CAETI[1], centro de investigación dependiente de la Facultad de Tecnología Informática de la UAI (Universidad Abierta Interamericana). Los proyectos radicados en el CAETI se clasifican en tres líneas de investigación: Ingeniería de Software, Sociedad del Conocimiento y Automatización y Robótica.

1. INTRODUCCIÓN

El término "ciberdefensa desplegable" se refiere a la capacidad de proyectar equipos o recursos de ciberseguridad/ciberdefensa de forma expedicionaria, móvil, fuera de instalaciones centrales para operar in situ, en el lugar donde sean requeridos (poner una referencia). También se pueden encontrar términos equivalentes o estrechamente relacionados en la literatura, como ciberdefensa móvil, ciberfuerzas proyectables o ciberoperaciones expedicionarias, y en inglés, deployable cyber defense, expeditionary cyberspace operations o expeditionary cyber teams.
Se investigaron tanto la evolución conceptual (definiciones, marcos doctrinales, aportaciones teóricas) como la

[1] CAETI: https://caeti.uai.edu.ar/ .

evolución práctica (implementaciones, casos de uso a nivel militar y de defensa). Así mismo, se identifican los principales contextos de aplicación desde militar/nacional hasta operaciones de coalición (OTAN, UE) y se reconocen las regiones geográficas en las que se ha debatido o implementado la ciberdefensa desplegable. Por último, se evalúan la madurez del campo, los principales vacíos de investigación y si es un área emergente o todavía en primer plano e investigación escasa.

1.1. ORÍGENES DEL CONCEPTO

A mediados de la década de 2010, la idea de proyectar capacidades de ciberdefensa en el teatro de operaciones comenzó a arraigarse en los círculos militares de Estados Unidos y la OTAN. Una primera señal se produjo en 2009, cuando la OTAN reconoció la necesidad de contar con «una capacidad de ciberdefensa desplegable» para ayudar a los Estados miembros que sufrieran ataques [1][2]. Poco después, en 2010, el Ejército de los Estados Unidos creó un comando cibernético (USCYBERCOM/ARCYBER) y comenzó a estudiar formas de incorporar las operaciones cibernéticas a nivel táctico. Un hito teórico de gran importancia fue la publicación en 2013 del Land Warfare Paper No. 95 por parte de la Asociación del Ejército de los Estados Unidos (AUSA) [4][5][7], que introdujo formalmente el concepto de operaciones expedicionarias en el ciberespacio en el contexto de las operaciones de contrainsurgencia en Afganistán.

1.2. CONSOLIDACIÓN Y ADOPCIÓN.

Entre 2016 y 2020, el concepto comenzó a ganar terreno tanto en la doctrina como en los ejercicios militares, especialmente en Occidente [6].

Un paso significativo en la conceptualización formal se produjo en 2020 con la publicación por parte de la Junta Interamericana de Defensa (JID) de la Guía Militar de Ciberdefensa [4][5][7]. Esta guía doctrinal (para las fuerzas armadas de las Américas) aborda temas que conciernen específicamente a la ciberdefensa desplegable, articulando su definición y esbozando sus requisitos. La JID también afirmó que, en ciertos casos, por razones tecnológicas, operativas o de procedimiento, las acciones de ciberdefensa no pueden llevarse a cabo de forma remota... se hace necesario desplegar parte de sus capacidades para hacer frente a situaciones específicas [4][5][7]. Además, las fuerzas cibernéticas deben contar con capacidades de ciberdefensa desplegables compuestas por personal, medios y procedimientos personalizables y transportables, entrenados y listos para desplegarse donde sea necesario (teatro de operaciones, ejercicios, instalaciones, infraestructura crítica, etc.) en el menor tiempo posible para aumentar las ciberdefensas existentes o reaccionar ante situaciones críticas urgentes [4][5][7].

1.3. FORMALIZACIÓN DOCTRINAL Y AVANCES RECIENTES

El período comprendido entre 2021 y 2025 fue testigo del establecimiento de normas doctrinales formales para la ciberdefensa y las operaciones cibernéticas expedicionarias en los Estados Unidos. La actualización de 2022 de la Publicación

Conjunta 3-12 del Departamento de Defensa estableció una definición formal de las operaciones expedicionarias en el ciberespacio, que define estas operaciones como necesarias cuando es preciso desplegar fuerzas físicas para operaciones en el ciberespacio porque el acceso remoto no es viable. El entorno militar recibió el reconocimiento oficial del concepto operativo a través de esta definición, que estableció un lenguaje militar estandarizado para su aplicación operativa.

El ejército estadounidense ha logrado avances mediante la implementación de programas de entrenamiento y despliegues operativos que muestran a todas las unidades militares dedicadas a la guerra cibernética, incluido el 11.º Batallón de Guerra Cibernética, a través de sus operaciones de seguridad. La Fuerza Aérea, la Armada y el Cuerpo de Marines han investigado diferentes formas de combinar sus capacidades cibernéticas que pueden utilizarse en sus futuras operaciones militares.

La alianza de la OTAN desarrolló sus equipos de reacción rápida (RRT), que operan como unidades de ciberdefensa capaces de responder en 24 horas para ayudar a las naciones aliadas que se enfrentan a situaciones de ciberataque, creando así un sistema operativo conjunto para sus operaciones de ciberdefensa. Naciones europeas como España establecieron sus unidades tácticas de ciberdefensa, que colaboran con las fuerzas militares aliadas de alta disponibilidad.

La comunidad académica de investigación de América Latina ha comenzado a producir estudios académicos sobre el desarrollo de capacidades de ciberdefensa, a pesar de que esta región aún no cuenta con un desarrollo doctrinal establecido.

Las prácticas y doctrinas militares de múltiples naciones y organizaciones establecerán el concepto como su procedimiento operativo estándar para 2025, a pesar de la limitada producción de investigación académica que actualmente se centra en el desarrollo doctrinal y militar profesional.

1. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta es una línea de investigación dentro del proyecto "Ciberseguridad en Redes Operacionales" radicado en el CAETI, que busca mejorar las condiciones de ciberseguridad en redes operacionales. En especial, en infraestructuras críticas, que operan tecnología operacional. El objetivo principal del proyecto consiste en el diseño y desarrollo de un escáner especializado para entornos industriales basado en PLC, concebido como un dispositivo desplegable orientado al análisis de redes de Tecnología Operacional (OT). La solución propuesta permite su conexión en un puerto de switch previamente autorizado por el administrador de red, habilitando la ejecución de un proceso de escaneo pasivo sobre los dispositivos PLC presentes en la red.

El sistema funciona en modo pasivo para descubrir e identificar activos industriales a través de sus procedimientos estándar sin producir ningún tráfico de red disruptivo. El dispositivo utiliza la información de activos establecida para comparar los datos recopilados con la base de datos de vulnerabilidades existente, que contiene

riesgos de seguridad conocidos para combinaciones específicas de marcas, modelos y versiones de firmware.

El dispositivo admite un modo de escaneo activo que habilita funciones adicionales que incluyen el análisis y la captura de tráfico en tiempo real, así como el mapeo de dispositivos, que muestra todos los dispositivos conectados y diversas pruebas de conectividad de la red.

Los diferentes métodos de análisis funcionan conjuntamente para crear una base técnica que permite elaborar informes de evaluación de la seguridad que muestran el estado de seguridad de la red OT a través de su inventario de activos y la identificación de vulnerabilidades conocidas y la evaluación de la exposición. La propuesta establece un marco de ciberdefensa que permite realizar evaluaciones rápidas sobre el terreno, al tiempo que mejora la conciencia situacional de los sistemas de infraestructura industrial.

2. CONCLUSIÓN

El análisis realizado permite afirmar que la ciberdefensa desplegable ha evolucionado desde una noción incipiente asociada a necesidades operativas específicas hacia un concepto formalmente incorporado en marcos doctrinales y estructuras militares de múltiples países y alianzas. Desde sus primeros antecedentes en la OTAN y en el ámbito doctrinal estadounidense, hasta su consolidación en publicaciones oficiales como la Joint Publication 3-12 y la Guía Militar de Ciberdefensa de la JID, el concepto ha alcanzado un nivel significativo de legitimidad operativa.

El período 2021–2025 marca un punto de inflexión, caracterizado por la formalización conceptual, la estandarización terminológica y la institucionalización de capacidades expedicionarias en ciberdefensa. Las experiencias de Estados Unidos, la OTAN y diversos países europeos evidencian que la proyección física de capacidades cibernéticas constituye una respuesta necesaria frente a escenarios donde el acceso remoto no es viable o resulta insuficiente. Asimismo, la creación de equipos de reacción rápida y unidades tácticas especializadas confirma la transición del concepto desde el plano teórico hacia su implementación práctica.

No obstante, el desarrollo académico continúa siendo limitado en comparación con el avance doctrinal y operativo. En América Latina, si bien comienzan a surgir aportes que reconocen la necesidad de capacidades desplegables, el campo aún se encuentra en una etapa temprana de consolidación investigativa.

En este contexto, la línea de investigación presentada —orientada al diseño de un escáner desplegable para redes OT basadas en PLC— se alinea con las tendencias internacionales identificadas, trasladando el concepto de ciberdefensa expedicionaria al ámbito industrial. La propuesta contribuye a materializar el enfoque desplegable mediante una solución técnica concreta que permite realizar evaluaciones in situ, mejorar la visibilidad de activos industriales y detectar vulnerabilidades conocidas sin afectar la operación crítica.

En síntesis, la ciberdefensa desplegable puede considerarse un campo emergente con validación doctrinal y operacional, pero aún con amplias oportunidades de

desarrollo científico y tecnológico. La articulación entre doctrina, práctica militar y aplicaciones en entornos industriales representa un espacio fértil para futuras investigaciones que fortalezcan la resiliencia de infraestructuras críticas en escenarios de creciente complejidad cibernética.

3. FORMACIÓN DE RECURSOS HUMANOS

El proyecto está dirigido por el Mg. Lic. Jorge Kamlofsky quien está finalizando un Doctorado en Ingeniería, abordando en su tesis, cuestiones de ciberseguridad operacional. Daniel Manrique desarrolla esta línea para su tesis doctoral. También participan de esta línea de trabajo José Castro Tramontina, Jonatan Schmidt, Gonzalo y Valentina Heinen, Mateo Wrobel y Mallid Dagil. José está finalizando su doctorado en Tecnología Informática, y es oficial del Ejército Argentino. Jonatan está finalizando la carrera de Ingeniería en Sistemas y busca aquí ideas para su desarrollo posgraduación. Por otro lado, Gonzalo, Valentina, Mallid y Mateo como alumnos de los años iniciales de Ingeniería en Sistemas, adquirirán conocimientos destacados para formación

5. BIBLIOGRAFIA

[1] NATO. (2009, abril 3). *Press briefing by NATO Spokesman*.

[3] Pendall, D. W., Wilkes, R., & Robinson, T. J. (2013). *Cybersecurity Operations in Support of Counterinsurgency Operations*.

Land Warfare Paper No.95, AUSA.

[4] Junta Interamericana de Defensa – JID

(2020). Guía de Ciberdefensa: Orientaciones para el diseño, planeamiento, implantación y desarrollo de una ciberdefensa militar. OEA: Washington, D.C.

[5] Ravera, M. (2024). *El rol y las capacidades cibernéticas de las Fuerzas Armadas Argentinas en el marco de los conflictos futuros*. Tesis de Maestría, IESE (Argentina).

[6] Pomerleau, M. (2023). "New DoD doctrine officially outlines and defines 'expeditionary cyberspace operations'". DefenseScoop, 12 May 2023.

[7] Chávez Salazar, E. (2022). *Ciberdefensa y ciberseguridad en el Perú: realidad y retos en torno a la capacidad de las FF. AA. para neutralizar ciberataques que atenten contra la seguridad nacional*. Revista de Ciencia e Investigación en Defensa.

Ciberseguridad en Redes Operacionales

Jorge Kamlofsky, José Castro Tramontina, Daniel Manrique, Jonatan Schmidt

CAETI – Facultad de Tecnología Informática – Universidad Abierta Interamericana
Buenos Aires, Argentina

Jorge.Kamlofsky@uai.edu.ar
JoseFederico.CastroTramontina@uai.edu.ar
daniel.manrique@alumnos.uai.edu.ar
jonatan.schmidt@alumnos.uai.edu.ar

Resumen

En los últimos siglos, las revoluciones industriales facilitaron el mayor desarrollo económico, tecnológico y humano de la historia de la humanidad. Estamos transitando hoy los inicios de una cuarta revolución, llamada Industria 4.0. Se apoya sobre tecnologías digitales disruptivas como ser Inteligencia Artificial, Big Data, Impresión 3D, y sobre nuevos protocolos de comunicación. Industria 4.0 propone una interconexión de redes tanto horizontal como vertical, lo que provocaría altísimos niveles de eficiencia y eficacia en la producción.

Pero integrar a los sistemas de control industrial (ICS según sus siglas en inglés) con otras redes dejaría a los ICS expuestos a vulnerabilidades, pudiéndose afectar su funcionamiento. Por otro lado, volver al aislamiento, se contrapone con la propuesta básica de Industria 4.0. Por eso, Industria 4.0 propone requerimientos de ciberseguridad para poder lograr sus propósitos.

En este trabajo se presentan abordajes para incorporar robustas herramientas de ciberseguridad en las redes industriales: Inteligencia Artificial, Forensia en vivo, Criptografía y Ciberdefensa desplegable.

Palabras clave: ciberseguridad en OT, Criptografía en OT, Forensia en OT, ciberdefensa en infraestructuras críticas, ciberdefensa desplegable, IA en OT.

Contexto

Este proyecto está radicado en el CAETI[1], centro de investigación dependiente de la Facultad de Tecnología Informática de la UAI (Universidad Abierta Interamericana). Los proyectos radicados en el CAETI se clasifican en tres líneas de investigación: Ingeniería de Software, Sociedad del Conocimiento y Automatización y Robótica. Este proyecto se enmarca dentro la línea de Automatización y Robótica. Se inició en abril de 2014.

Introducción

Desde mediados del siglo XVIII, las revoluciones industriales han producido una explosión demográfica: creció fuertemente la esperanza de vida y se redujo notoriamente la pobreza, el analfabetismo y la mortalidad infantil. Estos cambios se lograron gracias al incremento de la mayor disponibilidad de bienes económicos y de calidad [1].

[1] CAETI: https://caeti.uai.edu.ar/ .

En estos tiempos, el sistema productivo global está transitando una nueva fase que se caracteriza por la digitalización y la conectividad [2]. Se presenta como un nuevo paradigma para la industria llamada Industria 4.0 [3]. Industria 4.0 se considera ya como la *Cuarta Revolución Industrial*. Contempla la introducción de las tecnologías digitales en la industria de la fabricación: IoT, cloud, big data, IA, sensores inalámbricos, dispositivos móviles, sistemas embebidos, entre otros [Ybzunza,2017].

Industria 4.0 se basa en la integración entre la tecnología de operación de planta (OT según sus siglas en inglés) con las tecnologías de la información (IT según sus siglas en inglés) presentes en las redes corporativas, con las tecnologías digitales mencionadas, así también con las redes de clientes y proveedores. Su interconexión propone grandes ventajas competitivas que puede lograr niveles de eficiencia inéditos. Se espera también, que Industria 4.0 ayude a resolver algunos desafíos mundiales tales como la eficiencia energética, producción urbana y el cambio demográfico [3].

Pero las redes OT se crearon para funcionar en aislamiento físico: su interconexión con otras redes expondría sus vulnerabilidades, lo que supone serias consecuencias [4, 5, 6]. Para evitar la materialización de estos riesgos, Industria 4.0 requiere la investigación de estándares, procesos y soluciones de ciberseguridad robustas [3]. Mientras que en OT se apelaba a la ilusión de la seguridad por ocultamiento dada por el aislamiento físico, en IT se poseían antecedentes robustos en ciberseguridad: estándares como ISO/IEC27000 [7], NIST SP800-30 [8], criptografía, forensia informática, anti-virus, entre otros.

Es inmediato, entonces, cuestionarse la validez de llevar directamente la experiencia de IT a OT. Un reparo a este cuestionamiento puede hallarse en que en IT, la ciberseguridad se basa en tres pilares: disponibilidad, integridad y confidencialidad, mientras que en OT, importa la disponibilidad por sobre los otros pilares [5]. Un segundo reparo, más técnico que filosófico, es el hecho de que en el despliegue de un sistema OT, dado que se requiere interactuar con una vasta cantidad de dispositivos, se requiere validar una gran cantidad de librerías. A partir de ese momento, cualquier actualización en cualquier librería podría causar fallos o paradas inesperadas de planta. A esto se lo denomina *la rigidez de actualización*, que se contrapone con los estándares y buenas prácticas de IT. Entonces, por lo expuesto, no puede trasladarse directamente a OT las experiencias de IT,

Se desarrollaron estándares, buenas prácticas y herramientas específicas para OT: ISA/IEC62443 [9], NIST-SP800-82r3 [10], Forensia para OT (NIST-IR8428) [11], protocolo de aplicación seguro (Modbus Secure Protocol) [12]. Sin embargo, a pesar de su robustez, los niveles de adopción de estas medidas de ciberseguridad son muy bajos [13].

En este proyecto se estudian abordajes que podrían incrementar la ciberseguridad OT: uso de IA y forensia en vivo para tratar malware o anomalías ante la rigidez de actualización, implementación de criptografía para incorporar confidencialidad e integridad agregando así seguridad por diseño, y ciberdefensa desplegable para prevención y tratamiento de ataques o incidentes.

Líneas de Investigación, Desarrollo e Innovación

En el proyecto se identifican cuatro líneas de investigación:

La primera línea de investigación se denomina Uso de Inteligencia Artificial para la asistencia de Ciberseguridad en ICS. Dos abordajes de la ciberseguridad en OT surgen de esta línea: Uso de Modelos de Aprendizaje Automático para la detección de ransomware en instancias de preataque (se basa en el proyecto Infoscopia [14] más detalles en [15]) y uso de modelos de IA para la implementación de una DMZ industrial asistida por IA para soportar la Integración IT/OT.

La segunda línea de trabajo se denomina Forensia en OT. A diferencia del objetivo general de la Forensia en IT, que busca obtener evidencias acerca de qué es lo que sucedió, aquí se pone énfasis en el uso de herramientas y técnicas de Forensia en vivo (ya probado en [16]) para mejorar la prevención de incidentes frente a la imposibilidad de actualizar sistema operativo o software de seguridad sin arriesgar la disponibilidad del sistema.

La tercera línea de investigación busca la implementación efectiva, segura y conveniente de Criptografía en redes OT. Para ello, se propone la implementación de una infraestructura criptográfica de clave pública (PKI según sus siglas en inglés) securizada con factores adicionales y descentralizada eliminando así las principales debilidades de las PKI.

La cuarta línea de trabajo se denomina Ciberdefensa Desplegable: busca proponer dispositivos y métodos que puedan desplegarse en campo para atender incidentes y/o ataques en OT.

Resultados y Objetivos

Desde sus inicios en 2024 hasta la actualidad, el proyecto ha adquirido resultados remarcables. Aquellos anteriores a 2024 se detallan en [17].

Desde 2024 se destacan los siguientes hechos que marcaron los resultados recientes: discusiones acerca de Ciberdefensa Desplegable, securización de PKI, el desarrollo de un banco de pruebas OT y su posterior presentación en la feria Ekoparty[2] [18]

A continuación se detallan los resultados recientes.

- Publicaciones: en [15] se publicaron detalles en el abordaje del uso de IA para la detección temprana de Ransomware. En [19] se presentan detalles de un banco de pruebas OT portable desarrollado. En 2026 se planea la publicación de al menos media docena de artículos en: WICC, JAIIO, CYBER.AR y en algún Journal de Criptografía aún no definido.
- Tesis y trabajos finales aprobados: Se aprobaron los trabajos finales de Scussolin [20] y Bernardi [21]. Se avanzó en las tesis doctorales de Kamlofsky, Castro Tramontina y Manrique.
- Participación en Ekoparty (la conferencia de ciberseguridad más importante de Latinoamérica): entre los días 22 y 24/10/2025 se presentó un Stand con el banco de pruebas desarrollado y presentado en [19], en el sector SecureOT. Se incluyó demo de ataques. El día 24/10 se hizo una presentación en la sala C: “Ciberseguridad en Ambientes Industriales: Desafíos y Aportes” (disponible en Youtube[3] [22]).
- Convenios generales de colaboración (en proceso de negociación) entre la UAI con: 2 universidades nacionales, una empresa líder en automatización, una empresa líder del complejo siderúrgico.
- Desarrollos: se desarrolló un ICS portable para demo y experimentación.

El objetivo general del proyecto es:

[2] Ekoparty: https://ekoparty.org/

[3] Youtube; https://www.youtube.com/

- Proponer procesos y soluciones para mejorar la ciberseguridad en redes OT.

Los objetivos específicos más destacados son:

- Aplicar modelos de IA para identificar intentos de ataques por Ransomware o conexiones no autorizadas.
- Probar el uso de herramientas de Forensia en vivo para ayudar en la prevención de ataques en ICS.
- Implementar soluciones criptográficas que aseguren confidencialidad e integridad sin comprometer la disponibilidad.
- Desarrollar procesos y soluciones que permitan desplegarse en instalaciones industriales afectadas por incidentes de ciberseguridad.

Formación de Recursos Humanos

El proyecto está dirigido por el Mg. Lic. Jorge Kamlofsky quien está cursando un Doctorado. Los resultados colaborarán con el desarrollo de su tesis doctoral.

En el proyecto participan los siguientes investigadores: Daniel Manrique, José Castro Tramontina, Hernán Bottinelli, Jonatan Schmidt, Oscar Romero y los Auxiliares de investigación (alumnos): Gonzalo y Valentina Heinen, Mateo Wrobel y Juan Arias.

José Castro Tramontina y Daniel Manrique avanzan en el desarrollo de sus respectivas tesis doctorales. Oscar Romero y Hernán Bottinelli se encuentran realizando sus respectivas tesis de maestría.

Por otro lado, los auxiliares de investigación adquieren experiencias en el proyecto colaborando con desarrollos e implementaciones y tareas diversas.

Otras actividades han aportado a la capacitación de recursos humanos:

- Dictado de Cursos, Charlas y/o Seminarios de Criptografía y ciberseguridad.
- Entrevista realizada por el departamento de comunicaciones de la UAI en el marco de la presentación del proyecto en la Ekoparty.
- Clases Magistrales en la Diplomatura en Ciberseguridad (UAI).

Referencias

[1] Montagut Contreras, E. (2017). *La transición demográfica en la Revolución Industrial.* Los ojos de Hipatía.

[2] Basco, A. I., Beliz, G., Coatz, D., & Garnero, P. (2018). *Industria 4.0: fabricando el futuro* (Vol. 647). Inter-American Development Bank.

[3] Kagermann H., Wahlster W. & Helbig J. *Recommendations for implementing the strategic initiative INDUSTRIE 4.0*. Acatech.

[4] Sanchez, P. *Sistema de Gestión de la Ciberseguridad Industrial*. Trabajo Final de Master. Univ. Oviedo, (2013).

[5] Kamlofsky, J., Colombo, H., Sliafertas, M. y Pedernera, J. "Un Enfoque para Disminuir los Efectos de los Ciber-ataques a las Infraestructuras Críticas." III Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI, 2015), ISSN: 2346-9927. (2015).

[6] Ybzunza Cortes C. El Entorno de la Industria 4.0: Implicaciones y Perspectivas Futuras. En Conciencia Tecnológica n°54, 2017.

[7] ISO/IEC. (2018). Information technology — Security techniques — Information security management systems — Overview and vocabulary (ISO/IEC 27000:2018). International Organization for Standardization.

[8] NIST. "Special Publication 800 - 30, revision 1." Information Security. National Institute of Standards and Technology, U.S. Departament of Commerce, (2012).

[9] International Standard of Automation (ISA), ISA/IEC 62443 Series of Standards, (2020). En línea: https: // www.isa.org / standards – and – publications / isa – standards / isa – iec – 62443 – series – of - standards. [Consultado: 26/02/2024].

[10] Stouffer, K., Pease, M., Tang, C., Lightman, V., Cherdantseva, Y., & Pillitteri, V. (2023). Guide to Operational Technology (OT) Security (NIST Special Publication N.o 800-82,

Revision 3). National Institute of Standards y Technology.
[11] Salfati, E., & Pease, M. (2022, junio). Digital Forensics and Incident Response (DFIR) Framework for Operational Technology (OT). NIST Interagency/Internal Report N° 8428. National Institute of Standards y Technology (U.S.) Gaithersburg, MD.
[12] Modbus Organization. (2021). Modbus Security Protocol (Technical Specification).
[13] Control System Cyber Security Association International (CS)2 AI & KPMG. (2024, abril). Control System Cybersecurity Annual Report 2024. (CS)2 AI / KPMG.
[14] Liporace, Julio César, et al. "Metodología para el análisis de incidentes de ciberseguridad o ciberataques durante las acciones de ciberdefensa de las infraestructuras críticas de la defensa nacional–infoscopia–." *XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)*. 2019.
[15] Heinen, G., Milano, M. A., & Kamlofsky, J. (2025). Uso de técnicas de machine learning para la detección temprana de Ransomware. In *XXVII Workshop de Investigadores en Ciencias de la Computación (WICC). Mendoza, 10 y 11 de abril de 2025*.
[16] Kamlofsky, J., & Romero, R. O. (2022). Live Forensic Analysis on an ICS / SCADA. Sexta Conferencia Nacional de Informática Forense 2022, 30-38.
[17] Kamlofsky, J., Colombo, H. R., Milio, C., Romero, O., & Hecht, P. (2024). Ciberdefensa en sistemas operacionales. In *XXVI Workshop de Investigadores en Ciencias de la Computación (WICC). Puerto Madryn, 18 y 19 de abril de 2024*.
[18] Kamlofsky, J., Castro Tramontina J. & Manrique, D. (2025). Ciberseguridad en Ambientes Industriales: Desafíos y Aportes. Ekoparty 2025. En Researchgate.net
[19] Heinen, G. H., & Heinen, V. L. (2025). Prototipado de una Plataforma SCADA Portátil para el Análisis de Vulnerabilidades en Redes OT. *Revista Abierta de Informática Aplicada*, *9*(1), 124-140.
[20] Scussolin, L. A. (2024). Implementación de infraestructura de clave pública (PKI) en sistemas SCADA. Universidad Abierta Interamericana.
[21] Bernardi, F. A. (2025). Seguridad de entornos SCADA basada en infraestructura de clave pública PKI. Universidad Abierta Interamericana.

[22] Manrique, D., Castro, J. & Kamlofsky J. (2025). Ciberseguridad en Ambienes Industriales: Desafíos y Aportes, Video en Youtube: https://www.youtube.com/watch?v=Ukw2KRfa3js

Detección de Información de Identificación Personal (PII) en Argentina a través del Modelos de Inteligencia Artificial optimizados para el español

Zea Cardenas Milagros Ninfa, Barrionuevo Ileana Maricel, Fabian A. Gibellini, German N. Parisi, Analia L. Ruhl, Juliana M. Notreni, Leonardo R. Ciceri, Marcelo J. Auquer, Federico Bertola, Ignacio J. Sanchez Balzaretti

GISSIC - Grupo de Investigación en Seguridad de Sistemas de Información y Ciberseguridad. / Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas / Universidad Tecnológica Nacional / Facultad Regional Córdoba
Cruz Roja S/N, 5016

milyzc@gmail.com, ilebarrionuevo@gmail.com, fabiangibellini@gmail.com, germannparisi@gmail.com, analialorenaruhl@gmail.com, julinotreni@gmail.com, leonardorciceri@gmail.com, marcelo.auquer@gmail.com, fedebertola@gmail.com, ignaciojsb@gmail.com

Resumen

En un contexto global donde la ciberseguridad se ha consolidado como una prioridad estratégica, la protección de la Información de Identificación Personal (PII) se posiciona como uno de los principales desafíos para las organizaciones. La creciente digitalización ha generado un volumen exponencial de datos, entre los cuales la PII representa un subconjunto particularmente sensible, cuya exposición puede ocasionar consecuencias legales, económicas y reputacionales de gran envergadura.

En Argentina, la Ley 25.326 regula el tratamiento de datos personales, estableciendo principios y obligaciones orientadas a garantizar su resguardo. Sin embargo, el cumplimiento efectivo de esta normativa resulta complejo, especialmente cuando la PII se encuentra distribuida en documentos no estructurados, almacenados en formatos diversos, y, en ocasiones, inferible a partir de combinaciones de datos contextuales.

En este marco, el presente trabajo propone el desarrollo de un modelo de Inteligencia Artificial (IA) que identifique información sensible contenida en documentos organizacionales, adaptado al idioma español y alineado con la legislación argentina. Además, se plantea el desarrollo de una herramienta de benchmarking que permita evaluar el desempeño del modelo en términos de precisión, sensibilidad y aplicabilidad en entornos reales.

Palabras clave: PII, seguridad, IA, NLP, protección de datos, ciberseguridad, Argentina.

Contexto

A nivel internacional, el tratamiento y la protección de PII han sido abordados por múltiples normativas sectoriales, como HIPAA [1] o el GDPR [2] . Estas regulaciones han impulsado el desarrollo de herramientas tecnológicas —muchas de ellas basadas en técnicas de procesamiento de lenguaje natural y aprendizaje

automático— capaces de detectar y clasificar automáticamente datos sensibles. Investigaciones recientes exploran modelos de detección aplicables a contextos multilingües, aunque la mayoría de ellos se encuentran orientados a la lengua inglesa y a marcos normativos foráneos [3].

En el caso argentino, si bien la Ley 25.326 y su decreto reglamentario [4] constituyen un marco jurídico vigente y claro, no se observa aún una adopción extendida de herramientas automatizadas de detección de PII adaptadas al idioma español y al contexto legal local. Según Cimpanu [5] esto genera un vacío tecnológico que complica la trazabilidad de los datos personales y dificulta la respuesta ante incidentes de seguridad, en un entorno donde las filtraciones de datos se vuelven cada vez más frecuentes [6].

La identificación de PII en documentos organizacionales enfrenta obstáculos adicionales vinculados a la variedad de formatos (textos planos, PDFs, imágenes escaneadas), al uso de jergas sectoriales, y a la presencia de datos indirectos que permiten inferencias sensibles (Splunk). En este escenario, resulta fundamental avanzar hacia soluciones locales que integren conocimiento normativo, lingüístico y organizacional.

Introducción

Según el reporte 2023/2024 del Foro Internacional de Economía, a corto plazo (2 años), la inseguridad digital se coloca como el cuarto riesgo con impacto negativo en la economía y, a largo plazo (10 años, 2034), esta preocupación ocuparía el puesto 8 en su ranking de riesgos [7].

De acuerdo al mismo informe, la mayoría de los riesgos tecnológicos destacados a corto plazo, Desinformación e Inseguridad digital, bajan de ranking, pero permanecen en el top 10 a largo plazo, en el quinto y octavo lugar, respectivamente. El Foro Internacional de Economía define Inseguridad Digital como el uso de armas y herramientas cibernéticas para llevar a cabo guerra cibernética, espionaje y delitos cibernéticos para obtener control sobre una presencia digital y/o causar interrupción operativa.

Como ejemplos de fuga de datos se puede mencionar los 700 millones de datos filtrados de LinkedIn, los cuales incluían datos de ubicación de usuarios (ciudad, provincia, país), nombre de usuarios y correo electrónico [5].

El proceso de digitalización en el que se encuentran inmersas las organizaciones ha traído consigo beneficios significativos en eficiencia, productividad y accesibilidad a la información. Sin embargo, este avance también ha generado nuevos riesgos en materia de seguridad de la información, siendo uno de los más críticos la exposición de Información de Identificación Personal (PII).

Toda organización que maneja información de identificación personal (PII) tiene el deber de cuidar a sus usuarios y clientes para asegurarse de que la PII no quede expuesta al mundo exterior y, al mismo tiempo, cumplir con los requisitos normativos, de seguridad y de cumplimiento para mantenerla interna e intacta. El desafío que se presenta es que la PII aparece en lugares inesperados en los archivos de registro de una organización e identificarla puede ser complicado. La PII directa, como direcciones de correo electrónico o números de tarjetas de crédito, puede ser bastante sencilla de identificar observando la forma en que está formateada. Pero la PII también puede ser indirecta y menos fácil de identificar a partir de datos sin procesar, como información médica,

historial laboral o información descriptiva sobre un individuo [8].

La PII se define como cualquier dato que pueda ser utilizado para identificar directa o indirectamente a una persona física. Esto incluye no solo datos obvios como el nombre, la dirección o el número de documento, sino también datos contextuales o combinados, como direcciones IP, registros médicos o datos biométricos [1]. La exposición de estos datos puede tener consecuencias legales, financieras y sociales de gran alcance, como el robo de identidad, el fraude financiero o la pérdida de confianza en instituciones [9].

La región ha experimentado incidentes de fuga de datos de gran escala en los últimos años, que han afectado tanto a organismos públicos como a empresas privadas, evidenciando la urgencia de fortalecer las capacidades técnicas para proteger la PII; dentro de estas fuga de datos se puede mencionar la de RENAPER en el 2023 y 2024 [10] [11] [12]

En este contexto, desarrollar soluciones que integren Inteligencia Artificial (IA) aplicadas a la ciberseguridad se vuelve una necesidad estratégica para resguardar la privacidad de los ciudadanos y el cumplimiento normativo de las organizaciones [7][13].

Esta aseveración es parte del objetivo propuesto en el proyecto "Desarrollo de un modelo de IA para detección de Personal Identificable Information (PII)"; con código COSIEC611, en el cual esté inserto el presente trabajo. Parte de las pruebas de concepto de este proyecto incluye probar/catalogar herramientas actuales de detección de patrones de datos que sean de código abierto. Siendo la primera Presidio [14], y del cual se exhiben a continuación los avances sobre el mismo utilizando el idioma español.

Este trabajo propone un enfoque híbrido basado en Procesamiento de Lenguaje Natural (NLP) en español, contenedorización con Docker y reconocedores personalizados en Microsoft Presidio, adaptado al contexto argentino. El Procesamiento del Lenguaje Natural (PLN) es una rama de la Inteligencia Artificial y la Lingüística, dedicada a hacer que las computadoras comprendan las declaraciones o palabras escritas en lenguajes humanos [15].

Se busca responder a una brecha tecnológica clara: la falta de soluciones que integren requerimientos normativos locales, características lingüísticas del español rioplatense y escalabilidad técnica para entornos organizacionales.

Líneas de Investigación y Desarrollo

Esta investigación se basa en los conocimientos más avanzados y contemporáneos en el campo de la ciberseguridad, con un enfoque específico en la protección de Información de Identificación Personal (PII). La línea de investigación principal será seguridad informática, pero integra además de forma transversal técnicas y algoritmos provenientes del ámbito de Machine Learning, convirtiendo en un proyecto multitemático que combina disciplinas clave para abordar esta problemática.

El enfoque metodológico es empírico, basado en el método científico para la recolección y análisis de datos necesarios para el desarrollo del modelo de Inteligencia Artificial (IA) destinado a la detección automatizada de PII. Este modelo busca garantizar la trazabilidad y protección de datos sensibles, cumpliendo con las normativas locales y adaptándose a las particularidades del idioma español y los entornos organizacionales argentinos.

Este trabajo se desarrolla bajo los principios de Investigación, Desarrollo e

Innovación (I+D+i), integrando una visión estratégica que combina la teoría y la práctica con un enfoque orientado a resultados. Las actividades se estructuran abarcando: síntesis y teoría; exploración, hipótesis y clarificación; diseño, desarrollo y prueba; implementación; y mejora continua basada en la evaluación de la eficacia del producto final.

Resultados y Objetivos

La solución desarrollada se basa en la integración de Microsoft Presidio [14], una plataforma open-source de detección y anonimización de PII, con un entorno contenerizado de servicios mediante Docker [16]. La arquitectura de esta prueba de concepto consta de tres componentes principales:

- Motor NLP: encargado del preprocesamiento y análisis lingüístico de textos, utilizando utilizando modelos de Procesamiento de Lenguaje Natural (NPL). El modelo seleccionado para la prueba de concepto es es_core_news_md [17] ya que es uno de los modelos ya entrenados y disponibilizados por spaCy [18] para aplicarlo sobre el idioma español.
- Servicio de detección (Presidio Analyzer y Anonymizer): realiza la identificación de entidades sensibles.
- API REST: interfaz que expone los servicios a aplicaciones externas mediante endpoints simples y escalables [19].

Otra de las contribuciones técnicas de este trabajo es la creación de reconocedores específicos para el contexto argentino, en particular para DNI y CUIL/CUIT.

Mientras que Microsoft Presidio cuenta con detectores genéricos (emails, números de teléfono internacionales, tarjetas de crédito), fue necesario ampliar la base de patrones para incluir patrones adaptados de:

- DNI
- CUIL/CUIT

Estos reconocedores fueron desarrollados en Python utilizando la clase PatternRecognizer y probados con documentos de prueba que simulan contratos laborales, registros administrativos y formularios digitales.

Es importante destacar que las contribuciones técnicas presentadas en el presente son anexadas a este componente antes de su ejecución para su correcto procesamiento de los patrones buscando, teniendo en cuenta su aplicabilidad sea dentro del territorio argentino.

Formación de Recursos Humanos

Los involucrados en el proyecto son docentes e investigadores de la Universidad Tecnológica Nacional - Facultad Regional Córdoba (UTN - FRC). También son integrantes del Grupo de Investigación en Seguridad de Sistemas de Información y Ciberseguridad (GISSIC). Participan además alumnos becarios de la carrera de Ingeniería en Sistemas de Información. Los cuales pueden iniciar su formación en investigación científica y tecnológica.

Se organizan charlas, talleres y conferencias tanto a nivel provincial como nacional, con el objetivo de compartir los resultados del proyecto.

Referencias

[1] NIST SP 800-122. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). NIST, 2010.

[2] General Data Protection Regulation GDPR. (n.d.). GDPR-Info.eu.

https://gdpr-info.eu/. Última visita 10 de febrero de 2026.
[3] Silva, P., Gonçalves, C., Godinho, C., Antunes, N., & Curado, M. (2020, julio). Using NLP and machine learning to detect data privacy violations. En IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (pp. 972–977). IEEE. https://estudogeral.uc.pt/bitstream/10316/93821/1/WORKSHOP_ON_SECURITY_AND_PRIVACY_IN_BIG_DATA__Camera_Ready.pdf (Última visita: 10 de febrero de 2026).
[4] Gobierno de Argentina. (2000). Ley 25.326 - Protección de datos personales. https://www.argentina.gob.ar/normativa/nacional/ley-25326-64790/actualizacion (Última visita: 10 de febrero de 2026).
[5] Cimpanu, C. (2021). Hackers leak LinkedIn 700 million data scrape. The Record. https://therecord.media/hackers-leak-linkedin-700-million-data-scrape (Última visita: 10 de febrero de 2025)
[6] Microsoft Data Breaches: Full Timeline Through 2024. M. X. Heiligenstein. Febrero 2024. https://firewalltimes.com/microsoft-data-breach-timeline/, última visita 28/04/2024
[7] World Economic Forum. The Global Risks Report 2024. 19th Edition. https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf, última visita 28/04/2024.
[8] Defining and detecting Personally Identifiable Information (PII) in log data. Splunk Platform, última visita 8/02/2026.
[9] ISO/IEC 27001. Information Security Management Systems – Requirements. ISO, 2013.
[10] RENAPER solo expone la deficiencias del Estado Argentino en Ciberseguridad. https://blog.segu-info.com.ar/2024/04/renaper-solo-expone-la-deficiencias-del.html?m=0, última visita 28/08/2025
[11] Fuga de datos en el RENAPER. https://revistainnovacion.com/nota/12307/fuga_de_datos_en_el_renaper, última visita 28/08/2025.
[12] Descargar todas las fotos del Padrón Electoral Argentino. https://blog.segu-info.com.ar/2013/10/descargar-todas-las-fotos-del-padron.html#axzz2kLhdUSq1, última visita 28/08/2025
[13] CISCO. Data Privacy Benchmark Study 2023.
[14] Microsoft Presidio. Open-Source PII Protection Tool. https://github.com/microsoft/presidio, última visita 28/08/2025.
[15] Khurana, D., Koli, A., Khatter, K., & Singh, S. (2017). Natural Language Processing: State of the Art, Current Trends and Challenges (arXiv:1708.05148) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1708.05148, última visita 28/08/2025.
[16] Docker. Documentation and Containerization Best Practices. https://docs.docker.com, última visita 28/08/2025
[17] Modelo es_core_news_md https://spacy.io/models/es#es_core_news_md, última visita 28/08/2025
[18] spaCy. Industrial-Strength NLP in Python. https://spacy.io/, última visita 28/08/2025
[19] FastAPI. Modern Web Framework for APIs. https://fastapi.tiangolo.com/, última visita 28/08/2025

Disruptive Compliance Toolkit: Framework para Programas de Compliance de Ciberseguridad en Contextos Disruptivos

Mauro Stepanoski; Juan Pablo Estelles;

CAETI – Universidad Abierta Interamericana
DV. Montes de Oca 725 – Buenos Aires - Argentina
{mastepanoski@gmail.com; estellesjp@gmail.com}

RESUMEN

La adopción acelerada de tecnologías disruptivas —como la inteligencia artificial, la computación cuántica y los vehículos autónomos— evidencia una limitación estructural en los enfoques tradicionales de compliance, apoyados en incidentes históricos y controles estáticos. En este contexto, el **Disruptive Compliance Toolkit (DCT)** se propone como un framework para diseñar y operar programas de gestión de riesgos de ciberseguridad orientados específicamente a tecnologías emergentes sin precedentes claros.

El eje metodológico del DCT es la metodología **Disruptive Scenario Builder (DSB)**, que permite modelar escenarios de riesgo posibles mediante análisis sistémico de entradas, procesos y salidas, ejercicios de simulación de escenarios tecnológicos emergentes combinando análisis cualitativo y estimaciones probabilísticas, en lugar de limitarse a proyectar el pasado hacia el futuro. Sobre esta base, el framework articula líneas de trabajo que incluyen la construcción de taxonomías dinámicas de riesgos, modelos de madurez y métricas de efectividad adaptativas, variantes de implementación según tamaño organizacional e interoperabilidad con estándares internacionales como ISO 27001, ISO 37301 y marcos NIST.

Como resultado, el DCT ofrece un enfoque práctico, escalable y anticipatorio que busca reducir el **time-to-compliance** y mejorar la capacidad de las organizaciones para enfrentar riesgos asociados a tecnologías disruptivas, manteniendo la trazabilidad normativa y la validación humana como elementos centrales.

Palabras clave: compliance, ciberseguridad, gestión de riesgos, tecnologías disruptivas

CONTEXTO

El presente trabajo se inscribe en una iniciativa de investigación aplicada desarrollada por profesionales del sector privado con foco en ciberseguridad y compliance tecnológico. El proyecto surge de la práctica concreta de consultoría en organizaciones de distintos tamaños y sectores en Argentina, donde se identificó la brecha existente entre los marcos normativos internacionales y la capacidad operativa real de las organizaciones para adoptarlos frente a tecnologías emergentes. El trabajo se presenta en el contexto del WICC 2026 como una contribución al campo de la gestión de riesgos en tecnologías disruptivas, buscando el intercambio con la comunidad académica y científica

1. INTRODUCCIÓN

La digitalización acelerada de procesos críticos ha impulsado la adopción de tecnologías en sectores

tan diversos como finanzas, salud, logística e industria. Este avance viene acompañado de un panorama de amenazas mucho más dinámico, donde aparecen vectores de ataque que no tienen antecedentes claros en los registros históricos de incidentes de ciberseguridad, como los escenarios Harvest-Now-Decrypt-Later asociados a la criptografía post-cuántica o los ataques de prompt injection sobre modelos de lenguaje de gran escala. En este contexto, los marcos clásicos de gestión de riesgos, fuertemente apoyados en estadísticas pasadas y controles relativamente estáticos, muestran limitaciones para anticipar y tratar riesgos verdaderamente nuevos [1][2].

En paralelo, los estándares y marcos normativos internacionales han comenzado a incorporar lineamientos específicos para el tratamiento de riesgos en tecnologías emergentes, tanto desde la ciberseguridad como desde la perspectiva de compliance y responsabilidad organizacional. Entre ellos se destacan las normas de gestión de riesgos y de seguridad de la información (ISO/IEC 27005, ISO/IEC 27001), los marcos de gestión de riesgos de NIST, y las regulaciones recientes vinculadas a inteligencia artificial y protección de datos personales, como el AI Act europeo y el GDPR [11][14][15]. Sin embargo, su adopción práctica en organizaciones pequeñas y medianas suele verse obstaculizada por la complejidad de los requisitos, la falta de recursos especializados y la dificultad para traducir principios generales en decisiones operativas concretas.

El **Disruptive Compliance Toolkit (DCT)** se propone como una respuesta a esta brecha entre la teoría normativa y la práctica cotidiana de gestión de riesgos en tecnologías disruptivas. El proyecto combina una metodología específica para la construcción de escenarios de riesgo sin precedentes —**Disruptive Scenario Builder (DSB)**— con una plantilla operativa de programa de compliance que integra evaluación de riesgos, diseño de controles, gobernanza, capacitación, monitoreo y comunicación. El objetivo es ofrecer a las organizaciones un marco práctico, adaptable al tamaño y nivel de madurez, que permita identificar riesgos emergentes de manera estructurada, priorizar inversiones en seguridad con base en evidencia y mantener la alineación con estándares y regulaciones internacionales en constante evolución [3][4].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El Disruptive Compliance Toolkit (DCT) se estructura en torno a seis líneas principales de investigación y desarrollo interrelacionadas, diseñadas para abordar los desafíos emergentes en la gestión de riesgos de tecnologías disruptivas-

2.1 Metodología de Diseño de Escenarios Disruptivos (DSB):

Constituye el núcleo metodológico del DCT. Se centra en el desarrollo y refinamiento de técnicas para modelar riesgos emergentes sin precedentes históricos, utilizando análisis sistémico de entradas-procesos-salidas (EPO), simulaciones de escenarios adversariales y estimación bayesiana de probabilidades. La investigación aborda la identificación de riesgos únicos en tecnologías como Inteligencia Artificial, Computación Cuántica y Vehículos Autónomos, generando escenarios hipotéticos validables mediante simulación y análisis de expertos [1][2].

2.2 Taxonomía y Caracterización Dinámica de Riesgos Tecnológicos Emergentes:

Se investiga la clasificación sistemática de amenazas específicas de tecnologías disruptivas,

diferenciando entre riesgos transversales y riesgos específicos de cada dominio tecnológico [3][4]. Esta línea desarrolla matrices de riesgos actualizables que incluyen vectores de ataque novedosos como Harvest-Now-Decrypt-Later (HNDL) en criptografía post-cuántica [5], Prompt Injection en modelos de lenguaje [6], y ataques de fusión sensorial en sistemas autónomos. El objetivo estratégico es crear un framework con capacidades de Inteligencia Artificial para monitorear continuamente avances tecnológicos y proponer actualizaciones automáticas a la taxonomía de riesgos.

2.3 Modelos de Madurez, Métricas de Efectividad y Compliance Adaptativo:

Desarrolla frameworks de evaluación continua que permitan a las organizaciones medir su progreso desde estados iniciales de ausencia de controles hasta niveles de resiliencia anticipatoria. Se investigan indicadores clave de desempeño (KPIs) dinámicos, indicadores de riesgo (KRIs) predictivos con capacidades de machine learning y métricas de compliance automatizables mediante integración con plataformas GRC, SIEM y herramientas de threat intelligence [7][8].

2.4 Adaptación de Compliance a Contextos Organizacionales Diversos:

Se investigan estrategias de implementación diferenciadas según características organizacionales, con énfasis en startups tecnológicas, PYMEs y corporaciones multinacionales. Esta línea aborda la simplificación de procesos de compliance, el diseño de arquitecturas de gobernanza ágiles compatibles con metodologías DevOps/DevSecOps, y la integración con principios Zero Trust [9][10].

2.5 Interoperabilidad con Estándares Internacionales y Marcos Regulatorios:

Mapea sistemáticamente el DCT con estándares globales de compliance [11][12][13] y regulaciones emergentes [14][15]. Los objetivos incluyen la reducción del time-to-compliance organizacional entre 30-40% mediante plantillas prevalidadas y la creación de mapeos de trazabilidad automatizados entre controles del DCT y requisitos normativos específicos, facilitando auditorías y certificaciones.

2.6 Inteligencia Artificial para Orquestación y Autogobierno de Programas de Compliance:

Explora el desarrollo de sistemas de IA capaces de integrar las líneas 1-5 para crear programas de compliance autónomos y auto-actualizables [4] [7]. Los agentes inteligentes monitorearan continuamente el ecosistema tecnológico global, identificarán tecnologías emergentes relevantes y generarán propuestas de actualización de políticas y controles. La validación humana permanece como requisito crítico mediante mecanismos de human-in-the-loop, donde el Comité de Compliance aprueba, modifica o rechaza las propuestas automatizadas antes de su implementación.

3. RESULTADOS Y VALIDACIÓN

El desarrollo del DCT representa un avance teórico-metodológico en la gestión de riesgos de tecnologías disruptivas, con resultados consolidados en el diseño del framework y estrategias de validación empírica colaborativa propuestas para desarrollo futuro.

3.1 Resultados metodológicos consolidados

La metodología DSB ha sido formalizada como un proceso estructurado de cinco fases que

permite modelar riesgos emergentes sin precedentes históricos [1][2], integrando análisis sistémico EPO, simulación de escenarios tecnológicos emergentes combinando análisis cualitativo y estimación bayesiana. Se ha completado el desarrollo de una taxonomía de riesgos tecnológicos emergentes actualizada a 2025, que incorpora vectores de ataque como ataques HNDL en criptografía cuántica, inyección de prompts en modelos de lenguaje, ataques de fusión sensorial en sistemas autónomos y compromiso de estaciones terrestres en infraestructura satelital.

El marco de interoperabilidad con estándares internacionales ha sido desarrollado mediante análisis comparativo con ISO 37301, ISO/IEC 27001, NIST Cybersecurity Framework 2.0 y regulaciones emergentes [11][12][14][15], incluyendo análisis de brechas y guías de adaptación. Los casos de estudio ilustrativos demuestran la aplicación conceptual del DCT en cuatro dominios: IA para scoring crediticio, Computación Cuántica para datos genómicos, Vehículos Autónomos para logística y Comunicaciones Satelitales para operaciones mineras, con proyecciones de ROI entre 8x-35x y reducción de time-to-compliance del 30-40% —valores pendientes de validación empírica—.

3.2 Estrategia de validación colaborativa

Una vez consolidado el framework, se prevé su distribución como recurso de acceso abierto, incluyendo documentación, plantillas operacionales y herramientas de soporte para facilitar su adopción. La estrategia busca generar retroalimentación de múltiples implementaciones en contextos organizacionales diversos, lo que permitirá refinar la metodología con base en evidencia empírica real. Se propone el establecimiento de mecanismos estructurados para capturar experiencias de adopción, identificar patrones de éxito y fracaso, y documentar adaptaciones exitosas a industrias o tecnologías no contempladas inicialmente.

Como línea futura, se identifica el desarrollo de sistemas de IA para la orquestación de programas de compliance [4][7]. Su avance dependerá de la consolidación previa de validaciones empíricas que provean los datos necesarios para el entrenamiento de modelos y la verificación de las hipótesis fundamentales del framework

FORMACIÓN DE RECURSOS HUMANOS

Mauro Stepanoski es Ingeniero de Sistemas graduado en la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN) y cuenta con una Diplomatura en Compliance en Seguridad Informática otorgada por la Universidad Abierta Interamericana (UAI), con foco en análisis y gestión de riesgos y aplicación de normas ISO/IEC. Se desempeña en el sector privado como consultor estratégico en ciberseguridad, arquitectura de software y gestión tecnológica, con más de 20 años de experiencia y participación en más de 400 proyectos. Es co-fundador de TAI Dynamics y creador de la iniciativa Cumplimiento Digital + Ciberseguridad Argentina. Sus áreas de especialización incluyen compliance en ciberseguridad, modelado de negocios, arquitectura de sistemas, inteligencia artificial y gestión de riesgos tecnológicos.

Juan Pablo Estelles es Analista de Sistemas y estudiante avanzado de Ingeniería en Sistemas Informáticos. Se desempeña profesionalmente como Backend Software Developer en el sector privado. Su formación y práctica se orientan a la ciberseguridad y el compliance tecnológico, áreas en las que ha desarrollado trabajos de

investigación aplicada. Previo a su trayectoria en tecnología, fue atleta de alto rendimiento y representó a Argentina en competencias internacionales, incluyendo los Juegos Olímpicos de Río 2016.

El proyecto no registra tesis de posgrado directamente asociadas. Los autores desarrollan el framework desde la práctica profesional, con el objetivo de aportar instrumentos operativos validables a la comunidad académica y a las organizaciones que gestionan riesgos en tecnologías disruptivas.

BIBLIOGRAFIA

[1] ISO/IEC 27005:2022. Information security, cybersecurity and privacy protection — Information security risk management. International Organization for Standardization.

[2] NIST. (2023). Risk Management Framework for Information Systems and Organizations (SP 800-37 Rev. 2). National Institute of Standards and Technology.

[3] ENISA. (2025). ENISA Threat Landscape 2025. European Union Agency for Cybersecurity.

[4] NIST. (2024). Artificial Intelligence Risk Management Framework (AI RMF) 2.0. National Institute of Standards and Technology.

[5] KPMG. (2025). Quantum security report 2025. KPMG International.

[6] Daniel, G. (2023). Prompt injection: The new SQL injection. OWASP.

[7] MITRE. (2024). MITRE ATT&CK Framework. MITRE Corporation.

[8] NIST. (2025). Security and Privacy Controls for Information Systems and Organizations (SP 800-53 Revision 6). National Institute of Standards and Technology.

[9] NIST. (2020). Zero Trust Architecture (SP 800-207). National Institute of Standards and Technology.

[10] NIST. (2025). Implementing a Zero Trust Architecture (SP 1800-35). National Institute of Standards and Technology.

[11] ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection — Information security management systems — Requirements. International Organization for Standardization.

[12] ISO 37301:2021. Compliance management systems — Requirements with guidance for use. International Organization for Standardization.

[13] ISO/IEC 23894:2023. Information technology — Artificial intelligence — Guidance on risk management. International Organization for Standardization.

[14] European Parliament and Council. (2024). Regulation (EU) 2024/1689 on artificial intelligence (AI Act). Official Journal of the European Union, L 252/1.

[15] European Parliament and Council. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data (GDPR). Official Journal of the European Union, L 119/1

Evaluación de vulnerabilidades y desempeño en redes IoT en un entorno experimental controlado

Brian Llamocca[1], Aldana Tedesco[1], Paula Venosa[2], Sofía Martin[2], Patricia Bazán[2], Agustín Candia[2]

1- UNLP-Facultad de Informática, 2- LINTI UNLP-Facultad de Informática

llamoccas.brian@gmail.com, aldanatedesco28@alu.ing.unlp.edu.ar, pvenosa@info.unlp.edu.ar, pbaz@info.unlp.edu.ar, smartin@linti.unlp.edu.ar, acandia@linti.unlp.edu.ar

RESUMEN

El trabajo presenta una línea de investigación orientada al estudio de la seguridad en entornos IoT mediante la construcción de un entorno experimental controlado y de bajo costo. El estudio se justifica por el crecimiento sostenido de ataques a estos dispositivos, la baja visibilidad de los mismos en redes organizacionales y la heterogeneidad de tecnologías y fabricantes, que incrementan la superficie de ataque.

Se implementa un ecosistema basado en dispositivos Bluetooth Low Energy (BLE) en modo beacon y tecnologías LoRa/LoRaWAN, con el fin de analizar tráfico, identificar vulnerabilidades y ensayar mecanismos de mitigación.

Los resultados alcanzados confirman la viabilidad del entorno propuesto para realizar evaluaciones de seguridad en redes IoT y promover la definición de estrategias de ciberseguridad y su implementación.

CONTEXTO

La línea de investigación que se presenta, propone profundizar la investigación de tácticas, técnicas y procedimientos de los atacantes a nivel regional mediante el uso de honeypots. En este contexto, se abrió una línea de trabajo para construir un marco teórico y analizar el estado del arte del uso de honeypots y honeynets como herramienta para estudiar comportamiento malicioso poniendo énfasis, en las amenazas que afectan a los dispositivos de IoT (Internet de las Cosas) y a sus protocolos. Este trabajo presenta las primeras etapas de investigación en torno a IoT, con el objetivo final de contar con un ecosistema prototipo de una red de sensores con comunicación inalámbrica sin internet. Este ecosistema se utilizó para analizar tráfico, evaluar vulnerabilidades existentes y establecer estrategias de ciberseguridad.

1- INTRODUCCIÓN

Con el crecimiento de IoT y la proliferación de dispositivos, es fundamental analizar las tendencias de los ataques que tienen a éstos como objetivo.

Según un informe de Forrester Consulting[1], el 69% de las empresas tienen más dispositivos IoT en sus redes que computadoras, el 84% de los profesionales de la seguridad creen que los dispositivos IoT son más vulnerables que las computadoras, el 67% de las empresas han experimentado un incidente de seguridad de IoT, solo el 16% de los gerentes de seguridad empresarial dicen que tienen una visibilidad adecuada de los dispositivos IoT en sus entornos y el 93% de las organizaciones

[1] https://www.iot-now.com/wp-content/uploads/2025/06/BR_Locking-in-Value-with-IoT-Security_Final50.pdf

planean aumentar su gasto en seguridad para IoT y dispositivos no administrados.

El informe de Beecham Research del año 2025[2] destaca un aumento interanual superior al 400% en ataques de malware dirigidos a IoT, con un incremento del 107% solo en los primeros cinco meses de 2024.

Esta tendencia, sumada a que el 54% de las organizaciones reporta intentos de ciberataques semanales a sus dispositivos, subraya la necesidad crítica de reforzar las estrategias de seguridad.

Con IoT es posible interconectar objetos físicos a través de redes de comunicación. Estos objetos se comunican a través de protocolos específicos y responden a un modelo de capas que incluye el medio físico, el transporte y el nivel de aplicación.

Algunas de las características de las redes de IoT son: (1) limitaciones de recursos y configuración en la nube, (2) problemas con la captura de tráfico y registro de datos, (3) riesgos de seguridad y contención, (4) complejidades de configuración y despliegue de honeypots personalizados, (5) problemas técnicos específicos, entre otras [5].

La heterogeneidad de protocolos y fabricantes, así como las limitaciones de seguridad en las capas inferiores del modelo IoT, facilitan la aparición de vulnerabilidades explotables. Entre las amenazas más comunes se destacan la intercepción de comunicaciones, la suplantación de identidad de dispositivos y la manipulación de datos transmitidos [6].

Dispositivos IoT y protocolos

Los dispositivos IoT de bajo costo, disponibles en el mercado —como los basados en Beacon BLE (Bluetooth Low Energy) y LoRa (Long Range)—, ofrecen una alternativa accesible para el desarrollo de prototipos, proyectos educativos y entornos de experimentación. No obstante, la limitada capacidad de procesamiento y almacenamiento de estos dispositivos, junto con la simplicidad de sus protocolos, pueden implicar vulnerabilidades críticas [7] [8].

Bluetooth[3] es un estándar de comunicación inalámbrica universal diseñado para el intercambio de voz y datos a corta distancia, operando en la banda ISM (Industrial, Científica y Médica) de 2.4 GHz. Funciona como una Tecnología de Red de Área Personal Inalámbrica (WPAN), permitiendo la conexión de hasta siete dispositivos esclavos a un dispositivo maestro.

En el modo beacon, que se incorpora en la versión 4.0, es un modo de funcionamiento de Bluetooth Low Energy (BLE), un dispositivo actúa como emisor de señales periódicas (*advertising packets*), sin necesidad de establecer una conexión activa con otro dispositivo y además se amplía la cantidad de dispositivos esclavos que pueden conectarse.

LoRa (Long Range) es una tecnología inalámbrica que define únicamente la capa física de una red LPWAN. Desarrollada por Semtech en sus inicios en la actualidad está administrada pLoRa Alliance[4]. Se diseñó con el objetivo de poder establecer una conexión punto a punto con grandes áreas de cobertura y bajo consumo asumiendo que las tasas binarias serían reducidas.

2- LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las principales líneas de trabajo en las que el grupo de investigación desarrolla sus actividades y consolida su formación, resultan de la continuación del trabajo realizado en

[2] https://www.beechamresearch.com/

[3] Bluetooth es marca registrada de Bluetooth Special Interest Group (Bluetooth SIG)

[4] https://lora-alliance.org/

torno a estrategias de ciberseguridad distribuida [3] y el análisis de técnicas de ataque y comportamientos maliciosos [2].

A partir de estas líneas se plantea su continuidad en torno a dispositivos de IoT, para evaluar la seguridad y rendimiento de redes de IoT de bajo costo en un entorno controlado para simular casos de uso reales y realizar pruebas de seguridad éticas.

3- RESULTADOS ESPERADOS/OBTENIDOS

El objetivo de la línea de investigación presentada es desarrollar y evaluar un entorno experimental de red IoT utilizando dispositivos de bajo costo que implementen protocolos de comunicación unidireccionales (Beacon, BLE) y bidireccionales (LoRa), con el propósito de conformar un ecosistema funcional a pequeña escala que permita simular casos de uso representativos, analizar su comportamiento operativo y realizar ensayos de seguridad orientados a la detección de vulnerabilidades y la propuesta de medidas de mitigación.

Pruebas con BLE

Las pruebas se realizaron sobre módulos ESP32[5] ampliamente utilizados por su estabilidad y por incluir soporte nativo para BLE 4.2. Estos dispositivos permiten evaluar en detalle el comportamiento del modo beacon, el intervalo de transmisión y la capacidad de cifrado del payload en tiempo real.

Junto con ellos se empleó una Raspberry Pi 5 (8 GB de RAM) como gateway BLE. Por su adaptador Bluetooth integrado y su mayor capacidad de procesamiento, la Raspberry pudo operar como un receptor continuo de beacons, demostrando un rendimiento adecuado incluso en escenarios con múltiples dispositivos transmitiendo simultáneamente.

Se comenzaron las pruebas con la placa Raspberry Pi actuando como gateway BLE, capturando los mensajes y mostrándolos sin procesamiento, mientras que el ESP32 transmitía beacons sin ningún tipo de cifrado. Esta primera etapa permitió validar el funcionamiento del enlace y del proceso de *advertising*. Se evidenció una vulnerabilidad crítica debido a la transmisión de texto abierto.

A partir de esta observación, se avanzó hacia un segundo conjunto de pruebas centrado en la seguridad. Para ello se definió una Master Key compartida, almacenada tanto en el *firmware* del ESP32 como en el *gateway*. El objetivo inicial era evaluar distintos mecanismos para rotar o utilizar esta clave en el cifrado de los mensajes, garantizando que solo dispositivos autorizados pudieran generar información válida.

En una primera aproximación, este esquema criptográfico se implementó sobre GATT[6] - BLE el cual define dos perfiles de comunicación entre dispositivos: GAP y GATT. Cada uno cumple un rol distinto dentro del proceso de descubrimiento y transferencia de datos, aprovechando su canal de comunicación estructurado y su capacidad de gestionar características cifradas. Si bien GATT permitió validar el uso de la Master Key y del cifrado en un entorno controlado su utilización implicaba la necesidad de establecer conexiones BLE, aumentando significativamente el consumo energético y la complejidad del sistema.

[5] El ESP32 es un sistema en chip (SoC) de bajo costo y consumo de energía con Wi-Fi y Bluetooth integrado, desarrollado por Espressif Systems. Datasheet: https://documentation.espressif.com/esp32_datasheet_en.pdf

[6] GATT (Generic Attribute Profile) es el perfil de Bluetooth Low Energy que define la estructura y el intercambio de datos entre dispositivos mediante servicios y características, utilizando el protocolo ATT (Attribute Protocol) como base

Por ese motivo, las pruebas evolucionaron hacia el *modo beacon*, manteniendo la misma lógica de clave compartida, pero eliminando la necesidad de conexiones persistentes. Este cambio permitió conservar el bajo consumo característico del advertising y, al mismo tiempo, preservar la confidencialidad del mensaje.

Finalmente, se avanzó con la etapa más madura del diseño: cifrar directamente el payload del beacon, incorporando no solo el cifrado simétrico basado en la Master Key, sino también un mecanismo de integridad (HMAC). Con esta solución se logró un esquema de comunicación liviano y unidireccional, que mejora el nivel de seguridad del ecosistema prototipo de red de sensores analizado.

Pruebas con Lora/LoRaWAN

Las pruebas realizadas permitieron obtener una comprensión detallada del funcionamiento del protocolo LoRaWAN, tanto desde el punto de vista operativo como de seguridad. quA partir de la transmisión periódica de un mensaje con valores secuenciales, se logró evaluar el flujo completo de comunicación, desde el proceso de activación del nodo hasta la recepción final en el backend.

Uno de los resultados más relevantes fue verificar que la comunicación LoRaWAN opera bajo un esquema de cifrado robusto, basado en la generación dinámica de claves de sesión mediante el mecanismo OTAA[7]. Se pudo observar que los mensajes que llegaban al servidor estaban correctamente cifrados y solo eran interpretables una vez procesados por el Application Server que disponía de la AppSKey correspondiente. Esto confirmó que un atacante que interceptara los paquetes —incluso teniendo acceso al tráfico desde el gateway— no podría descifrar ni manipular su contenido. La imposibilidad práctica de ejecutar un ataque del tipo Man in the Middle quedó así demostrada.

En cuanto al control de acceso, se comprobó que el servidor únicamente acepta tráfico proveniente de dispositivos dados de alta previamente, utilizando el DevEUI registrado. Durante las pruebas, los intentos de transmisión desde dispositivos no autorizados o mal configurados fueron sistemáticamente descartados, lo que demuestra la eficacia del mecanismo de autenticación.

El hecho de que el UID del dispositivo esté definido tanto en el firmware como en el servidor garantiza que solo los equipos legítimos pueden establecer una sesión válida y generar claves de sesión. Esto aporta una capa adicional de protección ante intentos de suplantación o inyección de paquetes.

Las pruebas confirmaron que LoRaWAN proporciona un nivel de seguridad sólido por diseño, que el sistema de autenticación basado en DevEUI y claves de sesión funciona correctamente y que el proyecto puede operar con confianza sobre esta infraestructura, complementándola con capas adicionales de cifrado cuando sea necesario para la protección de datos sensibles.

Actualmente se está desarrollando el servidor backend, encargado de recibir los mensajes provenientes de ChirpStack vía HTTP, descifrar el contenido utilizando la clave de la aplicación y procesar la información de manera segura. Este backend también gestionará la validación de integridad, el almacenamiento de datos, la autenticación de dispositivos y la eventual integración con sistemas externos.

[7] OTAA: Over-The-Air-Activation

4- FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación trabaja un grupo de docentes/investigadores del LINTI (Laboratorio de Investigación en Nuevas Tecnologías Informáticas) de la Facultad de Informática de la UNLP (Universidad Nacional de La Plata), dando continuidad al trabajo en ciberseguridad iniciado hace muchos años en el Laboratorio [4].

Las líneas de investigación del proyecto permitieron producir dos propuestas de grado en el marco de las Prácticas Profesionales Supervisadas, para la carrera de Ingeniería en Computación de la Facultad de Informática de la UNLP, de Brian Llamocca y Aldana Tedesco, bajo la tutoría de Paula Venosa y Sofía Martin, cuyos objetivos son desarrollar y evaluar un entorno experimental de red IoT utilizando dispositivos de bajo costo para conformar un ecosistema funcional a pequeña escala.

Además, se concluyó el Trabajo Final Integrador para la Especialización en Redes y Seguridad, de Pablo Germán Maddalena Kreff cuyo objetivo es relacionar a los honeypots como fuente de información con una herramienta de Cyber Threat Intelligence (CTI) [1].

El trabajo de investigación presentado se enmarca en el proyecto de I+D+i "Redes de Honeypots: alcances, implementaciones y utilidad en la ciberseguridad de las organizaciones" de la Facultad de Informática de la UNLP llevado a cabo durante el año 2024.

5- BIBLIOGRAFÍA

1. Venosa, P., Bazán, P. A., Martin, S., Del Rio, N., Magdalena Kreff, P. G., Gagliardi, P., & Díaz, F. J. (2025). Honeypots como fuente de inteligencia de ciberamenazas. In *XXVII Workshop de Investigadores en Ciencias de la Computación (Mendoza, 10 y 11 de abril de 2025)*.
2. Venosa, P. (2021). Detección de ataques de seguridad en redes usando técnicas de ensembling (Doctoral dissertation, Universidad Nacional de La Plata). http://sedici.unlp.edu.ar/handle/10915/120856. https://doi.org/10.35537/10915/120856
3. Gallardo Urbini I. Estrategia de Ciberseguridad distribuida, aplicando el concepto de Operación de Inteligencia. Tesis Doctoral. La Plata 2023. http://sedici.unlp.edu.ar/handle/10915/152594
4. Diaz J, Venosa P, Macia N. Investigación en Ciberseguridad en un año de pandemia. WICC 2021. https://sedici.unlp.edu.ar/handle/10915/119490
5. Washofsky, A. D. (2021). Deploying and analyzing containerized honeypots in the cloud with t-pot (Doctoral dissertation, Monterey, CA; Naval Postgraduate School).
6. Roa Parra, E. F. (2019). Análisis de seguridad de tecnologías inalámbricas de comunicación basado en radio definida mediante software (Bachelor's thesis).
7. Arregui Caballero de Tineo, A. (2016). Beacons BLE (Bluetooth Low Energy) en el sector turístico, control de afluencia y servicios de valor añadido.
8. Saltzstein, W. (2020). Bluetooth Wireless Technology Cybersecurity and Diabetes Technology Devices. Journal of Diabetes Science and Technology, 14(6), 1111-1115. https://doi.org/10.1177/1932296819864416

Modelo de seguridad para Internet de las Cosas basado en Inteligencia Artificial Distribuida y Aprendizaje Federado

Jorge Eterovic; Edith García; Luis Torres.

Instituto de Investigación en Ciencia y Tecnología
Vicerrectorado de Investigación y Desarrollo
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{jorge.eterovic; edith.garcia; luisantonio.torres} @usal.edu.ar

RESUMEN

El crecimiento acelerado del Internet de las Cosas (IoT) ha incrementado la superficie de ataque de los sistemas distribuidos, exponiendo dispositivos heterogéneos y con recursos limitados a amenazas cada vez más sofisticadas. En este contexto, los enfoques tradicionales de seguridad centralizados resultan insuficientes para garantizar escalabilidad, baja latencia y preservación de la privacidad de los datos.

Este trabajo presenta un modelo de seguridad para entornos IoT basado en Inteligencia Artificial distribuida y aprendizaje federado, orientado a la detección de anomalías y ataques sin necesidad de centralizar la información sensible generada por los dispositivos. La arquitectura propuesta integra agentes livianos en dispositivos IoT, nodos de edge computing (computación en el borde) con modelos locales de detección de anomalías y un servidor central de agregación segura que consolida el conocimiento global mediante algoritmos de aprendizaje federado.

El modelo incorpora mecanismos de agregación segura, privacidad diferencial y técnicas de IA híbridas, combinando aprendizaje profundo y métodos estadísticos, con el objetivo de mejorar la precisión de detección y la robustez frente a ataques adversariales. La propuesta se fundamenta en el análisis del estado del arte y en resultados reportados en la literatura científica reciente, mostrando que el aprendizaje federado constituye una alternativa viable y eficaz para fortalecer la seguridad en entornos IoT e IIoT.

Palabras Clave:
Internet de las Cosas; Seguridad; Inteligencia Artificial; Aprendizaje Federado; Detección de Anomalías; Edge Computing.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndola como un servicio a la comunidad y entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias Unidades Académicas se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándolos a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e internacionales, como así también, apoyo y orientación de recursos para la investigación.

A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto denominado

"Análisis del aporte de la Inteligencia Artificial a la mejora de la seguridad en ambientes de Internet de las Cosas", con una duración de 2 años (2025-2026) y que ya ha sido evaluado y aprobado para su realización por Disposición N° 228/2024.

1. INTRODUCCIÓN

Internet de las Cosas (IoT) es el paradigma en el cual objetos físicos dotados de sensores, capacidad de procesamiento y conectividad pueden interactuar entre sí y con sistemas externos a través de Internet, generando y consumiendo grandes volúmenes de datos [1]. La adopción masiva de IoT en ámbitos domésticos, industriales, sanitarios y urbanos ha transformado procesos productivos y servicios, pero también ha introducido riesgos significativos en materia de seguridad.

La falta de estandarización, el uso de credenciales débiles, la limitada capacidad de actualización y las restricciones de recursos convierten a los dispositivos IoT en objetivos atractivos para los ciberatacantes [2]. A medida que crece la cantidad de dispositivos conectados, aumentan los ataques distribuidos, las botnets IoT y las intrusiones orientadas a la exfiltración de datos sensibles.

En este contexto, la Inteligencia Artificial (IA) emerge como una tecnología clave para evolucionar desde modelos de seguridad reactivos hacia enfoques proactivos y predictivos, capaces de adaptarse dinámicamente a entornos cambiantes y amenazas desconocidas [3].

Con tantos dispositivos IoT ingresando al mercado, es necesario que haya más funciones de seguridad integradas. Además, las capacidades informáticas de los dispositivos IoT están en duda. Se ha demostrado que los dispositivos IoT son vulnerables a las ciber-amenazas, pero también es fundamental conocer la naturaleza de estas amenazas [4]. Ellas son:

- **Gran variedad de dispositivos**. Innumerables fabricantes producen dispositivos IoT en varios tamaños, formas y capacidades. El sistema operativo y los protocolos de comunicación que utilizan son diversos. La falta de uniformidad plantea un gran problema a la seguridad de los dispositivos de IoT.
- **Limitaciones de recursos**. Los fabricantes de IoT inundan el mercado con dispositivos baratos con baja capacidad de almacenamiento, memoria mínima y bajas capacidades de procesamiento. Debido a estas limitaciones, es importante desarrollar soluciones de seguridad innovadoras.
- **Autenticación deficiente**. Los dispositivos IoT son vulnerables al acceso no autorizado debido a contraseñas extremadamente débiles y predeterminadas. La falta de un mecanismo de autenticación fuerte deja a estos dispositivos a merced de los delincuentes informáticos.
- **Rápido crecimiento y escala**. La cantidad de dispositivos IoT está creciendo a un ritmo explosivo, pero sin seguridad. Cuanto mayor sea el número de dispositivos IoT, más débil será la seguridad. Un número cada vez mayor de dispositivos IoT significa una posibilidad cada vez mayor de ciber-ataques exitosos.
- **Dispositivos desactualizados**. Hay una gran cantidad de dispositivos IoT antiguos que son funcionales, pero necesitan una actualización por parte del fabricante. Sin parches de seguridad, estos dispositivos se convierten en una importante amenaza para la seguridad, pero en muchos casos actualizarlos es costoso.

El rápido crecimiento en la cantidad de dispositivos conectados a Internet ha llevado a la necesidad de buscar soluciones de seguridad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La autenticación y la autorización son dos factores claves para asegurar a los dispositivos de la Internet de las Cosas. Se pueden definir de la siguiente manera:

- **Autenticación**: Se refiere a la verificación de la identidad de usuarios o dispositivos.

- **Autorización**: Consiste en la asignación de permisos y acceso a los propietarios, a los usuarios, a los datos sensibles y a la gestión de la red y de los dispositivos.

La seguridad en IoT ha sido abordada desde diferentes perspectivas, incluyendo marcos normativos, criptografía liviana, mecanismos de autenticación y arquitecturas de defensa en profundidad. El NIST propone lineamientos específicos para la ciberseguridad de dispositivos IoT, enfatizando la gestión de identidades, la actualización segura y la protección de datos [5].

Diversos autores han analizado el uso de técnicas de aprendizaje automático y profundo para la detección de intrusiones y anomalías en redes IoT, demostrando mejoras significativas frente a soluciones basadas en firmas estáticas. En particular, Mazhar et al. destacan el potencial de la IA para enfrentar ataques complejos y desconocidos, aunque señalan desafíos asociados al consumo de recursos y la explicabilidad de los modelos [6].

Nuestros trabajos previos de investigación han explorado tecnologías complementarias como blockchain [7], criptografía liviana [8], contratos inteligentes [9] y tokens no fungibles (NFT) [10] aplicados a la seguridad en IoT, sentando las bases para una integración con soluciones basadas en IA.

Las amenazas en IoT han evolucionado desde ataques simples de fuerza bruta hacia estrategias sofisticadas que incluyen botnets distribuidas, ataques de denegación de servicio (DDoS), movimientos laterales dentro de redes IoT y explotación de vulnerabilidades de día cero. Asimismo, se observa un uso creciente de técnicas de IA por parte de los atacantes para automatizar y optimizar los ataques.

Esta evolución obliga a replantear los mecanismos defensivos tradicionales, incorporando soluciones adaptativas capaces de aprender del comportamiento del sistema y anticipar incidentes de seguridad.

3. MODELO DE SEGURIDAD.

El diseño de modelos de seguridad para entornos de Internet de las Cosas (IoT) basados en Inteligencia Artificial (IA) enfrenta desafíos asociados a la heterogeneidad de dispositivos, la escalabilidad de la red, la privacidad de los datos y las limitaciones de recursos computacionales. En este contexto, el aprendizaje federado (Federated Learning, FL) ha emergido como un enfoque clave para el entrenamiento distribuido de modelos sin la necesidad de centralizar datos sensibles, preservando la privacidad y reduciendo los costos de comunicación [11], [12].

Diversos estudios han demostrado que la aplicación de aprendizaje federado en sistemas de detección de intrusiones (IDS) para IoT permite alcanzar niveles de precisión comparables o superiores a los enfoques centralizados tradicionales, al tiempo que mejora la protección de la información local de los dispositivos [13]. En particular, el uso de esquemas de agregación como Federated Averaging (FedAvg) y sus variantes optimizadas ha mostrado mejoras significativas en la convergencia del modelo y en la estabilidad del aprendizaje en escenarios con datos no IID (no idénticamente distribuidos) [14].

Arquitectura general del modelo propuesto

El modelo de seguridad IoT basado en IA propuesto en este trabajo adopta una arquitectura multicapa inspirada en enfoques recientes reportados en la literatura científica [15], [16]. Dicha arquitectura se compone de:

1. **Capa de dispositivos IoT:** Incluye sensores y actuadores que generan tráfico y eventos de red. Cada dispositivo ejecuta agentes livianos encargados de recolectar características relevantes para la detección de anomalías, minimizando el consumo de recursos.
2. **Capa de edge computing:** En esta capa se despliegan modelos locales de detección de anomalías basados en IA, entrenados con datos propios del entorno. La ejecución en el borde permite reducir latencia y detectar ataques en tiempo casi real [17].

3. **Capa de aprendizaje federado:** Los modelos locales comparten únicamente los parámetros entrenados con un servidor agregador central, evitando la transmisión de datos crudos. Este enfoque preserva la privacidad y cumple con principios de seguridad por diseño [11], [13].
4. **Servidor de agregación segura:** El servidor central combina los parámetros recibidos mediante protocolos de agregación segura y puede incorporar mecanismos de privacidad diferencial para mitigar ataques por inferencia [18].

En la Fig. 1 se muestra el modelo de seguridad para IoT con aprendizaje federado:



Figura 1. Arquitectura de seguridad

Integración de modelos de IA

La literatura reciente evidencia que los modelos híbridos superan en desempeño a los enfoques individuales. En este sentido, se han reportado resultados destacados al combinar técnicas de aprendizaje profundo (como LSTM y autoencoders) con esquemas de aprendizaje federado para capturar patrones temporales complejos del tráfico IoT [16], [19].

Asimismo, enfoques más livianos como Isolation Forest integrados en arquitecturas federadas (FLiForest) han demostrado ser altamente efectivos en entornos edge, alcanzando precisiones cercanas al 96 % con requerimientos mínimos de memoria y procesamiento, lo que los hace adecuados para dispositivos con recursos limitados [20].

Privacidad, robustez y escalabilidad

Uno de los principales aportes del modelo propuesto es la incorporación de mecanismos avanzados de protección de la privacidad y robustez frente a ataques. Estudios recientes han demostrado que la combinación de aprendizaje federado con técnicas de agregación segura, privacidad diferencial y aprendizaje por refuerzo mejora la resistencia del sistema ante ataques adaptativos y escenarios adversariales [21].

Además, arquitecturas jerárquicas de aprendizaje federado (mist–edge–fog) han mostrado una mayor tolerancia a la heterogeneidad de dispositivos IoT, mejorando la escalabilidad y manteniendo altos niveles de precisión en entornos distribuidos de gran tamaño [22].

En conjunto, el modelo propuesto se alinea con el estado del arte y consolida una solución de seguridad IoT basada en IA que combina detección precisa de anomalías, preservación de la privacidad y viabilidad práctica, posicionándose como una alternativa sólida para despliegues reales en entornos IoT e IIoT.

4. RESULTADOS OBTENIDOS/ESPERADOS.

A partir del modelo de seguridad IoT basado en Inteligencia Artificial distribuida y aprendizaje federado propuesto en este trabajo, se esperan una serie de resultados alineados con los reportados en la literatura científica reciente y con las características propias de la arquitectura planteada.

En primer lugar, se espera que el modelo alcance altos niveles de precisión en la detección de anomalías y ataques, comparables a los obtenidos por enfoques federados reportados en trabajos previos, con valores de precisión superiores al 95 % y una reducción significativa de la tasa de falsos positivos respecto de soluciones centralizadas tradicionales. La combinación de modelos híbridos de IA y aprendizaje federado permitiría capturar patrones complejos de comportamiento del tráfico IoT, incluso en

escenarios con datos heterogéneos y no idénticamente distribuidos.

En segundo lugar, el uso de aprendizaje federado permitiría una mejora sustancial en la preservación de la privacidad, ya que los datos sensibles generados por los dispositivos IoT no se transmiten ni almacenan de forma centralizada. Este enfoque reduciría la exposición de la información frente a ataques por inferencia o filtraciones de datos, manteniendo la confidencialidad de los entornos IoT e IIoT.

Asimismo, se espera que la arquitectura propuesta presente mejoras en términos de escalabilidad y eficiencia, al distribuir el entrenamiento de los modelos entre múltiples nodos de borde. Esta característica permitiría la incorporación progresiva de nuevos dispositivos sin degradar el rendimiento global del sistema, adaptándose a escenarios de crecimiento dinámico típicos de entornos IoT reales.

Desde el punto de vista operativo, la ejecución de modelos livianos en la capa de edge computing permitiría reducir la latencia en la detección de amenazas, facilitando respuestas más rápidas ante comportamientos anómalos. Esto resulta especialmente relevante en aplicaciones críticas, donde la detección temprana de incidentes de seguridad es un factor clave.

Finalmente, se espera que el modelo propuesto sirva como base para futuras implementaciones experimentales y validaciones cuantitativas, permitiendo evaluar su desempeño en distintos escenarios y con diversos datasets representativos de entornos IoT. En este sentido, el enfoque presentado sienta las bases para trabajos futuros orientados a la evaluación experimental, optimización de modelos y despliegue en entornos reales.

5. FORMACIÓN DE RECURSOS HUMANOS.

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas de la Facultad de Ingeniería, específicamente al área de Seguridad Informática, de la Universidad del Salvador.

A este proyecto, se incorporaron dos docentes investigadores con amplia trayectoria académica, un docente investigador con muchos años de desempeño en la industria de TI y tres alumnos de la carrerea de Ingeniería en Informática.

Esto redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes investigadores y alumnos, como así también conformará las bases para la investigación a futuro desarrollando esta línea de trabajo en la Facultad de Ingeniería.

6. BIBLIOGRAFÍA.

[1] Ciuta, S; "Securing the Internet of Things (IoT): Cybersecurity of Connected Devices"; Ed.: Independently published. 2023.

[2] Cheruvu, S; Kumar, A; Smith, N; Wheeler, D.M; "Demystifying Internet of Things Security: Successful IoT Device/Edge and Platform Security Deployment". 1st Edition. Ed.: Apress Open. 2020..

[3] Chahid, I; Benabdellah, M; "Artificial Intelligence approaches for IoT Security: State of Art," *Int. J. Adv. Eng. Res. Sci.*, vol. 9, no. 6, pp. 1–10, 2022.

[4] Eterovic, J; García, E; Torres, L; "Análisis del aporte de la Inteligencia Artificial a la mejora de la Seguridad en ambientes de Internet de las Cosas"; XXVII Workshop de Investigadores en Ciencias de la Computación; Libro de Actas WICC 2025; ISBN 978-987-575-267-2; Argentina. 2025.

[5] Fagan, M; Megas, K; Marron, J; Brady Jr, K; Cuthill, B; Herold, R; Lemire, D; Hoehn, B; "NIST SP 800-213A: IoT Device Cybersecurity Guidance for the Federal Government: IoT Device Cybersecurity Requirement Catalog"; Ed.: NIST (National Institute of Standards and Technology). 2022.

[6] Mazhar, T.; Talpur, D.B.; Shloul, T.A.; Ghadi, Y.Y.; Haq, I.; Ullah, I.; Ouahada, K.; Hamam, H. "Analysis of IoT Security Challenges and Its Solutions Using Artificial

Intelligence". Brain Sci. 2023, 13, 683. https://doi.org/10.3390/10.3390/brainsci1304 0683. 2023.

[7] Eterovic, J; Cipriano, M; García, E; Torres, L; "Marco de referencia para aplicaciones seguras de Blockchain"; 8 º Congreso Nacional de Ingeniería Informática Sistemas de Información CONAIISI 2020; Libro de Actas CoNaIISI 2020; ISBN 978-950-42-0202-8. Argentina. 2020.

[8] Cipriano, M; Eterovic, J; García, E; Torres, L; "Criptografía Liviana para Internet de las Cosas e Internet de las Cosas Industrial"; XXIII Workshop de Investigadores en Ciencias de la Computación; Libro de Actas WICC 2021; ISBN 978-987-24611-3-3; Argentina. 2021.

[9] Eterovic, J; Cipriano, M; García, E; Torres, L; "Aplicación de los Contratos Inteligentes en Internet de las Cosas"; XXIII Workshop de Investigadores en Ciencias de la Computación; Libro de Actas WICC 2021; ISBN 978-987-24611-3-3; Argentina. 2021.

[10] Eterovic, J; Cipriano, M; García, E; Torres, L. "Propuesta de uso de tokens inteligentes no fungibles (NFTs) para representar dispositivos de Internet de las Cosas". 11° CoNaIISI 2023 - Congreso Nacional de Ingeniería Informática / Sistemas de Información. Actas del CoNaIISI. 2023. ISBN 978-987-4417-73-2. 2023.

[11] Chowdhury, R; Hassan, M.M; Hossain, M; Laila, J; "Federated learning for intrusion detection in IoT environments: a privacy-preserving strategy," Discover Internet of Things, vol. 5, art. no. 72, 2025.

[12] Nagaiah, B; "Use of Federated Learning in IIoT Security," Int. J. Artif. Intell. Data Sci. Mach. Learn., vol. 6, no. 4, 2025.

[13] Vidhya, R; Lognathan, D; Saranya, S; Periyasamy, P; Sumathi, S; "Anomaly Detection in IoT Networks Using Federated Machine Learning Approaches," Int. J. Comput. Exp. Sci. Eng., vol. 11, no. 3, 2025.

[14] Pecherle, G; Győrödi, R; Győrödi, C; "Federated Learning-Based Intrusion Detection in Industrial IoT Networks," Future Internet, vol. 18, no. 1, p. 2, 2026.

[15] Zhang, Y; Suleiman, B; Alibasa, M; "FedGroup: A Federated Learning Approach for Anomaly Detection in IoT Environments," in Mobile and Ubiquitous Systems: Computing, Networking and Services, Springer, 2023.

[16] Reis, M. et al.; "Edge-FLGuard+: A Federated and Lightweight Anomaly Detection Framework for Securing 5G-Enabled IoT in Smart Homes," Future Internet, vol. 17, no. 8, 2025.

[17] Ahmad, M; Cide, F; "AI-Driven Anomaly Detection Framework for Industrial IoT Using Edge-Enabled Wireless Sensor Networks," J. Wirel. Sensor Netw. IoT, vol. 3, no. 1, pp. 33–39, 2025.

[18] Desai, C; "Secure Aggregation Protocols for Federated Learning in IoT Intrusion Detection," Int. J. Comput. Technol. Electron. Commun., vol. 4, no. 4, pp. 3808–3812, 2021.

[19] Alzubaidi, A; "Cyber Attack Classification on IoT Devices Using Federated Machine Learning Infrastructure and AI," Eng. Technol. Appl. Sci. Res., vol. 15, no. 6, pp. 30470–30476, 2025.

[20] Xiang, H; Zhang, X; Xu, X; Beheshti, A; Qi, L; Hong, Y; Dou, W; "Federated Learning-Based Anomaly Detection with Isolation Forest in the IoT-Edge Continuum," Comput. Security, vol. 155, 2025.

[21] Asiri, F; Malwi, W; Masood, F; Alshehri, M; Zhukabayeva, T; Shah; S. et al.; "Privacy Preserving Federated Anomaly Detection in IoT Edge Computing Using Bayesian Game Reinforcement Learning," Comput. Mater. Contin., vol. 84, no. 2, pp. 3943–3960, 2025.

[22] Izadi, S; Komasi,S; Salimi, A; Rezaei, A; Ahmadi, M; "Mist-Assisted Federated Learning for Intrusion Detection in Heterogeneous IoT Networks," arXiv preprint arXiv:2511.00271, 2025.

OTP-Vote: Avances en Seguridad, Auditoría y Verificabilidad End-to-End hacia su Implementación

Silvia Bast[1] Germán Montejano[2] MarioBerón[2]

[1]Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-425166– Int. 28
silviabast@exactas.unlpam.edu.ar – web: http://exactas.unlpam.edu.ar/

[2]Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
Tel.: +54-2652-424027 – Int. 251
[gmonte, mberon]@unsl.edu.ar – web: http://www.unsl.edu.ar

RESUMEN

Este trabajo presenta los avances más recientes del proyecto de investigación OTP-Vote, enfocado en la especificación integral de un sistema de voto electrónico basado en criptografía One Time Pad. Durante el último año se ha profundizado la formalización del modelo, fortaleciendo sus mecanismos de seguridad, auditoría y verificabilidad End-to-End (E2E), con el objetivo de consolidar su viabilidad como sistema de votación electrónica robusto, auditable y transparente.

Los aportes recientes incluyen mejoras sustanciales en la definición de procesos críticos, la trazabilidad de operaciones, la protección de datos sensibles y la formalización de mecanismos de verificabilidad individual y universal. Estos avances fueron integrados en la tesis doctoral, presentada en 2025 para su evaluación, y han dado lugar a publicaciones científicas en ámbitos nacional e internacional.

El trabajo contribuye a reducir la brecha entre los modelos teóricos de voto electrónico y su implementación práctica, proponiendo un enfoque que combina seguridad criptográfica, auditabilidad operativa y preservación estricta del secreto del voto.

Palabras clave: Voto electrónico; Seguridad; Verificabilidad End-to-End; Auditoría electoral; Criptografía One Time Pad; Anonimato.

CONTEXTO

El trabajo se enmarca en el proyecto de investigación "Especificación Integral del Sistema OTP-Vote Orientada a su Implementación" (Resolución N° 55/22, FCEyN – UNLPam), que profundiza y amplía la línea de investigación en voto electrónico basada en criptografía OTP desarrollada previamente en el proyecto "Aspectos de Seguridad en Proyectos de Software" (Resolución N° 488/14, FCEyN – UNLPam).

Asimismo, la investigación se articula con el proyecto vigente en la Universidad Nacional de San Luis "Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y

Migración de Sistemas en la Nube" (PROICO 2020, Resolución CS N° 138/2020), contribuyendo a consolidar su proyección hacia implementaciones concretas.

Durante 2025, el trabajo alcanzó un punto de consolidación clave, integrando en un marco unificado los aportes desarrollados en años previos y profundizando especialmente en los mecanismos de auditoría y verificabilidad.

La investigación se desarrolla de manera conjunta entre la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa y la Facultad de Ciencias Físico-Matemáticas y Naturales de la Universidad Nacional de San Luis. El proyecto es dirigido por el Doctor Germán Montejano y el Doctor Mario Berón (UNSL) y tiene como objetivo transformar el modelo OTP-Vote en un sistema integral, verificable y técnicamente implementable.

1. INTRODUCCIÓN

Los procesos electorales pueden entenderse como sistemas de información críticos, ya que de su correcto funcionamiento depende la legitimidad de las decisiones públicas que de ellos se derivan. En este contexto, los sistemas utilizados para la emisión y el recuento de votos deben garantizar no solo la integridad del proceso, sino también la confianza del electorado. Si bien el avance tecnológico habilita la digitalización de múltiples actividades, el voto presenta un desafío particular: a diferencia de otras transacciones digitales, debe impedir cualquier vínculo entre el votante y su elección.

Los sistemas de voto electrónico han sido objeto de debate debido a la tensión entre verificabilidad, auditabilidad y secreto del voto, sumada a experiencias previas poco satisfactorias en distintos contextos. Sin embargo, los sistemas manuales actualmente utilizados tampoco están exentos de limitaciones, incluyendo vulnerabilidades operativas, prácticas irregulares y elevados costos logísticos, así como demoras significativas en el escrutinio, especialmente en elecciones de gran escala.

Desde un punto de vista formal, un sistema de votación puede modelarse como una función que transforma las preferencias de los electores en un resultado electoral [1]. No obstante, los sistemas de voto electrónico requieren propiedades adicionales a las de otros sistemas informáticos. Diversos autores han identificado requisitos fundamentales como el secreto del voto, la elegibilidad y la integridad del recuento [2–4], junto con la necesidad de mecanismos de auditabilidad y resistencia a manipulaciones. En los últimos años, revisiones sistemáticas han analizado comparativamente estos enfoques criptográficos y sus propiedades de seguridad [5].

En este marco surge la verificabilidad End-to-End (VE2E), cuyo objetivo es garantizar que el resultado electoral refleje fielmente los votos emitidos, proporcionando evidencia verificable a todas las partes involucradas, independientemente del software utilizado [6]. Este enfoque ha sido ampliamente desarrollado en la literatura [7–11] y se complementa con el concepto de independencia del software [12]. Investigaciones recientes [13] continúan explorando variantes arquitectónicas y mecanismos criptográficos para fortalecer la verificabilidad y la resistencia a manipulaciones, lo que evidencia que la conciliación entre verificabilidad, auditabilidad y secreto del voto sigue constituyendo un desafío abierto en el diseño de sistemas electorales digitales.

El modelo OTP-Vote, presentado originalmente en [14], se basa en la premisa de que un sistema de voto electrónico debe garantizar privacidad indefinida del votante y protección robusta de los datos durante el proceso electoral. Para ello, combina anonimato incondicional con seguridad computacional configurable, apoyándose en el uso de claves One Time Pad [15], el esquema de almacenamiento Múltiples Canales Datos

Único (MCDU) [16] y redundancia estructurada en los datos [17].

Este artículo presenta los avances recientes en la especificación integral del modelo OTP-Vote, con foco en el fortalecimiento de sus mecanismos de seguridad, auditoría y VE2E, orientados a su implementación práctica.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación y desarrollo se orientan hacia el refinamiento integral del modelo OTP-Vote, con el objetivo de transformar sus bases teóricas en un sistema robusto e implementable. El proyecto avanza sobre dos ejes principales de desarrollo:

- El primer eje se centra en la optimización del modelo base, focalizándose en el fortalecimiento de las funcionalidades existentes. Esta línea incluye ajustes y mejoras que buscan incrementar el desempeño y la robustez del sistema, manteniendo sus características fundamentales de seguridad y confiabilidad.

- El segundo eje aborda la incorporación de nuevos elementos al modelo, con especial énfasis en los mecanismos de auditoría y VE2E. Estas contribuciones resultan críticas para aumentar la transparencia del sistema sin comprometer el anonimato del votante, aspecto fundamental para generar confianza en el electorado.

Ambas líneas de desarrollo convergen en el objetivo común de demostrar la confiabilidad del sistema OTP-Vote, abordando de manera sistemática cada uno de los aspectos que requieren mayor precisión y especificación. Las propuestas desarrolladas en cada eje se complementan para construir un sistema que no solo sea técnicamente sólido, sino también prácticamente implementable.

3. RESULTADOS Y OBJETIVOS

Durante el último año se lograron avances significativos que consolidan el modelo OTP-Vote como una propuesta integral. Estos resultados se incorporaron formalmente en la tesis doctoral y en nuevas publicaciones científicas.

1. Profundización en la especificación formal de procesos.
Se completó la especificación detallada de procesos críticos en las tres etapas del sistema:

- Configuración de la elección
- Emisión del voto
- Cierre y recuento

Cada proceso fue definido con:

- Requisitos de acceso
- Condiciones de integridad
- Validación de firmas digitales
- Registro en logs inmutables

Esto permitió fortalecer la trazabilidad completa del sistema, reduciendo ambigüedades operativas y mejorando la capacidad de auditoría.

2. Fortalecimiento del modelo de datos y estructuras críticas.
Se avanzó en la clasificación formal de los datos del sistema según su nivel de inmutabilidad, distinguiendo entre: datos estrictamente inalterables y datos modificables bajo condiciones controladas.

Además, se refinó la semántica de las tuplas utilizadas en los archivos críticos (ABV y CD), incorporando: i) Bits de control, ii) Configuraciones dispersas de información, iii) Parámetros generados automáticamente a partir de la configuración electoral.

Estos mecanismos incrementan la resistencia frente a intentos de manipulación y mejoran la coherencia interna del modelo.

3. Consolidación del sistema de auditoría. Uno de los avances más relevantes del período fue la ampliación del sistema de auditoría, que ahora incluye: registro estructurado de eventos críticos, detección de inconsistencias y accesos no autorizados, validación de integridad de claves y datos base y posibilidad de activar recuentos manuales ante anomalías. El archivo Log pasó a tener un rol central como fuente de evidencia técnica, fortaleciendo la transparencia del sistema sin comprometer el secreto del voto.

4. Avances en Verificabilidad End-to-End (E2E). Se profundizó la especificación de mecanismos que permiten satisfacer simultáneamente: verificabilidad individual: verificabilidad universal y secreto del voto.

Se fortaleció el mecanismo que vincula cada voto con un elemento verificable público que preserva el anonimato del elector.

Además, se avanzó en el análisis probabilístico de escenarios de posible compromiso del anonimato, mostrando que las probabilidades de vulneración son extremadamente bajas en contextos reales de votación.

5. Integración y validación académica de los resultados.
Los avances alcanzados fueron integrados en la versión final de la tesis doctoral presentada para su evaluación en 2025 y difundidos mediante publicaciones científicas. Esto representa un paso clave en la validación externa del modelo y en su consolidación como aporte científico en el área de voto electrónico verificable.

Impacto de los Avances

Los desarrollos recientes permiten afirmar que OTP-Vote ha evolucionado desde un modelo criptográfico teórico hacia una arquitectura electoral integral, que combina:

- Seguridad criptográfica fuerte
- Modelado sistemático de los procesos electorales
- Auditoría operativa
- Verificabilidad End-to-End
- Preservación del secreto del voto

Este equilibrio constituye uno de los principales aportes del proyecto, ya que aborda simultáneamente las principales críticas dirigidas a los sistemas de voto electrónico.

Trabajos Futuros

Las próximas líneas de trabajo se orientarán a:

- Validación empírica del modelo mediante simulaciones y pruebas controladas.
- Mejora del nivel de detalle y automatización de los archivos Log, fortaleciendo la trazabilidad del proceso electoral.
- Optimización de procesos críticos para escenarios de gran escala y alta concurrencia de votantes.
- Desarrollo de herramientas de apoyo para auditores y autoridades electorales, que faciliten la comprensión y supervisión del proceso.
- Análisis de nuevos escenarios de riesgo en contextos de verificabilidad End-to-End (VE2E), explorando configuraciones aún no consideradas mediante simulaciones y enfoques estadísticos y exploratorios que permitan estimar probabilidades de compromiso del anonimato a partir de métricas como entropía, desviación estándar y patrones de distribución, así como identificar patrones emergentes asociados a posibles escenarios de desanonimización bajo configuraciones heterogéneas de votos.

Estos desarrollos permitirán continuar la transición del modelo OTP-Vote desde la especificación formal hacia entornos de prueba más cercanos a implementaciones reales.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de este proyecto, Silvia Bast presentó en 2025 su tesis doctoral titulada: “Especificación Integral del Modelo OTP-Vote orientada a su implementación”, para optar al grado de Doctora en Ingeniería Informática por la Universidad Nacional de San Luis.

La tesis se encuentra finalizada y presentada para su evaluación, e integra, sistematiza y amplía los resultados obtenidos en el proyecto, formalizando los procesos de las distintas etapas del sistema electoral OTP-Vote y consolidando los mecanismos de seguridad, auditoría y verificabilidad desarrollados.

5. BIBLIOGRAFÍA

[1] F. Hao, P. Ryan. “Real-world electronic voting: Design, analysis and deployment.” CRC Series in Security, Privacy and Trust. CRC Press, 2016. ISBN-13: 978-1498714693. ISBN-10: 1498714692. 2017.

[2] J. Epstein, “Electronic Voting” in Computer, vol. 40, no. 8, pp. 92-95, Aug 2007. doi: 10.1109/MC.2007.271.

[3] K. M. Rokibul Alam and S. Tamura, “Electronic voting - Scopes and limitations”, International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, Bangladesh, pp. 525–529, 2012. doi: 10.1109/ICIEV.2012.6317324.

[4] J.van de Graaf, C. Henrich and J. Müller-Quade, “Requirements for secure voting”, Work Notes 2011.

[5] Y.-X. Khoo, S.-H. Heng, and J.-J. Chin, “A review of cryptographic electronic voting,” Symmetry, vol. 14, no. 5, p. 858, 2022.

[6] P. Ryan, S. Schneider, y V Teague. Prˆet `a voter—the evolution of the species. In Real-World Electronic Voting, pages 325–358. Auerbach Publications, 2016.

[7] M. Rabin and R. Rivest. Efficient end to end verifiable electronic voting employing split value representations. EVOTE 2014.

[8] Awad, Mohammed, and E. Leiss. "End-to-End Cryptography: Spreading Democracy." International Journal of Applied Engineering Research 11.11 (2016): 7391-7394.

[9] D. Chaum, A. Essex, R. Carback, J. Clark, S. Popoveniuc, A. Sherman, and P. Vora. “Scantegrity: End-to-end voter verifiable optical-scan voting”. IEEE Security & Privacy, 6(3):40–46, 2008.

[10] D. Chaum, R. Carback, J. Clark, A. Essex, S. Popoveniuc, R. Rivest, P. Ryan, E. Shen, A. Sherman and P. Vora. “Scantegrity II: End-to-end verifiability by voters of optical scan elections through confirmation codes”. IEEE transactions on information forensics and security, 4(4), 611-627. 2009.

[11] J. Benaloh, R. Rivest, P. Ryan, P. Stark,V. Teague, & P. Vora,(2015). End-to-end verifiability. arXiv preprint arXiv:1504.03778.

[12] R. Rivest and M. Virza. Software independence revisited. In Real-World Electronic Voting, pages 19–34. Auerbach Publications, 2016.

[13] S. Park et al., “zkVoting: Zero-knowledge proof based coercion-resistant and E2E verifiable e-voting system,” Cryptology ePrint Archive, Report 2024/1003, 2024.

[14] S. Bast. “Optimización de la integridad de datos en sistemas de e-voting”. Master’s thesis, Universidad Nacional de San Luis, Dic 2016.

[15] C. E. Shannon, "Communication theory of secrecy systems," in The Bell System Technical Journal, vol. 28, no. 4, pp. 656-715, Oct. 1949, doi: 10.1002/j.1538-7305.1949.tb00928.x.

[16] P. M. García,“Optimización de un Protocolo Dining Cryptographers Asíncrono,” Master’s thesis, Universidad Nacional de San Luis (UNSL) / Universidade Federal de Minas Gerais (UFMG), 2013.

[17] P. García, G. Montejano, S. Bast, E. Fritz, "Codificación de Sufragios con Detección de Colisiones en NIDC con Canales Paralelos de Slots” Congreso Nacional de Ingeniería en Informática / Sistemas de Información. CoNaIISI 2016.

Tecnología Informática Aplicada en Educación

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

ALFABETIZACIÓN DIGITAL Y TRANSFORMACIÓN DIGITAL: ACCIONES UNIVERSITARIAS PARA LA SOCIEDAD

Sonia I. Mariño, Mirta G. Fernandez, Pedro L. Alfonzo, Viviana Bercheñi, Silvana Armana, Juan E. Jara, Walter G. Barrios
soniaimarinio@gmail.com, mirta.fernandez@comunidad.unne.edu.ar, plalfonzo@hotmail.com, viviana.bercheni@comunidad.unne.edu.ar, sil_armana@yahoo.com.ar, juanemiliano.jara@gmail.com, waltergbarrios@comunidad.unne.edu.ar

Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Corrientes, Argentina

RESUMEN

En el camino a la transformación digital, la alfabetización digital es clave. En el trabajo se resumen las líneas de investigación que tratan la alfabetización digital en contextos diversos. Desde capacitaciones al análisis de los datos con fines de apoyar la toma de decisiones informadas.

Palabras Claves: Alfabetización digital, Educación en tecnologías, Tecnologías en educación, Educación Superior.

CONTEXTO

En el Grupo de Investigación Sistema de información y TI: modelos, métodos y herramientas acreditado por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste, se desarrollan actualmente, 2 proyectos de investigación: Sistemas informáticos y Gestión del conocimiento: modelos, métodos y herramientas (PI 22F025, Res. Nº 0931/22 CS UNNE) y El Observatorio de Tecnologías Aplicadas a la Educación de la FaCENA-UNNE (21F010 –SGCyT – UNNE).

1. INTRODUCCIÓN

La UNESCO (2021) sostiene que la transformación digital implica un cambio profundo en la cultura institucional y en el aprendizaje profesional.

La transformación digital se enfrenta a necesidades de alfabetización digital. Mejorar la alfabetización tecnológica repercute positivamente en la equidad e inclusión social (Comisión Europea, 2020; Mendoza Cuzcano et al., 2026). Asimismo, la alfabetización digital se operacionaliza mediante el desarrollo de competencias digitales.

La alfabetización digital trata las competencias en TIC, la búsqueda crítica de datos, la comunicación en línea y la seguridad digital (Gilster, 1997). Se diferencia de la alfabetización informática, que se centra en aspectos técnicos, al enfatizar el uso ético y reflexivo de la tecnología (Martin & Grudziecki, 2006; Ris et al, 2019).

La alfabetización digital es un camino para la inclusión social, la equidad, y para emprender (Reyes et al., 2021; Xin et al., 2025; Flores et al, 2026).

Respecto a la alfabetización digital en adultos mayores, esta se aborda desde el envejecimiento activo y la inclusión social (Xin et al., 2025), promoviendo el acceso equitativo a las tecnologías y el fortalecimiento de las competencias digitales para mejorar la calidad de vida de este grupo etario. Por otra parte, diversos estudios señalan que la articulación entre alfabetización digital y educación emprendedora en la educación superior y en la ciudadanía. Formar para emprender potencia la creatividad, la autonomía y la capacidad de transferencia de conocimientos hacia contextos productivos y sociales (UNESCO, 2018; Comisión Europea, 2016; OECD, 2021).

Según Florez Cueva et al. (2026), la alfabetización digital prepara a los estudiantes para superar barreras económicas, promover la inclusión social y crear negocios sostenibles mediante el aprovechamiento de las tecnologías digitales. El aporte de la alfabetización digital en la generación del espíritu emprendedor radica en que brinda a los estudiantes alternativas de creatividad y acceso a los recursos que la tecnología ofrece para desarrollar iniciativas emprendedoras innovadoras.

Este trabajo presenta líneas de I+D desarrolladas desde la educación superior, orientadas a promover la alfabetización digital en diversos contextos sociales y educativos, combinando intervenciones formativas con análisis sistemático de datos para la mejora continua.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto se estructura en torno a líneas de investigación que abordan la alfabetización digital desde una perspectiva multidimensional, atendiendo a diversos contextos educativos y productivos, y se complementa con el análisis de datos para fundamentar la toma de decisiones basada en evidencias.

2.1. Líneas temáticas principales

Alfabetización digital contextualizada. Se desarrollan propuestas diferenciadas que reconocen las particularidades de cada grupo destinatario:

- Alfabetización digital para adultos mayores: Diseño e implementación de estrategias formativas adaptadas a las necesidades y características de este grupo etario, promoviendo la inclusión digital y el acceso equitativo a las tecnologías.
- Alfabetización digital para la formación emprendedora no universitaria: Fortalecimiento de competencias digitales orientadas al desarrollo del talento emprendedor en contextos de formación técnica y profesional.
- Alfabetización digital para la formación emprendedora universitaria: Integración de competencias digitales con el espíritu emprendedor en la educación superior, fomentando la innovación y el desarrollo de soluciones tecnológicas viables.

Analítica de datos educativos. Una vez avanzadas las acciones formativas y disponiendo de fuentes de datos primarios provenientes de las diferentes intervenciones educativas, se propone su análisis sistemático para generar información útil que apoye la toma de decisiones pedagógicas e institucionales, mediante:

- Estudio y adecuación de modelos de aceptación tecnológica. Los modelos de aceptación tecnológica constituyen un área clave en los sistemas de información: El Technology Acceptance Model (TAM), propuesto por Davis (1989), permite comprender y predecir la adopción tecnológica a partir de la utilidad percibida y la facilidad de uso. Como extensión, Venkatesh et al. (2003) desarrollaron el Unified Theory of Acceptance and Use of Technology (UTAUT), integrando factores adicionales como la expectativa de desempeño, el esfuerzo, la influencia social y las condiciones facilitadoras. Recientemente, la integración del aprendizaje automático ha permitido detectar patrones complejos no evidentes con métodos estadísticos tradicionales (Chao, 2019; Marikyan & Papagiannidis, 2023), fortaleciendo la capacidad predictiva de estos modelos en contextos educativos.
- Estudio y adaptación de técnicas de analíticas de aprendizaje: Desarrollo e implementación de modelos computacionales que combinan técnicas de Inteligencia Artificial con heurísticas expertas para evaluar y predecir competencias digitales en estudiantes (Fernandez, 2024; Fernandez y Mariño, 2025; Candia López, 2023).

2.2. Enfoques metodológicos transversales

Las líneas de investigación propuestas se sustentan en un conjunto de enfoques metodológicos rigurosos y actuales:

- Revisión sistemática de la literatura:

Siguiendo estándares reconocidos internacionalmente como las propuestas PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) o el método de Kitchenham, se realizan revisiones exhaustivas que permiten conocer el estado del arte, identificar brechas de conocimiento y determinar nichos de investigación que viabilicen contribuciones originales.

- Metodologías ágiles: Se adopta el paradigma de la agilidad como estándar metodológico, adaptándolo a las particularidades de cada línea de I+D. Las metodologías ágiles se combinan con métodos específicos propuestos en la literatura, permitiendo desarrollos iterativos, adaptativos y centrados en las necesidades de los usuarios finales.
- Herramientas y tecnologías de la información: Se identifican, evalúan e implementan herramientas TI apropiadas para viabilizar las propuestas teóricas y metodológicas en relación con los temas prioritarios de investigación. Esto incluye plataformas educativas, sistemas de analítica de aprendizaje, entornos de desarrollo y herramientas de visualización de datos, entre otras.

  Objetivos de desarrollo sostenible (ODS). Se identifica el abordaje desde el cuarto ODS: "Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos" (UNESCO, 2017; Arroyave et al., 2024).

3. RESULTADOS ESPERADOS/OBTENIDOS

En esta sección se describen los resultados alcanzados durante el periodo 2024 a 2025.

- **Alfabetización digital para adultos mayores.** Se diseñaron e implementaron seis capacitaciones en el marco del Programa UPAMI (RES 2024-564-CD-EXA y RES 2025-265/622-CD-EXA). Las formaciones incorporaron recursos interactivos que favorecieron el aprendizaje significativo, abordando temáticas como uso del teléfono celular, redes sociales, WhatsApp y otras aplicaciones móviles. Los adultos mayores participaron activamente en múltiples ediciones, manifestando alto interés en continuar su formación digital. Se encuentra en desarrollo un instrumento de evaluación del impacto de estas intervenciones.
- **Alfabetización digital para la formación emprendedora no universitaria**. Se implementó la Diplomatura Universitaria en Gestión de Unidades de Negocios mediadas por TIC - DUGUNTIC (RES 0761/21-CD y modificatorias), con una carga horaria de 210 horas. La propuesta fortaleció competencias digitales orientadas al talento emprendedor mediante la plataforma Moodle institucional, contando con la participación de docentes especialistas que elaboraron materiales digitales específicos (Reyes et al., 2021; Mariño et al., 2025).
- **Alfabetización para la formación emprendedora universitaria**. Se promovió el espíritu emprendedor mediante capacitaciones dictadas por profesionales especializados y la realización de un *hackathon institucional*. Esta metodología ágil permitió a estudiantes diseñar, prototipar y defender soluciones tecnológicas integradas a planes de negocio viables (Vivanco-Galván et al., 2018).

En estas distintas modalidades de alfabetización digital se favorecen metodologías activas y participativas, el aprendizaje experiencial, el trabajo colaborativo y la resolución de problemas reales, en donde el pensamiento crítico y la ética son fundamentales. Se incorporan actividades concernientes a la producción de materiales educativos digitales, la integración de tecnologías digitales y la formación de recursos humanos

4. FORMACIÓN DE RECURSOS HUMANOS

La formación de recursos humanos se viabiliza a través de la incorporación de estudiantes,

docentes e investigadores en la participación como disertantes en las capacitaciones. Se elaboran tesis de posgrado para analizar los modelos de aceptación tecnológica y la aplicación de técnicas de analítica de aprendizaje sobre datos recuperados de aulas virtuales implementadas en sitios institucionales en la educación superior.

5. BIBLIOGRAFÍA

[1] Arroyave, J., González-Pardo, R., Campis Carrillo, R. M., & Repiso, R. (2024). Producción científica sobre Objetivos de Desarrollo Sostenible en los países hispanoamericanos: similitud e impacto. Investigación Bibliotecológica, 38(100), eib01005890210. https://doi.org/10.22201/iibi.24488321xe.2024.100.58902

[2] Campbell Rodríguez, V. M. (2025). Revolucionando la Educación: Integración de Inteligencia Artificial en Sistemas de Gestión del Aprendizaje. RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo, 15(30), e808. https://doi.org/10.23913/ride.v15i30.2242

[3] Candia López, J. C. (2023). Competencias digitales en la educación superior. Horizontes Revista De Investigación En Ciencias De La Educación, 7(29), 1548-1563. https://doi.org/10.33996/revistahorizontes.v7i29.612

[4] Chao, C. M. (2019). Factors determining the behavioral intention to use mobile learning: An application and extension of the UTAUT model. Frontiers in Psychology, 10, 1652. https://doi.org/10.3389/fpsyg.2019.01652

[5] Comisión Europea (2016a). A New Skills Agenda for Europe: Working together to strengthen human capital, employability and competitiveness. European Commission.

[6] Comisión Europea (2016b). EntreComp: The entrepreneurship competence framework. Publications Office of the European Union. https://publications.europa.eu/

[7] Comisión Europea (2020). Digital Education Action Plan 2021-2027: Resetting education and training for the digital age. European Commission.

[8] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-340. https://doi.org/10.2307/249008

[9] Fernández, M. G. (2024). Modelo de Analítica de Aprendizaje centrado en Competencias Digitales para estudiantes universitarios [Plan de tesis doctoral no publicada]. Universidad Nacional del Nordeste.

[10] Fernández, M. G., & Mariño, S. I. (2025). Modelo de Analítica de Aprendizaje: Competencia Digital inteligente para estudiantes universitarios. En XXXI Congreso Argentino de Ciencias de la Computación CACIC 2025. https://d3r3nlpo48l22w.cloudfront.net/uploads/CACIC-2025_finalOCR.pdf

[11] Florez Cueva, M. Y., Carazas Durand, C. R., & García Soller, T. M. (2026). Implicancias del emprendimiento digital: una revisión sistemática. Revista InveCom, 6(1), 1-10. https://zenodo.org/records/15362337

[12] George Reyes, C. E., & Avello-Martínez, R. (2021). Alfabetización digital en la educación. Revisión sistemática de la producción científica en SCOPUS. Revista de Educación a Distancia, 21(66). http://dx.doi.org/10.6018/red.444751

[13] Gilster, P. (1997). Digital literacy. Wiley.

[14] Marikyan, D., & Papagiannidis, S. (2023). Unified Theory of Acceptance and Use of Technology: A review. In S. Papagiannidis (Ed.), TheoryHub Book. https://open.ncl.ac.uk/theories/1/utaut/

[15] Mariño, S. I., Valesani, M. E., & Armana, S. (2023). TIC y adultos mayores. Formación en "Introducción a la informática". En 2° Congreso de Educación y Tecnologías del Mercosur.

[16] Mariño, S. I., Alfonzo, P. L., Bercheñi, V. R., & Armana, S. V. (2025). Trayecto formativo para gestionar unidades de negocios en emprendimientos.

Competencias digitales como estrategia de inclusión. Campus Virtuales, 14(2), 165-178. https://doi.org/10.54988/cv.2025.2.1618

[17] Mariño, S. I., & Bercheñi, V. R. (2025). ODS 4 y brechas digitales: Estudio de caso en dos asignaturas universitarias. Evento Universidad Nacional del Nordeste. Secretaría General de Extensión Universitaria.

[18] Martin, A., & Grudziecki, J. (2006). DigEuLit: Concepts and tools for digital literacy development. Innovation in Teaching and Learning in Information and Computer Sciences, 5(4), 249-267. https://doi.org/10.11120/ital.2006.05040249

[19] Mendoza Cuzcano, J. S., Zúñiga Peña, L. M., Trujillo Robles, P. L., & Sallo Accostupa, V. (2026). Revisión sistemática de la transformación digital en Hispanoamérica: retos, tendencias y perspectivas. Revista InveCom, 6(1), e601060. https://doi.org/10.5281/zenodo.15421144

[20] OECD (2021). OECD Skills Outlook 2021: Learning for Life. OECD Publishing. https://doi.org/10.1787/0ae365b4-en

[21] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. EDUCAUSE Review, 46(5), 30-40.

[22] UNESCO (2017). Educación para los Objetivos de Desarrollo Sostenible: objetivos de aprendizaje. https://unesdoc.unesco.org/ark:/48223/pf0000252423

[23] UNESCO (2018). Skills for a Connected World. UNESCO.

[24] UNESCO (2021). Reimagining our futures together: A new social contract for education. UNESCO.

[25] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425-478. https://doi.org/10.2307/30036540

[26] Vivanco-Galván, O. A., Castillo-Malla, D., & Jiménez-Gaona, Y. (2018).

[27] HACKATHON multidisciplinario: fortalecimiento del aprendizaje basado en proyectos. Revista Electrónica Calidad en la Educación Superior, 9(1), 119-135.

[28] Xin, Y., Weina, H., & Yan, D. (2025). Digital literacy impacts quality of life among older adults through hierarchical mediating mechanisms. Scientific Reports, 15, 19288. https://doi.org/10.1038/s41598-025-04472-9.

ALFADATIZANDO: Las humanidades digitales como puente hacia una alfabetización digital crítica en datos e inteligencia artificial desde la educación abierta.

Alejandra B. Lliteras [1,2][0000-0002-4148-1299], Alejandro Artopoulos [2,3,4][0000-0003-4554-2412],
Alejandro Fernández [1,2][0000-0002-7968-6871]

[1] UNLP, Facultad de Informática, LIFIA, [2] CICPBA,

[3] UDESA, Escuela de Educación, [4] UBA

alejandra.lliteras@lifia.info.unlp.edu.ar, alepoulos@udesa.edu.ar,
alejandro.fernandez@lifia.info.unlp.edu.ar

RESUMEN

El trabajo presenta una línea de investigación en curso que se enmarca en un tema doctoral con el fin de desarrollar alfabetización digital crítica en datos e inteligencia artificial (IA), en el nivel secundario. Este trabajo integra las humanidades digitales (HD) como puente metodológico y se sustenta en los principios de la educación abierta (EA) para fomentar una ciudadanía digital reflexiva.

La integración de las humanidades digitales permite diseñar estrategias de enseñanza que incorporan métodos computacionales para el análisis de objetos culturales y sociales. En este sentido, la educación abierta no solo garantiza el acceso democratizado a estos recursos y saberes, sino que promueve una praxis transparente: al abrir los procesos de procesamiento y análisis de datos, se fomenta una reflexión profunda sobre la mediación tecnológica. De este modo, se busca que, en el nivel secundario, se trascienda el uso instrumental de la tecnología digital para cuestionar cómo se construye el conocimiento a gran escala, promoviendo una ciudadanía digital con agencia para realizar análisis críticos en disciplinas de humanidades y ciencias sociales.

CONTEXTO

Esta línea de investigación profundiza la propuesta presentada en este Workshop en el año 2025 [Lliteras et al., 2025] donde se abordó el desarrollo del pensamiento computacional y la alfabetización en datos, en nivel secundario, considerando las humanidades digitales. Se presentó el Recurso Educativo Abierto Alfadatizando, una plataforma educativa que integra estrategias de visualización de datos para llevar adelante lo que se conoce como lectura distante y análisis de redes sociales y se abordó la importancia de la reflexión crítica en el uso de la tecnología y la mediación algorítmica, construyendo desde el ámbito educativo una participación digital activa.

La temática se enmarca en un proyecto de doctorado en curso [Lliteras, 2020].

1- INTRODUCCIÓN

La alfabetización digital crítica, desde la perspectiva de Loveless y Williamson, propone una visión de la alfabetización digital que va más allá de la de adquisición de habilidades técnicas y del uso funcional de herramientas o plataformas digitales. Este enfoque sostiene que la tecnología educativa no debe verse como un conjunto de dispositivos o de software neutrales, sino como un sistema sociotécnico, donde lo tecnológico y lo social se constituyen mutuamente. Por lo tanto, la alfabetización implica reconocer que estos artefactos simbolizan valores políticos, económicos y culturales que no pueden separarse de su uso educativo [Loveless & Williamson, 2017].

A finales del siglo XX y principios del XXI, la evolución tecnológica permitió que las interacciones digitales y aspectos de la realidad que antes no se consideraban información se transformaran en datos cuantificados, un

fenómeno popularizado como "datificación" [Mayer-Schönberger & Cukier, 2013].

De este modo, la necesidad de alfabetización critica en datos se evidencia al observar cómo los sistemas digitales comenzaron a operar extrayendo datos mediante procesos algorítmicos opacos que actúan como cajas negras. Estas dinámicas generan asimetrías de poder e información, erosionando la autonomía y el libre albedrío de las personas [Pangrazio & Selwyn, 2023].

Por otro lado, con el avance de la tecnología, la alfabetización digital crítica en inteligencia artificial emerge como una capa integradora que requiere comprender tanto la infraestructura técnica como las implicaciones éticas y sociales de su uso. Según Velander et al. esta alfabetización debe trascender la competencia funcional para que los estudiantes y docentes no solo sean capaces de construir tecnología, sino también de deconstruirla para así poder analizar críticamente cómo estos sistemas moldean la vida cotidiana y cómo las estructuras de poder influyen en su diseño y despliegue, buscando de esta manera, el empoderamiento de los ciudadanos [Velander et al., 2024a] [Velander et al., 2024b].

Adicionalmente, la alfabetización crítica en IA debe centrarse en el desarrollo de la agencia humana y la autonomía frente a la automatización de la toma de decisiones, entre otros aspectos. Esto significa educar a las personas no solo para que sean operarios eficientes de sistemas inteligentes, sino para que mantengan su capacidad de actuar libremente y de evitar delegar decisiones éticas a algoritmos que operan bajo lógicas estadísticas carentes de sentido social. Así, la alfabetización crítica en IA se configura como la capacidad de evaluar cuándo y cómo la tecnología aumenta la cognición humana y cuándo amenaza con reducir la educación a procesos de condicionamiento operante, asegurando que la tecnología sirva al desarrollo del sujeto [Holmes & Tuomi, 2022] [Tuomi, 2023] [Tuomi, 2024].

Bajo este marco, la alfabetización crítica en datos e inteligencia artificial requiere, por un lado, comprensión teórica acerca de la opacidad algorítmica y por otro, de una praxis metodológica que permita, a los docentes y estudiantes, interactuar con la escala y la complejidad de los datos. Es en esta intersección donde las Humanidades Digitales aportan las herramientas necesarias para materializar dicha alfabetización.

Las HD son un campo multidisciplinar que investiga y aplica métodos digitales en la producción de conocimiento en humanidades y ciencias sociales, y que combina métodos cualitativos y cuantitativos. Las HD requieren del pensamiento computacional [Berry & Fagerjord, 2017] para estudiar los medios digitales y para la creación y análisis de soluciones críticas.

Complementariamente, esta propuesta se fundamenta en los principios de la educación abierta, bajo la premisa de que el conocimiento sobre el funcionamiento de estas tecnologías no debe constituir una caja negra, sino un bien común accesible [UNESCO, 2021]. Dado que los datos masivos y la inteligencia artificial han transformado la producción cultural y el desarrollo del pensamiento humano, la apertura resulta indispensable para trabajar en comunidad. De este modo, la educación abierta actúa como un marco ético que permite analizar, reflexionar y proponer estrategias ante los nuevos desafíos educativos, garantizando que la alfabetización digital crítica se desarrolle con foco en la equidad, la igualdad y la diversidad.

En el escenario presentado hasta el momento, la lectura distante [Moretti, 2013] surge como una estrategia plausible para el desarrollo de la alfabetización crítica en datos e inteligencia artificial. A diferencia de la lectura tradicional o “cercana”, la lectura distante permite trabajar con grandes volúmenes de datos utilizando algoritmos y visualizaciones que permiten identificar patrones, tendencias y analizar sesgos que podrían ser imperceptibles de otra manera. Al integrar la lectura distante en el nivel secundario, es posible 'desarmar' la caja negra de los datos, pasando de una recepción pasiva de la información a un análisis basado en la evidencia computacional. Así, el uso de estos

métodos expande las capacidades analíticas en las humanidades y las ciencias sociales, y se constituye como una práctica donde docentes y estudiantes interrogan a los datos y analizan las lógicas que subyacen a los mismos y, en consecuencia, a los modelos de inteligencia artificial.

2- LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A partir de los avances presentados en las ediciones anteriores de este Workshop, esta línea de investigación expande su alcance hacia la alfabetización digital crítica en datos e inteligencia artificial y adopta conceptos de la educación abierta. El objetivo es fortalecer la agencia de ciudadanos digitales en el nivel secundario a través de dos ejes integrados.

- Definición de un marco metodológico para el diseño de Recursos Educativos Abiertos (REA) que permita empoderar a los docentes, transformando la tecnología de una "caja negra" a un bien común.
- Desarrollo de la plataforma educativa Alfadatizando como el entorno que permite materializar el marco metodológico del eje anterior y que da soporte a una comunidad de práctica [Wenger, 1998]. El diseño se orienta a facilitar el intercambio de saberes, la transparencia en el análisis de datos y la co-creación de recursos educativos, promoviendo una innovación educativa basada en la equidad y la soberanía tecnológica.

3- RESULTADOS OBTENIDOS

Como parte de los resultados obtenidos se cuenta con Alfadatizando 2.0[1] una plataforma educativa abierta que facilita trabajar con visualización de datos en áreas de humanidades y ciencias sociales y que promueve y alberga una comunidad de práctica. Desde esta plataforma es posible crear y considerar las 5R *("Retain"," Reuse", "Revise", "Remix", "Redistribute"*) de los recursos educativos abiertos [Wiley & Hilton Iii, 2018]. Desde Alfadatizando, es posible acceder a un sitio de soporte del proyecto llamado Alfadatizando.Info[2] Dicho sitio presenta recursos educativos abiertos como secuencias didácticas, tutoriales y tips para que docentes y estudiantes encuentren un espacio de consulta e intercambio.

Por otro lado, desde la formación de recursos humanos, se presentó el informe final y se defendió una tesina de grado titulada "Visualización de datos en el marco de las humanidades digitales para estudiantes secundarios".

Como parte del trabajo realizado con esta línea de investigación se realizaron diversos tipos de publicaciones.

Respecto a las publicaciones realizadas, las mismas se listan a continuación:

Artículos en revistas [Lliteras& Artopoulos, 2025c] Artículos en congresos [Lliteras& Artopoulos, 2025a], [Lliteras& Artopoulos, 2025b]. Artículos en workshops [Lliteras et al., 2025]. Ponencia en congreso [Lliteras & Artopoulos, 2025d].

Esta línea de investigación fue presentada en los talleres[3] ofrecidos por el Hub Ibero Americano de GO-GN en enero de 2026 en la ciudad de Monterrey México, donde la primera autora asistió becada por dicho grupo. Por otra

[1] https://alfadatizando.okd.lifia.info.unlp.edu.ar/

[2] https://alfadatizandoinfo.wordpress.com/

[3] Se desarrollaron en el marco del Proyecto Iberoamericano GO-GN Hub, coordinado por el Instituto para el Futuro de la Educación en Europa, España. GO-GN y sus centros regionales están financiados actualmente a través del programa OER de la Fundación William y Flora Hewlett y administrados por el Centro de Investigación en Educación Abierta del Instituto de Tecnología Educativa de la Universidad Abierta del Reino Unido.

parte, la autora, fue aceptada como miembro de la red de estudiantes doctorales de GO-GN[4].

4- RESULTADOS ESPERADOS

A partir de las líneas de investigación planteadas, se proyectan los siguientes resultados:

- Evolución Tecnológica (Alfadatizando 3.0): Se prevé la mejora de la plataforma en su versión 3.0, integrando módulos de ingeniería de software que den soporte a la metodología de co-creación de REA. Este desarrollo se formalizará a través de una tesina de grado que profundizará en la arquitectura técnica necesaria para sostener procesos colaborativos.
- Validación en Dominios Específicos: Se espera obtener resultados preliminares sobre la efectividad de la alfabetización digital crítica aplicada al campo de las lenguas extranjeras (inglés), mediante el avance de una tesis de maestría vinculada. Esto permitirá validar la versatilidad del marco metodológico en diferentes áreas de las humanidades.
- Avance en la Investigación Doctoral: Finalmente, se espera avanzar en la tesis doctoral que motiva esta línea de trabajo, mediante la sistematización del marco teórico-metodológico y la publicación de artículos científicos

5- FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación permitio consolidar un espacio para la formación de recursos humanos en distintos niveles académicos, integrando perfiles de grado y posgrado en el estudio de la alfabetización digital crítica y la ingeniería de software. Actualmente, la formación se distribuye de la siguiente manera:

- Nivel de Grado: Se cuenta con una (1) tesinas de grado aprobada que sentaron las bases técnicas de la plataforma y que permitió la formación de dos estudiantes. En la actualidad, se encuentran dos (2) estudiantes enfocados en la evolución tecnológica de Alfadatizando, abordando aspectos de ingeniería de software y metodologías de diseño para la creación de recursos educativos abiertos.
- Nivel de Posgrado: Se dirige una (1) tesis de maestría en curso, cuyo foco es la alfabetización crítica a la enseñanza de lenguas extranjeras. Asimismo, la línea se sustenta en el avance de una (1) tesis doctoral en curso, la cual provee el marco teórico-metodológico integrador para los demás trabajos de investigación.
- Otras actividades: Se prevé continuar con la incorporación de estudiantes de títulos intermedios y pasantes, orientados a tareas de desarrollo, testeo de usabilidad y soporte técnico de la plataforma educativa, fomentando una participación en proyectos de investigación desde los inicios de la formación universitaria.

6- BIBLIOGRAFÍA

[Berry & Fagerjord, 2017] Berry, D. M., & Fagerjord, A. Digital humanities: Knowledge and critique in a digital age. John Wiley & Sons (2017).

[Loveless & Williamson, 2017] Loveless, A. y Williamson, B. Nuevas identidades de aprendizaje en la era digital: creatividad, educación, tecnología, sociedad. Narcea. (2017)

[Holmes & Tuomi, 2022] Holmes, W., & Tuomi, I. State of the art and practice in AI in education. European

[4] https://go-gn.net/

journal of education, 57(4), 542-570. (2022). https://doi.org/10.1111/ejed.12533

[Lliteras, 2020] Lliteras A., “Aprendo Sociales con Datos. Con el fin de construir conocimiento, en Ciencias Sociales, a partir del análisis de conjuntos de datos usando modelización, visualización y razonamiento algorítmico para escuelas secundarias de Argentina.”. Propuesta doctoral. UNLP. Facultad de Informática. Diciembre de 2020.

[Lliteras & Artopoulos, 2025a] Lliteras A., Artopoulos A: Alfabetización Critica en la Enseñanza de las Humanidades Digitales.XXXI Congreso Argentino de Ciencias de la Computación - CACIC 2025. Octubre. Sede Atlántica de la Universidad Nacional de Río Negro. Argentina. pp.276-285. ISBN 978-987-8258-99-7.

[Lliteras & Artopoulos, 2025b] Lliteras A., Artopoulos A: Enseñanza de las Humanidades Digitales con Datos: Un Enfoque Metodológico. XXXI Congreso Argentino de Ciencias de la Computación - CACIC 2025. Octubre. Sede Atlántica de la Universidad Nacional de Río Negro. Argentina. pp.296-305. ISBN 978-987-8258-99-7.

[Lliteras & Artopoulos, 2025c] Lliteras, A. B., & Artopoulos, A. (2025). Explorando la enseñanza de las humanidades digitales en secundaria: alfabetización crítica en datos e inteligencia artificial a través de actividades educativas con lectura distante. Publicaciones De La Asociación Argentina De Humanidades Digitales, 6, e071. https://doi.org/10.24215/27187470e071

[Lliteras & Artopoulos, 2025d] Lliteras A.B., Artopoulos A. (2025). Alfabetización digital crítica en inteligencia artificial: las humanidades digitales como puente metodológico en la enseñanza. 3º Congresso Internacional de Pesquisadores e Acadêmicos da Ibero-América, promovido pela RedeIPA, Revista Veritas e Alianza Red, realizado entre os dias 08 e 10 de outubro de 2025. URL: https://www.youtube.com/watch?v=_0cFN74g-qQ. *Último acceso: 2026-02-24*

[Lliteras et al., 2025] Lliteras A, Artopoulos A, Fernandez A. Alfadatizando: Enseñanza en nivel secundario de Pensamiento Computacional para Humanidades Digitales. Workshop de investigadores en ciencias de la computación WICC 2025: libro de actas / Marisa Fabiana Haderne; Compilación de Marisa Fabiana Haderne. - 1a ed. - Mendoza: Universidad Nacional de Cuyo. Secretaría de Investigación, Internacionales y Posgrado, 2025. pp.655-658. ISBN 978-987-575-267-2 http://sedici.unlp.edu.ar/handle/10915/184413

[Mayer-Schönberger & Cukier, 2013] Mayer-Schönberger, V., & Cukier, K.. Big data: la revolución de los datos masivos. Turner. (2013)

[Moretti, 2013] Moretti, F. Distant reading. Verso Books (2013).

[Pangrazio & Selwyn, 2023] Pangrazio, L., & Selwyn, N. Critical data literacies: rethinking data and everyday life. MIT Press. (2023).

[Tuomi, 2023] Tuomi, I. A framework for socio-developmental ethics in educational AI. In Proceedings of the 56th Hawaii International Conference on System Sciences, January 2023. (2023). https://doi.org/10.24251/HICSS.2023.752

[Tuomi, 2024] Tuomi, I. Beyond mastery: Toward a broader understanding of AI in education. International Journal of Artificial Intelligence in Education, 34(1), 20-30. (2024). https://doi.org/10.1007/s40593-023-00343-4

[UNESCO, 2021] Global, O. E. Recomendación sobre los Recursos Educativos Abiertos (REA). *Annotating the UNESCO Recommendation on OER*. (2021).

[Velander et al., 2024a] Velander, J., Milrad, M. y Otero, N. Exploring AI literacy in Swedish K-12 education: A study of preexisting beliefs, comprehension, and learning outcomes. En H. Crompton y D. Burke (Eds.), Artificial Intelligence Applications in K-12 (pp. 193-214). Routledge. (2024). https://doi.org/10.4324/9781003440192-12

[Velander et al., 2024b] Velander, J., Otero, N. y Milrad, M. What is critical (about) AI literacy? Exploring conceptualizations present in AI literacy discourse. En A. Buch Y. Lindberg y T. Cerrato-Pargmant (Eds.), Framing futures in postdigital education: Critical concepts for data-driven practices (pp. 139–160). Springer Nature Switzerland. (2024). https://doi.org/10.1007/978-3-031-58622-4_8

[Wenger, 1998] Wenger, E. *Communities of practice: Learning, meaning, and identity.* Cambridge University Press. (1998). https://doi.org/10.1017/CBO9780511803932

[Wiley & Hilton Iii, 2018] Wiley, D., & Hilton Iii, J. L. Defining OER-enabled pedagogy. The International Review of Research in Open and Distributed Learning, 19(4). (2018).

https://doi.org/10.19173/irrodl.v19i4.3601

Análisis de la prosodia emocional japonesa mediante Reconocimiento de Emociones del Habla (SER)

Alejandro Vega[1] , Martín Bilbao[1], Marcelo A. Falappa[1,2]

(1) Grupo de Investigación en Inteligencia Artificial Facultad de Ingeniería
Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB)
Ruta 1 Km 4 Comodoro Rivadavia - Chubut
alejandrovega@unpata.edu.ar, martinbilbao@ing.unp.edu.ar,
(2) Departamento de Ciencias e Ingeniería en Computación
Universidad Nacional del Sur (UNS)
San Andrés 800 – Campus de Palihue - Bahía Blanca
mfalappa@cs.uns.edu.ar

Resumen

Esta investigación explora la evolución de las emociones durante el aprendizaje del japonés como segunda lengua (L2) en hispanohablantes nativos. Utilizando tecnología de Reconocimiento de Emociones en el Habla (SER), se analizan las características prosódicas al enfrentar una lengua caracterizada por un sistema de *pitch accent* (acento tonal) frente al acento de intensidad del español. Nuestra hipótesis postula que la prosodia evoluciona con la competencia: los niveles iniciales muestran marcadores de ansiedad por la elevada carga cognitiva, mientras que los avanzados tienden a la estabilidad prosódica de su lengua materna.

Palabras claves: Japonés como segunda lengua, reconocimiento de emociones en el habla, prosodia emocional, carga cognitiva, acento tonal.

Contexto

El proyecto de investigación se desarrolla bajo convenio de actividad interinstitucional entre el Departamento de Ciencias e Ingeniería en Computación de la UNS y el Grupo de Investigación en Inteligencia Artificial de la UNPSJB.

Introducción

La influencia de los factores afectivos en la adquisición de una L2 es un área de investigación consolidada. La ansiedad lingüística actúa como un filtro que puede bloquear el procesamiento del *input* y mermar el rendimiento académico (Fonseca Mora, 2005; Botes et al., 2020). Estudios contemporáneos, influenciados por la psicología positiva, sugieren que el disfrute y la autoconfianza son predictores clave del éxito en el aula (Derakhshan, 2022; Wang et al., 2021).

Históricamente, la medición de estos estados dependía de autoinformes, pero el Reconocimiento de Emociones en el Habla (SER) ofrece una alternativa objetiva mediante el análisis de la señal de voz (Joshi & Zalte, 2013). En el caso del japonés para hispanohablantes, la carga cognitiva se eleva por la distancia tipológica entre una lengua de ritmo silábico y una de ritmo moraico con acento tonal (Ueda, 1978; Martínez Martínez, 2002). Esta discrepancia genera una fatiga que se correlaciona negativamente con el logro académico (Teimouri et al., 2019).

Líneas de Investigación, Desarrollo e Innovación

La mayoría de los sistemas SER actuales son monolingües o se prueban en lenguas cercanas. Nuestro trabajo aborda la vacancia de aplicar SER en un par de lenguas tipológicamente distantes (Español-Japonés)(Han, 2025;Farooq, 2022). El objetivo es validar si estas distorsiones prosódicas pueden servir como biomarcadores fiables del estado afectivo, permitiendo a los docentes identificar bloqueos emocionales y fomentar un clima de aula positivo que recupere el compromiso académico(Wang, 2021; Ebn-Abbasi).

- ***Análisis del Conflicto Prosódico***: Investigación de la transferencia del ritmo silábico al moraico y su impacto en la carga cognitiva del estudiante.
- ***Modelado SER Translingüe***: Evaluación de la robustez de modelos como Sense Voice ante la interferencia del tono léxico japonés (Han et al., 2025).
- ***Aprendizaje Profundo Aplicado a la Educación***: Implementación de arquitecturas CNN-Transformer para la clasificación precisa de estados afectivos (Li, 2025).

Resultados y Objetivos

Se diseñó un corpus limitado con desafíos fonéticos específicos para evaluar la producción bajo estrés. Los resultados preliminares, comparados con bases de datos de referencia (Garcia-Cuesta et al., 2024), muestran que mientras la línea base en L1 refleja estabilidad emocional, la producción en L2 para niveles iniciales arroja la etiqueta UNKNOWN. Esto evidencia un "bloqueo prosódico" donde el esfuerzo por gestionar el pitch accent satura la capacidad expresiva del hablante.

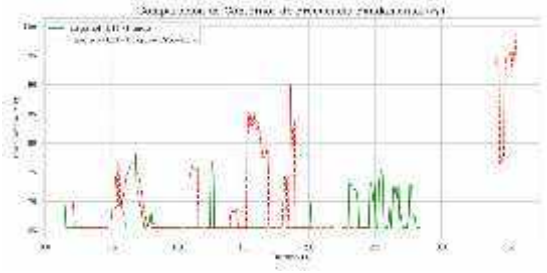


Fig. 1. Comparación de Contornos de F0. La línea roja (L2) muestra inestabilidad severa con picos erráticos (jitter) y rupturas de continuidad, en contraste con la estabilidad de la línea base en L1 (verde).

Análisis de Frecuencia Fundamental F0: como se observa en la Fig. 1. La producción en español (línea verde) mantiene una estabilidad relativa en el rango de 65-75 Hz. Por el contrario, la producción en japonés (línea roja) presenta un comportamiento caótico caracterizado por picos abruptos que no corresponden a patrones de entonación léxica, sino a inestabilidad laríngea provocada por el estrés.

Como trabajo en progreso, proponemos el desarrollo de una aplicación móvil diseñada para **recopilar** muestras de habla de estudiantes ante desafíos fonéticos específicos, con el fin de consolidar un corpus translingüe robusto. Estos datos permitirán optimizar nuestro módulo de detección dentro de un **'Tutor Empático'** interactivo. El sistema emplea la etiqueta UNKNOWN como un indicador proactivo para iniciar estrategias de **andamiaje (scaffolding)** y síntesis de voz con entonación tranquilizadora, cerrando el ciclo afectivo y restaurando la seguridad del aprendiz.

Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por tres investigadores, dos docentes investigadores y un alumno del Doctorado en Ciencias de la Ingeniería de la UNPSJB, cuya tesis de posgrado en desarrollo explora, entre otros, la temática abordada en esta línea de trabajo.

Referencias

Fonseca Mora, M.C. (2005). El componente afectivo en el aprendizaje de lenguas.

Botes, E., Dewaele, J. M., & Greiff, S. (2020). The foreign language classroom anxiety scale and academic achievement: An overview of the prevailing literature and a meta-analysis. The Journal for the Psychology of Language Learning, 2(1), 26-56.

Derakhshan, A. (2022). Revisiting Research on Positive Psychology in Second and Foreign Language Education: Trends and Directions. Language Related Research,13(5), 1-43.

Wang, H., Peng, A., & Patterson, M. M. (2021). The roles of class social climate, language mindset, and emotions in predicting willingness to communicate in a foreign language. System, 99, 102529.

Joshi, D. D., & Zalte, M. B. (2013). Speech Emotion Recognition: A Review. IOSR Journal of Electronics and Communication Engineering, 4(4), 34-37.

Ueda, H. (1978). Estudio contrastivo de los sonidos españoles y japoneses (2): consonantes.

Martínez Martínez, I. (2002). Un estudio contrastivo español - japonés. Fundación Canon.

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. Studies in Second Language Acquisition, 41(2), 363-387.

Han, Z., Yuan, J., Geng, T., Richmond, K., Feng, H.,Li,Y.(2025). Cross-Lingual Speech Emotion Recognition: Humans vs.Self-SupervisedModels.arXiv:2409.16920v2.

Farooq, M. U., & Hain, T. (2022). Investigating the Impact of Cross-lingual Acoustic-Phonetic Similarities on Multilingual Speech Recognition. Interspeech 2022, 3849-3853.

Ebn-Abbasi, F., Fattahi, N., Sayyahi, M. J., \& Nushi, M. (n.d.). Language Learners' Mindset and Their Academic Engagement in Online Classrooms: The Mediating Role of Achievement Emotions.

Li, B. (2025). Speech Emotion Recognition Based on CNN-Transformer with Different Loss Function. Journal of Computer and Communications, 13(3), 103-115.

Garcia-Cuesta, E., Salvador, A. B., \& Gachet Páez, D. (2024). EmoMatchSpanishDB: study of speech emotion recognition machine learning models in a new Spanish elicited database. Multimedia Tools and Applications, 83, 13093-13112.

APLICACIÓN DE MODELOS DE INTELIGENCIA ARTIFICIAL PARA LA ACCESIBILIDAD DE CONTENIDO EDUCATIVO

Javier Díaz, Alejandra Schiavoni, Paola Amadeo, Ivana Harari,
Alejandra Osorio
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata
Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528
jdiaz@unlp.edu.ar , ales@info.unlp.edu.ar , pamadeo@linti.unlp.edu.ar, iharari@info.unlp.edu.ar,
aosorio@cespi.unlp.edu.ar

RESUMEN

La discapacidad surge de la fricción entre las capacidades de una persona y las barreras de su entorno. En el ámbito educativo digital, una de las barreras más evidentes es la falta de acceso a los recursos educativos, lo que deja a estudiantes con discapacidad en una clara desventaja. Las personas mayores y las personas con discapacidad pueden interactuar con el contenido web utilizando distintos medios y tecnologías de hardware o software denominados productos de apoyo. Por este motivo, resulta fundamental que tanto las páginas web, las plataformas y recursos educativos se implementen utilizando las mejores prácticas en diseño y accesibilidad, para facilitar el acceso a la información. El proceso de accesibilizar la información ya existente en la web es en general manual y lento debido a la ausencia de herramientas específicas que faciliten la realización de esta tarea. Este artículo presenta las líneas de estudio e investigación para el desarrollo de una herramienta de apoyo tecnológico para automatizar la accesibilidad de material educativo, utilizando modelos de Inteligencia Artificial (IA). Durante la implementación se realizan testeos con material educativo utilizado en cátedras de las carreras de la Facultad de Informática de la UNLP.

Palabras clave: recursos educativos, accesibilidad, inteligencia artificial

CONTEXTO

En la Facultad de Informática de la Universidad Nacional, particularmente en el LINTI, Laboratorio de Investigación en Nuevas Tecnologías Informáticas, desde el año 2002 se lleva adelante un trabajo sostenido en torno a la accesibilidad que ha permitido construir una perspectiva que coloca la inclusión como un paradigma que significa construir oportunidades para todos y todas [1]. También se institucionalizó su abordaje mediante la creación de una Dirección de Accesibilidad desde el año 2010. En este marco se desarrollaron tesinas, trabajos de cátedra al respecto, proyectos de innovación y desarrollo con estudiantes de Informática y diferentes proyectos de extensión acreditados por la Universidad Nacional de La Plata. Los proyectos de los últimos años son “Trabajando por una Web Accesible”, “Por una Web Inclusiva”, “Por una Web inclusiva: abordaje en escuelas secundarias”, que entró en el Programa de promoción de

derechos y fortalecimiento de la organización comunitaria [2].

Cabe mencionar también que se llevó a cabo el proyecto del Kit tecnológico para ciegos, en el marco de la convocatoria de Proyectos de Extensión Universitaria, "Universidad, Cultura y Territorio", Área Ampliación y Consolidación de Derechos, que se realizó junto con escuelas técnicas de la ciudad de La Plata y alrededores (Berisso y Ensenada) en los periodos 2022-2023 y 2024-2025.

Las líneas de investigación que se mencionan en este artículo se desarrollan en el LINTI de la UNLP y están enmarcadas en el proyecto "Tecnologías digitales para la Inclusión, la Equidad y la Sostenibilidad" a partir del 2025, bajo la dirección del Lic. Javier Díaz. Las líneas que aquí se presentan continúan con lo presentado en WICC 2025, abordando la temática de accesibilidad y usabilidad de aplicaciones utilizando [3].

1. INTRODUCCIÓN

Nos encontramos en la era de las Tecnologías de la Información y la Comunicación (TIC) en la cual se ha constituido y acrecentado un modelo de Universidad Digital [4]. Se establece una modalidad de enseñanza virtual donde las universidades ofrecen nuevas oportunidades y posibilidades a personas que antes tenían dificultades en acceder a una formación superior.

Las personas mayores y las personas con discapacidad pueden interactuar con el contenido web utilizando distintos medios y tecnologías de apoyo. Por este motivo, resulta fundamental que las páginas web se implementen utilizando las mejores prácticas en diseño y accesibilidad, para facilitar el acceso a la información [5]. Las instituciones de educación superior tienen la obligación de conseguir una igualdad de acceso a las oportunidades educativas, así como asegurar que los estudiantes en situación de discapacidad tengan la posibilidad de acceder a todos sus recursos, sin ningún tipo de discriminación [6]. Sin embargo, la mayoría de sitios web presentan barreras que dificultan y en ocasiones imposibilitan a algunas personas el poder usarlos [7]. Estas barreras electrónicas obstaculizan el progreso de un usuario cuando navega por una web y accede a sus recursos utilizando productos de apoyo.

Cuando nos referimos a materiales educativos accesibles, estamos hablando de cualquier recurso para la enseñanza diseñado o adaptado para ser utilizado por el mayor número posible de estudiantes, sin discriminar entre aquellos que están transitando una situación con discapacidad o no. Estos materiales aseguran que los estudiantes con distintas necesidades, ya sean auditivas, cognitivas, etc. puedan acceder a los contenidos y participar de manera equitativa en el proceso de aprendizaje y dependiendo del tipo de material, es necesario contemplar las normas W3C WCAG aplicables para contenido textual, para videos, para presentaciones del tipo slideshow, entre otras. La generación de material accesibilizado es un punto destacado en cuanto a la accesibilidad académica [8]. Entre las principales barreras podemos mencionar problemas de estructura y navegación, incompatibilidad con tecnologías asistivas, contrastes inadecuados, falta de capacitación y conciencia sobre accesibilidad digital.

Desde hace ya muchos años han comenzado a realizarse estudios y desarrollar técnicas y herramientas basadas en IA, que permiten cambiar la forma en que aprendemos, interactuamos y nos comunicamos. También en el contexto de accesibilidad, la IA puede ser

aplicada para mejorar la calidad de vida de las personas en situación de discapacidad, promoviendo la independencia y la accesibilidad física y académica. Se trata de democratizar la IA teniendo una mayor injerencia en las prácticas laborales cotidianas, motivadas por la preocupación de que la capacidad de utilizar técnicas de IA no quede en un campo de desarrollo minoritario. En base a esto se están implementando herramientas que ayudan a hacer testeos y validaciones de accesibilidad y usabilidad de distintos recursos digitales que reducen significativamente el trabajo manual.

En las próximas secciones se desarrollarán las líneas de investigación y desarrollo que se llevan a cabo sobre este tema, haciendo una revisión sistemática de la literatura y el estado del arte para analizar cómo utilizar la IA para mejorar el proceso de accesibilización de los recursos educativos.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Las líneas de investigación y desarrollo continúan con el proyecto sobre accesibilidad que se viene llevando a cabo en el LINTI, y está enmarcado en el proyecto de Incentivos "Tecnologías digitales para la Inclusión, la Equidad y la Sostenibilidad".

En estas líneas de investigación, las tareas que se llevan a cabo son las siguientes:

- Estudiar y analizar aplicaciones orientadas a la ayuda e inserción de personas con alguna discapacidad.
- Analizar la accesibilidad en documentos, considerando su estructura, y con especial atención en el impacto que tiene sobre el grupo de personas con discapacidad visual y ceguera.
- Estudiar la interacción con los recursos educativos que realiza una persona con discapacidad que utiliza diferentes productos de apoyo.
- Estudiar las herramientas existentes que utilizan IA en el campo de la accesibilidad, tanto para testeos, como para la corrección de posibles errores.
- Investigar el trabajo llevado a cabo en la tesina "Accesibilidad en documentos para personas con discapacidad visual y ceguera en CILSA" [9]
- Investigar los modelos transformadores para el procesamiento de lenguaje natural [10]
- Estudiar los modelos generativos multimodales que permiten combinar texto e imagen en una misma arquitectura, habilitando tareas como el image captioning (descripción automática de imágenes), Visual Question Answering (VQA), y diálogo visual.

Considerando estas líneas de investigación es importante mencionar que desde diferentes espacios académicos e institucionales desplegamos prácticas educativas y de trabajo en relación con la accesibilidad. Todos los procesos de formación que se vienen impartiendo tienen como eje el reconocer la relevancia de trabajar la inclusión y accesibilidad, entendiendo que la enseñanza es la

herramienta para plantear nuevos problemas que desafían la formación profesional.

3. RESULTADOS Y OBJETIVOS

Continuando con los análisis de accesibilidad tanto de las plataformas como de los recursos educativos, y con la evaluación de herramientas que permiten la detección e identificación de errores, se comenzó a desarrollar una solución que ayuda en la sistematización del proceso de accesibilización de documentos. En base a esto, se plantean los resultados logrados y los objetivos a seguir:

- Análisis de la accesibilidad en documentos respecto de los aspectos disciplinares y su composición respecto a elementos tales como textos, imágenes, y distintas estructuras.
- Análisis del proceso de evaluación de accesibilidad, llevado a cabo en los recursos de CILSA [9].
- Investigación y comparación de la efectividad de diferentes modelos de IA visión-lenguaje (como BLIP, LLaVA, GPT, entre otros) para interpretar y describir con precisión el contenido de imágenes complejas de carácter educativo, como diagramas y gráficos.
- Diseño y desarrollo de un prototipo de software que ayuda a automatizar la tarea de generar descripciones y textos alternativos para las imágenes contenidas en los documentos.
- Testeo del prototipo desarrollado con recursos educativos de cátedras de las carreras de la Facultad de Informática.
- Validación de los resultados obtenidos por la herramienta en la generación de textos asociados analizando la precisión de los mismos.
- Validación centrada en el usuario con la participación de alumnos y docentes involucrados con el material analizado
- Continuar con el estudio de técnicas de IA para mejorar la automatización en cuestiones de accesibilidad, entre ellas las que abordan las funciones de reconocimiento de voz para conversión a texto, procesamiento del lenguaje natural, texto predictivo, generación automática de textos alternativos para las imágenes.

La accesibilización de los recursos educativos digitales permiten la interacción de los usuarios independientemente de sus condiciones y preferencias, lo que potencia el proceso de enseñanza-aprendizaje. Es importante pensar en accesibilidad no sólo desde el punto de vista de la discapacidad de una persona, sino desde los diferentes elementos que pueden influir para que se genere una inclusión exitosa y el recurso educativo sea para todos [11].

Entre los objetivos que se proponen también se encuentra continuar interactuando con organizaciones no gubernamentales, asociaciones civiles, centros de ayuda con el objetivo de interpelar sobre posibles soluciones que puedan dar respuesta a sus necesidades y requerimientos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por tres profesoras de amplia trayectoria en el campo de la investigación, que trabajan en el área de ambientes virtuales de aprendizaje y accesibilidad web. Además, una alumna becaria de la carrera de Licenciatura en Informática y un alumno tesista de la carrera de Licenciatura en Sistemas de la Facultad de Informática.

La participación en eventos de la especialidad, ha permitido seguir estableciendo canales de comunicación con otros investigadores que trabajan en las mismas áreas.

En la Facultad de Informática se realizan cada año las *Hackaton por la Inclusión*, que reúne a alumnos con el objetivo de trabajar en conjunto, en pos de lograr desarrollos accesibles orientados a la comunidad.

Las Jornadas de Accesibilidad que se realizan todos los años permiten intercambiar experiencias entre los participantes del curso y exponer los trabajos realizados.

5. REFERENCIAS

[1] Díaz, J. Harari, I. Amadeo, P. Schiavoni, A. Osorio, A, Gomez, S. (2022). Higher education and virtuality from an inclusion approach. Computer Science - CACIC 2021. 27th Argentine Congress, CACIC 2021, Salta, Argentina, October 4?8, 2021, Revised Selected Papers.

[2] Díaz, J. Harari, I. Amadeo, P., Ungaro Ana, Osorio, A. y Carmona B. (2024). Recovering Voices to Build Inclusive Computing: From Accessibility, Gender, and Rights to Education. Laclo 2024. Uruguay.

[3] Díaz, J. Harari, I. Amadeo, P. Schiavoni, A. Osorio. WICC 2025 - XXVII Workshop de Investigadores en Ciencias de la Computación, 1era Edición, Mendoza, Universidad Nacional de Cuyo, Secretaría de Investigación, Internacionales y Posgrado, ISBN 978-987-575-267-2.

[4] Laviña, J. y L. Mengual, Libro blanco de la Universidad digital 2010, Ariel, Fundación Telefónica Madrid, España (2008)

[5] Máñez Carvajal, Carlos. Evaluación de accesibilidad web de las universidades chilenas. Formación Universitaria vol.13 no.5 La Serena oct. 2020, versión On-line ISSN 0718-5006. http://dx.doi.org/10.4067/S0718-50062020000500069

[6] Fernández-Morales, F. H. y J. E. Duarte, Retos de la inclusión Académica de Personas con Discapacidad en una Universidad Pública Colombiana, doi:10.4067/S0718-50062016000400011, Formación universitaria, 9(4), 95-104, (2016).

[7] Harper, S. y A. Q. Chen, Web accessibility guidelines, doi:10.1007/s11280-011-0130-8 World Wide Web, 15(1), 61-88, (2012)

[8] Katz, Sandra. 2020. Políticas De Discapacidad / Accesibilidad en la UNLP. Elaboración del diseño de un documento para la definición de una política institucional de discapacidad / accesibilidad en la UNLP. Programa de Posgrado Gestión de la Educación Superior (EGES) https://sedici.unlp.edu.ar/handle/10915/115288

[9] Huertas Godoy, D. Accesibilidad en documentos para personas con discapacidad visual y ceguera en CILSA, Tesina de Grado, Licenciatura en Sistemas, 2023. http://sedici.unlp.edu.ar/handle/10915/158032

[10] Vaswani, A, Shazeer, N, Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł., Polosukhin, I. (December 2017). "Attention is All you Need". https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[11] UNESCO. (2013). Enfoques Estratégicos sobre las TICS en educación en América Latina y el Caribe. Enfoque Estratégico sobre Tics en educación en América Latina y El Caribe, 62. Retrieved from http://www. unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/images/ticsesp.pdf

APORTES A LA TRANSFORMACION DE LA EDUCACIÓN DIGITAL. USABILIDAD E INTERACCIÓN INMERSIVA CON TECNOLOGÍAS INFORMÁTICAS EMERGENTES

Sosa Mabel, Digión Leda, Maldonado Ivanna, Silva Carmen, Velázquez Isabel
FCEyT, Universidad Nacional de Santiago del Estero
{litasosa@unse.edu.ar}

RESUMEN

Los paradigmas actuales de interacción, que incluyen tecnologías de realidad virtual, realidad aumentada y experiencias en el metaverso, ofrecen oportunidades inéditas para innovar en los procesos de enseñanza y aprendizaje, promoviendo experiencias de aprendizaje más inmersivas, significativas y participativas. Sin embargo, a pesar de la creciente disponibilidad de recursos y plataformas digitales, la integración de estas tecnologías en la educación no siempre se traduce en mejoras sustantivas en el aprendizaje. Con frecuencia, su uso se limita al acceso a la información o a la reproducción de prácticas tradicionales, sin aprovechar plenamente su potencial para estimular la construcción activa del conocimiento. Para que cumplan efectivamente estos propósitos, su diseño debe atender de manera integrada aspectos vinculados a tres dimensiones esenciales: la usabilidad, la interacción y la experiencia de usuario (UX), garantizando así entornos de aprendizaje más inclusivos, motivadores y enriquecidos.

En ese sentido, el proyecto de investigación propone diseñar y desarrollar herramientas informáticas inmersivas que respondan a las demandas de los contextos educativos en permanente transformación, integrando los principios de usabilidad, interacción y UX. Se busca generar conocimientos y aportes teórico-prácticos que orienten la creación de entornos digitales centrados en el estudiante, capaces de promover experiencias de aprendizaje significativas, colaborativas y activamente comprometidas. Asimismo, el proyecto pretende explorar cómo los paradigmas de interacción emergentes, tales como realidad virtual, realidad aumentada y metaversos, pueden contribuir a innovar las prácticas pedagógicas, fortaleciendo la relación entre tecnología, enseñanza y aprendizaje desde una perspectiva tecnopedagógica integral.

PALABRAS CLAVES: entornos educativos digitales, tecnologías informáticas emergentes, interacción, usabilidad, experiencia de usuario

CONTEXTO

Este trabajo se enmarca en el proyecto de investigación titulado *"Herramientas informáticas emergentes en contextos educativos en permanente transformación: usabilidad, interacción y experiencia de usuario inmersiva"*, que inició en enero de 2026 en el ámbito del Instituto de Investigaciones de Informática y Sistemas de Información, en la Facultad de Ciencias Exactas y Tecnologías, UNSE.

La investigación se inscribe en el campo aplicado a la educación, dentro del área de las Ciencias de la Computación. Su marco disciplinar articula aportes de la Informática Educativa y de la Interacción Persona-Computadora, con el propósito de analizar y desarrollar herramientas tecnológicas que favorezcan experiencias de uso significativas en entornos educativos en constante transformación.

El equipo de trabajo está conformado por profesionales de las áreas de Informática y Ciencias de la Educación, lo que posibilita un abordaje interdisciplinario del objeto de estudio.

INTRODUCCIÓN

En la actualidad, los entornos educativos se encuentran en un proceso de transformación constante, impulsado por los avances tecnológicos, los cambios sociales y las nuevas dinámicas culturales en la producción y gestión del conocimiento. Las tecnologías digitales han modificado las prácticas

pedagógicas tradicionales, promoviendo modalidades de enseñanza y aprendizaje más flexibles, colaborativas y personalizadas (Area & Adell, 2021). Este contexto demanda una revisión profunda de los modelos educativos y de las herramientas tecnológicas que median la práctica docente, dado que la mera incorporación de tecnología no garantiza una mejora en los procesos formativos (Cabero Almenara, 2020). Frente a esta realidad, resulta pertinente la elaboración y construcción de modelos y herramientas informáticas orientadas a optimizar la experiencia educativa, considerando principios de diseño centrado en el usuario y de interacción significativa.

Desde una perspectiva tecnológica, el desarrollo de herramientas digitales para la educación requiere atender tres dimensiones clave: usabilidad, interacción y experiencia de usuario (UX). La usabilidad, entendida como la medida en que un sistema puede ser utilizado con eficacia, eficiencia y satisfacción por usuarios específicos (Nielsen, 1993; ISO, 2019), resulta esencial para garantizar que los entornos de aprendizaje sean accesibles y fáciles de usar. Una buena usabilidad reduce la carga cognitiva y permite que los estudiantes centren su atención en el aprendizaje más que en la herramienta (Preece, Rogers, & Sharp, 2019). La interacción, por su parte, comprende los procesos de comunicación entre el usuario y el sistema, así como entre los propios participantes del entorno educativo, promoviendo experiencias de aprendizaje activas, colaborativas y constructivas (Cabero & Marín, 2018).

Finalmente, la experiencia de usuario amplía esta perspectiva al considerar las emociones, percepciones y niveles de satisfacción que emergen del uso de la tecnología, influyendo directamente en la motivación y el compromiso del estudiante (Hassenzahl & Tractinsky, 2006).

En este contexto, las herramientas para el aprendizaje inmersivo surgen como una extensión natural de las transformaciones digitales en educación, al permitir experiencias de interacción multisensorial que amplían los límites del aula tradicional. Las tecnologías como la realidad virtual (RV), la realidad aumentada (RA) y los entornos de metaverso educativo ofrecen espacios tridimensionales donde el estudiante puede explorar, manipular objetos, participar en simulaciones o colaborar con otros usuarios en tiempo real, potenciando así procesos de aprendizaje experiencial y significativo (Cabero & Fernández, 2018).

En este sentido, los entornos de realidad virtual, aumentada y metaverso educativo abren nuevas posibilidades para explorar cómo la interacción inmersiva y la experiencia sensorial ampliada pueden impactar en la construcción del conocimiento y en el desarrollo de competencias metacognitivas, especialmente en la educación superior (Cabero & Fernández, 2018).

Desde la perspectiva pedagógica, los entornos inmersivos promueven una interacción activa, colaborativa y reflexiva, alineada con las teorías del aprendizaje constructivista y sociocultural. El estudiante deja de ser un receptor pasivo de información para convertirse en agente protagonista de su propio aprendizaje, capaz de experimentar, equivocarse, tomar decisiones y reflexionar sobre su desempeño dentro de un entorno seguro y controlado. Además, el componente emocional y motivacional de la experiencia de usuario adquiere especial relevancia en estos contextos, ya que la sensación de presencia, la empatía con los escenarios y la retroalimentación inmediata favorecen estados de flujo, motivación intrínseca y desarrollo de competencias metacognitivas (Makransky & Lilleholt, 2018).

Integrar estas tres propiedades en el diseño de soluciones informáticas educativas contribuye a la creación de entornos de aprendizaje más inclusivos, interactivos y centrados en el estudiante, en consonancia con los enfoques pedagógicos contemporáneos que promueven el aprendizaje activo, situado y significativo (Coll & Monereo, 2010). Su desarrollo exige integrar criterios rigurosos de usabilidad, que aseguren accesibilidad y facilidad de uso; de interacción, que posibiliten una comunicación significativa con el entorno y con otros usuarios; y de experiencia de usuario, que promuevan placer, motivación y aprendizaje

sostenido. Avanzar en la investigación de estas dimensiones permitirá no solo optimizar el diseño de entornos inmersivos educativos, sino también contribuir a la consolidación de modelos pedagógicos innovadores que sitúen la tecnología al servicio de la experiencia humana y del aprendizaje significativo (Cabero & Marín, 2018; Radianti et al., 2020; Area, 2022).

De esta manera, la investigación orientada al desarrollo de modelos conceptuales y herramientas digitales que prioricen la usabilidad, la interacción y la experiencia de usuario se configura como una vía estratégica para responder a las demandas de la educación actual, en la que la tecnología no solo actúa como un recurso, sino como un mediador esencial del conocimiento, la comunicación y la construcción colectiva del aprendizaje (Area, 2018).

Problemas y objetivos definidos

Los entornos educativos actuales atraviesan una transformación acelerada impulsada por las tecnologías digitales, la globalización de la información y los nuevos modos de comunicación y producción del conocimiento. Si bien estos cambios amplían las posibilidades de acceso, colaboración y personalización del aprendizaje, también generan desafíos pedagógicos y tecnológicos que configuran una problemática compleja para la educación contemporánea. En este contexto de cambio permanente, la integración de tecnologías en los procesos formativos no siempre se acompaña de una reflexión pedagógica profunda ni de un diseño centrado en las necesidades de los usuarios. En efecto, la incorporación acelerada de herramientas digitales suele producir una brecha entre su disponibilidad y su aprovechamiento educativo efectivo.

A ello se suman dificultades vinculadas a la usabilidad e interfaces poco intuitivas que pueden dificultar el aprendizaje en lugar de potenciarlo (Area & Adell, 2021). De manera complementaria, la expansión de los entornos virtuales e híbridos plantea la necesidad de nuevas competencias digitales en docentes y estudiantes. Sin una formación sólida, el uso de las herramientas tecnológicas se vuelve superficial y rutinario, limitando su potencial transformador.

Por consiguiente, la transformación educativa actual no solo implica incorporar tecnología, sino también exige repensar la calidad de la experiencia de aprendizaje que esta posibilita. Resulta necesario diseñar herramientas informáticas accesibles, interactivas y significativas, que desde la experiencia de usuario (UX) promuevan la motivación, la participación y la construcción de sentido (Hassenzahl & Tractinsky, 2006).

En este marco, el problema que orienta el presente trabajo radica en la necesidad de diseñar y desarrollar herramientas informáticas que integren de manera articulada criterios de usabilidad, interacción y experiencia de usuario, con el fin de responder a las demandas de una educación inclusiva, significativa y adaptable a contextos dinámicos. En consecuencia, el objetivo general del estudio es contribuir al diseño y evaluación de entornos y recursos tecnológicos centrados en el usuario, que favorezcan experiencias de aprendizaje motivadoras, accesibles y pedagógicamente pertinentes.

A partir de los avances logrados en trabajos previos sobre usabilidad y experiencia de usuario en sistemas educativos (Leguizamón y Sosa, 2023), esta investigación busca profundizar en el diseño de modelos y herramientas inmersivas que integren una experiencia de usuario centrada en lo motivacional, orientada a incrementar la implicación emocional, la autorregulación y la calidad del aprendizaje. Así, el presente estudio aporta una continuidad innovadora al campo de la tecnología educativa, al proponer una perspectiva tecnopedagógica que vincula la usabilidad, la interacción y la experiencia de usuario con los procesos de aprendizaje inmersivo, contribuyendo al desarrollo de modelos educativos más significativos, inclusivos y sostenibles.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se organiza en líneas de investigación diferenciadas pero interrelacionadas.

Una primera línea aborda la usabilidad y la experiencia de usuario en entornos educativos digitales, con el propósito de definir criterios, indicadores y métodos de evaluación que orienten el diseño centrado en el estudiante. Esta línea funciona como eje metodológico transversal, aportando fundamentos para el desarrollo y validación de las demás propuestas tecnológicas.

Una segunda línea se focaliza en la realidad virtual y la realidad aumentada, explorando su potencial pedagógico para generar experiencias de aprendizaje inmersivas, interactivas y contextualizadas. Aquí se analizan sus implicancias didácticas y se desarrollan aplicaciones experimentales orientadas a mejorar la comprensión y la forma de integración.

Una tercera línea se centra en los metaversos educativos, estudiando sus posibilidades para la colaboración, la construcción social del conocimiento en escenarios híbridos.

Finalmente, una línea integradora articula los aportes de las anteriores para la construcción de un modelo conceptual de referencia, que vincule fundamentos tecnológicos, criterios de usabilidad y estrategias pedagógicas con el propósito de consolidar un andamiaje conceptual integrador que permita fundamentar y validar decisiones de diseño instruccional en contextos digitales complejos y cambiantes.

En conjunto, estas líneas se nutren mutuamente y permiten avanzar de manera coherente hacia el diseño y evaluación de herramientas inmersivas pertinentes para contextos educativos en transformación.

RESULTADOS ESPERADOS

En general, se espera que los resultados obtenidos aporten al fortalecimiento de la innovación educativa, ofreciendo un modelo y un prototipo que pueden ser aplicados en universidades, institutos y centros de formación docente. En concreto, se espera obtener los siguientes resultados:

- Elaboración de un estado del arte sistematizado sobre usabilidad, interacción y experiencia de usuario inmersiva en educación.
- Diseño de un marco de referencia que articule dimensiones tecnológicas, metodológicas y educativas para la creación de herramientas informáticas que den soporte a experiencias de aprendizaje inmersivas.
- Desarrollo de un prototipo funcional que materialice el modelo conceptual en un entorno educativo concreto.
- Obtención de resultados empíricos a partir de pruebas con usuarios, que evidencien efectividad del aprendizaje inmersivo y posibles mejoras en la motivación, comprensión y construcción del aprendizaje.
- Formulación de un conjunto de buenas prácticas para el diseño e implementación de herramientas inmersivas en instituciones educativas.

FORMACIÓN DE RECURSOS HUMANOS

Se prevé el asesoramiento y la capacitación de estudiantes de grado y posgrado de carreras vinculadas con la Informática, promoviendo una articulación teórico-práctica orientada a la formación de recursos humanos. En el nivel de grado, se contempla la participación de estudiantes del Profesorado en Informática; y en el nivel de posgrado, de la Especialización y la Maestría en Informática Educativa, así como de la Especialización en Enseñanza en Tecnologías y otras carreras de la FCEyT de la UNSE.

Asimismo, en continuidad con proyectos anteriores, se incorporarán estudiantes en el marco de becas de investigación y prácticas profesionales supervisadas, en función de las líneas de trabajo y las demandas específicas del proyecto. Esta participación contribuirá al fortalecimiento de competencias investigativas, al desarrollo de capacidades

para la producción científica y tecnológica, y a la consolidación de equipos interdisciplinarios en el ámbito institucional.

BIBLIOGRAFIA

Area Moreira, M. (2018). Hacia la universidad digital: ¿dónde estamos y a dónde vamos? RIED. Revista Iberoamericana de Educación a Distancia, 21(2), pp. 25-30. doi: http:// dx.doi.org/10.5944/ried.21.2.21801

Area, M. (2022). *Educación digital y transformación pedagógica: Retos y oportunidades en la era posdigital*. Ediciones Morata.

Area, M. y Adell, J. (2021). Tecnologías digitales y cambio educativo. Una aproximación crítica. REICE. Revista Iberoamericana sobre Calidad, Efi cacia y Cambio en Educación, 19(4), 83-96. https://doi.org/10.15366/reice2021.19.4.005

Coll, C., & Monereo, C. (2010). Psicología de la educación virtual: aprender y enseñar con las tecnologías de la información y la comunicación. Ediciones Morata.

Cabero Almenara, J., y Fernández Robles, B. (2018). Las tecnologías digitales emergentes entran en la Universidad: RA y RV. RIED. Revista Iberoamericana de Educación a Distancia, 21(2), pp. 119-138. doi: http://dx.doi.org/10.5944/ ried.21.2.20094

Cabero Almenara, J., y Marín-Díaz, V. (2018). Blended learning y realidad aumentada: experiencias de diseño docente. RIED. Revista Iberoamericana de Educación a Distancia, 21(1), pp. 57-74. doi: http://dx.doi.org/10.5944/ ried.21.1.18719

Cabero-Almenara, J. (2020). Tecnología y enseñanza: retos y nuevas tecnologías y metodologías. CITAS, 6(1). https://doi.org/10.15332/24224529.6356

Hassenzahl, M. y Tractinsky, N. (2006). Experiencia de usuario: una agenda de investigación. Behaviour & Information Technology, 25 (2-3), 91–97. https://doi.org/10.1080/01449290500330331

Leguizamon, Fernanda; Sosa, Mabel (2023) Metacognicion propiciada por los atributos de usabilidad en software educativos. Actas CINAIC 2023. doi: 10.26754/Cinaic.2023.0071

International Organization for Standardization. (2019). ISO 9241-210: Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. ISO.

Nielsen, J. (1993). Usability engineering. Academic Press.

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educational Technology Research and Development, 66(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

Preece, J., Rogers, Y., & Sharp, H. (2019). Interaction design: Beyond human-computer interaction (5th ed.). Wiley.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers & Education, 147, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Apropiación de tecnologías digitales en Río Negro. Desafíos educativos en contextos de cambio socioeconómico y productivo

Edith Lovos[1]; Lucía Cantamutto[4], Alejandra Marin Aranda[1], Martín Goin[2], Pablo Fabian Carranza[3], Jose Ignacio Muñoz[1], Mariana Roche[4], Emanuel Llancapan [1]

[1] Universidad Nacional de Río Negro, Sede Atlántica, Viedma, Río Negro
[2] Universidad Nacional de Río Negro, Sede Andina, Bariloche, Río Negro
[3] Universidad Nacional de Río Negro, Sede Alto Valle y Valle Medio, Allen, Río Negro
[4]CONICET, Universidad Nacional de Río Negro, CURZA -Universidad Nacional del Comahue
Centro Interdisciplinario en Derechos, Inclusión y Sociedad (CIEDIS)
{elovos, lcantamutto,mmarin,mgoin,pcarranza,jmunoz}@unrn.edu.ar

RESUMEN

En este trabajo se presentan en forma resumida algunos resultados alcanzados en la segunda etapa de un proyecto de investigación cuyo objetivo es indagar sobre los procesos de apropiación de tecnologías digitales móviles, por parte de los estudiantes de los diferentes niveles educativos y modalidades que tienen lugar en la provincia de Río Negro, y que los mismos adoptan en su quehacer cotidiano.

Palabras clave: TECNOLOGÍAS, PROCESOS DE APROPIACIÓN, INSTITUCIONES EDUCATIVAS, RÍO NEGRO .

CONTEXTO

Se describe brevemente el proyecto de investigación denominado "Procesos de apropiación y uso de tecnologías móviles en Río Negro (Argentina): Desafíos para las instituciones educativas" (PI-UNRN40C1149). El mismo ha sido aprobado y financiado por la Universidad Nacional de Río Negro (UNRN) para desarrollarse en forma bianual (2024-2026), en el marco de las actividades de investigación que tienen lugar en la unidad ejecutora Centro Interdisciplinario de Estudios sobre Derechos Inclusión y Sociedad (CIEDIS) perteneciente a la Sede Atlántica de UNRN, específicamente dentro de la línea de investigación sobre epistemología, investigación e innovación docente.

1. INTRODUCCIÓN

Desde el proyecto, se busca producir conocimiento sobre los procesos de apropiación de tecnologías (AT) digitales móviles, por parte de los estudiantes de los diferentes niveles educativos y modalidades que tienen lugar en la provincia de Río Negro, y que los mismos adoptan en su quehacer cotidiano. Para ello se han tomados diferentes aportes teóricos, tal es el caso de Sandoval (2020), quien propone observar la AT como un proceso "*histórico en el cual un ingenio técnico (hardware, software o conjuntos complejos de ambos) pasa de la inexistencia a ser incluido en la cotidianidad de personas y grupos, para lo cual deben atravesarse un conjunto de fenómenos socioculturales y condensarse una multitud de sentidos*" (pp.36). En este sentido la AT, posibilita la intervención social de determinadas organizaciones o colectivos, a partir de la

creación y uso de tecnologías (Lago Martínez et al., 2021). Por su parte, Susana Morales (2019) propone no solo observar la relación sujeto-tecnología en términos de prácticas de uso, sino extenderla buscando comprender " *¿quién se apropia de qué, cómo y con qué consecuencias?*" (pp. 38) buscando abordar con ello, el trabajo sobre los derechos digitales y la regulación de internet. En palabras de la autora, "*la supervivencia de nuestras democracias, el ejercicio pleno de nuestra condición de ciudadanos, depende cada vez más de ello.*" (Morales, 2019, p.48)

Así también es posible mencionar al menos 2 líneas de investigación sobre AT :

- el caso de infancias y adolescencias, en virtud del impacto que pueden tener en su desarrollo cognitivo, emocional, o físico entre otros. Algunos aportes previos al contexto de pandemia como los estudios de Cabello y Alonso (2019), y Bedacarratx et al.,(2022) abordan la temática en niños de escolaridad primaria en instituciones públicas, por una parte en el Gran Buenos Aires y por otra en la Patagonia Austral. En el primer caso los autores señalan, que si bien el acceso a las TIC es alto no es igualitario, por otra parte en ambos estudios identifican usos dominantes principalmente vinculados al entretenimiento (juegos / videos) con un efecto que Cabello y Alonso denominan "*sobrecarga recreativa*", poniendo en evidencia la necesidad de mediación adulta y pedagógica.
- otra línea de investigación se enfoca en actividades productivas y laborales, así un estudio realizado por Muñoz, Balmaseda y Melzigo (2025), pone en discusión el rol fundamental de la transformación digital como impulsora del desarrollo económico, social y político para la región de América Latina y el Caribe, y reconoce que *"las habilidades digitales son esenciales para permitir a los individuos interactuar de manera efectiva con tecnologías digitales y destapar todo el potencial de las nuevas tecnologías para la región."* (pp.69). Así también identifican la necesidad de *"invertir en el desarrollo de habilidades —con especial énfasis en aquellas que permiten acceder al empleo avanzado y experto— no es solo una prioridad educativa o social; es una palanca económica esencial para alcanzar un crecimiento más dinámico, inclusivo y sostenible" (pp. 75).*

Los aportes de estos grupos de estudios resultan de interés y relevancia a la luz de los cambios en la matriz productiva que se están generando en la Provincia de Río Negro, asociada principalmente a proyectos del sector energético tales como: el Gas Natural Licuado en la zona del Golfo San Matías, y la instalación de plantas de licuefacción flotantes (FLNG) para la exportación de gas natural licuado (GNL) desde Vaca Muerta. Este tipo de proyectos demandan de mano de obra con habilidades para intervenir en escenarios productivos altamente tecnificados y en permanente transformación, reafirmando la importancia de explorar y analizar las prácticas educativas, en particular las de formación técnica y de educación superior en relación a la apropiación de tecnologías digitales en general, y al desarrollo del pensamiento computacional en particular. Por otra parte, este tipo de proyecto no están exentos de

preocupaciones y conflictos[1][2], entre ellos, el impacto en el ecosistema marítimo y sus consecuencias (Acacio y Wyczkier, 2024, Christiansen y Varsanyi, 2024). Y estos cambios que atraviesan al territorio rionegrino, reafirman la idea la apropiación de tecnologías digitales como herramienta para el ejercicio de la ciudadanía (Morales, 2019).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Desde el proyecto se propuso avanzar sobre:

- El diseño de propuestas pedagógico - didácticas situadas en el contexto de las infancias y adolescencias de la provincia de Río Negro, que permitan desarrollar conocimientos y habilidades para comprender e intervenir las tecnologías digitales en general y las disruptivas en particular
- Revisión de los diseños curriculares vinculados tecnologías y educación, en especial en la formación de formadores
- Implementación de actividades de transferencia situadas en las instituciones públicas y privadas de nivel medio y primario de la provincia de Río Negro

El desarrollo de proyecto se lleva adelante como un estudio de tipo exploratorio-descriptivo, siguiendo el enfoque mixto predominantemente cualitativo sin descartar aportes cuantitativos.

3. RESULTADOS OBTENIDOS/ESPERADOS

A continuación se presentan los resultados alcanzados hasta el momento en la segunda etapa del proyecto:

- Relevamiento de datos (encuestas y focus group) sobre apropiación de TIC por parte de estudiantes de nivel superior de diferentes carreras de la UNRN y del IFDC de Sierra Grande durante 2025 e ingreso 2026. Un análisis preliminar de esta información denominado "Procesos de apropiación y uso de tecnologías móviles en Río Negro mapeando las tecnologías en la provincia", se presentó en las "Jornadas Reflexiones sin distancia" EAD - UNRN a finales de septiembre de 2025.
- Diseño e implementación de un recurso didáctico basado en IA para acompañar el estudio de matemáticas con estudiantes universitarios (Carranza et al., 2025)
- Diseño y dictado de un taller sobre programación visual en el marco de las actividades de educación continúa que ofrece la UNRN. El mismo, estuvo destinado a docentes de educación Primaria de la provincia de Río Negro, y contó con la declaración de interés por el Consejo Provincial de Educación (2025)
- Diseño e implementación de actividades de difusión y divulgación entre ellas la Charla abierta: "Humanos y máquinas: ejercicios de imaginación" realizada en el marco de la Semana Nacional de la Ciencia. UNRN - Sede ATL y la conferencia "De la idea a la práctica: la educación frente a la inteligencia artificial" a cargo de las investigadoras del CONICET Guadalupe Álvarez y Lourdes Morán (2025) .
- En mayo de 2025 se aprobó e inició el Proyecto de Extensión "Alfabetización En Datos, Tecnología & Diabetes. Hagamos La Diferencia" (CT354-40)

[1] https://www.rionegro.com.ar/politica/comunidades-mapuches-en-alerta-por-una-encuesta-relacionada-con-el-oleoducto-vmos-4253582/

[2] https://dataportuaria.ar/nota/23077/rio-negro-cobrara-bono-ambiental-a-los-proyectos-de-gnl/

aprobado y financiado por la UNRN (Res. Rectoral 25/0354), enfocado en la alfabetización en datos sobre diabetes en la población de la comarca Viedma - Patagones en particular en aquellos contextos vinculados al sector educativo (escuelas, institutos, universidades, entre otros). En el marco del mismo se llevan adelante Prácticas de Trabajo Social con estudiantes avanzados de la Licenciatura en Sistemas.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo del proyecto incluye docentes del área de matemáticas, literatura, programación de computadoras y tecnología educativa que llevan adelante la práctica docentes en las diferentes sedes y modalidades que ofrecen las instituciones educativas de nivel superior: UNRN, UNCOMA e Institutos de formación docente de la provincia de Río Negro. La conformación interdisciplinar e interinstitucional del equipo de trabajo, es un aporte a la construcción e interpretación de los fenómenos socioeducativos sobre los que se enfoca la investigación

En cuanto a la formación de recursos humanos, durante esta segunda parte del proyecto, 2 tesistas de posgrado finalizaron y defendieron sus trabajos vinculados por una parte, a la inclusión de TIC como mediadoras en la formación superior y por otra, a la construcción de una cartografía aumentada para abordar relatos sobrenaturales del contexto local (Buratti, 2025). Sumado a ello, una integrante inició su beca Doctoral del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) vinculada con el tema "reconfiguración de prácticas comunicativas digitales en contexto escolar en pandemia/pospandemia: análisis de repertorios de medios y estrategias pragmáticas". Por su parte, un estudiante avanzado de la Lic. en Comunicación Social de la Sede, está desarrollando actividades de formación en investigación a través de una beca CIN, y por último continúan las prácticas de trabajo social de estudiantes de carreras de la UNRN (Lic. en Sistemas y Lic. en Administración).

5. REFERENCIAS

Acacio, J. A. y Wyczykier, G. (2024). Un mar azul para un futuro común. Conflictos y resistencias frente al avance fósil en la costa argentina. Fundación Ambiente y Recursos Naturales.

Buratti, M. E. (2025). Relatos sobrenaturales de la ciudad de Viedma: una experiencia artística de cartografía mediada por el uso de realidad aumentada (Bachelor's thesis, Universidad Nacional del Comahue. Centro Universitario Regional Zona Atlántica y Sur).

Cabello, R., & Alonso, J. M. (2019). Encuesta sobre apropiación de tecnologías por parte de niños y niñas de 6 a 8 años. RevCom, (9), e025-e025.

Carranza, P., Lovos, E., & Marin Aranda, A. (2025). Retomar la clase de matemáticas a través de un chatbot. In XX Congreso de Tecnología en Educación y Educación en Tecnología (TE & ET)(UNDAV, 12 y 13 de junio de 2025).

Christiansen, F. C., & Varsanyi, A. D. R. (2024). Megaproyecto Vaca Muerta: más allá de la noción de impacto (o sobre cómo analizar la degradación socioambiental que genera). Revista de Historia, (25), 192-217

Consejo Provincial de Educación de Río Negro. (2025, 10 de noviembre). *Resolución N.º 5675/25: Declaración de interés educativo "Taller de Programación usando Scratch"*. Provincia de Río Negro.

Universidad Nacional de Río Negro, Sede Atlántica. (2025, 27 de noviembre). *Disposición UNRN-ATL N.º 1455/2025: Aval*

a la conferencia "De la idea a la práctica: la educación frente a la inteligencia artificial". Universidad Nacional de Río Negro.

Lago Martínez, S. B., Gendler, M. A., & Méndez, A. (2021). Movimientos sociales, apropiación de tecnologías digitales y transformaciones en los procesos de acción colectiva.

Morales, Susana (2019). Derechos digitales y regulación de Internet. Miradas críticas de la apropiación en América Latina, 35.

Muñoz, V., Balmaseda, M., & Melguizo, Á. (2025). Habilidades para la economía digital: Impacto económico y social de la formación para el presente y el futuro del trabajo en América Latina y el Caribe. Madrid: Fundación Carolina.

Sandoval, Luis Ricardo (2020). La apropiación de tecnologías como proceso: Una propuesta de modelo analítico en Roberto Canales Reyes y Consuelo Herrera Carvajal (eds.) Acceso, democracia y comunidades virtuales: apropiación de tecnologías digitales desde el Cono Sur. Ciudad Autónoma de Buenos Aires: CLACSO; Red de Investigadores sobre Apropiación de Tecnologías Digitales, pp. 33-49.

Asistentes Virtuales de Enseñanza basados en Generación Aumentada por Recuperación

Juan Pablo Tessore [1 2], Claudia Russo [2 3], Hugo Ramón [2 3], Leonardo Esnaola [2], Tamara Ahmad [2]

Instituto de Investigación y Transferencia en Tecnología (ITT)
Centro asociado a la Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

juanpablo.tessore@itt.unnoba.edu.ar, claudia.russo@itt.unnoba.edu.ar, hugo.ramon@itt.unnoba.edu.ar, leonardo.esnaola@itt.unnoba.edu.ar,, tamara.ahmad@itt.unnoba.edu.ar

Resumen

Los avances recientes en los Grandes Modelos de Lenguaje (LLM) los han consolidado como una tecnología transformadora en inteligencia artificial. Entrenados con grandes volúmenes de texto, aprenden patrones del lenguaje que les permiten generar contenido coherente, responder preguntas, resumir, traducir, programar y asistir en tareas cognitivas complejas. En educación, su potencial es especialmente relevante: pueden actuar como tutores virtuales, adaptar explicaciones al nivel del estudiante, proponer ejercicios, corregir actividades y acompañar procesos de aprendizaje personalizado, favoreciendo la accesibilidad y la retroalimentación inmediata.

No obstante, su aplicación efectiva requiere abordar el problema de las alucinaciones (es decir, la generación de información incorrecta o no verificada con apariencia plausible) así como integrar contenido institucional específico. En este sentido, la Generación Aumentada por Recuperación (RAG, Retrieval-Augmented Generation, por sus siglas en inglés) [1] combina un modelo generativo con un sistema externo de recuperación de información, lo que permite consultar documentos y repositorios verificados antes de producir una respuesta.

El presente trabajo propone medir empíricamente la mejora en la calidad de las respuestas de un LLM base frente a una variante con enfoque RAG en contextos educativos. Asimismo, busca elaborar una metodología para guiar la carga de contenido académico con dicha arquitectura.

***Palabras clave*:** grandes modelos del lenguaje, RAG, asistentes virtuales.

Contexto

Esta línea de investigación forma parte de los proyectos “Técnicas y algoritmos avanzados para la optimización de procesos: seguridad, sostenibilidad e innovación” e “Innovación y Desafíos en la Educación Digital” presentados en el marco de la convocatoria a Subsidios de Investigación Bianuales (SIB 2025) de la Secretaría de Investigación, Desarrollo y Transferencia de la Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA). A su vez, se enmarca en el contexto de un plan de trabajo aprobado por la Comisión Nacional de Investigaciones Científicas y Técnicas, en el marco de la

[1] Becario Postdoctoral CONICET. [2] Docente Investigador ITT. [3] Inv. Asociado Adjunto sin director CIC.

convocatoria “Becas Postdoctorales 2022 Temas Estratégicos”.

El proyecto se desarrolla en el Instituto de Investigación y Transferencia en Tecnología (ITT) dependiente de la mencionada Secretaría y se trabaja en conjunto con la Escuela de Tecnología de la UNNOBA.

El equipo está constituido por docentes e investigadores pertenecientes al ITT y a otros institutos de investigación, así como también estudiantes de las carreras de Informática de la Escuela de Tecnología de la UNNOBA.

Introducción

En los últimos años, los LLM han transformado de manera significativa la forma en que interactuamos con la información. Su desarrollo se basa en la arquitectura Transformer, introducida por Google [2], y posteriormente perfeccionada en modelos de gran escala como los desarrollados por OpenAI [3], Google [4] y Anthropic [5], entre otros. Dichos modelos han demostrado capacidades avanzadas para comprender y generar texto en lenguaje natural, responder preguntas, resumir contenidos, asistir en programación y apoyar procesos creativos. En el ámbito educativo, estas tecnologías permiten ofrecer tutorías personalizadas, generación automática de material didáctico, retroalimentación formativa inmediata y acompañamiento en procesos de aprendizaje autónomo. Su potencial radica en la capacidad de adaptarse al contexto del estudiante y responder en tiempo real a consultas diversas.

Sin embargo, pese a sus avances, los LLM presentan limitaciones relevantes. Una de las principales es el fenómeno de las alucinaciones. Dado que estos modelos se entrenan con grandes volúmenes de datos generales y no tienen acceso directo y actualizado a fuentes específicas, pueden producir respuestas inexactas o inventadas, especialmente cuando se les consulta sobre información institucional, reglamentos internos, programas académicos o contenidos específicos de una organización educativa. Esta limitación resulta crítica cuando se pretende implementar asistentes virtuales en contextos formales, donde la precisión y la confiabilidad de la información son fundamentales.

Ante esta problemática, existen distintas alternativas para implementar asistentes educativos. Una opción tradicional son los sistemas basados en reglas (rule-based systems), que operan mediante flujos predefinidos y árboles de decisión. Si bien ofrecen alto control sobre las respuestas y reducen el riesgo de errores factuales, presentan importantes desventajas: son rígidos, difíciles de escalar, costosos de mantener y poco flexibles ante consultas abiertas o imprevistas. Otra alternativa consiste en realizar fine-tuning de algún LLM con datos institucionales. Aunque esta estrategia permite adaptar el comportamiento del modelo a un dominio específico, implica costos computacionales elevados, requiere grandes volúmenes de datos curados, demanda procesos técnicos complejos y, además, no garantiza la actualización dinámica del conocimiento sin reentrenamientos periódicos.

En este contexto, el enfoque RAG emerge como una solución intermedia y eficiente. RAG combina la capacidad generativa de los LLM con un mecanismo de recuperación de información desde una base de conocimiento externa y controlada. Antes de generar una respuesta, el sistema consulta documentos relevantes (por ejemplo, reglamentos, programas de estudio o materiales institucionales) y utiliza ese contenido como contexto explícito para fundamentar la respuesta. De esta manera, se reduce significativamente la probabilidad de

alucinaciones y se asegura que la información provenga de fuentes verificadas y actualizadas, sin necesidad de reentrenar el modelo.

A continuación, se presentan diversas implementaciones basadas en RAG en el ámbito educativo, junto con los resultados obtenidos en términos de mejora en la calidad, precisión y pertinencia de las respuestas generadas.

En [6] se desarrolló un sistema de tutoría inteligente llamado *LPITutor* que integra un modelo de lenguaje con una arquitectura RAG y técnicas de prompt engineering para generar explicaciones adaptadas al nivel del estudiante. El sistema recupera información relevante desde fuentes externas antes de generar la respuesta, lo que permite ajustar la profundidad conceptual según la consulta. La evaluación reporta mejoras en la precisión y claridad de las respuestas, así como una reducción significativa de alucinaciones en comparación con un LLM sin recuperación, destacándose su capacidad para ofrecer explicaciones pedagógicamente alineadas con los contenidos académicos.

También en [7] se desarrolló un tutor basado en RAG llamado *SmartStudy*, en este caso para cursos universitarios de Machine Learning. El sistema procesa automáticamente los libros de texto del curso y construye una base de conocimiento que alimenta al modelo durante la generación de respuestas y la creación de exámenes de práctica. En una evaluación con estudiantes de ciencia de datos, se observaron mejoras en el desempeño en pruebas y efectos positivos en el aprendizaje percibido, lo que sugiere que la incorporación de recuperación contextual fortalece la comprensión conceptual y la retención de contenidos.

En [8] se implementó en el ámbito de la educación médica en la Geisel School of Medicine de Dartmouth, una arquitectura RAG alimentada con apuntes y bibliografía oficial de un curso de neuroanatomía, *NeuroBot*. El sistema registró una adopción cercana a un tercio de los estudiantes y fue utilizado especialmente en períodos previos a exámenes. Los estudiantes valoraron positivamente que las respuestas estuvieran fundamentadas en el material oficial, con niveles de utilidad percibida moderados pero consistentes. La integración de RAG permitió anclar las respuestas al contenido curricular, aumentando la confianza en el sistema.

Otro tutor [9], *TutorLLM*, fue desarrollado en la Universidad de Southampton y combinó RAG con técnicas de knowledge tracing para personalizar la tutoría según el estado de aprendizaje del estudiante. El sistema recuperaba fragmentos relevantes del curso y ajustaba la generación en función del progreso individual. En evaluaciones comparativas, el enfoque con RAG mostró mejoras aproximadas del 5% en resultados de evaluación y cerca de un 10% en satisfacción estudiantil frente a un LLM sin recuperación, evidenciando beneficios tanto en desempeño como en experiencia de uso.

Finalmente, en [10], también se implementaron chatbots restringidos exclusivamente al material de cada curso en dos universidades. Los resultados mostraron altos niveles de compromiso y motivación en los estudiantes, así como una tasa muy baja de respuestas incorrectas. La arquitectura RAG redujo de manera sustancial las alucinaciones respecto a versiones sin recuperación, demostrando que el acceso controlado a fuentes académicas específicas mejora la precisión y la pertinencia pedagógica de las respuestas en entornos universitarios.

De los trabajos analizados se evidencia que la arquitectura RAG no solo mejora el desempeño de los asistentes virtuales basados en LLM en términos de reducción de

alucinaciones, sino también en cuanto a la satisfacción de los usuarios. Resulta pertinente, en consecuencia, profundizar la investigación con esta arquitectura.

Líneas de Investigación, Desarrollo e Innovación

La presente investigación se encuadra dentro de los siguientes proyectos SIB 2025:

- *"Técnicas y algoritmos avanzados para la optimización de procesos: seguridad, sostenibilidad e innovación"*, cuyo objetivo general es *"Identificar procesos en actividades económicas susceptibles de ser mejorados y optimizados mediante la aplicación de tecnologías innovadoras, que podrían integrar las señales analizadas con un enfoque multimodal para potenciar aún más estos procesos"*.
- *"Innovación y Desafíos en la Educación Digital",* cuyo objetivo general es *"Investigar, diseñar y desarrollar soluciones informáticas innovadoras aplicadas a la educación digital, basadas en inteligencia artificial, tecnologías emergentes y analítica de datos, entre otras, con el propósito de transformar los procesos educativos en todas sus modalidades."*

En este sentido, la presente investigación sobre asistentes virtuales basados en RAG contribuye a dichos proyectos SIB al abordar los siguientes aspectos:

- Investigar arquitecturas, técnicas y herramientas que faciliten la implementación de asistentes virtuales basados en RAG, tales como bases de datos vectoriales, grandes modelos de lenguaje, estrategias de segmentación e indexación de contenido, etc.
- Definir métricas que permitan evaluar el impacto en el desempeño de los asistentes basados en RAG en comparación con una línea base sin RAG, así como medir su utilización e impacto educativo en cursos de nivel universitario.
- Realizar pruebas piloto en cursos universitarios con el objetivo de evaluar el desempeño de los asistentes implementados, utilizando las métricas definidas en el punto anterior.

Resultados y Objetivos

En el marco de este trabajo, se espera diseñar e implementar un asistente virtual de enseñanza basado en una arquitectura RAG, orientado a cursos universitarios. El sistema integraría un LLM de propósito general con una base de conocimiento compuesta por programas de estudio, reglamentos, apuntes de cátedra, trabajos prácticos y bibliografía oficial, procesados mediante estrategias de segmentación e indexación en una base de datos vectorial.

Como objetivo principal, se buscaría evaluar comparativamente el desempeño de un LLM base frente a su variante con RAG en escenarios educativos reales. Para ello, se definirían métricas para evaluar la calidad de las respuestas generadas por el asistente. Asimismo, se analizará la satisfacción estudiantil, la percepción de utilidad y el nivel de confianza en las respuestas generadas.

Entre los resultados esperados, se anticipa que la arquitectura RAG podría reducir significativamente la generación de respuestas incorrectas o no fundamentadas, al anclar la producción textual en documentos institucionales verificados. Se proyecta que el asistente basado en RAG mostraría mejoras en precisión y adecuación contextual respecto a la línea base, especialmente en consultas relacionadas con contenidos específicos del curso o normativa institucional.

Adicionalmente, se prevé que la implementación piloto permitiría obtener evidencia empírica sobre el impacto del asistente en un escenario educativo real.

Finalmente, el trabajo tendrá como objetivo elaborar una guía metodológica para la incorporación de contenido académico en arquitecturas RAG, contemplando criterios de curación, actualización y validación de la base de conocimiento. Dicha metodología podría servir como referencia para futuras implementaciones institucionales, contribuyendo a la escalabilidad de asistentes virtuales en el ámbito de la educación superior.

Formación de Recursos Humanos

En esta línea de I+D se ha obtenido y se encuentra desarrollando actualmente una beca postdoctoral del CONICET. Asimismo, se espera desarrollar una tesina de grado y/o PPS, dirigida por miembros de este proyecto.

Bibliografía

1. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). *Retrieval-augmented generation for knowledge-intensive NLP tasks.* arXiv. https://arxiv.org/abs/2005.11401
2. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need.* In I. Guyon et al. (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates. http://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
3. OpenAI. (2025). *ChatGPT* (versión 12 de mayo de 2025) [Modelo de lenguaje de IA]. https://chat.openai.com/
4. Anthropic. (2025). *Claude* [Modelo de lenguaje de IA]. https://www.anthropic.com/claude
5. Google AI. (2025). *Gemini* [Modelo de lenguaje de IA]. https://gemini.google.com/
6. Liu, Z., Agrawal, P., Singhal, S., Madaan, V., Kumar, M., & Verma, P. K. (2025). *LPITutor: An LLM based personalized intelligent tutoring system using retrieval-augmented generation and prompt engineering. PeerJ Computer Science, 11*, e2991. https://doi.org/10.7717/peerj-cs.2991
7. Vu-Minh, A., Nguyen-The, Q., Nguyen, N., Pham, X.-L., & Nguyen, A. (2025). *Designing a course-grounded AI tutor with retrieval-augmented generation: A DSR approach to technical education. HICSS-59 Proceedings.* https://doi.org/10.2139/ssrn.5607971
8. Thesen, T., & Park, S. H. (2025). *A generative AI teaching assistant for personalized learning in medical education. npj Digital Medicine, 8*, 627. https://doi.org/10.1038/s41746-025-02022-1
9. Li, Z., Yazdanpanah, V., Wang, J., Gu, W., Shi, L., Cristea, A. I., Kiden, S., & Stein, S. (2025). *TutorLLM: Customizing learning recommendations with knowledge tracing and retrieval-augmented generation* (Version 2502.15709v1) [Preprint]. arXiv. https://arxiv.org/abs/2502.15709
10. Németh, R., Tátrai, A., Szabó, M., Zaletnyik, P. T., & Tamási, Á. (2025). *Exploring the use of retrieval-augmented generation models in higher education: A pilot study on artificial intelligence-based tutoring. Social Sciences & Humanities Open, 12*, 101751. https://doi.org/10.1016/j.ssaho.2025.101751

Autoeficacia en línea y Grit como atributos predictivos en Sistemas de Aprendizaje Adaptativo en Educación Superior Virtual: resultados de un estudio piloto y arquitectura conceptual

Monferrato, Marcelo. Facultad de Tecnología Informática. Universidad Abierta Interamericana
marcelo.monferrato@uai.edu.ar

RESUMEN

Esta línea de investigación explora la autoeficacia en línea (OLSE) y la pasión y perseverancia ante las metas a largo plazo (Grit) como variables psico-cognitivas medibles para enriquecer sistemas de aprendizaje adaptativo (SAA) en educación superior virtual. Un estudio piloto con N = 72 estudiantes de la carrera Analista Programador (a distancia), en la Facultad de Tecnología Informática de la UAI identificó correlaciones significativas ($r = 0,361$; $p < 0,01$), con la Perseverancia del Esfuerzo como predictor destacado de la autorregulación del aprendizaje ($r = 0,410$; $p < 0,001$). Se propone una arquitectura conceptual de SAA que integra estos constructos como atributos de clasificación (features) de perfiles estudiantiles (bajo/medio/alto riesgo de abandono) al inicio del cursado, activando trayectorias personalizadas y alertas tempranas.

Palabras clave: aprendizaje adaptativo, autoeficacia en línea, Grit, educación virtual, personalización.

CONTEXTO

Este trabajo se inscribe en la línea de investigación "Sociedad del Conocimiento y Tecnologías Aplicadas a la Educación" del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). Aborda la personalización del aprendizaje en educación superior a distancia y se nutre directamente de las investigaciones previas del autor sobre OLSE y Grit, así como de su revisión sistemática sobre autoeficacia en línea en educación superior virtual para el período 2016-2025 [Monferrato, 2026].

La expansión de la educación virtual en Argentina plantea una demanda creciente de sistemas capaces de adaptarse a las diferencias individuales de los estudiantes. Los SAA actuales se apoyan principalmente en indicadores cognitivos de rendimiento y raramente incorporan variables psico-cognitivas de forma sistemática. Estudios recientes confirman que la integración de SAA con LMS mejora significativamente el desempeño y la satisfacción estudiantil [Andhika et al., 2024; Contrino et al., 2024; Lim et al., 2023], mientras que variables no cognitivas moderan su eficacia [Liu et al., 2017].

1. INTRODUCCIÓN

Los sistemas de aprendizaje adaptativo (SAA) transforman el entorno educativo de un modelo uniforme a uno flexible, ajustando trayectorias de contenido y retroalimentación en función de las características del estudiante [Andhika et al., 2024]. La evidencia sobre su eficacia en educación superior es promisoria aunque heterogénea, dependiente del diseño instruccional y de la riqueza de variables en el modelo del usuario [Lim et al., 2023].

Liu et al. (2017) demostraron que características no cognitivas —entre ellas la ansiedad y la motivación— moderan significativamente el aprovechamiento de los SAA, lo que respalda la hipótesis de que variables como el Grit y la autoeficacia en línea pueden enriquecer el modelo del usuario en estos sistemas.

La escala del test para autoeficacia en línea OLSES [Zimmerman & Kulikowich, 2016] evalúa la confianza del estudiante en cuatro dimensiones: gestión del aprendizaje, comunicación, autorregulación y manejo tecnológico. A su vez, el test simplificado Grit-S mide la perseverancia del esfuerzo (PE) y la consistencia de interés (CI) [Duckworth et al., 2007]. Ambos generan datos estructurados y cuantificables, condición necesaria para su uso como features en modelos de clasificación computacional.

La revisión sistemática del autor [Monferrato, 2026] confirma que, desde 2016, la OLSE ha madurado como constructo predictivo de autorregulación, compromiso y bienestar en entornos virtuales universitarios, trascendiendo enfoques meramente tecnológicos. La presente propuesta integra estos instrumentos como señales de entrada en un SAA integrado al LMS institucional, permitiendo la segmentación de perfiles estudiantiles al inicio del cursado y activar trayectorias de soporte diferenciado o alertas tempranas para el cuerpo docente.

La arquitectura conceptual propuesta contempla tres componentes principales: (1) captura inicial de señales psico-cognitivas (puntajes OLSES y Grit-S administrados al inicio del cursado); (2) modelo de clasificación de perfiles (bajo/medio/alto riesgo de abandono, mediante umbrales por terciles o futuros algoritmos de clustering); y (3) motor de adaptación y alertas (generación automática de recomendaciones de trayectoria, refuerzos adicionales, recordatorios intensificados, sesiones sincrónicas focalizadas y notificaciones tempranas al docente basadas en analytics del LMS). Esta estructura se alinea con diseños instruccionales adaptativos validados en entornos universitarios (Contrino et al., 2024), pero incorpora de forma novedosa variables no cognitivas estandarizadas como predictores proactivos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los ejes principales de investigación y desarrollo son:

- Validación y caracterización de los constructos OLSES y Grit-S en población estudiantil de educación superior virtual argentina.
- Exploración de la capacidad predictiva de estos puntajes sobre indicadores de desempeño y retención mediante correlaciones, regresiones y análisis de clustering.
- Diseño de una arquitectura conceptual de SAA que integre ambos constructos como features de entrada en modelos de clasificación de perfiles estudiantiles (bajo/medio/alto riesgo de abandono).
- Desarrollo e implementación de un módulo prototipo con generación de alertas tempranas automáticas para docentes y evaluación de su impacto sobre retención y desempeño.

3. RESULTADOS OBTENIDOS / ESPERADOS

3.1 Resultados obtenidos

Se administraron los instrumentos a N = 72 estudiantes (73,6 % masculino, 26,4 % femenino; 68 % empleados a tiempo completo o parcial; 94 % con acceso residencial a internet). La Tabla 1 presenta los estadísticos descriptivos principales.

Tabla 1. Estadísticos descriptivos (N = 72)

Variable	M	DE	Mín	Máx
Grit-S Total	3,70	0,61	2,00	4,75
Perseverancia (PE)	3,97	0,61	—	—
Consistencia (CI)	3,44	0,72	—	—

OLSES Total	4,15	0,40	3,18	4,86
Autorregulación	3,86	0,66	—	—
Gestión	4,36	0,44	—	—

Nota: Escala 1–5. Ítems inversos de Grit-S corregidos previamente.

La Tabla 2 presenta las correlaciones de Pearson entre ambos constructos.

Tabla 2. Correlaciones de Pearson — Grit-S y OLSES (N = 72)

Par de variables	r	p
Grit Total ↔ OLSES Total	0,361	0,0018 **
Grit PE ↔ OLSES Total	0,370	0,0014 **
Grit PE ↔ OLSES Autorregulación	0,410	< 0,001 ***
Grit CI ↔ OLSES Total	0,296	0,0117 *

*Nota: * p < 0,05; ** p < 0,01; *** p < 0,001. PE = Perseverancia del Esfuerzo; CI = Consistencia de Interés.*

El análisis por terciles de Grit reveló un gradiente consistente: estudiantes con Grit alto presentan OLSES promedio de 4,31 frente a 3,99 en Grit bajo (diferencia de 0,32 puntos). La subescala Autorregulación OLSES es la más sensible al nivel de Grit (r = 0,410; p < 0,001), lo que refuerza su relevancia como variable objetivo para el sistema adaptativo. Estos hallazgos preliminares alinean con evidencia de que variables no cognitivas moderan el impacto de los SAA [Liu et al., 2017] y de que su integración con LMS eleva significativamente el desempeño [Contrino et al., 2024; Andhika et al., 2024].

3.2 Resultados esperados

- Ampliación de la muestra y análisis de clustering (k-means o similar) para segmentación robusta de perfiles de riesgo.
- Desarrollo de un diagrama de arquitectura conceptual (flujo de datos desde captura OLSES/Grit hasta alertas y trayectorias).
- Prototipo de módulo en LMS institucional que integre OLSES/Grit-S como señales de entrada con alertas tempranas y trayectorias adaptativas.
- Evaluación del impacto sobre retención y desempeño en al menos dos asignaturas virtuales, comparando grupos experimental y control (siguiendo diseños cuasi-experimentales validados en Lim et al., 2023).
- Difusión de resultados en eventos y revistas de Informática Aplicada a la Educación, incluyendo posible extensión a otras carreras de la UAI.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está integrado por un investigador-docente de la UAI con formación en Ciencias Matemáticas, Ciencias de la Computación e Inteligencia Artificial Generativa Aplicada a la Educación (UNIR, 2024–2025).

Producción en curso: revisión sistemática sobre autoeficacia en educación superior virtual (UAI, 2026). Antecedente directo: presentación con distinción sobre LLM locales aplicados a Cálculo Infinitesimal.

Se prevé la incorporación de estudiantes de grado y posgrado en las etapas de prototipado, implementación y evaluación del SAA, generando al menos una tesina de grado y una tesis de maestría asociadas a la línea durante 2026–2027.

5. BIBLIOGRAFÍA

[1] Andhika et al. (2024). The Effectiveness of Adaptive Learning Systems Integrated with LMS in Higher Education. Jurnal KomtekInfo, 11(2), 49–56. https://doi.org/10.35134/komtekinfo.v11i2.505

[2] Contrino, M. F. et al. (2024). Using an adaptive learning tool to improve student performance and satisfaction in online and face-to-face education. Smart Learning Environments, 11, 6. https://doi.org/10.1186/s40561-024-00292-y

[3] Duckworth, A. L. et al. (2007). Grit: perseverance and passion for long-term goals. Journal of Personality and Social Psychology, 92(6), 1087–1101.

[4] Lim, L. et al. (2023). Efficacy of an Adaptive Learning System on Course Scores. Systems, 11(1), 31. https://doi.org/10.3390/systems11010031

[5] Liu, M. et al. (2017). Investigating the effect of an adaptive learning intervention on students' learning. Educational Technology Research and Development, 65(6), 1605–1625. https://doi.org/10.1007/s11423-017-9542-1

[6] Monferrato, M. (2026). Autoeficacia en línea en educación superior virtual: una revisión sistemática de la literatura (2016–2025). Trabajo final, Curso Cultura de la Investigación, UNIR (España).

[7] Zimmerman, W. A. & Kulikowich, J. M. (2016). Online learning self-efficacy in students with and without online learning experience. American Journal of Distance Education, 30(3), 180–191. https://doi.org/10.1080/08923647.2016.1193801

Data Warehouse Soft Skills: Modelo Diagnóstico Integral de Competencias mediante Business Intelligence e Inteligencia Artificial

Agustín Bressan, AgustinIgnacio.Bressan@alumnos.uai.edu.ar, (1)
Gastón Ruiz, gaston.ruiz1@alumnos.uai.edu.ar, (2)
Nahuel Cardú, nahuelsebastian.cardu@uai.edu.ar, (3)
Silvia Victoria Poncio, silvia.poncio@uai.edu.ar, (4)
Cintia Cuña Giménez, cintia.cuna@uai.edu.ar, (5)
Briseida Hernández Sánchez, brizeida@usal.es, (6)
(1) (2) (3) (4) (5)Universidad Abierta Interamericana. Facultad de Tecnología informática. Centro de Altos Estudios en Tecnología Informática. Rosario, Argentina
(6)Facultad de Psicología, Universidad de Salamanca. Miembro del SIN- SENACY. Salamanca, España

RESUMEN

Este trabajo presenta los avances de una línea de investigación sobre el diagnóstico y fortalecimiento de soft skills en educación superior. La integración del Modelo Diagnóstico de Competencias Universitarias (MDCU), Business Intelligence (BI) e Inteligencia Artificial (IA), se propone como un ecosistema digital para transformar constructos cualitativos en dimensiones observables y gestionables. El estudio longitudinal del año 2022 al 2024, basado en estudiantes de Ingeniería en Sistemas, analiza la relación entre la actitud empática y la cooperación entre pares. Utilizando procesos ETL, un Data Warehouse y dashboards interactivos, se identificó una brecha estructural entre la autopercepción empática y la ejecución conductual. La aplicación de fine-tuning en modelos de lenguaje permitió una clasificación semántica precisa de los indicadores. Los resultados indican que, pese a una alta disposición emocional, la transferencia a la cooperación requiere implementaciones estratégicas y pedagógicas sistemáticas para atender a las demandas de empleabilidad global.

PALABRAS CLAVES: Soft skills, inteligencia artificial, business intelligence, cooperación, datawarehouse.

CONTEXTO

Esta línea de investigación se desarrolla en la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI), específicamente en su sede de Rosario, Argentina. El proyecto se encuentra radicado en el Centro de Altos Estudios en Tecnología Informática (CAETI) y se inscribe en la línea de investigación "Sociedad del Conocimiento y Tecnologías Aplicadas a la Educación".

El marco normativo que sustenta el estudio se basa en las competencias definidas por el Consejo Federal de Decanos de Ingeniería CONFEDI [1] a través del denominado "Libro Rojo", documento que establece los estándares de segunda generación para la acreditación de carreras de ingeniería en la República Argentina. Estos estándares obligan a las instituciones a garantizar la formación en competencias sociales, políticas y actitudinales como parte del perfil de egreso profesional. Asimismo, el proyecto sigue las recomendaciones de organismos internacionales como la Organización para la Cooperación y el Desarrollo Económicos, OCDE [2] respecto a la importancia de las habilidades socioemocionales para la resiliencia

laboral.

1. INTRODUCCIÓN

En el contexto actual, donde la automatización y la digitalización transforman el mercado laboral, el desarrollo de soft skills se convierte en un factor determinante para el éxito profesional. Si bien las hard skills siguen siendo esenciales, especialmente en áreas técnicas como la ingeniería, las competencias interpersonales y de adaptación han cobrado un protagonismo cada vez mayor. Según las investigaciones del Estudio de la Agencia Australiana de Empleo y el Instituto Monarch [3], se indica que un 85% de las competencias más valoradas por las empresas corresponden a habilidades socioemocionales, mientras que solo el 15% se relaciona con conocimientos técnicos, lo cual refuerza la necesidad de integrarlas en la formación académica.

El desarrollo de competencias como la comunicación efectiva, la resolución de conflictos, la adaptabilidad y, fundamentalmente, la empatía y la cooperación, se ha posicionado como un factor crítico. En áreas como la ingeniería en sistemas, donde el trabajo en equipos interdisciplinarios es pilar del diseño de software, estas habilidades dejan de ser periféricas para volverse esenciales. Sin embargo, diagnosticar estas capacidades es un desafío, ya que los enfoques tradicionales de enseñanza han puesto énfasis histórico en la adquisición de conocimientos técnicos, relegando a un segundo plano el desarrollo de capacidades sociales.

En respuesta a esta problemática, el presente estudio propone la implementación del Modelo Diagnóstico de Competencias Universitarias MDCU [4]asistido por tecnologías de frontera. La integración de Business Intelligence permite la visualización dinámica de la evolución competencial, mientras que la Inteligencia Artificial aporta capacidades analíticas superiores para la clasificación de datos cualitativos. Este enfoque se alinea con la visión de Cook [5], quien afirma que la empatía desempeña un papel fundamental al guiarnos en la creación de tecnologías diseñadas para mejorar la calidad de vida de las personas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La investigación se estructura sobre ejes fundamentales que articulan la teoría pedagógica con la arquitectura de datos y el procesamiento inteligente.

El MDCU, propuesto originalmente por Poncio y Trottini [5], es un marco evaluativo integral diseñado para la observación sistemática de competencias genéricas. En la presente línea de investigación, el modelo ha sido ampliado para integrar dimensiones de la inteligencia emocional, permitiendo que el diagnóstico se convierta en un insumo estratégico para la mejora del currículo académico.

Para gestionar los datos generados por el diagnóstico longitudinal del año 2022 al 2024, se ha diseñado un ecosistema de BI cuyo núcleo es un Data Warehouse. Como mencionan López-Meneses et al. [6] En el que destacan que la minería de datos, combinada con el modelado predictivo, es fundamental para analizar el comportamiento del estudiante en entornos digitales. La visualización se realiza mediante dashboards que, según Bodily et al [7], están diseñados para hacer visibles patrones de aprendizaje y áreas de mejora. Además, se utiliza la funcionalidad de drill down que, como expresan Shabaninejad et al. [8] Permite profundizar desde una visión general hasta niveles detallados de información para optimizar la toma de decisiones pedagógicas.

Dada la naturaleza cualitativa de la información, se incorporan técnicas de IA para el procesamiento del lenguaje natural. Ouyang et al. [9] Han demostrado cómo el entrenamiento de modelos de lenguaje con retroalimentación humana potencia la capacidad de seguir instrucciones complejas. En este proyecto, se implementa la técnica de fine-tuning para ajustar modelos de lenguaje a datos específicos del dominio educativo, optimizando la clasificación semántica de los

ítems de diagnóstico y permitiendo identificar patrones de aprendizaje ocultos en los registros educativos.

3. RESULTADOS OBTENIDOS

El análisis longitudinal sobre la muestra de 200 estudiantes ha permitido extraer datos reveladores mediante el uso de los dashboards desarrollados por Poncio et al. [10] Los resultados mostraron niveles de empatía generalmente positivos, con un promedio global del 80%. El análisis por cohorte indicó que en 2022 el valor fue del 69%, ascendiendo al 88% en 2023 y situándose en un 81% en 2024. Estos datos permiten validar en cada grupo su enfoque específico, demostrando que la sensibilidad hacia las perspectivas ajenas fluctúa según el contexto de la cohorte.

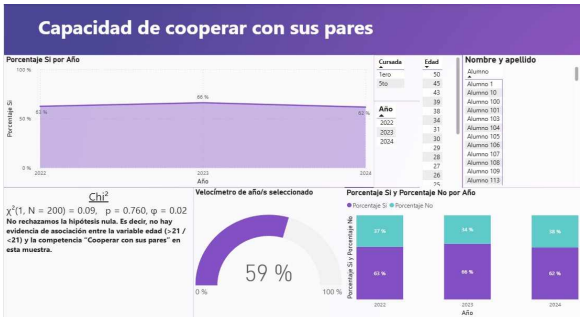


Fig.1: Análisis de la Dimensión Cooperación

La capacidad de cooperar presenta niveles más moderados. Según el estudio de Poncio et al. [11] El promedio global de logro alcanza el 59%. Los valores anuales se mantuvieron estables: 63% en 2022, 66% en 2023 y 62% en 2024, lo que refleja una base colaborativa presente pero aún no consolidada.

Brecha Estructural y Análisis Estadístico

Se ha identificado una brecha estructural de aproximadamente el 16% entre la autopercepción empática (80%) y la ejecución conductual de la cooperación (59%). Este diferencial sugiere que la sensibilidad emocional no se traduce automáticamente en eficiencia colaborativa. Para profundizar, se realizó un análisis inferencial utilizando la prueba de chi cuadrado, obteniendo los siguientes valores: $\chi^2(1, N = 200) = 0.09$; $p = 0.760$ y $\phi = 0.02$. Al ser el p-valor mayor a 0.05, se confirma que no existe asociación entre la variable edad y la competencia de cooperación, lo que refuerza que estas habilidades deben ser cultivadas sistemáticamente con independencia de la madurez cronológica.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto constituye un espacio de formación interdisciplinaria bajo la dirección de Silvia Victoria Poncio, cuya Tesis Doctoral se centra en el diagnóstico de competencias blandas mediante herramientas de BI. El equipo se completa con la investigadora colaboradora Cintia Cuña Giménez, Briseida Hernández Sánchez y los investigadores en formación Nahuel Sebastián Cardú, Gastón Ruiz y Agustín Ignacio Bressan. Esta estructura promueve la formación académica y científica de estudiantes en análisis de datos educativos y aplicación de tecnologías emergentes, fomentando el trabajo cooperativo y la difusión de hallazgos en la comunidad científica.

5. BIBLIOGRAFÍA

[1] Consejo Federal de Decanos de Ingeniería (CONFEDI), *Libro Rojo: Estándares de segunda generación para la acreditación de carreras de ingeniería en la República Argentina*. Universidad FASTA Ediciones, 2018.

[2] Organización para la Cooperación y el Desarrollo Económicos (OCDE), “Colaborar con los empleadores en la contratación de personas refugiadas: Un plan de acción de 10 puntos”, 2018 Disponible: https://www.oecd.org

[3] D. B. de Campos, L. M. M. de Resende y A. B. Fagundes, "The Importance of Soft Skills for the Engineering," *Creative Education*, vol. 11, pp. 1504–1520, 2020. [En línea]. Disponible:

https://doi.org/10.4236/ce.2020.118109
[4] S. Poncio y A. M. Trottini, “Modelo Diagnóstico de Competencias Universitarias (MDCU): Aprender a Emprender” en *VI Congreso Internacional de Emprendimiento AFIDE*. Boca del Río, Veracruz, México, 2018.
[5] T. Cook, *Commencement address at Duke University*. Duke University, May 13, 2018 Disponible: https://commencement.duke.edu
[6] E. López-Meneses, P. C. Mellado-Moreno, C. Gallardo Herrerías, y N. Pelícano-Piris, “Educational Data Mining and Predictive Modeling in the Age of Artificial Intelligence: An In-Depth Analysis of Research Dynamics,” *Computers*, vol. 14, no. 2 , p. 68, 2025 Disponible: https://doi.org/10.3390/computers14020068
[7] R. Bodily , T. K. Ikahihifo, B. Mackley y C. R. Graham, “The design, development, and implementation of student-facing learning analytics dashboards,” *Journal of Computing in Higher Education*, 30(3), 572-598, 2018 Disponible: https://doi.org/10.1007/s12528-018-9186-0
[8] S. Shabaninejad, H. Khosravi, S. J. J. Leemans, S. Sadiq y M. Indulska, “Recommending Insightful Drill-Downs Based on Learning Processes for Learning Analytics Dashboards,” en *Artificial Intelligence in Education*. Springer, 2020 pp. 486-499 Disponible: https://doi.org/10.1007/978-3-030-52237-7_39
[9] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin y J. Leike, "Training language models to follow instructions with human feedback," en Advances in Neural Information Processing Systems (NeurIPS), vol. 35, 2022, pp. 27730–27744.
[10] S. V. Poncio, C. N. Cuña Giménez, N. Cardú y G. Ruiz, "Data Warehouse Soft Skills: Modelo Diagnóstico para la observación de la actitud empática", Centro de Altos Estudios en Tecnología Informática (CAETI), Universidad Abierta Interamericana, 2024.
[11] S. V. Poncio, C. Cuña Giménez, N. Cardú, G. Ruiz y A. Bressan, "Data Warehouse Soft Skills: Diagnóstico de la capacidad de cooperar con otros", Centro de Altos Estudios en Tecnología Informática (CAETI), Universidad Abierta Interamericana, 2024.

Desarrollo de competencias en Inteligencia Artificial aplicada a la Educación: Resultados y prospectiva de un Programa de Formación

María Viviana Godoy Guglielmone (mvgodoy@exa.unne.edu.ar), Beatriz Castro Chans (beatrizcc@exa.unne.edu.ar), Lucas Oviedo (lucas.oviedo@exa.unne.edu.ar), Sonia I. Mariño (msonia@exa.unne.edu.ar), Guillermo A. Arduino (gaarduino@exa.unne.edu.ar)

Grupo de Investigación Sistema de información y TI: modelos, métodos y herramientas
PI Observatorio de Tecnologías Aplicadas a la Educación FaCENA – UNNE
PI Sistemas informáticos y gestión del conocimiento. Modelos, métodos y herramientas
Departamento de Informática, Facultad de Ciencias Exactas y Naturales y Agrimensura,
Universidad Nacional del Nordeste

Eje: Tecnología Informática Aplicada en Educación

Resumen
En el contexto de la educación superior contemporánea, la innovación educativa ha dejado de ser una opción pedagógica para transformarse en un imperativo institucional. No obstante, para que esta transición no sea meramente instrumental o tecnológica, requiere de un sustento epistémico sólido. Es aquí donde los grupos de investigación especializados en tecnologías aplicadas a la educación emergen como motores estratégicos de cambio.
En tal sentido, los estudios y producciones realizados por el Grupo de Investigación Sistema de información y TI: modelos, métodos y herramientas de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional del Nordeste (FaCENA – UNNE) contribuyen al desarrollo de un Programa de Formación en IA aplicada a la educación, que abarca actividades formativas y de divulgación destinadas a integrantes de la propia comunidad educativa, docentes e investigadores, profesionales interesados en la temática, etc.
Entre sus líneas, el Grupo de Investigación aborda el análisis y desarrollo de tecnologías aplicadas a la educación como herramientas de innovación educativa y, en los últimos tiempos, en particular, se enfocó en analizar, los modelos de lenguaje de inteligencia artificial como asistentes en las prácticas de enseñanza.
Este trabajo aborda los fundamentos y las líneas de acción planteados en el Programa de Formación en IA aplicada a la educación de la FaCENA - UNNE, sus resultados y prospectiva.

Palabras Clave: educación superior, tecnologías educativas, inteligencia artificial, investigación

Contexto:
El Grupo de Investigación *Sistema de información y TI: modelos, métodos y herramientas* acreditado por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste alberga actualmente, 2 proyectos de investigación: El Observatorio de Tecnologías Aplicadas a la Educación de la FaCENA-UNNE (21F010 – SGCyT – UNNE) y Sistemas informáticos y gestión del conocimiento. Modelos, métodos y herramientas (PI 22F025 SGCyT- UNNE).

1. Introducción:

Las tecnologías como mediadoras en el proceso de enseñanza y de aprendizaje son un tema de interés para las instituciones educativas, en particular, para las universidades. (Lion, Kap y Ferrareli, 2023). Aunque la Inteligencia Artificial (IA) lleva más de medio siglo de desarrollo, en los años más recientes, las IA Generativas expusieron estos modelos de lenguaje exponencial y lograron un alcance masivo irrumpiendo en

los sistemas educativos. Se plantea así la urgencia de reflexionar sobre *"el rol que las instituciones educativas en general y en especial las de Educación Superior deben afrontar para preparar a las personas en un contexto de continuos cambios, y distintos campos de actuación: estudios, trabajo, esparcimiento, relaciones personales."* (García-Peñalvo, Faraón Llorens-Largo y Javier Vidal, 2024, p.11)

Además, se requiere abordar las paradojas y tensiones que presenta el uso de las IA, tales como sesgos, su influencia en la toma de decisiones, el uso de la información compartida por usuarios, burbujas de información, sustitución de la mano de obra humana, entre otros riesgos y desafíos de su desarrollo y uso.

Estos son algunos de los temas que aborda el Grupo de Investigación *Sistema de información y TI: modelos, métodos y herramientas*. La relación entre grupos de investigación y la gestión de programas de formación docente no debe entenderse como una secuencia lineal, sino como un proceso sinérgico y circular. En este ecosistema, el producto de la investigación —hallazgos sobre aprendizaje mediado por tecnologías, inteligencia artificial generativa o analíticas de aprendizaje— no permanece estanco en el repositorio académico, sino que retroalimenta dinámicamente la toma de decisiones en la gestión académica de las instituciones.

La gestión de la formación docente, cuando es nutrida por la investigación, asegura que la innovación sea sostenible, pertinente y, sobre todo, centrada en la mejora continua de los procesos de enseñanza-aprendizaje.

2. Líneas de Investigación y Desarrollo

De manera articulada entre el Grupo de Investigación (GI) y el equipo de gestión de la FaCENA UNNE, en 2023, se realizó el Ateneo *"Oportunidades y desafíos del uso de inteligencia artificial en carreras de la FACENA-UNNE*" (Resolución N° 1564/23 D) para conocer el estado de situación y expansión sobre el uso de la IA y los interrogantes respecto de su empleo en las asignaturas de grado y postgrado.

En 2024 se implementó la Diplomatura Universitaria en Inteligencia Artificial aplicada a la Educación (Nivel básico). Esta propuesta formativa consta de 5 módulos que abordan fundamentos teóricos y la IA como asistente para resolver actividades propias del ámbito académico.

En 2025, juntamente con el Área de Educación Virtual se diseñó el Programa de Formación en Inteligencia Artificial aplicada a la educación, con miras a sistematizar las acciones a desarrollar en ese campo.

Los objetivos del Programa son:

- Facilitar el aprendizaje y la comprensión de los principios y aplicaciones de la inteligencia artificial en el ámbito educativo. (Zawacki-Richter et al., 2019).
- Promover el intercambio de conocimientos y experiencias entre los participantes, acerca del uso de la IA en diferentes niveles educativos.
- Fomentar la apropiación y la incorporación efectiva de herramientas y metodologías basadas en IA para mejorar los procesos de enseñanza y aprendizaje.

El Consenso de Beijing sobre la inteligencia artificial (2019) reconoce la complejidad del abordaje del uso de IA y educación y las agrupa en distintas*: "planificación de la IA en las políticas educativas; la IA para la gestión y la impartición de la educación; la IA para apoyar la docencia y a los docentes; la IA para el aprendizaje y la evaluación del aprendizaje; el desarrollo de valores y competencias para la vida y el trabajo en la era de la IA; la IA para ofrecer oportunidades de aprendizaje permanente para todos."* (García Peñalvo, Llorens-Largo y Vidal, 2024, p. 37)

En este contexto, un programa de formación se vuelve fundamental por las siguientes razones:

- Comprender la naturaleza de la IA en la educación a fin de colaborar con los docentes a comprender sus fundamentos y alcances en el ámbito educativo.
- Analizar la IA como amenaza y oportunidad: en tanto, ofrece oportunidades para la personalización del aprendizaje, la automatización de tareas rutinarias y la mejora de los resultados de

aprendizaje; así como, también, plantea amenazas relacionadas con la calidad del aprendizaje, la privacidad de los datos, la dependencia tecnológica, el desarrollo del pensamiento crítico y la originalidad de los trabajos. El programa de formación se plantea como estrategia institucional para capacitar a los educadores en la identificación de los riesgos, y el aprovechamiento de las potencialidades de la IA de manera ética y efectiva

- Redefinir el rol docente: La capacidad de la IA para resumir, explicar conocimiento e incluso desarrollar argumentos implica una evolución del rol del profesor. En este sentido, el programa de formación busca preparar a los docentes para centrarse en aspectos de la enseñanza que la IA no puede replicar, como las relaciones humanas, la motivación personalizada, el acompañamiento emocional, el desarrollo de habilidades integrales y la promoción del pensamiento crítico.
- Explorar herramientas y estrategias pedagógicas con IA: para el aprendizaje personalizado, tutoría inteligente, calificación automatizada, procesamiento del lenguaje natural, creación de contenido inteligente, asistentes virtuales, etc.
- Fomentar la investigación y la reflexión crítica: incentivar a los docentes e investigadores a explorar el impacto real de la IA en el aula y a desarrollar prácticas basadas en evidencia.
- Considerar las perspectivas de los estudiantes: sensibilizar a los educadores sobre las mismas y fomentar un uso de la IA que responda a las necesidades y experiencias de los estudiantes.

Como señalan García Peñalvo, Llorens-Largo y Vidal, (2024, p. 11) *"cuando se habla de IA en educación se suele pensar, fundamentalmente, en la personalización del aprendizaje (Zhang et al., 2020). Así, se habla de tutores inteligentes (Yilmaz et al., 2022), asistentes virtuales (Gubareva y Lopes, 2020), experiencias de aprendizaje inmersivas e interactivas (Chng et al., 2023), utilización de los datos para mejorar el desempeño de los estudiantes (Vázquez-Ingelmo et al., 2021), entre otros tópicos."*.

En definitiva, la fundamentación de este programa de formación radica en la necesidad de preparar a los profesionales de la educación para navegar el complejo y vertiginoso panorama de la IA, transformando los desafíos en oportunidades para enriquecer los procesos de enseñanza y aprendizaje y para seguir siendo figuras esenciales en la construcción del conocimiento en la era digital.

3. Resultados:

El Programa incluye actividades de: **a) capacitación :** cursos de actualización, trayectos formativos (diplomaturas, ciclos) y actividades de posgrado (seminarios, talleres, cursos); b) a**ctividades de divulgación:** eventos académicos: congresos, seminarios, charlas y jornadas como espacios de intercambio y gestión del conocimiento y producción de publicaciones: sistematización de experiencias, artículos y ponencias; y c) **transferencia y vinculación:** asesoramiento a equipos docentes e instituciones educativas con las cuales FACENA-UNNE establece acuerdos de trabajo/cooperación.

En tal sentido, como resultado del estudio sistemático de literatura realizado entre 2018 y 2025, Escudero Cristobal, Buitron Bruno y Rodríguez Rojas (2025, p.299 concluyen que: *"la integración efectiva de la IA en la educación superior requiere un enfoque pedagógico crítico, que se sustente en principios éticos, políticas educativas sostenibles y estrategias formativas que fortalezcan la ciudadanía digital. Esta integración no debe limitarse a la adopción instrumental de tecnologías, sino orientarse hacia la construcción de ecosistemas universitarios que promuevan aprendizajes significativos, justicia educativa y transformación social."*

Bajo dichos presupuestos, se estableció este programa, entre los años 2024 y 2025 se implementaron cuatro ediciones de la Diplomatura Universitaria en Inteligencia Artificial aplicada a la Educación (Nivel Básico) (RES - 2024 - 341 - CD-EXA # UNNE) totalizando 415 inscriptos.

Además, se dictaron cursos virtuales

autoasistidos y de corta duración de los que participaron 352 docentes de distintos niveles educativos. Los temas de los cursos fueron: 1) De la instrucción al resultado: redacción de prompts eficaces para proyectos educativos (RES-2025-224-CD); 2) IA Generativa para Docentes: diseño de cuestionarios de evaluación efectivos en Moodle (RES-2025-246-CD-EXA#UNNE); 3) Construcción de chatbots para potenciar la educación con Inteligencia Artificial Generativa (RES - 2025 - 264 - CD-EXA) y 4) Evaluación efectiva con rúbricas: diseño asistido por IA e implementación en Moodle (RES- 2025 - 333 - CD-EXA # UNNE).

También se realizaron actividades de capacitación en articulación con el Sistema de Educación a Distancia de la UNNE: un MOOC (Massive Open Online Course) a través de la plataforma UNNE Global "Fundamentos de IA en la educación" del cual participaron 2142 personas Y una micro capacitación virtual "Diseño de juegos de escape educativos con IA para transformar el aprendizaje "en el marco de los Nanoencuentros con 218 inscriptos.

Asimismo, se organizaron actividades de divulgación con la participación de expertos externos a la Unidad Académica: charla virtual: "¿Cómo funciona el cerebro durante el aprendizaje? Claves para mejorar la enseñanza universitaria" desarrollada por la Dra. Andrea Goldin, neurocientífica investigadora de UNNE) y una charla híbrida: "La Inteligencia Artificial y su utilización en la educación: Experiencia internacional" (RES - 2025 - 1690 - D-EXA # UNNE), dictada por el Dr. Guillermo Ramírez Cáceres (Univ.de Soka, Japón).

Atendiendo a diferentes requerimientos y la diversidad de campos en los cuáles se utiliza la IA, integrantes del GI desarrollaron cursos de posgrado, tales como: "Nociones de Sostenibilidad en procesos de formación emprendedora mediadas por IA", así como, "Redes Neuronales Artificiales", "Inteligencia Artificial" y "Tecnologías para la enseñanza". en el marco de la Maestría en Informática y Computación (UNNE-UNaM)

En otro orden, se realizaron actividades destinadas a alumnos de nivel medio en el marco de "FaCIENCIA" Jornada de Ciencia y Tecnología de la FaCENA - UNNE: ¿Quién entrena a un chatbot? (2024) y "El secreto de la IA para transformar tus imágenes (2025).

En estas líneas de acción antes mencionadas, es clave la articulación con las actividades docentes. Es así como en el marco de la asignatura Inteligencia Artificial de la Licenciatura en Sistemas de Información (FaCENA – UNNE), se fomenta en los estudiantes la identificación y el análisis de requerimientos para el diseño y desarrollo de asistentes inteligentes orientados a resolver problemáticas del contexto. (Arias-Chávez, Ramos-Quispe & Cangalaya Sevillano, 2024).

Asimismo, el trabajo del Grupo en concordancia con autoridades y equipo de gestión de la Unidad Académica dio lugar a publicaciones y presentaciones a eventos científicos - académicos.

En 2026 está planificado dar continuidad a las actividades formativas y profundizar en nuevas líneas (UNESCO, 2025), tanto en lo que respecta a formación y divulgación, como a través del desarrollo de aplicaciones en diferentes ámbitos, especialmente en instancias de formación de recursos humanos, tales como pasantías y proyectos de tesis de grado y posgrado.

4. Conclusiones:

El Programa de Formación en IA aplicada a la educación en la FaCENA-UNNE, surge como una respuesta institucional ante la irrupción de las IA Generativas en la educación superior. Este programa trasciende lo meramente instrumental para proponer un sustento epistémico sólido, abordando desde la alfabetización digital y la redacción de *prompts* hasta el desarrollo de *chatbots* y la redefinición del rol docente. La iniciativa se estructura a través de líneas de capacitación formal, divulgación y transferencia, buscando que los educadores gestionen de manera ética y efectiva las potencialidades y riesgos de estas tecnologías.

La importancia de la articulación entre los grupos de investigación y la gestión universitaria radica en la creación de un ecosistema sinérgico y circular, donde la

investigación basada en evidencia nutre dinámicamente la toma de decisiones institucionales. Esta unión asegura que la innovación educativa no sea una respuesta reactiva, sino un proceso sostenible y pertinente que retroalimenta la formación docente con hallazgos actualizados sobre el aprendizaje mediado por tecnologías. De este modo, la gestión deja de ser sólo administrativa para convertirse en un motor estratégico de cambio sustentado en la producción científica del propio grupo.

Hacia el futuro, el desafío principal reside en transitar de la adopción de herramientas hacia una integración pedagógica crítica y sistémica que promueva el acceso a la educación y la transformación social. La prospectiva para 2026 plantea la profundización en la formación de recursos humanos de alto nivel y el desarrollo de aplicaciones específicas que den respuesta a problemas surgidos en el contexto. El éxito dependerá de la capacidad institucional para fortalecer este tipo de articulaciones, sistematizar las experiencias, evaluar el impacto de las competencias adquiridas y mantener una actualización continua en un entramado educativo y social de cambio tecnológico acelerado.

5. Formación de Recursos Humanos

El Grupo de investigación está compuesto por: una Mgter. en Informática y Doctora en Cs. Cognitivas, una Magister en Informática, una Magister en Ciencias Sociales y Humanas, un Mgter. en Ing. de Software, una Dra. en Cs. Económicas, una Mgter. en Tecnologías de la Información, un Doctorando en Informática, una Especialista en Educación Superior y una Lic. en Sistemas cursando la Maestría en Informática, una Profesora de Matemática y Especialista en Educación con Tecnologías, una Lic. en Relaciones Laborales realizando su trabajo final de Especialización en Educación con Tecnologías, una Magíster en Educación y una Licenciada en Sistemas. Además, se han dirigido tesis de grado de la Lic. Sistemas de Información, así como de postgrado, de la Maestría en Informática ya defendidas y otras en proceso de ejecución.

Bibliografía:

Arias-Chávez , D., Ramos-Quispe, T., & Cangalaya Sevillano, L. M. (2024). Analysis and trends in the use of chatbots and conversational agents in education: a bibliometric review. *Innovaciones Educativas*, 26(41), 242–260. https://doi.org/10.22458/ie.v26i41.5135

Escudero Cristobal, W. J., Buitron Bruno, C. R., & Rodríguez Rojas, M. (2025). Competencias digitales y uso de inteligencia artificial en la educación superior: una revisión sistemática. *Revista Tribunal*, 5(13), 278-302. https://doi.org/10.59659/revistatribunal.v5i13.270

García Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. RIED-Revista Iberoamericana de Educación a Distancia, 27(1), 9–39. https://doi.org/10.5944/ried.27.1.37716

Lion, C., Kap, M. y Ferrareli, M. (2023) Universidades desafiadas: alfabetismos fluidos, hibridaciones y nuevas estrategias de enseñanza. *Revista Educación Superior y Sociedad,* 2023, VOL.35 N.º 2, 130-155. https://doi.org/10.54674/ess.v35i2.768

UNESCO. (2025) *Marco de competencias para docentes en materia de IA.* https://unesdoc.unesco.org/ark:/48223/pf0000393813

Zawacki-Richter, O., Marín V. I., Bond M. & Gouverneur F. (2019). Systematic review of research on artificial intelligence applications in higher education. *Journal of Educational Technology in Higher Education*, Nº. 16, 2019

Desarrollos tecnológicos para apoyar el aprendizaje y la enseñanza en grupos colaborativos

Rosanna Costaguta, María de los Ángeles Menini, Daniela Misio, Silvina Unzaga, Nevelin Salazar, Pablo Santana-Mansilla, Carlos Galván, Fernando Aramayo, Valeria Acosta, Walter Anriquez-Atia, Guillermo González, Facundo Gerez, Denise Díaz, María Luciana Figueroa, Nicolás Torres, Sofía Nieto-Piccoli, Juan Gómez, Luis López, Rodolfo Valentini.
Universidad Nacional de Santiago del Estero (UNSE)
Instituto de Investigación en Informática y Sistemas de Información (IIISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
rosanna@unse.edu.ar; marameni@unse.edu.ar

RESUMEN

El Aprendizaje Colaborativo Soportado por Computadora (ACSC) es un tipo de aprendizaje en el que los estudiantes desarrollan actividades colaborativas organizados en pequeños grupos, generalmente de entre tres a cinco integrantes, utilizando la computadora como medio de comunicación, cooperación, coordinación y colaboración, siendo orientados y asistidos por un docente o e-tutor.

Si bien el ACSC presenta ventajas reconocidas, se sabe que no puede asegurarse que los aprendizajes sean logrados realmente. Existen determinadas variables que condicionan que el resultado obtenido sea el esperado, y es aquí donde desarrollos informáticos específicos pueden marcar una diferencia significativa.

Este artículo presenta un proyecto de investigación que se enfoca en la problemática planteada, donde utilizando técnicas de inteligencia artificial; técnicas de aprendizaje de máquina; tecnologías actuales como: realidad aumentada y agentes conversacionales; enfoques promisorios como: gamificación, personalización y adaptación, con la finalidad de desarrollar aplicaciones que brinden apoyo al ACSC.

Las aplicaciones creadas serán utilizadas por estudiantes y docentes universitarios en sesiones colaborativas especialmente diseñadas. La recopilación de los datos experimentales y su posterior análisis permitirán realizar las validaciones correspondientes.

Palabras clave: *Inteligencia artificial, aprendizaje de máquina, agentes conversacionales, gamificación, realidad aumentada, objetos de aprendizaje, aprendizaje colaborativo soportado por computadora.*

CONTEXTO

El proyecto se titula "*Desarrollos tecnológicos para apoyar el aprendizaje y la enseñanza en grupos colaborativos",* se ejecutará en el período 2026-2029, con acreditación y financiamiento de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT - UNSE). El proyecto responde a una de las líneas prioritarias de investigación del IIISI, vinculada con la creación y uso de tecnologías informáticas destinadas a mejorar los procesos de enseñanza y de aprendizaje. Por otro lado, el proyecto continúa una línea de investigación iniciada en 2010-2011 por el proyecto 23/C089 "Fundamentos Conceptuales y Soportes Tecnológicos de la Informática Educativa", continuada en 2012-2016 por el proyecto 23/C098 "Sistemas de Información Web Basados en Agentes para Promover el Aprendizaje Colaborativo Soportado por Computadora", en 2017-2021 por el proyecto 23/C138 "Mejorando

escenarios de aprendizaje colaborativo soportado por computadora", y en 2022-2025 por el proyecto 23/C176-A-2022 "Desarrollo de aplicaciones para colaboración en e-learning". Todos los proyectos citados fueron acreditados y financiados por SECyT – UNSE.

INTRODUCCIÓN

Johnson, Johnson y Holubec [1] plantean el Aprendizaje Cooperativo como el trabajo conjunto de un grupo de personas que distribuye la realización de las tareas para concretar una meta; pero, es menester señalar que tal distribución no involucra *per se* una verdadera colaboración e implicación en el trabajo grupal, cuestión esta que sí es esencial en el Aprendizaje Colaborativo (AC), ya que requiere, de los estudiantes, interacción y acuerdos para el logro de una meta.

Bruffee [2] reconoce que entre el Aprendizaje Cooperativo y el AC existe una linealidad en la que se daría primero el Aprendizaje Cooperativo y luego, cuando el grupo está más afianzado, se abriría el espacio del AC. No obstante, los integrantes de este equipo de investigación no concuerdan con este planteo, por considerar que se trata de una concepción diferente y hasta disruptiva, ya que ambos tipos de aprendizaje grupal no tienen los mismos objetivos, ni la misma concreción. Pensar en Aprendizaje Cooperativo es pensar en tareas, organización y planificación indicadas y estructuradas por un docente [3], y por lo tanto, hay una dependencia en el trabajo parcializado de los compañeros.

En el AC la situación es diferente, se requiere de una interacción necesaria y verdadera para que el aprendizaje se produzca, esta interacción como ya se explicó, no sólo demanda acuerdos y negociaciones entre distintos puntos de vista, sino también un tiempo necesario para el trabajo grupal y especialmente, una organización interna en el grupo para el logro de los objetivos. Se necesita, además, el cambio de función del e-tutor y/o docente, donde este pasa de ser conductor de la tarea, a ser, podría decirse, un asesor, que orienta, asiste y está disponible para atender las dudas; pero que espera que sea el mismo grupo el que se organice para realizar su propia labor, aunque esto no siempre signifique o garantice el logro del resultado esperado.

Cuando las situaciones anteriormente desarrolladas en AC implican, además, el uso de medios computacionales, se habla de Aprendizaje Colaborativo Soportado por Computadora (ACSC), en inglés *Computer Supported Collaborative Learning* (CSCL), que comparte las características esenciales del AC, pero, adiciona el beneficio de la incorporación de un soporte tecnológico –en este caso computadoras o celulares inteligentes– particularmente útiles para este tipo de aprendizaje, posibilitando que los participantes pueden contribuir al grupo desde cualquier lugar y en cualquier momento.

Las experiencias de ACSC apuntan, entonces, a entender el aprendizaje como un proceso social de construcción de conocimiento colaborativo en el que la computadora actúa como un mediador que permite la potencia de trabajar con otros, debatir, dialogar y disponer de estrategias de colaboración productivas y efectivas, que posibilita pasar de la elaboración individual y/o cooperativa guiada y dividida, a una forma de trabajar y producir entre todos, con una verdadera participación, implicancia, intercambio y negociación de los integrantes del grupo, esto es, un trabajo grupal significativo y genuino [3].

Sin embargo, no basta con la mera conformación de grupos de estudiantes y la asignación de actividades a desarrollar mediante la mediación del e-tutor, para alcanzar una construcción adecuada de aprendizajes. Sin duda, crear grupos e indicar a sus miembros que resuelvan consignas de manera colaborativa, no garantiza que la experiencia sea exitosa; de hecho, el éxito de las experiencias en ACSC depende de la conjunción de una serie de variables que caracterizan el comportamiento tanto de los e-tutores como de los estudiantes que participan de dichas experiencias, siendo esta

conjunción fundamental para alcanzar los propósitos de enseñanza y los objetivos de aprendizaje que se prevean.

Nuestras investigaciones previas nos permiten afirmar que en el ACSC resulta fundamental poner especial atención sobre aspectos tales como: la conformación de los grupos, el comportamiento y las preferencias o inclinaciones individuales de los integrantes en el trabajo colaborativo, la consideración del desempeño del grupo como un todo, las habilidades de los e-tutores para actuar como tales, los materiales o recursos didácticos digitales disponibles y las tareas que se asignan para ser realizadas de manera colaborativa [4, 5]. Por lo expuesto, se consideran relevantes, entre otras, las siguientes variables de análisis dentro de un grupo de ACSC: habilidades colaborativas de los estudiantes, roles de equipo desempeñados, estilos de aprendizaje, estilos de personalidad, habilidades de los e-tutores, recursos de enseñanza y de aprendizaje disponibles.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO/OBJETIVOS

La hipótesis de investigación que guía la ejecución del proyecto que se presenta es la siguiente: *"En los sistemas de ACSC es posible mejorar el desempeño académico de los estudiantes universitarios y el desempeño de los docentes, así como la satisfacción de usuario, mediante el uso de técnicas de inteligencia artificial y aprendizaje de máquina, realidad aumentada, objetos de aprendizaje, agentes conversacionales, y/o gamificación*".

Considerando la hipótesis formulada, se fijaron los siguientes objetivos específicos:

- Determinar el estado actual de conocimiento y desarrollo de sistemas que apliquen técnicas provenientes de la inteligencia artificial y del aprendizaje de máquina, realidad aumentada, objetos de aprendizaje, agentes conversacionales, y/o gamificación en el ámbito del ACSC.
- Desarrollar aplicaciones que utilicen técnicas de inteligencia artificial y/o de aprendizaje de máquina, y/o realidad aumentada, y/u objetos de aprendizaje, y/o agentes conversacionales, y/o gamificación, considerando, características de los estudiantes tales como: habilidades de colaboración, roles de grupo, estilos de aprendizaje y estilos de personalidad; y características de los docentes, tales como: habilidades para desempeñarse como e-tutores, estilos de enseñanza, entre otras.
- Evaluar el impacto producido por la inclusión de técnicas de inteligencia artificial y/o de aprendizaje de máquina, y/o realidad aumentada, y/u objetos de aprendizaje, y/o agentes conversacionales, y/o gamificación, en los procesos de enseñanza y de aprendizaje, así como la satisfacción de usuario, en el ámbito de los sistemas de ACSC en contextos universitarios.
- Fomentar la participación de los integrantes del proyecto en actividades de extensión y transferencia.

Para alcanzar estos objetivos, primero se realizará la búsqueda de antecedentes vinculados con las temáticas abordadas. Analizados los antecedentes recopilados, se elaborará un estado del arte actualizado. Dicho estado del arte permitirá idear y construir aplicaciones informáticas que promuevan mejoras en los procesos de enseñanza y de aprendizaje, en el ámbito de los sistemas de ACSC. Para desarrollar estas aplicaciones se indagarán las características presentes en los estudiantes y docentes universitarios (por ejemplo: habilidades colaborativas de los estudiantes, roles de equipo desempeñados, estilos de aprendizaje, estilos de personalidad, habilidades de los e-tutores, estilos de enseñanza, entre otras). Además, será necesario diseñar las estructuras de datos que permitan almacenar la

información recopilada y también elaborar los algoritmos que permitan gestionarla. Estos algoritmos utilizarán técnicas y métodos provenientes de inteligencia artificial, aprendizaje de máquina. También se incorporarán nuevas tecnologías y enfoques como: realidad aumentada, objetos de aprendizaje, agentes conversacionales, y/o gamificación, seleccionando aquella o aquellas que resulten más adecuadas para mejorar los procesos de enseñanza y de aprendizaje en grupos colaborativos.

Para validar la hipótesis de investigación, se desarrollarán diferentes experiencias de uso de las aplicaciones creadas, en las que participarán docentes y estudiantes de la UNSE, recopilándose todas las interacciones manifestadas. Mediante la realización de encuestas también se recabará información de los actores involucrados, la que se sumará a los datos experimentales recolectados. Todos los datos recabados se procesarán para evaluar el impacto provocado por las aplicaciones creadas, en el desempeño tanto de los estudiantes como de los docentes. También se evaluará la satisfacción de usuario utilizando encuestas especialmente diseñadas, técnicas estadísticas y métricas de evaluación específicas del área.

La participación del equipo en actividades de extensión y transferencia se concretará a través del dictado de cursos, charlas, y conferencias, la elaboración de artículos para revistas indexadas, y presentaciones en congresos, simposios, seminarios, etc.

RESULTADOS ESPERADOS

Algunos de los resultados a los que se espera arribar al concretar este trabajo son:

- Estado del arte vinculado con el uso de inteligencia artificial, aprendizaje de máquina, realidad aumentada, objetos de aprendizaje, agentes conversacionales, y/o gamificación, en sistemas de ACSC.
- Especificación de las características consideradas para los estudiantes y también para los docentes en sistemas de ACSC.
- Desarrollo de aplicaciones que incorporen inteligencia artificial, aprendizaje de máquina, realidad aumentada, objetos de aprendizaje, agentes conversacionales, y/o gamificación, para mejorar el desempeño de los estudiantes y también de los docentes en sistemas de ACSC.
- Instalación de las aplicaciones desarrolladas.
- Ejecución de actividades experimentales que permitan validar el funcionamiento de las aplicaciones concretadas y hacer los ajustes que fueran necesarios.
- Presentación de resultados en eventos científicos nacionales e internacionales y publicaciones en revistas indexadas.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación está compuesto por tres docentes investigadores formados, y tres en formación. El equipo también cuenta con tres integrantes que desarrollarán sus tesis de posgrado (Maestría en Informática Educativa) y con diez integrantes estudiantes que desarrollarán sus trabajos finales para obtener sus títulos de grado (Licenciatura en Sistemas de Información) en el marco de este proyecto.

REFERENCIAS

[1] Johnson, D., Johnson, R., & Holubec, E. (1999). *El aprendizaje cooperativo en el aula*. Barcelona: Paidós.

[2] Bruffee, K. (1999). *Collaborative learning: higher education, interdependence and the authority of knowledge.* Baltimore (Maryland, EEUU): Johns Hopkins University Press

[3] Zañartu Correa, L. (2003). *Aprendizaje colaborativo: una nueva forma de Diálogo*

Interpersonal y en Red. Revista Digital de Educación y Nuevas Tecnologías – Contexto Educativo, Vol. 28.

[4] Costaguta, R., Sarmiento, W., & Luna-García, H. (2023). Recommendations to improve collaborative learning of Latin American computer science students according to dominant learning styles. *Campus Virtuales*, *12*(2), 45-55.

[5] Costaguta, R., & Menini, M. (2020). Grupo de Investigación en Aprendizaje Colaborativo Soportado por Computadora. *Revista de la Asociación Interacción Persona Ordenador (AIPO)*, *1*(1), 97-100.

Detección Temprana del Rendimiento Estudiantil mediante Analítica Predictiva y Vibe Coding

Carrizo Claudio, Javier Saldarini, Raúl Navarro Peláez, Daiana Viotti
Facultad Regional San Francisco – Universidad Tecnológica Nacional
Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
{cjcarrizo77, saldarinijavier, ingrnavarropelaez, dayviotti2015}@gmail.com

RESUMEN

La deserción universitaria constituye uno de los principales desafíos que enfrentan las instituciones de educación superior. En el ámbito de una Universidad Pública, se han detectado problemas de rendimiento académico en instancias de evaluación parcial en materias del área de programación del primer nivel de la carrera de Ingeniería en Sistemas de Información. Para atender esta problemática, se viene desarrollando un proceso de analítica académica predictiva orientado a anticipar el rendimiento estudiantil, en instancias de evaluación parcial. El presente trabajo aborda el análisis comparativo de nueve algoritmos de aprendizaje automático supervisado, con el objetivo de identificar los algoritmos que ofrecen mayor poder predictivo sobre el conjunto de datos. A modo de innovación, se introducirá una nueva metodología de asistencia a la codificación denominada Vibe Coding, donde se utilizará GitHub Copilot como herramienta de Inteligencia Artificial Generativa. Los beneficios esperados incluyen la generación de alertas tempranas para docentes y gestores académicos, y la reducción de la tasa de deserción estudiantil mediante intervenciones pedagógicas oportunas.

Palabras clave: analítica predictiva, educación superior, inteligencia artificial generativa, rendimiento estudiantil, vibe coding

CONTEXTO

La presente línea de investigación se enmarca dentro del Proyecto de Investigación y Desarrollo (PID) denominado "Analítica Predictiva aplicada para la mejora del desempeño académico en UTN Facultad Regional San Francisco", el cual fue homologado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, bajo el código TEECSF0010247 y según la Disp. SCTyP Nº 32-2023. Cabe destacar que este PID se viene ejecutando desde inicios del año 2024.

1. INTRODUCCIÓN

En la actualidad, las organizaciones generan grandes volúmenes de datos que, por sí solos, no aportan valor (Adi & Isnanto, 2020). Aquellas organizaciones que no transformen estos datos en conocimiento accionable, desperdician uno de sus activos más importantes (Sumpala & Rasyid, 2019). La analítica de datos surge como respuesta a esta necesidad, integrando técnicas

provenientes de disciplinas como la estadística, la inteligencia de negocios, la minería de datos y el aprendizaje automático (Argonza, 2016; Burman & Som, 2019).

En el ámbito educativo, la analítica académica predictiva se ha consolidado como una herramienta fundamental para anticipar el rendimiento estudiantil y diseñar intervenciones pedagógicas oportunas (Viberg et al., 2018). La identificación temprana de estudiantes en riesgo de bajo rendimiento o deserción, es uno de los desafíos más relevantes en la educación superior, especialmente en carreras de ingeniería, donde el primer año representa una etapa crítica (Murnion & Helfert, 2013; Di Gresia, 2007).

En el ámbito de una Facultad de Ingeniería de una Universidad Pública Argentina, los datos estadísticos del Área de Gestión Académica evidencian una alta tasa de bajo rendimiento en instancias de evaluación parcial en materias del área de programación del primer nivel de la carrera de Ingeniería en Sistemas de Información, lo que frecuentemente deriva en la reprobación de dichas materias y, en muchos casos, en el abandono de la carrera (Carrizo et al., 2025). Esta problemática motivó el inicio de un proyecto de investigación, cuyo objetivo principal es construir un modelo predictivo que permita anticipar el desempeño académico de los estudiantes.

Durante el primer año del proyecto, se avanzó en la construcción del perfil del estudiante, el relevamiento de un dataset compuesto por 2335 registros y 33 atributos, la descripción de variables, tareas de pre-procesamiento en Python (limpieza, transformación, calidad de datos), el análisis exploratorio mediante estadística descriptiva, y la selección del conjunto de atributos para el modelo predictivo (Carrizo et al., 2025). En el segundo año, la investigación avanzó hacia la etapa de estudio, selección y evaluación de algoritmos de aprendizaje supervisado.

El aprendizaje automático supervisado constituye la familia de algoritmos de Machine Learning más utilizada para tareas de clasificación y predicción sobre datos educativos (Kotsiantis, 2012). Algoritmos como Árbol de Decisión, Random Forest, k-Nearest Neighbors (KNN), Regresión Logística, Support Vector Machine (SVM) y Naive Bayes han demostrado resultados efectivos en contextos similares (Dutt et al., 2017; Romero & Ventura, 2020). Sin embargo, la elección del algoritmo más apropiado depende de las características propias del conjunto de datos, y del problema a resolver, lo que hace necesario llevar a cabo un estudio comparativo sistemático.

Una innovación metodológica que se incorpora en esta etapa del proyecto es el uso de "Vibe Coding", un paradigma emergente de desarrollo asistido por Inteligencia Artificial Generativa, en el que el programador describe en lenguaje natural el problema o la tarea a resolver, y un modelo de IA genera el código correspondiente (Sarkar & Drosos, 2025). Esta práctica está transformando el modo en que los investigadores y desarrolladores abordan la implementación de soluciones complejas, reduciendo los tiempos de codificación, facilitando la exploración de múltiples alternativas algorítmicas y permitiendo que el investigador focalice su esfuerzo en el análisis crítico de los resultados. Herramientas como GitHub Copilot, Cursor, Claude y ChatGPT se encuentran entre las más utilizadas en este contexto (Chen et al., 2021; Sarkar et al., 2022; Barke et al., 2023).

La incorporación de Vibe Coding como metodología de asistencia a la codificación en proyectos de investigación académica representa un aporte novedoso, con potencial para mejorar la productividad del

equipo, reducir errores y acelerar la curva de aprendizaje en el uso de librerías especializadas como scikit-learn, pandas y matplotlib en Python.

En consecuencia, el presente trabajo propone realizar un estudio comparativo de nueve algoritmos de aprendizaje supervisado (Regresión Logística Multinomial, Árbol de Decisión, Random Forest, Gradient Boosting, XGBoost, k-Nearest Neighbors, Support Vector Machine, Naive Bayes y Redes Neuronales), aplicados al conjunto de datos, y asistido por Vibe Coding, con el fin de identificar el modelo predictivo más adecuado para anticipar el desempeño académico de estudiantes en materias de programación de primer nivel. Cabe mencionar que este trabajo es la continuación de la línea de investigación presentada en WICC 2025 (Carrizo, C., Saldarini, J., Navarro Peláez, R., Viotti, D., & Calloni, A., 2025).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los ejes de trabajo de la línea de investigación para esta etapa del proyecto son los siguientes:

- Estudio en profundidad de los principales algoritmos de aprendizaje supervisado aplicables a la predicción del desempeño académico.
- Incorporación del paradigma Vibe Coding como metodología de asistencia a la codificación e implementación de modelos en Python, mediante el uso de herramientas de IA generativa.
- Diseño del esquema de evaluación comparativa: definición de métricas de desempeño (exactitud, precisión, recall, F1-Score, AUC-ROC) y estrategia de validación cruzada.
- Implementación de cada uno de los algoritmos seleccionados sobre el conjunto de datos pre-procesados, utilizando la biblioteca scikit-learn de Python.
- Evaluación y comparación de los resultados obtenidos por cada modelo a través de las métricas definidas.
- Selección del modelo predictivo que ofrezca el mejor desempeño para la problemática propuesta.
- Documentación del proceso de Vibe Coding aplicado: prompts utilizados, herramientas de IA empleadas, calidad del código generado y correcciones realizadas.
- Comunicación de resultados al Área de Gestión Académica y docentes de materias del Área de Programación.

3. RESULTADOS

Resultados Obtenidos

- Construcción del perfil del estudiante de primer año, integrado por datos personales, geográficos, académicos y socioeconómicos
- Obtención de un dataset de cursado de estudiantes en materias del área de programación, compuesto por 33 atributos y 2335 registros.
- Descripción detallada de cada atributo del dataset, incluyendo tipo de dato y valores asociados.
- Exploración inicial de datos a través de herramientas de visualización.

- Pre-procesamiento del dataset utilizando Python: limpieza, transformación y verificación de calidad de datos.
- Análisis exploratorio de datos mediante estadística descriptiva.
- Selección del conjunto de atributos a incluir en el modelo predictivo.

Resultados Esperados

- Informe comparativo de los algoritmos de aprendizaje supervisado evaluados, con análisis de métricas de desempeño por cada modelo.
- Selección y justificación del algoritmo más apropiado para la construcción del modelo predictivo definitivo.
- Documentación metodológica del proceso de Vibe Coding aplicado en la investigación, como aporte a la comunidad científica.
- Construcción, entrenamiento y evaluación del modelo predictivo seleccionado.
- Comunicación de resultados al Área de Gestión Académica y docentes de las cátedras del Área de Programación, mediante una herramienta de visualización.
- Publicación de los resultados en congresos y jornadas científicas.

FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación se trabaja desde un Grupo de Investigación y Desarrollo de UTN Facultad Regional San Francisco.

El equipo de trabajo está constituido por: 1 Director de Proyecto, 1 Co-Director, 1 docente investigador de apoyo y 1 estudiante de nivel avanzado, todos de la especialidad de Ingeniería en Sistemas de Información.

En el contexto de la ejecución del proyecto, en donde se inserta la presente línea de investigación, se está desarrollando un trabajo final de carrera de un estudiante de Ingeniería en Sistemas de Información.

Cabe destacar, que este equipo de trabajo tiene antecedentes de haber desarrollado y obtenido resultados en PID de UTN, vinculados a la temática "Minería de datos educacional". También se han presentado resultados en jornadas, workshop y congresos científicos de la especialidad.

5. BIBLIOGRAFÍA

Adi, K., & Isnanto, R. R. (2020). Rice crop management expert system with forwarding chaining method and certainty factor. Journal of Physics: Conference Series, 1524(1), 012037.

Argonza, J. S. (2016). Big data en la educación. Revista de Investigación, 17(1).

Barke, S., James, M. B., & Polikarpova, N. (2023). Grounded copilot: How programmers interact with code-generating models. Proceedings of the ACM on Programming Languages, 7(OOPSLA1), 1–27.

Burman, I., & Som, S. (2019). Predicting students' academic performance using support vector machine. In 2019 Amity International Conference on Artificial Intelligence (AICAI) (pp. 756–759). IEEE.

Carrizo, C., Saldarini, J., Navarro Peláez, R., Viotti, D., & Calloni, A. (2025). Uso de Analítica Predictiva para Anticipar el Desempeño Académico en Materias del Área

Programación. En XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025). Argentina.

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. O., Kaplan, J., ... & Zaremba, W. (2021). Evaluating large language models trained on code.

Di Gresia, L. (2007). Rendimiento académico universitario. Asociación Argentina de Economía Política.

Dutt, A., Ismail, M. A., & Herawan, T. (2017). A systematic review on educational data mining. IEEE Access, 5, 15991–16005.

Sarkar, A., & Drosos, I. (2025). Vibe coding: Programming through conversation with artificial intelligence. In Proceedings of the 36th Annual Conference of the Psychology of Programming Interest Group (PPIG 2025).

Kotsiantis, S. B. (2012). Use of machine learning techniques in educational data mining: A review. IADIS International Journal on Computer Science and Information Systems, 7(2), 31–50.

Murnion, P., & Helfert, M. (2013). Academic Analytics in Quality Assurance Using Organisational Analytical Capabilities. UK Academy for Information Systems Conference Proceedings 2013.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. WIREs Data Mining and Knowledge Discovery, 10(3), e1355.

Sarkar, A., Gordon, A. D., Negreanu, C., Poelitz, C., Ragavan, S. S., & Zorn, B. (2022). What is it like to program with artificial intelligence? In Proceedings of the Psychology of Programming Interest Group (PPIG 2022).

Sumpala, A. T., & Rasyid, R. (2019). Expert system for corn plant disease diagnosis with the breadth-first search method. IOP Conference Series: Earth and Environmental Science, 382(1), 012001.

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in Human Behavior, 89, 98–110.

Diseño y validación de Objetos de Aprendizaje (OA) mediante el uso de modelos de Inteligencia Artificial Generativa: Un enfoque hacia la personalización de contenidos educativos en la educación superior

Minnaard Claudia[1], Torres Zulma[1], Minnaard Vivian[2]

[1]Instituto de Investigaciones de Tecnología y Educación (IITE)

Facultad de Ingeniería – Universidad Nacional de Lomas de Zamora

[2]Universidad FASTA

Buenos Aires, Argentina

minnaardclaudia@gmail.com , iite@ingenieria.unlz.edu.ar

RESUMEN

Esta investigación aborda la creación de objetos de aprendizaje optimizados para el ámbito de la ingeniería mediante el uso de Inteligencia Artificial. El proceso incluye una fase de implementación y una posterior auditoría de resultados para verificar su efectividad pedagógica.

Palabras claves: Inteligencia artificial, Objetos de aprendizaje, Ambientes virtuales de aprendizaje.

CONTEXTO

El proyecto se desarrolla en el Instituto de Investigaciones de Tecnología y Educación de la Facultad de Ingeniería de Universidad Nacional de Lomas de Zamora. Desde la fecha de su creación el Instituto, en 2009, resulta ser un ámbito de investigación y desarrollos concretos sobre ambientes virtuales y exploración de aplicaciones de software para la enseñanza en carreras tecnológicas, en particular la ingeniería.

La literatura académica reciente coincide en que el lanzamiento de ChatGPT en noviembre de 2022 marcó un punto de inflexión, desencadenando una adopción masiva de la inteligencia artificial en múltiples sectores. Si bien esta herramienta es solo una de las diversas tecnologías generativas existentes, su influencia ha sido particularmente disruptiva en el ecosistema educativo (García-Peñalvo, 2024; González, 2004; Heredia et al., 2023). Este fenómeno ha obligado a replantear las dinámicas de enseñanza y adquisición de conocimientos frente a la ubicuidad de estos sistemas.

Al profundizar en la naturaleza de esta tecnología, García-Peñalvo (2024) aclara que la IA generativa opera bajo una lógica probabilística de creación de contenidos. Estos modelos lingüísticos se fundamentan en el entrenamiento con vastos volúmenes de datos para predecir secuencias de elementos con alta correlación estadística. Por consiguiente, las respuestas generadas no son producto de un proceso cognitivo o de razonamiento, sino de un cálculo de verosimilitud dentro de su corpus de entrenamiento. Esta distinción es crítica en el contexto pedagógico: aunque los resultados presenten una coherencia gramatical y discursiva notable, carecen de una garantía intrínseca de veracidad. Por ello, su integración en la enseñanza exige una supervisión crítica que considere esta limitación estructural para evitar la validación de información errónea.

El Consenso de Beijing sobre la inteligencia artificial y la educación (UNESCO, 2019) destaca 44 recomendaciones, agrupadas en distintos aspectos que pueden ayudar a entender la magnitud de la tarea:

- planificación de la IA en las políticas educativas;
- la IA para la gestión y la impartición de la educación;
- la IA para apoyar la docencia y a los docentes;

- la IA para el aprendizaje y la evaluación del aprendizaje;

- el desarrollo de valores y competencias para la vida y el trabajo en la era de la IA;

- la IA para ofrecer oportunidades de aprendizaje permanente para todos;

- promoción del uso equitativo e inclusivo de la IA en la educación;

- IA con equidad de género e IA para la igualdad de género; y velar por el uso ético, transparente y verificable de los datos y algoritmos educativos

La UNESCO, a través del Consenso de Beijing (2019), aboga por una inteligencia artificial con enfoque humano que reduzca las brechas de conocimiento y desigualdad. La prioridad es evitar que esta tecnología profundice la distancia digital entre países, garantizando que los avances en innovación sean un bien común. En esencia, se busca que la revolución de la IA sea inclusiva, permitiendo que todas las sociedades se beneficien equitativamente de sus descubrimientos.

Según los lineamientos establecidos por la UNESCO (2023) en su guía dedicada a la educación e investigación, el panorama tecnológico global experimentó una transformación radical hacia finales del año 2022. Este cambio fue impulsado principalmente por la aparición de ChatGPT, una herramienta de inteligencia artificial generativa (IAGen) que marcó un hito histórico al posicionarse como la plataforma con el crecimiento de usuarios más acelerado desde que se tiene registro. El impacto de estas aplicaciones radica en su sorprendente capacidad para imitar facultades cognitivas humanas, permitiendo la creación automatizada de contenidos diversos que van desde textos complejos y líneas de código de programación hasta piezas musicales, imágenes y producciones audiovisuales.

Desde una perspectiva técnica, la IAGen no es un fenómeno aislado, sino que se sustenta en el Aprendizaje Automático (AA) o Machine Learning. Esta disciplina se basa en el uso de algoritmos diseñados para evolucionar de manera autónoma, optimizando su precisión y eficiencia a medida que procesan volúmenes masivos de datos. Dentro de este ecosistema, destacan las Redes Neuronales Artificiales (RNAs), las cuales han sido responsables de avances críticos como el reconocimiento facial. Estas redes están arquitectónicamente inspiradas en la biología humana, intentando replicar la dinámica de las conexiones sinápticas y el flujo de información que ocurre entre las neuronas del cerebro.

En el contexto universitario actual, el auge de estas tecnologías ha provocado un giro drástico en los modelos pedagógicos tradicionales. Investigaciones recientes (Aguilar et al., 2023) subrayan que el propósito de esta transición es analizar profundamente cómo se implementan estas herramientas en las aulas para anticipar las tendencias educativas del mañana. No se trata solo de adoptar tecnología, sino de comprender su rol en la construcción de nuevas metodologías de enseñanza-aprendizaje.

Finalmente, el funcionamiento práctico de estos sistemas depende de la interacción directa con el usuario a través de los denominados "prompts". Un prompt se define como el comando, pregunta o instrucción específica que una persona introduce en la interfaz de la IA. Es el punto de partida que orienta al sistema para producir un resultado coherente y útil, convirtiéndose en la pieza fundamental de comunicación entre la intención humana y la capacidad de procesamiento de la máquina.

1. INTRODUCCIÓN

En la actualidad, el auge de las Inteligencias Artificiales (IA) ha implicado un cambio significativo en las metodologías de enseñanza dentro del ámbito de la educación universitaria. El objetivo principal de esta transformación es explorar y comprender el desarrollo y la implementación de las IA en el contexto de la educación universitaria, así como anticipar el enfoque futuro que representará el uso de estas tecnologías emergentes en los procesos pedagógicos para

obtener una comprensión más clara y precisa del contexto actual. (Aguilar et al, 2023).

Los sistemas de inteligencia artificial generativos permiten crear textos o imágenes a través de las demandas de los usuarios. Estas demandas o instrucciones se canalizan a través de los prompts.

En el contexto de la IA, el prompt es un fragmento de texto que se proporciona a un modelo de lenguaje para influir en la generación de un texto. Por ejemplo, en un chatbot, el prompt es una solicitud que guía la interacción entre el usuario y el sistema de IA. Un prompt bien diseñado debe ser:

• Claro y conciso: Evita ambigüedades y formula tu solicitud de manera directa.
• Específico: Proporciona detalles concretos sobre lo que deseas que se genere.
• Contextualizado: Ofrece suficiente información de fondo para que el modelo comprenda tu petición.
• Bien estructurado: Organiza tus ideas de forma lógica y coherente.
• Orientado a resultados: Define claramente el objetivo que quieres alcanzar.

Aunque los GPTs y su capacidad para generar texto automáticamente han estado disponibles para los investigadores desde 2018, el lanzamiento de ChatGPT fue novedoso por su acceso gratuito a través de una interfaz de uso fácil, lo que significa que cualquier persona con acceso a Internet podría explorar la herramienta. El lanzamiento de ChatGPT causó conmoción en todo el mundo y rápidamente impulsó a otras empresas tecnológicas transnacionales a readecuarse al nuevo paradigma, sucediendo lo mismo con numerosas start-ups.

Uno de los pilares más críticos en el debate contemporáneo sobre la tecnología es la urgencia de instaurar la ética de la inteligencia artificial (IA) como un eje transversal en los planes de estudio de la educación superior. No se trata simplemente de aprender a utilizar herramientas tecnológicas, sino de comprender las implicaciones morales que su uso conlleva. En este sentido, la investigación presentada se enfoca en desmenuzar los peligros intrínsecos al uso masivo de estas herramientas en el entorno escolar. El propósito último es diseñar y sugerir un catálogo de competencias éticas que permitan a estudiantes y docentes adoptar estas innovaciones de una manera consciente, responsable y alineada con los valores sociales.

Para dotar de evidencia científica a esta propuesta, se llevó a cabo un proceso de investigación bajo un enfoque descriptivo, empleando un diseño no experimental y de corte transversal. El estudio no se limitó a una sola geografía, sino que buscó una perspectiva latinoamericana al integrar una muestra de 144 alumnos de grado provenientes de diversas instituciones de educación superior en México, Ecuador y Argentina. La selección de los participantes se realizó mediante un muestreo no probabilístico por conveniencia, lo que permitió capturar la realidad de distintos contextos académicos. Para la recolección de los datos, se utilizó un formulario digital distribuido en línea, facilitando una participación asincrónica y eliminando las barreras geográficas entre los investigadores y los sujetos de estudio.

Los datos obtenidos a través de este instrumento revelaron una serie de inquietudes profundas respecto al avance de la IA. Entre los hallazgos más significativos destacan:

- La deshumanización del aprendizaje: El temor a que la mediación tecnológica sustituya el vínculo humano esencial en la pedagogía.

- Vulnerabilidad de la privacidad: La preocupación por el manejo de datos personales y la seguridad de la información.

- Deterioro del pensamiento crítico: El riesgo de que los alumnos pierdan la capacidad de evaluar la veracidad y calidad de los contenidos generados automáticamente.

- Integridad académica: El incremento de prácticas ligadas a la deshonestidad y el plagio, así como el uso de

materiales protegidos por derechos de autor sin la debida autorización.

Como conclusión derivada de estos resultados, el estudio subraya que es imposible separar el uso de la IA de sus repercusiones éticas. Existe una necesidad imperiosa de que las universidades establezcan marcos competenciales sólidos. Estos marcos no deben ser únicamente técnicos, sino que deben lograr una armonía entre el dominio tecnológico, el compromiso con los valores humanos y la responsabilidad social. Solo mediante esta articulación será posible garantizar que la inteligencia artificial actúe como un potenciador del talento humano y no como un agente de erosión ética en la educación. (Balladares-Burgos, J. ; Minnaard, V.; Urrutia, ME.; Minnaard, C. ; Sánchez Martínez, F. (2026))

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En este contexto se presentan los objetivos planteados y se detallan los procedimientos implementados para alcanzarlos:

- *Relevar, indagar y seleccionar las apps desarrolladas utilizando IA que tengan aplicación en las distintas cátedras de la FIUNLZ.*

Se relevarán todas fuentes bibliográficas primarias y secundarias, a fin de reconocer las distintas herramientas de IA.

- *Diseñar objetos de aprendizaje utilizando herramientas de IA.*

Desarrollar e implementar propuestas pedagógicas adecuadas para las carreras de ingeniería. Diseño de rúbricas ajustadas a las propuestas.

- *Aplicar y medir el resultado de aprendizaje de los OA diseñados.*

Una vez desarrollados los OA, se aplican en las distintas cátedras y evalúa el resultado del aprendizaje a través de rúbricas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Desde la perspectiva del Instituto de Investigaciones de Tecnología y Educación (IITE), el presente proyecto de investigación, propone, desde un marco teórico inscripto en el paradigma de la innovación educativa y la enseñanza y evaluación por competencias, evaluar experiencias desarrolladas utilizando herramientas de IA aplicables a las distintas cátedras, para proponer un modelo que contribuya a la eficiencia de la enseñanza de por competencias en carreras de ingeniería, a través de modelos híbridos (presenciales y virtuales), favoreciendo el aumento de la retención.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto prevé la formación de un becario EVC-CIN para el período 2025 - 2026.

5. BIBLIOGRAFÍA

-Aguilar, G. M. F., Gavilanes, D. C. A., Freire, E. M. A., & Quincha, M. L. (2023). Inteligencia artificial y la educación universitaria: Una revisión sistemática. Magazine de las Ciencias: *Revista de Investigación e Innovación, 8(*1), 109-131.

- Balladares-Burgos, J. ; Minnaard, V.; Urrutia, ME.; Minnaard, C. ; Sánchez Martínez, F. (2026) Competencias éticas asociadas a la Inteligencia Artificial en la Educación Superior. Etic@net. Electronic Scientific Journal of Education and Communication at the Knowledge. Granada, España.

- Consenso de Beijing sobre la inteligencia artificial y la educación(UNESCO, 2019) En: https://www.unesco.org/es/digital-education/artificial-intelligence

-García Peñalvo, F. J., Llorens Largo, F., & Vidal García, F. J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. RIED. *Revista*

iberoamericana de educación a distancia.
- Guidance for generative AI in education and research (2023). Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. En : https://unesdoc.unesco.org/ark:/48223/pf0000389227
-Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.
-González, C. S. (2004). Sistemas inteligentes en la educación: una revisión de las líneas de investigación y aplicaciones actuales. RELIEVE. *Revista Electrónica de Investigación y Evaluación Educativa, 10*(1), 3-22.
-Heredia, M. O. T., Correa, Y. K. D., & Carballo, E. V. (2023). Inteligencia artificial y educación: nuevas relaciones en un mundo interconectado. *Estudios del desarrollo social: Cuba y América Latina*, *11*(2), 312-328.

DISEÑO, MANUFACTURA ADITIVA E INTEGRACIÓN DE HARDWARE DE UNA PLATAFORMA BÍPEDA MODULAR PARA INVESTIGACIÓN EN LOCOMOCIÓN Y CONTROL DINÁMICO

Axel Leonel Gasser, Nahuel Cerrano,
Néstor Adrián Balich, Franco Adrián Balich
Laboratorio de Robótica Física e Inteligencia Artificial (LRFIA)
Centro de Altos Estudios en Tecnología Informática (CAETI)
Universidad Abierta Interamericana. Informática (UAI)
Montes de Oca 745. Ciudad Autónoma de Buenos Aires, Argentina.
{ axelleonel.gasser, nahuel.cerrano } @alumnos.uai.edu.ar
{ nestor.balich, francoadrian.balich} @uai.edu.ar

RESUMEN

El trabajo se enmarca en la línea de investigación LRFIA del CAETI-UAI. Describe el diseño, manufactura aditiva e integración de hardware de una plataforma bípeda modular para investigar la locomoción y el control dinámico. El objetivo es desarrollar una plataforma de bajo costo que valide modelos teóricos de estabilidad y aprendizaje por refuerzo en entornos reales, superando las limitaciones de la simulación. Se abordó la materialización del tren inferior mediante impresión 3D (PLA/ABS), el uso de microcontroladores ESP32 y actuadores de alto torque. Como resultado, se obtuvo un kit funcional preparado para algoritmos basados en el Punto de Momento Cero (ZMP) y fusión sensorial mediante IMU.

Palabras clave: Robótica Bípeda, Manufactura Aditiva, Control Dinámico, ZMP, Formación en I+D.

CONTEXTO

El proyecto pertenece a la línea de investigación Sistemas Inteligentes y Robótica del CAETI y Tecnología Informática Aplicada en Educación del WICC, enfocada en la validación experimental de plataformas físicas para el estudio de locomoción y control dinámico. Dado que muchas investigaciones se limitan a simulaciones por los altos costos de los robots bípedos, esta plataforma modular de bajo costo permite contrastar modelos teóricos con variables físicas reales como fricción, vibraciones y restricciones energéticas.

1. INTRODUCCIÓN

La construcción de sistemas bípedos mediante manufactura aditiva representaría una solución eficiente para el prototipado de bajo costo en robótica dinámica. El presente trabajo se centra exclusivamente en la fase de diseño, manufactura y validación de componentes de un robot bípedo modular. A diferencia de otros proyectos que abordan únicamente la simulación, nuestro enfoque ha sido la materialización de los eslabones físicos mediante impresión 3D y la definición de la arquitectura de hardware. Se describirá la fabricación de un robot bípedo impreso en ácido poliláctico (PLA), diseñado con tres grados de libertad por extremidad: cadera, rodilla y tobillo. Este sistema se centraría exclusivamente en el movimiento dentro del plano sagital, utilizando servomotores MG996R para ejecutar las trayectorias de marcha.

Para garantizar la estabilidad, se

implementaría el criterio del **Punto de Momento Zero (ZMP)** [1] Este método, ampliamente validado en el diseño de controladores para robots bípedos, permite determinar la tendencia al vuelco analizando el punto de cruce de las fuerzas inerciales y gravitacionales con la superficie de apoyo, facilitando así la planificación de una marcha estable en superficies planas [2]. A pesar de la existencia de métodos alternativos como la **Dinámica Cero Híbrida (HZD)**, que ofrece ventajas en la regulación de velocidad mediante optimización compleja [3], el ZMP se seleccionaría por su compatibilidad con el modelo de Péndulo Invertido Lineal (LIPM), permitiendo un control predictivo del equilibrio basado en el soporte de los pies. También su implementación facilita la planificación de una marcha estable en superficies planas, constituyendo una base sólida para el desarrollo de algoritmos de locomoción. La gestión de la marcha se realizaría a través de una **máquina de estados finitos** que coordinaría la altura del Centro de Masa (CoM) y la longitud de zancada, inspirada en los principios de balance activo de Raibert [4].

Dado que el diseño prescindiría de sensores de presión en la planta de los pies utilizando en su lugar una base de ABS [8] para la absorción de vibraciones y rebotes , la estabilidad dependería críticamente de una unidad de medición inercial (IMU). Para la fusión de datos sensoriales del IMU [10], se implementará inicialmente un **Filtro Complementario**. Esta técnica ha sido seleccionada por su bajo costo computacional, permitiendo una estimación eficaz de la orientación en microcontroladores estándar mediante la combinación ponderada de acelerómetro y giroscopio (0.98/0.02) [5]. En etapas posteriores de locomoción dinámica, se prevé la migración al algoritmo de **Madgwick**, el cual ofrece una compensación superior de la deriva magnética y mayor precisión angular con una carga de procesamiento menor que un filtro de Kalman [5]. Este sensor, ubicado en la cadera cerca del CoM, proporcionaría los datos de inclinación necesarios para ajustar el ZMP y evitar el volcado. La eficacia de utilizar acelerómetros para medir la inclinación y ajustar la velocidad de los motores en tiempo real ha sido demostrada en plataformas móviles para compensar desniveles y mantener una trayectoria constante [6], principio que se adaptará aquí para la estabilidad bípeda.

La construcción de robots bípedos presenta desafíos únicos que van más allá del software; la estabilidad de la estructura y la precisión en el ensamblaje son prerrequisitos fundamentales antes de escribir cualquier línea de código de control. En el ámbito académico, disponer de un robot físico permite validar modelos teóricos con las restricciones reales de fricción, peso y torque.

Se describe la metodología utilizada para asegurar un ensamblaje correcto, minimizando holguras mediante el uso adecuado de rulemanes y tornillería de precisión, y garantizando una gestión limpia del cableado para evitar interferencias mecánicas durante el movimiento.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

2.1 Objetivo General

Desarrollar una plataforma bípeda modular de bajo costo mediante manufactura aditiva e integración de hardware embebido, orientada a la investigación experimental en locomoción, control dinámico y aprendizaje por refuerzo, que permita validar modelos teóricos de estabilidad en condiciones físicas reales. [7]

2.2 Objetivos Específicos

1. Diseñar el tren inferior de una plataforma bípeda con tres grados de libertad por extremidad (cadera, rodilla y tobillo), optimizando su arquitectura para facilitar la modularidad, el mantenimiento y la replicabilidad académica (figura 1).

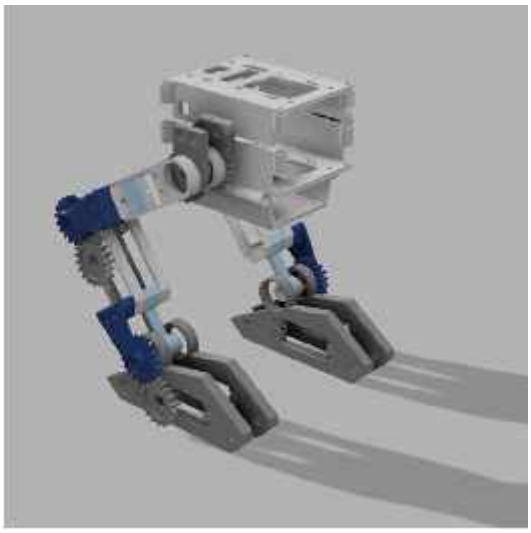


Fig 1. Modelo CAD completo ensamblado en Fusion 360.

2. Implementar la manufactura aditiva de los componentes estructurales utilizando materiales poliméricos (PLA y ABS), evaluando su desempeño mecánico y adecuación estructural para aplicaciones en robótica dinámica (figura 2).



Fig 2. Componentes físicos manufacturados mediante impresión 3D.

3. Seleccionar e integrar actuadores de alto torque (servomotor MG996R) y una arquitectura electrónica basada en microcontroladores ESP32, garantizando capacidad de procesamiento suficiente para futuras implementaciones de control en tiempo real.
4. Diseñar la arquitectura eléctrica y de alimentación del sistema, asegurando una distribución eficiente de energía y una gestión ordenada del cableado que minimice interferencias mecánicas durante el movimiento.
5. Incorporar sensores inerciales (IMU) para la futura implementación de estrategias de control en lazo cerrado basadas en el criterio del Punto de Momento Cero (ZMP) [1] y modelos dinámicos como el Péndulo Invertido Lineal (LIPM) [2].
6. Desarrollar un modelo digital del robot en entorno de simulación física

(PyBullet) que permita validar algoritmos de locomoción antes de su implementación en la plataforma real.

7. Generar un entorno formativo interdisciplinario que integre diseño mecánico, electrónica, programación embebida y simulación dinámica, promoviendo la formación de recursos humanos en investigación aplicada.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1 Resultados Obtenidos

En la etapa actual del proyecto se logró la materialización de los componentes físicos del tren inferior mediante manufactura aditiva, así como la definición y adquisición de la arquitectura de hardware, validando su factibilidad de forma teórica. Los principales resultados alcanzados incluyen:

• Impresión en PLA y ABS y posprocesado de la totalidad de los componentes estructurales y validación teórica mediante modelo CAD del encastre de los servomotores MG996R en las articulaciones de cadera, rodilla y tobillo, confirmando la viabilidad espacial del diseño modular.

• Definición y adquisición de la arquitectura de control, implementada mediante una placa WEMOS D1 R32 basada en ESP32, seleccionada por su mayor capacidad de procesamiento y conectividad inalámbrica integrada en comparación con microcontroladores tradicionales de 8 bits, características relevantes para futuras tareas de telemetría y control dinámico [9].

• Diseño teórico de un esquema de cableado ordenado y una distribución diferenciada de alimentación para lógica y actuadores. Esto permitirá a futuro minimizar interferencias eléctricas y garantizar que no se generen esfuerzos indebidos sobre los conductores durante el movimiento articular.

• Desarrollo inicial del modelo digital en entorno de simulación física, estableciendo correspondencia geométrica estricta entre las piezas reales impresas y su representación virtual.

3.2 Proyección y líneas futuras

Con los componentes impresos y el hardware electrónico (servos y placa WEMOS) a disposición, el proyecto avanza hacia el ensamblaje físico final y la etapa de implementación de control dinámico y locomoción autónoma. Las líneas de trabajo previstas incluyen:

• Implementación de algoritmos de estabilidad basados en el criterio del Punto de Momento Cero (ZMP) [1] y el modelo del Péndulo Invertido Lineal (LIPM) [2].

• Integración de sensores inerciales (IMU) en configuración cercana al centro de masa para control en lazo cerrado, empleando técnicas de fusión sensorial como filtro complementario y algoritmo de Madgwick [5].

• Desarrollo de una máquina de estados finitos para la coordinación del ciclo de marcha (gait cycle), regulando la altura del centro de masa, la velocidad y la longitud de zancada.

• Validación experimental comparativa entre la simulación en PyBullet y el comportamiento físico del robot una vez ensamblado, ajustando los parámetros dinámicos reales.

• Exploración futura de técnicas de aprendizaje por refuerzo para la optimización automática de los patrones de marcha.

La continuidad del proyecto permitirá generar producción científica en el área de locomoción bípeda y fortalecer la infraestructura experimental del laboratorio.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto integra a docentes y alumnos de Ingeniería en Sistemas en un espacio de formación en investigación aplicada. Diseño y manufactura: Modelado CAD (Fusion 360), ensamblajes mecánicos e impresión 3D (PLA/ABS). Diseño mecánico: Estructuras, distribución de cargas, estabilidad y uso de componentes técnicos. Electrónica y control: Arquitectura basada en ESP32, gestión de señales PWM, alimentación y cableado.

5. BIBLIOGRAFÍA

[1] M. Vukobratović y B. Borovac, "Zero-Moment Point — Thirty Five Years of Its Life," *International Journal of Humanoid Robotics*, vol. 1, no. 1, pp. 157–173, 2004. Disponible en: [World Scientific] [CrossRef].

[2] Lobo, M. F. V. (2013). *Control en Modo Deslizante con Estimación de la Perturbación Aplicado a la Marcha de un Robot Bípedo* (Doctoral dissertation, Instituto Nacional de Astrofísica, Óptica y Electrónica).

[3] M. J. Powell, A. Hereid, y A. D. Ames, "Speed Regulation in 3D Robotic Walking through Motion Transitions between Human-Inspired Partial Hybrid Zero Dynamics," en *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Karlsruhe, Alemania, 2013, pp. 2314–2321. Disponible en: [IEEE Xplore].

[4] M. H. Raibert, *Legged Robots That Balance*. Cambridge, MA, EE. UU.: MIT Press, 1986. Disponible en: [MIT Press].

[5] De Miguel Pueyo, P. (2024). ESTIMACIÓN DE ÁNGULOS DEL TREN INFERIOR MEDIANTE IMUS Y FILTRO MADGWICK PARA SU APLICACIÓN EN EXOESQUELETOS PARA REHABILITACIÓN. Info:eu-repo/semantics/studentthesis. https://hdl.handle.net/10115/39044

[6] Valdivieso, C., García, J., & Sánchez Castro, X. (2011). Aplicación con inclinómetro y acelerómetro para mantener control de movimiento de un robot.

[7] Murillo Herrera, S. (2019). Robot caminador bipedo semi-pasivo como plataforma educativa de bajo costo. Uniandes. Disponible en: https://hdl.handle.net/1992/44230

[8] Alvarado Carlos, M y Vilca Pari, A. (2022). Diseño de robot móvil para la detección del uso de mascarillas y sensado de temperatura corporal utilizando visión artificial en lenguaje Python y Raspberry Pi. Universidad Ricardo Palma - URP. Disponible en: https://hdl.handle.net/20.500.14138/5922

[9] López Valero, G. (2020). *Diseño, estudio y simulación de un robot bípedo en CoppeliaSim (V-REP) y construcción de un prototipo en impresión 3D* (Doctoral dissertation, Universitat Politècnica de València).

[10] Zúñiga Rodríguez, S., & Carmona Muñoz, J. E. (2025). (BIPRO) bipedal intelligent player robot.

EL PROMPT COMO UNIDAD CRÍTICA DE ACCESIBILIDAD EN LA INTERACCIÓN CON INTELIGENCIA ARTIFICIAL CONVERSACIONAL EN EDUCACIÓN SUPERIOR

María Andrea Guisen[1,2] [0000-0001-8704-0897], Luciano Nahuel Giorgi[1] [0009-0000-9121-4415], Pamela Estefanía Acosta[1] [0009-0008-6737-8119], Lisandro López Serra[1] [0009-0008-6176-684X], Alejandro Sartorio[1] [0000-0002-0735-1810], Claudia Pons[1] [0000-0003-1149-0976]

[1]Centro de Altos Estudios en Tecnología Informática

Facultad de Ingeniería en Sistemas Informáticos – Universidad Abierta Interamericana (UAI), Argentina

{mariaandrea.guisen, alejandro.sartorio, claudia.pons}@uai.edu.ar

{lucianonahuel.giorgi, pamelaestefania.acosta, Lisandro.LopezSerra}@alumnos.uai.edu.ar

[2]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

RESUMEN

La incorporación de Inteligencia Artificial (IA) conversacional en la educación superior ha redefinido la Interacción Humano–Computadora (HCI), consolidando al lenguaje natural como interfaz predominante entre los actores sociales del contexto académico y los Modelos de Lenguaje a Gran Escala (LLMs). En este escenario, el *prompt* se configura como unidad técnica crítica, dado que su formulación incide directamente en la precisión semántica, la carga cognitiva y las condiciones de acceso a la información generada. La presente investigación se orienta a construir una taxonomía de *prompts* que permita caracterizar sus dimensiones estructurales, funcionales, contextuales y expresivas en entornos académicos universitarios. Desde un enfoque analítico y aplicado, se elaboró una matriz conceptual preliminar que integra prompting y accesibilidad como dimensiones sociotécnicas interdependientes. El principal resultado consiste en la formulación de un modelo conceptual sistemático, con potencial de validación empírica y proyección científica, orientado a fundamentar estrategias de interacción más comprensibles y técnicamente robustas en educación superior.

Palabras clave: accesibilidad; inteligencia artificial (IA) conversacional; *prompting*; Interacción Humano–Computadora (HCI); educación superior.

CONTEXTO

El trabajo se inscribe en la línea de investigación institucional del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI) denominada "Sociedad del conocimiento y tecnologías aplicadas a la educación".

En el marco de esta línea de trabajo se desarrolla el estudio "Taxonomía de prompting para la optimización de la accesibilidad en sistemas de inteligencia artificial conversacional". La investigación propone problematizar la Interacción Humano–Computadora mediada por sistemas de Inteligencia Artificial (IA) conversacional, incorporando la accesibilidad como dimensión estructurante del análisis y avanzando en la construcción de un marco conceptual que permita caracterizar el *prompt* como unidad central de la interacción lingüística entre personas y modelos de lenguaje.

A partir de este punto de partida, el proyecto se orienta hacia una ampliación progresiva de su alcance, incorporando otros contextos académicos y profundizando el análisis de la accesibilidad tanto desde las estrategias de interacción desarrolladas por los usuarios como desde las funcionalidades y características de diseño que ofrecen las aplicaciones de IA conversacional,

considerando sus posibilidades y limitaciones en términos de accesibilidad.

1. INTRODUCCIÓN

La integración de la IA en la educación superior ha transformado las dinámicas de acceso y construcción del conocimiento [1]. En este escenario, la interacción humano–computadora experimenta un cambio de paradigma: el desplazamiento desde interfaces gráficas tradicionales hacia interacciones mediadas por lenguaje natural. Este cambio reconfigura competencias digitales y sitúa la formulación verbal y escrita en el centro de prácticas académicas vinculadas a la enseñanza, el aprendizaje y la investigación.

Para comprender este fenómeno, es necesario precisar que la IA generativa refiere a sistemas capaces de producir contenido nuevo —texto, imágenes o código— a partir de patrones aprendidos, mientras que la IA conversacional agrupa específicamente a aquellos modelos diseñados para interactuar con usuarios a través de diálogos fluidos. Estos sistemas se sustentan en Modelos de Lenguaje a Gran Escala (LLMs) basados en arquitecturas *transformer*, que procesan secuencias extensas de texto manteniendo la coherencia semántica [2].

En este ecosistema, el *prompt* se configura como unidad central de la interacción. Más que un comando técnico, constituye un enunciado en lenguaje natural que orienta la generación del modelo. Así, la pertinencia y calidad de los resultados producidos dependen en gran medida de la capacidad de formulación de quien interactúa, como señalan los estudios en ingeniería de *prompts* [3]. En consecuencia, el *prompting* emerge como una competencia estratégica en entornos mediados por IA conversacional.

Esta centralidad del lenguaje introduce una cuestión crítica: si la interacción eficaz requiere habilidades lingüísticas específicas, la disponibilidad tecnológica no asegura por sí misma condiciones equitativas de uso. En consonancia con los principios de Diseño Universal establecidos en la Convención sobre los Derechos de las Personas con Discapacidad (CDPD) [4], la accesibilidad en sistemas de IA implica analizar condiciones de interacción utilizables por la diversidad de perfiles sin requerir ajustes posteriores [5].

Desde esta perspectiva, la accesibilidad en la interacción humano–IA puede leerse en al menos dos planos complementarios: (i) el *input* que provee la persona al sistema, principalmente a través del *prompt*; y (ii) las decisiones de diseño funcional de la herramienta (formatos de respuesta, opciones de adaptación, modos de interacción), que también inciden en las condiciones de acceso. Abordar estas tensiones resulta necesario para evitar que la brecha en la formulación lingüística se traduzca en nuevas formas de desigualdad social en el ámbito educativo [6].

En este marco, y considerando el carácter incipiente del proyecto —iniciado a mediados de 2025—, la presente investigación se enmarca en un enfoque aplicado orientado a generar conocimiento transferible al ámbito académico universitario.

El objetivo general de esta investigación es construir una taxonomía de *prompts* que permita analizar sus dimensiones estructurales, funcionales, contextuales y expresivas, con el propósito de derivar estrategias de *prompting* orientadas a optimizar condiciones de accesibilidad en la interacción con sistemas de IA conversacional en educación superior.

Asimismo, esto sentará las bases para su continuidad en una segunda instancia orientada al análisis del diseño funcional de las plataformas de IA conversacional y su incidencia en las condiciones de accesibilidad, ampliando el foco hacia la configuración integral de la interacción humano–IA.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

• Interacción Humano–Computadora mediada por lenguaje natural. Se analizan escenarios en los que la interacción con sistemas informáticos se realiza mediante lenguaje natural, enmarcándose en el campo de HCI y su aplicación en educación superior.

• El *prompt* como unidad de interacción. Se estudia el prompt como unidad formal que estructura la interacción con modelos de lenguaje y configura los workflows conversacionales.

• Accesibilidad. Se incorpora la accesibilidad como dimensión estructurante del análisis, orientada a identificar condiciones de interacción comprensibles y adaptables a diversos perfiles de usuarios.

• IA conversacional. Se examinan herramientas de IA conversacional considerando sus funcionalidades actuales y su potencial adaptativo en contextos académicos.

3. RESULTADOS OBTENIDOS / ESPERADOS

En la primera fase de la investigación se desarrolló una revisión bibliográfica orientada a la construcción del estado del arte en torno al *prompting* y su relación con la accesibilidad en sistemas de IA conversacional aplicados a la educación superior. Las búsquedas se concentraron en los últimos tres años, priorizando artículos de congresos de la IEEE y revistas indexadas en bases de alto impacto como Scimago. Si bien el relevamiento continúa en proceso, esta primera ronda permitió identificar un núcleo conceptual recurrente: inteligencia artificial generativa y conversacional, arquitecturas *transformers*, *prompting* y accesibilidad.

A partir de este análisis se elaboró una matriz preliminar (Figura 1) que constituye un andamio conceptual que organiza dimensiones, variables e indicadores provisorios para orientar la construcción del instrumento analítico.

La dimensión estructural refiere a la organización interna del *prompt*, incluyendo su longitud, formato, nivel de detalle y presencia de ejemplos o plantillas.

La dimensión funcional examina la finalidad del *prompt* y el nivel de exigencia cognitiva de la tarea solicitada, en diálogo con la taxonomía revisada de Bloom [7], clasificando las demandas en niveles bajo, medio o alto.

La dimensión contextual caracteriza las condiciones situadas de uso —actor, dominio temático, escenario académico y herramienta tecnológica— que inciden en la eficacia de la interacción.

Finalmente, la dimensión expresiva analiza la modalidad de emisión, la complejidad lingüística y la secuencialidad comunicativa, aspectos que condicionan la claridad del intercambio con el modelo.

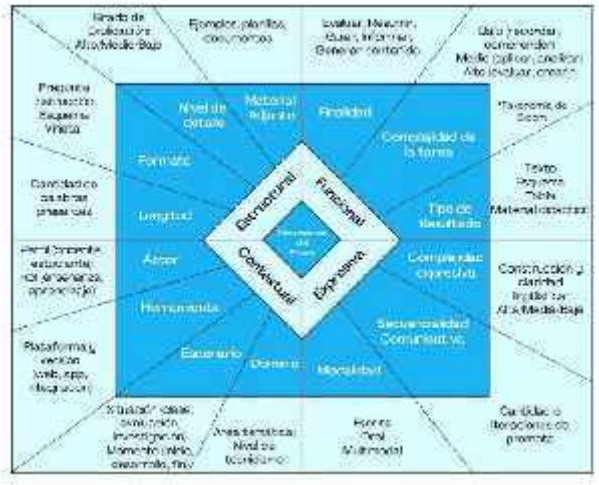


Figura 2. Sintetiza la estructura conceptual propuesta, representando en su núcleo las cuatro dimensiones del *prompt* y, en niveles sucesivos, las variables e indicadores asociados.

Como preparación de la segunda fase —de carácter empírico— se realizaron discusiones internas en modalidad de *focus group* con integrantes del equipo de investigación de la Universidad Abierta Interamericana (UAI), institución que constituye el referente empírico del estudio. Estas instancias permitieron delimitar actores sociales (estudiantes, docentes e investigadores), identificar sus principales tareas académicas (búsqueda, interpretación, sistematización y producción de información, tutoría, evaluación y generación de conocimiento) y reconocer las herramientas de IA de mayor presencia en el ámbito universitario.

Entre ellas se destacan ChatGPT, desarrollada por OpenAI; Perplexity, orientada a la búsqueda contextualizada con referencias; y NotebookLM, desarrollada por Google DeepMind para el análisis guiado de documentos. Asimismo, se identificaron

aplicaciones con funciones integradas de IA, como PDFelement**,** utilizada para lectura y extracción de información en archivos PDF.

Los resultados alcanzados constituyen un insumo exploratorio que orienta el diseño de los instrumentos de recolección de datos y la planificación del trabajo de campo. En términos conceptuales, la matriz preliminar aporta una estructura analítica para caracterizar los *prompts* desde una perspectiva de accesibilidad basada en principios de diseño universal, entendida como capacidad sociotécnica para interactuar eficazmente con sistemas de IA conversacional.

Si bien los resultados obtenidos hasta el momento son de carácter incipiente y corresponden a una etapa exploratoria, han permitido avanzar en la sistematización conceptual del problema y derivar productos académicos concretos.

En efecto, parte de estos avances ya ha sido publicada [8] [9], lo que da cuenta de la relevancia y proyección del enfoque propuesto.

La siguiente etapa se centrará en el análisis empírico de interacciones reales en el ámbito universitario, con el objetivo de refinar la taxonomía propuesta y derivar estrategias de *prompting* que favorezcan usos más comprensibles, situados y equitativos de estas tecnologías en la educación superior.

4. FORMACIÓN DE RECURSOS HUMANOS

La investigación es dirigida por una investigadora del CONICET con sede en el CAETI (Rosario), especialista en accesibilidad comunicacional, que inicia una línea de profundización en IA aplicada a este campo.

El equipo se complementa con el asesoramiento académico de la Dirección del CAETI en Buenos Aires y Rosario, integrada por especialistas con experiencia consolidada en IA, quienes aportan su expertise para su aplicación en el diseño de estrategias orientadas a la accesibilidad.

A su vez, el equipo se integra por tres estudiantes de Ingeniería en Sistemas Informáticos (UAI, Rosario), quienes participan en tareas de análisis, sistematización y exploración de interacciones con IA conversacional, en el marco de una formación temprana en investigación aplicada.

5. BIBLIOGRAFÍA

[1] Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI Literacy and its Implications for Prompt Engineering Strategies. Computers and Education: *Artificial Intelligence*, 5, 100123.

[2] Liu, X., Wang, J., Sun, J., Yuan, X., Dong, G., Di, P., Wang, W., & Wang, D. (2023). Prompting Frameworks for Large Language Models: A Survey. *ACM Computing Surveys*, *1*.

[3] Schulhoff, S., Ilie, M., Balepur, N., Kahadze, K., Liu, A., Si, C., Li, Y., Gupta, A., Han, H., Schulhoff, S., Dulepet, P. S., Vidyadhara, S., Ki, D., Agrawal, S., Pham, C., Kroiz, G., Li, F., Tao, H., Srivastava, A., … Resnik, P. (2024). *The Prompt Report: A Systematic Survey of Prompt Engineering Techniques*.

[4] Naciones Unidas, *Convención sobre los Derechos de las Personas con Discapacidad (CDPD)*, Resolución A/RES/61/106, 13 de diciembre de 2006. [En línea]. Disponible en: https://www.un.org/disabilities/documents/convention/convoptprot-s.pdf. Accedido: Oct. 30, 2025.

[5] Ayala, S. (2023). ChatGPT as a Universal Design for Learning Tool Supporting College Students with Disabilities. *Educational Renaissance*, *12*, 22–41.

[6] Torres Molina, K., & Blanco Pérez, I. (2023). Arquitectura de prompts académicos para el uso de inteligencias artificiales (AI) en Areandina: desafíos, experiencias y obstáculos. *Revista Investigaciones Andina*, 25(46), 62–81.

[7] Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

[8] Giorgi, L. N., Acosta, P. E., López Serra, L., & Guisen, M. A. (2025). Hacia una taxonomía de prompting accesible en la educación superior: Avances preliminares en la construcción de un modelo aplicado. *En Actas del Congreso Iberoamericano en Inteligencia Artificial Aplicada e Innovación (CIIAE 2025)*. Universidad Nacional del Centro de la Provincia de Buenos Aires.

[9] Guisen, M. A., Giorgi, L. N., López Serra, L. & Acosta, P. E. (2025, 29 de diciembre). Taxonomía de prompts para optimizar la accesibilidad académica en sistemas de inteligencia artificial conversacional. *Revista Abierta de Informática Aplicada*, 9(1), 36-51.

ESTRATEGIAS DE LUDOEVALUACIÓN PARA LA OPTIMIZACIÓN DEL APRENDIZAJE Y REDUCCIÓN DE ANSIEDAD EN INGENIERÍA EN SISTEMAS

Brest Federico Nicolás[1]
ORCID 0009-0008-7225-2609
[1]Universidad Abierta Interamericana, Buenos Aires, Argentina
federiconbrest@gmail.com

RESUMEN

Esta línea de Investigación y Desarrollo (I/D) tiene como objetivo principal la formulación y validación de un marco metodológico de evaluación basado en la gamificación para carreras de ingeniería. Ante la problemática persistente de la ansiedad académica en contextos de alta complejidad técnica, la investigación propone la "ludoevaluación" como eje transformador de la acreditación de saberes. A través del estudio de mecánicas de juego aplicadas y la observación del estado de flujo cognitivo, se busca sistematizar estrategias que reduzcan el estrés percibido y potencien la asimilación profunda por sobre la memoria mecánica. Los avances actuales, realizados sobre un universo piloto de 65 alumnos, permiten proyectar la creación de protocolos de evaluación formativa transferibles, orientados a mejorar la satisfacción académica y el rendimiento integral de los estudiantes de sistemas.

Palabras clave: Gamificación, Ludoevaluación, Innovación Pedagógica, Ansiedad Académica, Ingeniería de Software.

CONTEXTO

La presente línea de I/D se inscribe en el programa de investigación educativa de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI). El proyecto está diseñado como un proceso de innovación incremental a implementarse en las cátedras de Análisis y Diseño de Sistemas I, Metodologías Ágiles y Auditoría Informática. La investigación busca establecer estándares de evaluación que se alineen con los nuevos modelos de acreditación por competencias, contando con el apoyo institucional del Departamento de Ingeniería en Sistemas para la futura transferencia de resultados a otras unidades académicas.

1. INTRODUCCIÓN

La evaluación tradicional en Ingeniería en Sistemas suele operar bajo un paradigma de presión que, en muchos casos, inhibe el desempeño real del estudiante. La ansiedad académica resultante no es solo un fenómeno emocional, sino un obstáculo cognitivo: niveles elevados de cortisol afectan la memoria de trabajo y la capacidad de resolución de problemas complejos (Ormazábal et al., 2023).

Esta línea de I/D parte de la premisa de que es posible investigar y diseñar instrumentos evaluativos que, lejos de ser punitivos, actúen como catalizadores del aprendizaje. Se postula que la integración de elementos de diseño de juegos (gamificación) en las instancias de acreditación puede generar un entorno de "seguridad psicológica" que favorezca el rendimiento (Calbacho Contreras, 2022). La investigación se propone no solo aplicar herramientas, sino estudiar cómo estas modifican la percepción del alumno sobre su propio proceso de aprendizaje, transformando la evaluación en una "comunidad de aprendizaje" (Borjas et al., 2019).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de I/D se articula en torno a tres ejes fundamentales de investigación activa:

2.1. Modelado de Evaluación Basada en Desafíos (EBD)

Se investiga el desarrollo de un sistema de evaluación híbrido. El objetivo es determinar la proporción óptima entre la evaluación técnica tradicional y el componente lúdico (Jeopardy, Ta-Te-Ti). Se busca establecer un modelo donde el 60% de la nota se obtenga mediante dinámicas que exijan toma de decisiones bajo presión controlada y colaboración, analizando si esta estructura favorece la retención conceptual a largo plazo.

2.2. Implementación de Micro-Evaluación Continua y Feedback Inmediato

Este eje se centra en el estudio de herramientas ludotecnológicas (como Kahoot!) para la evaluación puntual. La investigación busca validar si el feedback instantáneo reduce la brecha entre el error y la corrección, permitiendo que el estudiante autorregule su aprendizaje de manera dinámica. Se proyecta la creación de un banco de reactivos gamificados para materias del núcleo técnico.

2.3. Neuroeducación y Diseño de Experiencias de Evaluación

Se investiga el impacto de la gamificación en la motivación intrínseca y la reducción de la carga cognitiva extrínseca. El foco está puesto en cómo el diseño de niveles, las recompensas simbólicas y la competencia sana pueden inducir el "estado de flujo" (flow), permitiendo que el estudiante se concentre en la resolución de problemas técnicos sin la interferencia del estrés evaluativo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Aunque la línea de investigación es de carácter continuo, los avances preliminares permiten validar la dirección del proyecto:

3.1 Avances realizados:

Se ha completado una fase de validación con una muestra de 65 alumnos, donde se observó una preferencia metodológica superior al 90%. Los datos indican que la segmentación de la evaluación redujo significativamente la percepción de fatiga mental. Estos resultados sirven como línea base para la expansión de la investigación.

3.2 Resultados Cualitativos y Observacionales

Se evaluó la percepción de los alumnos sobre el uso de Kahoot! al inicio de cada clase mediante una encuesta de tipo cerrada, en la cual se debía seleccionar entre 5 opciones posibles (Estudié para responderlo, Jugaba a hacer "ta-te-tí", Me divertía competir con mis compañeros/as, Era una pérdida de tiempo, Me servía de repaso).

Esta herramienta, entendida como "evaluación puntual" (Conde de Felipe et al., 2020), no fue percibida como una carga, sino como un motor de competencia y repaso. Los datos reflejan que la mayoría de los estudiantes oscilaron entre la motivación por la competencia y la utilidad del repaso conceptual. La figura 1 ilustra los resultados obtenidos.

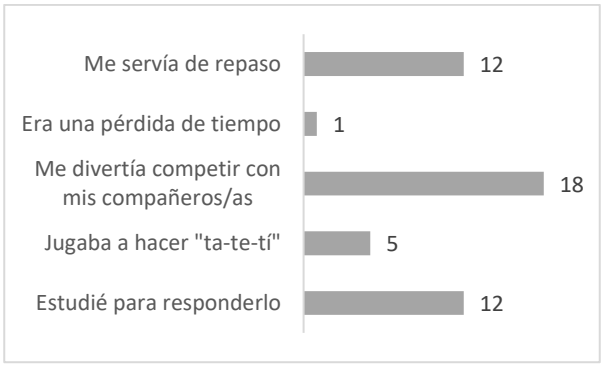


Figura 1. Percepción del alumnado respecto al uso de Kahoot

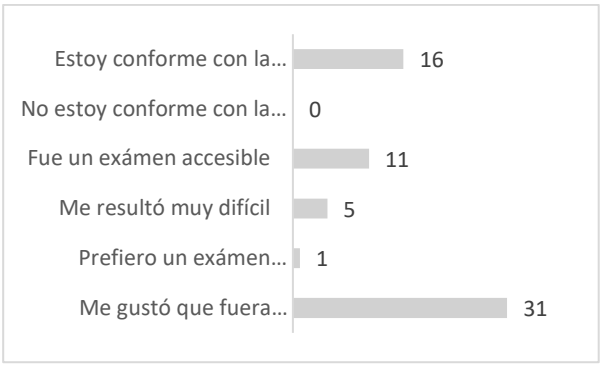


Figura 2. Percepción del alumnado respecto a evaluaciones semiestructuradas

3.2.1 Valoración de la Metodología de Exámenes Semiestructurados

Se indagó sobre la estructura del examen físico (que combinaba V/F, unir con flechas y juegos como Jeopardy o Ta-Te-Ti) mediante una encuesta cerrada con 6 opciones diveras: Me gustó que fuera semiestructurado (V/F, Unir con flechas, responder, dibujar, etc), Prefiero un exámen estructurado tradicional (todo responder), Me resultó muy difícil, Fue un exámen accesible, No estoy conforme con la calificación obtenida, Estoy conforme con la calificación obtenida

Los resultados muestran una inclinación masiva hacia formatos diversos que reducen la ansiedad del "papel en blanco" y permiten demostrar competencias desde diferentes ángulos cognitivos. La figura 2 ilustra la conformidad con la calificación obtenida, la cual refuerza la idea de que la ludoevaluación (Borjas et al., 2019) es percibida como un proceso justo y transparente.

3.3 Resultados y Objetivos en Curso

- **Desarrollo de un Marco de Evaluación**: Se busca formalizar la "Guía de Ludoevaluación para Ingeniería", un documento metodológico que estandarice las mecánicas de juego aplicables a la informática.
- **Análisis Comparativo:** Se proyecta la realización de estudios comparativos entre grupos que utilicen ludoevaluación frente a grupos con evaluación tradicional, midiendo no solo la nota final sino la persistencia del conocimiento en el tiempo.
- **Transferencia Educativa:** El objetivo final es la creación de un repositorio compartido de dinámicas de juego para docentes de la Facultad de Tecnología Informática.

4. FORMACIÓN DE RECURSOS HUMANOS

Como pilar fundamental de esta línea de I/D en la UAI, se prioriza la formación de nuevos investigadores y la actualización docente:

1. **Cuerpo Docente:** Se planea realizar talleres internos para que otros integrantes de la cátedra participen en

la co-creación de nuevos instrumentos evaluativos, promoviendo la discusión abierta entre pares.

2. **Participación Estudiantil:** Los alumnos de las cátedras actuarán como co-creadores de los contenidos de los juegos, fomentando la capacidad crítica.

3. **Liderazgo:** Se fomenta que los alumnos de años superiores participen como "Game Masters" o moderadores de las dinámicas en los años inferiores, desarrollando habilidades blandas y de gestión.

5. BIBLIOGRAFÍA

Borjas, M. P., et al. (2019). Experiencias ludoevaluativas en el contexto universitario: la evaluación desde una comunidad de aprendizaje. *Revista Investigación, Desarrollo e Innovación*, 10(1), 177-190.

Calbacho Contreras, V. P. (2022). *Gamificación como metodología de la enseñanza y el aprendizaje para el fomento de la motivación, la satisfacción y el mejoramiento del rendimiento académico* (Tesis Doctoral). UCC, Argentina.

Conde de Felipe, M., et al. (2020). *Evaluación puntual del aprendizaje mediante una técnica de gamificación*. VII Jornadas Iberoamericanas de Innovación Educativa.

Ormazábal, V., et al. (2023). El juego como herramienta de aprendizaje en educación superior. *Revista Electrónica de Investigación Educativa*, 25, e28.

Experiencias educativas innovadoras con tecnologías emergentes

Sanz Cecilia[1,2][0000-0002-9471-0008], Artola Verónica[1][0000-0003-0833-1077], Salazar Mesía Natalí[1][0000-0003-3946-8383], Iglesias Luciano[1][10000-0002-3284-6997], Buffarini Abril[1], Ibañez Bárbara[1], Bruno Laureano[1], Zeballos Matías[1], Del Gener Aldana[1][0009-0008-0024-3703], Segovia Maite[1], Nordio Mauricio[1], Astudillo Gustavo[3], Baldassarri Sandra[4][0000-0002-9315-6391]

[1]Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC.
Facultad de Informática – Universidad Nacional de La Plata
[2] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
[3] Facultad de Ciencias Exactas y Naturales - Universidad Nacional de La Pampa
[4]GIGA AffectiveLab, Universidad de Zaragoza, España
{csanz, vartola, nsalazar}@lidi.info.unlp.edu.ar,
li@info.unlp.edu.ar, abrilbuffarini@gmail.com, laureanobruno@alu.ing.unlp.edu.ar, matiasmzeb@gmail.com,
aldana.gamepad@gmail.com, maitesegovia.a@gmail.com, nordio.muricio@gmail.com,
astudillo@exactas.unlpam.edu.ar, sandra@unizar.es

RESUMEN

Este trabajo detalla los resultados alcanzados en la línea de investigación, desarrollo e innovación relacionada con el estudio de paradigmas de interacción persona – computadora y sus potencialidades para el escenario educativo. Esta línea se corresponde con el subproyecto "Creación de tecnologías digitales para el escenario educativo", que forma parte de uno de los proyectos acreditados del Instituto de Investigación en Informática LIDI (III-LIDI). Se describen avances en la investigación, así como procesos de creación de experiencias educativas innovadoras que aplican diferentes paradigmas de interacción. Se destacan los resultados en relación al uso de técnicas de captura de posicionamiento para crear experiencias de realidad aumentada. Se valora el trabajo interdisciplinario y la transferencia lograda a distintos espacios educativos, que permiten dar cuenta de una propuesta de investigación-acción e impacto.

Palabras clave: interacción persona-ordenador, interacción tangible, realidad aumentada, realidad virtual, entornos inmersivos, inteligencia artificial, computación afectiva, escenarios educativos, juegos serios.

CONTEXTO

Esta línea de I+D+i se desarrolla en el marco de un subproyecto del proyecto acreditado del Instituto de Investigación en Informática LIDI (III-LIDI) de la Facultad de Informática de la UNLP, titulado: "Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital". El subproyecto mencionado es el SP3: "Creación de tecnologías digitales para el escenario educativo". En este contexto se estudian diferentes paradigmas de interacción persona-ordenador y sus potencialidades para crear experiencias educativas innovadoras. En el subproyecto participan estudiantes, becarios, investigadores en formación y formados para llevar adelante las tareas propias de esta línea.

1. INTRODUCCIÓN

En la actualidad se ha incrementado el uso de diferentes tecnologías emergentes en escenarios educativos. Autores como Abrahamson et al. (2021), citado por Walkington et al. (2026), indican que participar en acciones físicas, sensoriales y simuladas impacta en el razonamiento conceptual. De este modo, no solo los pensamientos conducen a acciones, sino que la realización de acciones puede conducir a ideas novedosas, y a la reflexión (Nathan, 2022). Bajo esta consideración estudiar diferentes paradigmas de interacción persona – ordenador que promuevan los movimientos físicos, las acciones sobre el mundo y el aumento del escenario real con simulaciones y el agregado de información digital, se convierten en elementos centrales de la cognición y el aprendizaje (Candanedo Yau, 2026). Se definen las acciones como movimientos físicos utilizados para interactuar con algo del entorno o manipularlo. Los gestos, sin embargo, son actos comunicativos o generadores de significado que a menudo se producen junto con el habla y transmiten información espacial o relacional importante. Tanto los gestos como las acciones funcionan para descargar los procesos cognitivos en el cuerpo o en el mundo (Castro-Alonso et al., 2024), citado en Walkington et al. (2026). En el subproyecto mencionado del III-LIDI se vienen estudiando diferentes paradigmas de interacción con computadoras, y se crean diferentes tecnologías y entornos educativos inmersivos, aumentados (realidad aumentada –RA) (Artola, Sanz & Ibañez, 2024; Salazar & Sanz, 2023), y que promueven las acciones de los participantes con el fin de abordar objetivos concretos que fomenten la reflexión y el aprendizaje (Cabero-Almenara, Valencia-Ortiz, R., & Llorente-Cejudo, 2022). Al mismo tiempo, se desarrollan experiencias inmersivas vinculadas al paradigma de realidad virtual

(RV) para promover la simulación en entornos creados sintéticamente, pero válidos para acercar determinados aprendizajes (Martínez-Requejo, Lores-Gómez, y Ruiz-Lázaro, 2024). En este caso, se busca crear un entorno interactivo y que promueva la sensación de presencia para lograr la simulación deseada y que las acciones resulten significativas (Martínez-Requejo, S., Lores-Gómez, B., & Ruiz-Lázaro, 2024, citando a Lannutti, 2022).

En relación a las acciones en el mundo, también se viene trabajando con el paradigma de interacción tangible (IT) en el que se aborda el uso de objetos físicos para promover eventos en un sistema digital (Sanz & Artola, 2025; Hulusic et al., 2023). Estos objetos pueden funcionar tanto como dispositivos de entrada (el usuario los usa para generar acciones) como de salida (el sistema digital muestra acciones con o sobre los objetos). Esto último puede lograrse tanto con mapeo de proyecciones o con componentes electrónicos (actuadores) que permiten que el objeto se vuelva activo (Artola, Sanz & Baldassarri, 2023).

Cabe destacar que también, en los últimos años, se ha profundizado en el área de computación afectiva y de inteligencia artificial para enriquecer entornos educativos que combinen el mundo físico y digital. La computación afectiva aborda técnicas que posibilitan la identificación de emociones en los estudiantes y docentes, lo que puede ser beneficiosos en educación. En particular, en esta línea de investigación del Instituto, se ha avanzado con la detección de estrés y de aburrimiento, dos emociones que se relacionan de manera directa con el escenario educativo.

Para llevar adelante estas líneas temáticas se abordan metodologías de investigación –acción, revisiones sistemáticas de literatura, se prueban técnicas y herramientas que posibiliten el avance en cada línea, y se desarrollan y crean tecnologías específicas para el abordaje de necesidades educativas concretas. En el subproyecto se trabaja también en el diseño y desarrollo de juegos serios educativos que integran algunos de los paradigmas de interacción aquí mencionados. Los juegos serios resultan un área de investigación actual, con diferentes beneficios reportados en diversas áreas (Gubaro et al., 2024; Ding & Yu, 2024; Archuby, Sanz & Manresa-Yee, 2023). En el subproyecto, se han abordado diferentes tesis enfocadas en esta temática y se crean juegos con objetivos educativos, y otros enfocados en promover la actividad física (*exergames*).

Las líneas de investigación aquí presentadas guardan relación directa con las temáticas de la Maestría y Especialización en Tecnología Informática Aplicada en Educación de la misma Facultad, por lo que se desarrollan tesis y trabajos finales en estos temas. En 2025, se abrió una Diplomatura en Tecnologías digitales en escenarios educativos. También se cuenta con una agenda de cooperación a nivel nacional e internacional, que permite profundizar el trabajo que se lleva adelante en el III-LIDI.

2. LÍNEAS DE INVESTIGACIÓN / DESARROLLO

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Paradigmas de Interacción Persona-Ordenador en escenarios educativos.
- Diseño de actividades educativas basadas en diferentes paradigmas. Las posibilidades de la Realidad Aumentada, la Interacción Tangible, la Realidad Virtual en la configuración de actividades educativas.
- La Inteligencia artificial generativa en el escenario educativo.
- Técnicas con hologramas y mapeo de proyecciones para crear entornos aumentados e interactivos.
- Juegos serios como actividades educativas. Multisensorialidad a través de juegos y actividades educativas basadas en diferentes paradigmas de interacción.
- Juegos móviles usando aprendizaje situado y basado en la experiencia.
- *Exergames* para promover la actividad física.
- Diseño de Objetos Activos para IT. Tipos de *feedback* en objetos activos. Uso de actuadores y sensores.
- Entornos inmersivos e interactivos basados en realidad virtual y aumentada. *Embodied Interaction.*
- Sistemas educativos adaptativos y sistemas recomendadores para recursos educativos.
- Computación afectiva en entornos digitales para el escenario educativo.
- Salas de escape educativas.

Estas líneas se basan en el trabajo conjunto de docentes investigadores, becarios, tesistas, pasantes y estudiantes que participan. Al mismo, tiempo se aborda en muchos de estos proyectos un diseño participativo con docentes y estudiantes de diferentes niveles educativos que intervienen con propuestas y en el diseño.

3. RESULTADOS OBTENIDOS

Se presentan aquí los resultados alcanzados durante el año 2025 y los avances en esta línea de I+D+i. A continuación, se presentan los resultados organizados por temáticas.

3.1 Experiencias educativas basadas en IT, RA y RV

En 2024 se trabajó en el desarrollo de narrativas digitales interactivas orientadas a fomentar aprendizajes específicos. Se abordó un proyecto llamado el “reflector mágico” (Sanz et al., 2025), que consiste en un dispositivo creado ad-hoc, con un sensor acelerómetro y giroscopio. Este dispositivo simula ser un reflector, el cual, a partir de su movimiento, determina un área del entorno sobre la que se realiza una proyección. De esta manera, se genera una experiencia de entorno aumentado. Este proyecto es fruto de la investigación en los temas de RA y entornos interactivos. Desde el punto de vista educativo, se aprovechó esta tecnología para presentar personajes relacionados con el libro de “El

Principito" y reflexionar con estudiantes sobre las características de estos personajes (Figura 1). Se utilizó en diferentes exposiciones, y en talleres en el Centro de Innovación y Transferencia Tecnológica (CIyTT) de la Facultad de Informática. En 2025, se profundizó esta línea de trabajo mediante la creación de un nuevo dispositivo, llamado "linterna interactiva". La linterna cuenta con un *tracker* que trabaja en conjunto con una estación base, para detectar la posición del objeto linterna, y a partir de ésta realizar acciones situadas en el entorno (en este caso, mostrar determinadas proyecciones en una posición específica del entorno físico). Esta tecnología presenta mejores posibilidades para establecer una ubicación en un espacio físico cerrado, que las desarrolladas en el reflector mágico. Como resultado, con la linterna se pueden realizar movimientos más libres. Esta tecnología se aprovechó para extender un proyecto (TicTec) de mapeo de proyecciones, que ya se había creado en el marco de esta línea de investigación. TicTec muestra en un reloj en la pared la evolución en el tiempo de diferentes tecnologías que han aportado a la disciplina Informática. Se ilumina cada hora y se muestra una pequeña imagen y video sobre esa hora que representa un hito en la historia de la Informática. Con la incorporación de la linterna, se permite enfocar una hora determinada (Figura 2 A) y encontrar una animación que extiende la narrativa. De esta manera, el usuario participa de manera activa en la experiencia eligiendo dónde enfocar[1]. Este proyecto fue presentado en la Noche de los Museos 2025.

Al mismo tiempo, durante 2025 se desarrolló un cuadro interactivo utilizando un sensor de Kinect y modelos IAG para presentar a Ada Lovelace en el marco de la Sala de mujeres innovadoras de la Informática del CIyTT (Figura 2 B). El proyecto forma parte de las actividades con estudiantes en formación de la Facultad, y posibilitó explorar nuevos modelos de interacción. En este caso se utiliza la detección de las manos y gestos con éstas para provocar que el personaje de Ada Lovelace "despierte" en el cuadro e interactúe de diversas formas con el público, dando información sobre sus aportes. Así se relaciona con la temática de entornos interactivos para el aprendizaje.

Se continuaron realizando talleres con juegos de RV para acercar el patrimonio argentino, con estudiantes de escuelas. En particular, se desarrollaron talleres con las Casas del Niño de la Municipalidad de La Plata. Se obtuvieron resultados de interés para el proyecto a partir de la observación de prácticas concretas. Estos retroalimentan el diseño de otros proyectos. Por ejemplo, se está iniciando en el marco de una maestría, el desarrollo de un entorno de RV para abordar temas de cuidado del medioambiente (Lanata, Sanz & Rucci, 2024). También se está trabajando en una propuesta de tesis doctoral vinculada a los temas de RV.

[1] Enlace a video del prototipo de la linterna: https://drive.google.com/file/d/1DYoFqP94JrbaxsCiJrbovX2eqK6HYYsF/view?usp=drive_link

Además, en 2025 se creó un juego serio (llamado "Aventura Sin TACCtica") vinculado al tema de alimentación saludable y celiaquía. El juego, a través del movimiento de las manos, propone a los jugadores seleccionar entre alimentos proyectados en pantalla (Figura 3), aquellos que resulten más saludables y aptos para personas con celiaquía. La propuesta se estructura en tres etapas: (1) identificación de alimentos con y sin TACC; (2) reconocimiento de alimentos sin TACC nutritivos; y (3) resolución de preguntas sobre contaminación cruzada, orientadas a concientizar acerca de prácticas seguras en la manipulación de alimentos, utensilios y espacios. El juego ofrece retroalimentación inmediata sobre las decisiones y su clasificación. El desarrollo se realizó con JavaScript, empleando *TensorFlow.js* y modelos para el reconocimiento de manos y poses corporales. En el desarrollo participaron dos estudiantes y una nutricionista, quien aportó al diseño y elaboración de contenidos, los cuales fueron trabajados con estudiantes de prácticas profesionales de la carrera de Nutrición. Desde el punto de vista del subproyecto se destaca la formación de estudiantes, y la investigación de diferentes técnicas que permiten abordar este tipo de desarrollos.



Figura 1 – Estudiantes interactuando con el Reflector mágico, y viendo un personaje del Principito



Figura 2 – A-Un usuario utilizando la Linterna Mágica. B-Niño interactuando con el cuadro de Ada

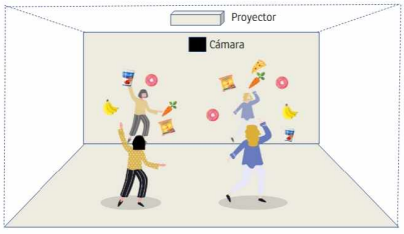


Figura 3 – Esquema del juego Aventura Sin TACCtica

Se destaca que se llevaron a cabo en estos temas vinculaciones con otros ámbitos de la universidad. En

este sentido se desarrolló un proyecto interactivo para el Museo de Ciencias Naturales de la UNLP, en conjunto con estudiantes de Informática, y docentes investigadores de la Facultad de Ciencias Naturales y Museo. El proyecto muestra posibles impactos del cambio climático en la vida de los invertebrados en la Antártida[2]. A través de un panel interactivo, se ven proyectadas en una instalación física diferentes imágenes y videos de posibles escenarios futuros en contraste con el presente (Figura 4).



Figura 4 – Instalación interactiva para aprender sobre posibles impactos del cambio climático

También se desarrolló una instalación interactiva que se orienta al tema transformación de energía para la muestra de Universo Energía del Centro de Artes de la UNLP. Se abordó una simulación de crecimiento de una planta a partir de energía radiante. Se utilizaron técnicas de mapeo de proyecciones a partir de la detección de una fuente lumínica utilizando sensores y un microcontrolador.

3.2 Sala de Escape Educativa

Durante 2024 se desarrolló una sala de escape educativa. Estos espacios permiten abordar dinámicas de interés para la educación, ya que se basan en la colaboración y la negociación entre los participantes para lograr resolver los desafíos. Al mismo tiempo, posibilitan el abordaje de un objetivo educativo concreto combinado con dinámicas lúdicas y de resolución de problemas, lo que despierta la motivación y se logra que el estudiante se involucre activamente. El objetivo educativo de la primera sala se centró en los temas de alimentación saludable y nutrición con una serie de desafíos para resolver en un tiempo máximo de 15 minutos. La sala incluye diferentes dispositivos interactivos que fueron creados en el marco del subproyecto y en cooperación con estudiantes. Además, la sala de escape se utilizó para un experimento vinculado a las investigaciones del área de computación afectiva para la detección de estrés en las personas que están jugando. Esto formó parte de una práctica profesional supervisada ya finalizada en 2025 (Zeballos & Sanz, 2025).
En 2025 se creó una segunda sala de escape con el fin de abordar la temática de identificación de emociones, practicar la escucha y atención del otro. Esta segunda sala se desarrolló durante todo el año con pruebas piloto y quedó disponible de forma completa en diciembre de 2025. Incluye una mesa interactiva sobre la que transcurren parte de los desafíos a resolver; los participantes interactúan a través de objetos del entorno con una aplicación que se ejecuta sobre esta mesa (Figura 4). El objetivo del juego es provocar las 6 emociones básicas en un personaje llamado Augutsen, que quiere destruir a los humanos, ya que están destruyendo el planeta Tierra. Para lograrlo es necesario poner en juego la empatía, la identificación de emociones y la escucha.



Figura 5 – Esquema del juego de la sala de escape

La sala es resultado de la investigación en las temáticas que aquí se exponen, recupera parte del marco teórico y técnico sobre interacción tangible, entornos interactivos y diseño de juegos serios educativos. Se trabajó con dos estudiantes de las carreras en Informática y una estudiante de Psicología que aportó inicialmente en el proyecto.

3.3 Proyectos interactivos con estudiantes de escuela media

En 2025 se continuó con el proyecto llamado “Semillero de innovadores” en el contexto de las actividades del CIyTT, destinado a estudiantes de escuela media (Sanz et al., 2024b). En el primer semestre se trabajó con 10 estudiantes de distintas escuelas que crearon aplicaciones móviles para controlar una mascota interactiva y dos maquetas de edificios. En el segundo semestre se repitió la experiencia con 10 estudiantes de escuela técnica. Los proyectos involucran formación a dos niveles: 1) estudiantes de los primeros años de las carreras de la Facultad quienes realizan prototipos y luego asumen roles de tutoría; 2) los estudiantes de escuelas secundarias que participan en las experiencias diseñadas. Al finalizar, se realiza una muestra de proyectos, abierta a la comunidad.

3.5 Otros avances en las investigaciones

Se profundizó en los temas de investigación con revisiones sistemáticas y análisis de diferentes técnicas. Se finalizó una tesis doctoral sobre la gamificación en entornos virtuales que cuenta con una dirección compartida con profesores de la Universidad de Zaragoza y la UNLP (Vera Mora, 2025). Se continúa

2 https://www.instagram.com/reel/DTThg0Uiqeu/?igsh=Y2FxaTUwY2hhcDho

con una tesis doctoral, desarrollada en co-tutela con la Universidad de Zaragoza, enfocada en la detección de la emoción de aburrimiento mientras se visualizan videos, en particular educativos (Astudillo et al., 2025). Se avanzó en la tesis de maestría en temas de realidad virtual (Lanata, Sanz & Rucci, 2024), que se vincula con el programa de extensión Huellas patrimoniales[3] (en el que se participa) (Elizalde, Possemato & Sanz, 2025).

En el proyecto se investiga sobre *exergames* y se viene trabajando en técnicas para la detección de movimientos y gestos del jugador. Se han desarrollado diferentes dispositivos, y se llevan a cabo experimentos analizando la performance. En 2025 se logró concretar un dispositivo guante, para detectar movimientos de la mano que representan acciones en el juego (Herrera Del Gener & Sanz, 2025; Del Gener, Sanz & Iglesias, 2023). Cabe aclarar, que todos estos resultados se desarrollaron en vinculación con las líneas aquí presentadas y los juegos y experiencias están disponibles para quienes participan de las actividades del CIyTT.

En cuanto a los proyectos vinculados y los acuerdos de cooperación, el III- LIDI participa en las siguientes acciones de colaboración:

- Se cuenta con un acuerdo de colaboración con la Universidad de Zaragoza y la Universidad de Islas Baleares.
- Se ha obtenido financiamiento en el marco de una convocatoria K171 de Erasmus+ para movilidades entre la Universidad de Zaragoza y la Universidad Nacional de La Plata en 2024. Se ha participado del Congreso de Unita Geminae 2025 en Zaragoza con una exposición y se ha ganado un premio de "Best Practice Award" en este marco con el proyecto de juegos serios vinculados a salud.
- La Dra. Sanz es miembro colaborador del grupo de Investigación en Interfaces Avanzadas (AffectiveLab).
- Se asesora en el proyecto de la Universidad de Costa Rica: "Juegos serios basados en interfaces tangibles con objetos activos" con código: 510-C2-321
- Se participa en la RedAUTI: Red temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva.
- Se participa del proyecto TEMOR: Ecosistema tecnológico para el reconocimiento del estado de ánimo de los pacientes durante el proceso de rehabilitación cardíaca. Proyecto de la Universidad de Zaragoza, aprobado en 2022.
- Se participa del proyecto: Experiencias Lúdicas con Agentes Sociales Interactivos y Robots: Aprendizaje Social y Comunicación Intergeneracional (PLEISAR) y su continuación. Dirigido por la Dra. Cerezo de la Universidad de Zaragoza.
- Se cuenta con los siguientes acuerdos de cooperación 2025: 1-con el Museo de Ciencias Naturales de La Plata; 2- Instituto Privado De Neuromedicina S.A. (IPNM) con quienes se trabaja en aplicaciones interactivas para adultos mayores.

4. FORMACION DE RRHH

Se trabaja en la formación de recursos humanos a través de la dirección de tesis de doctorado, maestría, trabajos de especialización y tesinas de grado; también con becas de diferentes organismos de ciencia y técnica. En 2025, se ha completado en estas líneas 1 tesis doctoral, y otra en el marco del proyecto. Se entregó una tesis de maestría vinculada a videojuegos serios, que se encuentra en evaluación. Se terminó un trabajo de Especialización en TIAE sobre juegos serios (Gubaro, 2025). Se trabajó en la formación de 6 estudiantes de grado. Tres de los estudiantes de grado que colaboran en el proyecto alcanzaron su título de Ingenieros en Computación.

5. BIBLIOGRAFÍA

- Abrahamson, D., & Abdu, R. (2021). Towards an ecological-dynamics design framework for embodied-interaction conceptual learning: The case of dynamic mathematics environments. Educational Technology Research & Development, 69(4), 1889–1923. https://doi.org/10.1007/s11423-020-09805-1
- Artola, V., Sanz, C. V., & Ibañez, B. (2024). EmpoderAR: un juego serio de realidad aumentada para acercar a jóvenes a la ciencia. Interacción, Revista de la Asociación Interacción Persona-Ordenador (AIPO), *5*(1), 6–13. https://revista.aipo.es/index.php/INTERACCION/article/view/107
- Archuby, F., Sanz, C., & Manresa-Yee, C. (2023). DIJS: Methodology for the Design and Development of Digital Educational Serious Games in IEEE Transactions on Games, doi: 10.1109/TG.2022.3217737. Early access.
- Artola, V., Sanz, C., & Baldassarri, S. (2023). "A Novel Tangible Interaction Authoring Tool for Creating Educational Activities: Analysis of Its Acceptance by Educators" in IEEE Trans. on Learning Technologies, vol. 16, no. 5, pp. 633-647, doi: 10.1109/TLT.2022.3216117
- Astudillo, G. J., Sanz, C. V., & Baldassarri, S. (2025). Emotionally-tagged video database: a systematic review. Data Technologies and Applications, 1-26.
- Cabero-Almenara, J., Valencia-Ortiz, R., & Llorente-Cejudo, C. (2022). Ecosistema de tecnologías emergentes: realidad aumentada, virtual y mixta. Revista Tecnología, Ciencia Y Educación, (23), 7–22. https://doi.org/10.51302/tce.2022.1148
- Candanedo Yau, R. M. (2026). Simbiosis digital: Redefiniendo la interacción persona-computador a través de Inteligencia Artificial como puente y compañero. Scientia, 36(1), 108–125. https://doi.org/10.48204/j.scientia.v36n1.a8433
- Castro-Alonso, J. C., Ayres, P., Zhang, S., De Koning, B. B., & Paas, F. (2024). Research avenues supporting

[3] https://www.huellaspatrimoniales.econo.unlp.edu.ar/

embodied cognition in learning and instruction. Educational Psychology Review, 36(1), 10. https://doi.org/10.1007/s10648-024-09847-4
- Centro de Innovación y Transferencia Tecnológica, FI, UNLP, https:/ciytt.info.unlp.edu.ar/
- Del Gener A., Sanz C., & Iglesias L. (2023). Propuesta de un gamepad para sensar movimientos del jugador y su integración a un exergame. Revista AIPO, 4(2), 68-77.
- Ding, A., Yu, Ch. Serious game-based learning and learning by making games: Types of game-based pedagogies and student gaming hours impact students' science learning outcomes, Computers & Education, Volume 218, 2024, 105075, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2024.105075
- Elizalde, M.P., Possemato, T., y Sanz C. (2025). Huellas Patrimoniales: tecnología digital para la inclusión. Prácticas integrales en la formación universitaria. Simposio Regional sobre Prácticas Universitarias Integrales. Disponible en: https://drive.google.com/file/d/16X2iifvtyWBszaafqRlV8mK-tSdgZopj/view
- Herrera del Gener, A. M., & Sanz, C. V. (2025). *Development of a Hand Motion Sensing Glove for Exergames: Design Evolution and Future Perspectives*. In XIII Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 24 al 26 de junio de 2025).
- Hulusic, V., Gusia, L., Luci. N. & Smith, M. (2023) Tangible user interfaces for enhancing user experience of virtual reality cultural heritage applications for utilization in educational environment. ACM J Comput Cult Herit 16(2):1–24
- Gubaro, M. (2025). Juegos Serios Educativos Digitales Orientados a la Enseñanza de la Matemática en Educación Secundaria. Estado del Arte. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, (41), 166-168.
- Gubaro, M., Sanz, C. V., & Artola, V. (2024). *Juegos Serios Educativos digitales orientados a la enseñanza de la Matemática en educación secundaria*. En *Libro de Actas del 30° Congreso Argentino de Ciencias de la Computación (CACIC 2024)* (pp. 379–389). Universidad Nacional de La Plata. https://doi.org/10.35537/10915/172755
- Lanata, L., Sanz, C. & Rucci, A. (2024). Un viaje a la era del hielo: un recorrido virtual educativo por el Parque Nacional los Glaciares para alumnos del ciclo básico del nivel secundario. Propuesta de tesis de maestría en TIAE.
- Lannutti, E. D. (2022). Robótica: De la ciencia ficción a la realidad científica (Vol. 13). EDIUNC.
- Martínez-Requejo, S., Lores-Gómez, B., & Ruiz-Lázaro, J. (2024). Efectividad de las tecnologías inmersivas para potenciar el aprendizaje en educación superior: una revisión sistemática. Edutec, Revista Electrónica y Tecnología Educativa, (90), 54-73. https://doi.org/10.21556/edutec.2024.90.3391
- Nathan, M. J. (2022). Foundations of embodied learning: A paradigm for education. Routledge.
- Salazar Mesía, N. A., & Sanz, C. V. (2023). Avances en la herramienta de autor AuthorAR para la creación de actividades educativas con geolocalización. In XVIII Congreso Nacional de Tecnología en Educación y Educación en Tecnología-TE&ET 2023.
- Sanz, C. V., & Artola, V. (2025). Experiencias con juegos y aplicaciones educativas usando diferentes paradigmas de IPO con escuelas secundarias. In XX Congreso de Tecnología en Educación y Educación en Tecnología-TE&ET (UNDAV, 12 y 13 de junio de 2025).
- Sanz, C. V., Artola, V., Violini, M. L., Zeballos, M., & Lombardo, S. (2025). Narrativas digitales con el Reflector Mágico. In XX Congreso de Tecnología en Educación y Educación en Tecnología-TE&ET (UNDAV, 12 y 13 de junio de 2025).
- Sanz, C. V., Artola, V., & Baldassarri, S. (2024a). Developing educational serious games as a proposal for university training in the area of Child Computer Interaction. International Journal of Child-Computer Interaction, 41, 100672.
- Sanz, C., Artola, V., Medina, S., Zeballos, M., & Lombardo, S. (2024b). Crear con tecnologías digitales para el desarrollo del pensamiento computacional. El caso del proyecto de semillero de innovadores.
- Vera Mora, G. R. (2025). Propuesta de un modelo de gamificación para entornos virtuales de enseñanza y aprendizaje en Instituciones de Educación Superior y su impacto en la presencia social, cognitiva y docente en dichos entornos (Tesis doctoral). UNLP. https://doi.org/10.35537/10915/179166
- Walkington, C., Washington-Henderson, J., Hunnicutt, J., Nathan, M. The impact of different collaboration formats on mathematical problem-solving in augmented reality, Computers & Education, Volume 242, 2026, 105491, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2025.105491.
- Zeballos, M., & Sanz, C. (2025). Informe final Práctica Profesional Supervisada: "Prototipo funcional de un firmware de registro de sensores fisiológicos para el análisis de emociones. El caso de una sala de escape inmersiva y educativa.

Framework para desarrollos de robótica educativa en espacios simulados

Gonzalo Zabala, Ricardo Moran

Centro de Altos Estudios en Tecnología Informática
Facultad de Tecnología Informática
Universidad Abierta Interamericana
Av. Montes de Oca 745, Ciudad Autónoma de Buenos Aires, República Argentina
(+54 11) 4301-5323; 4301-5240; 4301-5248
{Gonzalo.Zabala, Ricardo.Moran}@uai.edu.ar

RESUMEN

La robótica ha demostrado su utilidad como recurso didáctico y para el desarrollo de las habilidades para el siglo XXI. Sin embargo, es un recurso costoso, y necesita de una renovación y mantenimiento constante. Además, el material físico aún genera temor y rechazo en docentes que no están formados en el área. Por otro lado, el crecimiento de la potencia gráfica permite la creación de simuladores cada vez más realistas. Por lo tanto, los simuladores pueden ser un recurso para la inserción de la robótica educativa en las escuelas, con un costo menor a los materiales físicos, manteniendo similitud con la realidad y alcanzando metas de aprendizaje similares al mundo real.

Palabras clave: robótica educativa, simulación, programación visual, laboratorio virtual.

CONTEXTO

El presente proyecto está radicado en el Centro de Altos Estudios en Tecnología Informática (CAETI), dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana. El mismo se encuentra inserto en la línea de investigación "Sociedad del conocimiento y Tecnologías aplicadas a la Educación". El financiamiento es brindado por la misma Universidad Abierta Interamericana.

1. INTRODUCCIÓN

El uso de robots en las aulas provee a los estudiantes un espacio de aprendizaje reflexivo y experimental. Cuando en equipo diseñan, construyen, programan y documentan el diseño de un robot autónomo, no sólo aprenden de tecnología, sino que también ponen en práctica un conjunto de habilidades fundamentales para el mundo que encontrarán cuando salgan de su educación. Por otra parte, el mercado laboral requiere cada día más de recursos humanos vinculados al desarrollo de tecnología, en campos como la programación, el diseño y las ingenierías. El estudio de la robótica puede ser un recurso para despertar vocaciones en estas áreas, además de aumentar la retención de estudiantes mujeres que están subrepresentadas en el campo de la tecnología.

Por estos motivos, los robots se han convertido en un recurso didáctico valorado por los docentes de nivel primario y secundario. Sin embargo, aún es un recurso costoso para la mayoría de las escuelas y familias, y los docentes presentan algunos temores y rechazos para su incorporación en el aula. Por otro lado, en contextos donde la educación necesariamente debe realizarse en forma remota, como el acaecido por la pandemia del COVID 19, el acceso a laboratorios con material físico se hace imposible.

Asimismo, a partir del aumento de poder de cómputo de las placas gráficas, han surgido con fuerza diferentes ambientes de simulación de robots: CoderZ, Webots, VRT, Gazebo CoppeliaSim, entre otros. Lamentablemente, los que tienen por detrás un motor físico que representa lo más cabalmente posible al mundo real tienen alguno o algunos de los siguientes inconvenientes:

- Son muy costosos para nuestro país.

- Necesitan de un equipamiento medianamente sofisticado por el alto nivel de procesamiento para simular la física.
- Son complejos de instalar y operar.

En síntesis, hasta el momento, no existe aún una solución al problema de la robótica educativa sin material concreto, que presente las siguientes características:

- Acceso gratuito al software de modo que sea totalmente accesible para que lo utilicen los estudiantes.
- Que no posea altos requerimientos de cómputo.
- Facilidad de instalación para que este proceso no se convierta en una barrera de entrada.
- Interfaz de uso intuitivo.
- Versatilidad en el framework propuesto, de modo que pueda utilizarse tanto para la enseñanza de los fundamentos de la robótica como para la experimentación de temas de mayor complejidad, de forma tal que proporcione un espacio para que los estudiantes puedan explorar y probar.

Dejando de lado el primer punto, en el mundo físico ya se había logrado una solución para los siguientes requerimientos, principalmente con el surgimiento de Arduino y similares. Por lo tanto, un framework que permita el desarrollo de plataformas de simulación que cumplan con todas estas características, no sólo sería una herramienta fundamental en situaciones donde es necesaria la educación a distancia, sino que se convertiría en un recurso didáctico de gran utilidad por las siguientes razones:

- Costo bajo o nulo.
- Permite el trabajo con estudiantes que se encuentran alejados del acceso a los recursos materiales.
- Permite el trabajo colaborativo entre estudiantes que se encuentran distanciados entre sí.
- Sirve de entrada al mundo de la robótica educativa de una manera más sencilla que los kits materiales de bajo costo.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Respondiendo a los problemas no resueltos que se describieron en el punto anterior, la solución que se propone es la construcción de un framework para desarrollos online de robótica educativa, basada en un simulador 3D, que permita vivenciar una experiencia lo más similar posible a la realizada con material concreto. Este framework ofrecerá un espacio para la presentación de diferentes propuestas didácticas con diversas situaciones problema en distintos escenarios. Para resolver las situaciones, los estudiantes elegirán la arquitectura de robot adecuada y programarán su comportamiento con un lenguaje de programación visual y/o textual.

Con relación a las características del proyecto, en primer lugar, el uso de un simulador basado en un motor físico 3D se acerca a la experiencia de trabajar con kits físicos. Por otro lado, el costo material es muy bajo comparado con estos recursos físicos. Además, la ejecución en un browser disminuye el requerimiento de equipamiento, lo que baja aún más la necesidad de recursos. Por último, el desarrollo de una interfaz que simplifica el uso del motor 3D, el acompañamiento en la misma plataforma de propuestas didácticas y la incorporación de un lenguaje visual para la programación facilita la enseñanza de robótica en los niveles primario y secundario.

Ahora bien, en función de la simplicidad en el uso de la plataforma, se ha decidido recortar algunas funcionalidades:

- Construcción o importación de un robot no integrado en la plataforma.
- Uso de diversos lenguajes para la programación de los robots.
- Construcción de escenarios por parte del usuario.

De todos modos, dado que la plataforma es libre y de código abierto, están dadas las condiciones para que la comunidad académica y educativa tome el proyecto y lo extienda en

la dirección que desee. Por lo tanto, los límites comentados pueden superarse con trabajos posteriores basados en el desarrollo de esta investigación.

3. RESULTADOS OBTENIDOS/ESPERADOS

En primer término, se definieron los principales aspectos didácticos que debía presentar la implementación del framework. Al mismo tiempo, se hizo un análisis de diferentes simuladores, seleccionado Webots como plataforma para el desarrollo del framework.

Luego de la selección, comenzó la construcción de simulaciones de diferentes competencias tradicionales de robótica. El primer caso fue el simulador de Sumo. En el mismo, dos robots deben enfrentarse en una pista intentando empujar al otro robot fuera del espacio de lucha. Para ello, además del simulador, se construyó un supervisor árbitro, que permite la carga de los dos robots (programación y estructura), lanza la lucha, lleva cuenta del tiempo, reacomoda los robots cuando se produce una situación de traba y determina el ganador ante la salida de la pista. Además, se diseñó un customizador online de los robots que permite el reposicionamiento de sus sensores. Los desafíos encontrados fueron la construcción de una pista con un buen sistema de colisión, y una estructura de robot similar a la que usan los estudiantes en el mundo real, con una malla de colisión compleja que permitiera una simulación realista. El resultado de las primeras pruebas con estudiantes fue tan exitoso, que dio pie a la organización de un evento latinoamericano que se mantiene hasta el día de hoy.

También se adaptó un simulador de fútbol de tres robots contra tres, ya existente, para poder programarlo desde cualquier lenguaje usando sockets para la comunicación con el simulador.

Y por último, a partir del simulador Rescue Sim de la RoboCup, se desarrolló material de capacitación y cursos para docentes, para ingresar en el mundo del rescate simulado. Esta competencia también tuvo mucho éxito, y Argentina comenzó a competir a nivel mundial, obteniendo resultados excelentes en 2024 y 2025.

Además del trabajo con los simuladores, con el objetivo de analizar los beneficios y dificultades del uso de simuladores en el aprendizaje de la robótica, se realizó una experiencia con un grupo de estudiantes de control y otro experimental, comparando el uso de robots físicos con simulados. Para la simulación, se crearon pruebas exactamente iguales a las presentadas en el mundo físico, y se desarrolló un lenguaje de programación exactamente igual para ambos casos. Los desafíos por resolver eran exactamente los mismos.

Durante 5 encuentros, 20 estudiantes sin ningún tipo de conocimiento en robótica y programación realizaron los mismos ejercicios, 10 de ellos usando simuladores y los otros 10 usando robots reales. El resultado obtenido fue un mejor y más robusto avance en aquellos que trabajaron en ambientes simulados. Dado que el grupo estudiado fue muy chico, se deben realizar estudios con mayor volumen de estudiantes.

Por último, se entrenaron a los equipos que representaron a la Argentina en la RoboCup del 2024 y 2025 en las categorías de simulación, obteniendo el primer puesto en SuperTeams en el mundial de 2024, primer puesto en la general de RoboCup Americas 2025 y tercer puesto en la general de RoboCup Mundial 2025

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por un investigador adjunto del Centro de Altos Estudios en Tecnología Informática (CAETI) quien ejerce el rol de director del proyecto, un doctorando, y dos ayudantes de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana.

5. BIBLIOGRAFÍA

- Ayala, A., Cruz, F., Cruz, F., Campos, D., Rubio, R., Fernandes, B. J. T., & Dazeley, R. (2020). *A Comparison of Humanoid Robot Simulators: A Quantitative Approach.*
- Buckley, E., Monaco, J. D., Schultz, K. M., Chalmers, R. W., Hadzic, A., Zhang, K., Hwang, G. M., & Carr, M. D. (2022). An interdisciplinary approach to high school curriculum development: Swarming Powered by Neuroscience. *2022 IEEE Integrated STEM Education Conference (ISEC)*, 1-8.
- Buronova, G. (2022). Didactic Possibilities of Using Simulators in the Organization of Robotics Circles through Virtual Educational Platforms. *INTERNATIONAL JOURNAL OF INCLUSIVE AND SUSTAINABLE EDUCATION*, *1*(4), 199-203.
- Campos, F. M. (2023). Ardosia: Simulating circuits and robotic systems in a single learning platform. *IEEE Transactions on Learning Technologies*.
- Cañas, José M., Perdices, E., García-Pérez, L., & Fernández-Conde, J. (2020). A ROS-based open tool for intelligent robotics education. *Applied Sciences*, *10*(21), 7419. https://doi.org/10.3390/app10217419
- Cheng, L., Guo, S., & Wu, J. (2021). Innovation Education Research Based on Cross-object Interactive Robot Simulation Platform. *SHS Web of Conferences*, *123*, 01020.
- Chikurtev, D. S., Chikurteva, A. A., & Spasova, N. R. (2021). Information technologies for development of educational resources in robotics. *IOP Conference Series: Materials Science and Engineering*, *1031*(1), 012122.
- Conde, M. Á., Rodríguez Lera, F. J., Fernández-González, D., J. Rodríguez-Sedano, F., Guerrero-Higueras, Á. M., & Fernández, C. (2020). SUFFER–SimUlation Framework for Education in Robotics. *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*, 14-17.
- Fabro, J. A., Gutierrez, M. B., & Caetano, F. H. R. (2022). FluxProg 2.0-Teaching Introductory Programming Using Flowcharts with Real and Simulated Robots for The Brazilian Robotics Olympiad (OBR). *2022 Latin American Robotics Symposium (LARS), 2022 Brazilian Symposium on Robotics (SBR), and 2022 Workshop on Robotics in Education (WRE)*, 424-429.
- Gena, C., Mattutino, C., Maltese, W., Piazza, G., & Rizzello, E. (2021). *Nao_PRM: an interactive and affective simulator of the Nao robot*. 727-734. https://doi.org/10.1109/ro-man50785.2021.9515410
- Joventino, C. F., Silva, R. de A. A. e, Pereira, J. H. M., Yabarrena, J. M. S. C., & de Oliveira, A. S. (2023). A Sim-to-real Practical Approach to Teach Robotics into K-12: A Case Study of Simulators, Educational and DIY Robotics in Competition-based Learning. *Journal of Intelligent & Robotic Systems*, *107*(1), 14. https://doi.org/10.1007/s10846-022-01790-2
- Lima, J., Kalbermatter, R. B., Braun, J., Brito, T., Berger, G., & Costa, P. (2022). A realistic simulation environment as a teaching aid in educational robotics. *2022 Latin American Robotics Symposium (LARS), 2022 Brazilian Symposium on Robotics (SBR), and 2022 Workshop on Robotics in Education (WRE)*, 430-435. https://doi.org/10.1109/LARS/SBR/WRE56824.2022.9996083
- Lima, J., N. Martins, F., & Costa, P. (2022). Teaching Practical Robotics During the COVID-19 Pandemic: A Case Study on Regular and Hardware-in-the-Loop Simulations. *Iberian Robotics conference*, 538-549.
- Mistretta, S. (2022). Virtual robotics in hybrid teaching and learning. *New Updates in E-Learning*, 145.
- Nascimento, L. M. do, Neri, D. S., Ferreira, T. do N., Pereira, F. de A., Albuquerque, E. A. Y., Gonçalves, L. M. G., & Sá, S. T. de L. (2021). sBotics-Gamified Framework for

Educational Robotics. *Journal of Intelligent & Robotic Systems*, *102*(1), 1-20.

- Sarkar, A., Pandey, A., Kumar, A., Furtado, A., Karia, K., & Arya, K. (2022). Teaching complex skills through an online robotics competition during Covid-19 pandemic. *2022 International Conference on Advanced Learning Technologies (ICALT)*, 199-200.
- Souza, J. D., Silva, P. J., Pinto, A. H., Monteiro, F. F., & Teixeira, J. M. (2020). Assessing the Reality Gap of Robotic Simulations with Educational Purposes. *2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)*, 1-6.
- Stein, G., Jean, D., Brady, C., & Lédeczi, Á. (2023). Browser-based simulation for novice-friendly classroom robotics. *Frontiers in Computer Science*, *4*, 1031572.
- Stein, G., & Lédeczi, A. (2021). Enabling Collaborative Distance Robotics Education for Novice Programmers. *2021 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 1-5.
- Stoffova, V., Zboran, M., & Hyksová, H. (2021). Application of Simulation Tools in Educational Robotics. *EDULEARN21 Proceedings*, 9214-9221.
- Xiao, G. (2021). Research on Artificial Intelligence Robots and Simulation Software in Virtual Training Pendant. *Journal of Physics: Conference Series*, *1952*(4), 042042.
- Yue, H., Miao, J., Zhang, J., Fan, C., & Xu, D. (2021). Simulation for senior undergraduate education of robot engineering based on Webots. *Computer Applications in Engineering Education*, *29*(5), 1176-1190.
- Zhan, Z., Zhong, B., Shi, X., Si, Q., & Zheng, J. (2022). The design and application of IRobotQ3D for simulating robotics experiments in K-12 education. *Computer Applications in Engineering Education*, *30*(2), 532-549.

Gamificación como catalizador del pensamiento computacional en carreras STEM: diseño, implementación y evaluación

Fornari Javier[1], López Mariel[1], Deco Claudia[1], Burzacca Luciana[1]

[1]Universidad Nacional de Rafaela
Facultad de Tecnologías e Innovación para el Desarrollo
Bv. Roca y Artigas, (2300) Rafaela, Argentina
54-3492-422880

{javier.fornari, mariellopez, claudia.deco, luciana.burza}@unraf.edu.ar

RESUMEN

En la Universidad Nacional de Rafaela (UNRaf) se observa una creciente demanda del sector productivo regional por profesionales con competencias avanzadas en pensamiento computacional, incluyendo abstracción, descomposición y resolución algorítmica de problemas. Sin embargo, los enfoques pedagógicos tradicionales presentan limitaciones para desarrollar estas habilidades en estudiantes de carreras STEM, afectando la motivación y la calidad del aprendizaje. Este trabajo presenta un proyecto como propuesta educativa basada en gamificación orientada a mejorar el desarrollo del pensamiento computacional mediante experiencias formativas innovadoras. La metodología contempla un diagnóstico inicial del nivel de competencias en estudiantes de carreras STEM, seguido del diseño e implementación de actividades interdisciplinarias gamificadas utilizando herramientas digitales y recursos lúdicos contextualizados. El impacto será evaluado mediante un enfoque mixto, combinando mediciones pre y post intervención, encuestas de motivación y técnicas cualitativas. Se espera evidenciar mejoras significativas en competencias de pensamiento computacional, motivación y desempeño académico. Asimismo, el proyecto busca generar recursos didácticos validados y metodologías transferibles, fortaleciendo la articulación entre universidad, sistema educativo y sector productivo regional.

Palabras clave: pensamiento computacional, gamificación, educación STEM, innovación educativa.

CONTEXTO

La necesidad actual de las cátedras de algoritmos, estructuras de datos y programación de las carreras de Ingeniería en Computación y las Licenciaturas en Bioinformática y Agroinformática en la Universidad Nacional de Rafaela (UNRaf) está enmarcada en una realidad educativa que genera cambios en muy cortos plazos e involucra no sólo a docentes, sino a alumnos y a todo personal comprendido en el proceso de enseñanza-aprendizaje. El proyecto que se desarrolla actualmente se denomina "Mejorando el Pensamiento Computacional en Ambientes Educativos STEM mediante Proyectos de Gamificación" y es reconocido por la resolución

rectoral de UNRaf en su convocatoria de 2025, en el ámbito del Programa de Incentivos a los Docentes Investigadores en la Modalidad A en el ámbito de la Universidad.

1. INTRODUCCIÓN

El Pensamiento Computacional (PC) es definido como un conjunto de habilidades y técnicas que permiten resolver problemas complejos mediante procesos sistemáticos y estructurados, involucrando principalmente la abstracción, descomposición de problemas, reconocimiento de patrones y diseño de algoritmos (Wing, 2006). Se considera una competencia esencial en la educación del siglo XXI, dado que desarrolla habilidades fundamentales para afrontar desafíos tecnológicos y laborales actuales y futuros. Uno de los enfoques pedagógicos que ha demostrado ser particularmente eficaz para desarrollar el PC es la gamificación, que implica la incorporación de elementos y dinámicas propias de juegos en contextos educativos con el objetivo de aumentar la motivación, la participación y la efectividad del aprendizaje (Dicheva et al., 2015). La gamificación se basa en teorías psicológicas y pedagógicas que enfatizan la motivación intrínseca, la autorregulación y la retroalimentación constante como factores clave para un aprendizaje significativo (Ryan & Deci, 2000). Diversas investigaciones respaldan la aplicación exitosa de la gamificación en el desarrollo del PC mediante herramientas específicas como Scratch, ScratchJr y juegos educativos como Robot Turtles. Estos estudios muestran que las actividades gamificadas incrementan significativamente habilidades como el razonamiento algorítmico, la abstracción y la resolución de problemas complejos, particularmente en estudiantes jóvenes (Kalelioğlu & Gülbahar, 2020; Tsarava et al., 2018; Giannakoulas & Xinogalos, 2018). Adicionalmente, la integración curricular del PC utilizando gamificación promueve ambientes educativos colaborativos, incrementa la motivación intrínseca de los estudiantes y facilita el aprendizaje transversal en diferentes asignaturas escolares (Sáez-López et al., 2016; Grover et al., 2015). La evidencia sugiere también que implementar actividades unplugged (sin dispositivos digitales) y digitales gamificadas contribuye eficazmente al desarrollo cognitivo y emocional necesario para adquirir competencias tecnológicas avanzadas desde etapas educativas tempranas (Brackmann et al., 2017; Bers, 2018). En (Kazimoglu, et al, 2012) investiga cómo el juego digital influye en la adquisición de habilidades del pensamiento computacional. Los autores diseñaron un juego denominado "Program Your Robot", que requiere aplicar abstracción, pensamiento algorítmico y reconocimiento de patrones para resolver desafíos. A través de métodos cuantitativos y cualitativos, analizaron a estudiantes de secundaria divididos en grupos experimentales y de control. Los resultados mostraron que el grupo que utilizó el juego mejoró significativamente en habilidades algorítmicas, resolución de problemas y conceptualización abstracta comparado con el grupo que siguió métodos tradicionales. Además, las entrevistas indicaron una mayor motivación y participación activa en los estudiantes que participaron del juego. El estudio concluyó que integrar juegos digitales en la enseñanza de programación promueve el pensamiento computacional al involucrar a los estudiantes en actividades atractivas y desafiantes, confirmando la utilidad educativa de herramientas gamificadas para la enseñanza efectiva del pensamiento computacional. Por otro lado, en (Kalelioğlu, F. et al, 2014) se explora la influencia de diversas técnicas instruccionales, entre ellas juegos educa-

tivos digitales, en el desarrollo de habilidades de pensamiento computacional y crítico. Se utilizó un diseño cuasi-experimental, aplicando técnicas basadas en juegos en clases de programación durante un semestre escolar. Las actividades gamificadas fueron comparadas con estrategias tradicionales centradas en la teoría. El estudio identificó mejoras estadísticamente significativas en las habilidades analíticas, resolución algorítmica y evaluación crítica en los estudiantes expuestos a juegos. Entrevistas complementarias revelaron mayor motivación, colaboración grupal y compromiso cognitivo en contextos de juegos. Los autores concluyeron que los juegos educativos proporcionan entornos efectivos para estimular habilidades fundamentales del pensamiento computacional y crítico, especialmente en contextos educativos formales, sugiriendo que las técnicas de gaming deben integrarse regularmente en programas curriculares para mejorar la calidad educativa tecnológica y científica. Finalmente, en (Papadakis, S. et al, 2017) se analizó la efectividad de actividades gamificadas en tabletas y computadoras para desarrollar habilidades numéricas tempranas y pensamiento computacional en niños de preescolar. Utilizando juegos interactivos que involucran reconocimiento de patrones, secuencias numéricas y resolución algorítmica simple, se llevó a cabo una intervención durante un período de 14 semanas. Los datos cuantitativos indicaron mejoras significativas en habilidades numéricas y lógicas en estudiantes que participaron en sesiones respecto a grupos control tradicionales. Los docentes reportaron mayor motivación, atención y satisfacción de los niños en las tareas de aprendizaje numérico. Se observó además un impacto positivo en la colaboración entre estudiantes y en la retención a largo plazo del conocimiento adquirido. Los investigadores concluyeron que las actividades gamificadas basadas en tabletas son herramientas altamente eficaces para el desarrollo temprano del pensamiento computacional, recomendando su uso frecuente en entornos educativos iniciales para fortalecer conceptos matemáticos fundamentales desde edades tempranas.

Tomando como referencia la información oficial de la Universidad Nacional de Rafaela, las carreras de Ingeniería en Computación y las Licenciaturas en Bioinfomática y Agroinformática han registrado durante los últimos años una deserción que se ha incrementado. Aunque, si bien los datos demuestran poco interés en las carreras, tal como pasa con las carreras STEM, se puede afirmar de manera elocuente la importancia de las carreras en el ámbito de la región ya que son innovadoras.

La carrera de Ingeniería en Computación se creó en el año 2018, comenzando su dictado en 2019, mientras que las Licenciaturas en Bioinformática y Agroinformática se crearon en 2020 y se comenzó su dictado en 2021. La cantidad total de alumnos durante los últimos tres años, 2023, 2024 y 2025 fue de 124. El siguiente cuadro, tabla 1, muestra la cantidad de ingresantes, mientras que la tabla 2 se indican los alumnos cursando todos los años y egresados de la carrera en los últimos tres años:

Año/Carrera	2023	2024	2025
Ingeniería en Computación	76	65	52
Bioinformática	29	27	28
Agroinformática	8	12	17

Tabla 1: Datos relevados de las carreras.

Año	2023	2024	2025
Alumnos	166	161	155
Egresados	0	0	1

Tabla 2: Datos totales relevados.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto está pensado con un enfoque vinculado con diversas instituciones y actores clave del entorno local y regional de Rafaela, dado que una de sus fortalezas radica precisamente en la articulación estratégica con entidades públicas, privadas, educativas, empresariales y del tercer sector. En primer lugar, UNRaf por su rol central como agente de conocimiento y de desarrollo territorial, actúa como la principal entidad académica de referencia del proyecto, ofreciendo no solo el soporte académico y científico, sino también facilitando el acceso a docentes y estudiantes de carreras STEM. El proyecto contribuirá directamente al fortalecimiento académico y a la transferencia metodológica en las carreras tecnológicas y de ingeniería, lo que refuerza el compromiso institucional de formar profesionales innovadores, capaces de responder a las demandas tecnológicas del territorio. Adicionalmente, la articulación con el sistema educativo secundario es esencial. En este sentido, se establece un vínculo estrecho con diversas escuelas secundarias técnicas y orientadas a ciencias exactas y naturales de la región, donde estas instituciones aportarán la base estudiantil necesaria para implementar las actividades del proyecto en las aulas. Además, los docentes de estas instituciones serán capacitados en estrategias innovadoras de gamificación y pensamiento computacional, generando así un beneficio significativo en términos de calidad educativa regional. Finalmente, el proyecto prevé establecer redes académicas y científicas con otras instituciones universitarias y centros de investigación nacionales e internacionales que trabajan temáticas afines. Esto permitirá compartir experiencias, validar científicamente los resultados obtenidos y asegurar la mejora continua del proyecto mediante procesos colaborativos de innovación educativa.

3. RESULTADOS OBTENIDOS ESPERADOS

En primer término, se espera lograr una mejora significativa en las competencias de pensamiento computacional de los estudiantes, reflejada en un incremento tangible de habilidades específicas como abstracción, descomposición de problemas, reconocimiento de patrones y diseño algorítmico. Estos resultados serán medidos mediante pruebas que permitirán evaluar objetivamente el progreso de los estudiantes. A nivel general, se espera que al menos el 75% de los estudiantes involucrados en el proyecto demuestren mejoras significativas en estas habilidades críticas para las carreras STEM. Asimismo, otro resultado relevante será el incremento en la motivación intrínseca y el compromiso de los estudiantes hacia las disciplinas STEM mediante el uso de metodologías gamificadas. Tras la implementación de estas técnicas, se espera lograr un mayor interés, participación activa y compromiso sostenido por parte de los estudiantes, especialmente en aquellos provenientes de sectores menos interesados inicialmente en estas disciplinas. El impacto esperado es que las tasas de abandono o bajo rendimiento en áreas tecnológicas y científicas se reduzcan considerablemente. Un tercer resultado esperado es la generación de recursos didácticos innovadores, validados y adaptados a las realidades educativas del territorio. Esto incluye guías metodológicas, proyectos interdisciplinarios, y herramientas digitales específicas para la enseñanza del pensamiento computacional en contextos educativos locales.

Por un lado, el impacto social como la inclusión educativa y digital de sectores más vulnerables en Rafaela y su región, promoviendo igualdad de oportunidades en el acceso a una formación tecnológica avanzada. La implementación de metodologías innovadoras como la gamificación

asegurará que estudiantes provenientes de distintos contextos socioeconómicos desarrollen habilidades clave para su futuro profesional. Desde el punto de vista del impacto económico tendrá un efecto directo en la mejora de la competitividad regional al formar profesionales con sólidas competencias tecnológicas y cognitivas, altamente demandados por el sector productivo local. Y teniendo en cuenta el impacto tecnológico que fortalecerá el desarrollo de habilidades del pensamiento computacional, fundamentales para la adopción, comprensión y creación de nuevas tecnologías.

En este caso se prevé la elaboración y publicación de artículos científicos como productos que tendrán como objetivo compartir metodologías, resultados empíricos y experiencias concretas sobre la implementación de la gamificación y su impacto en el desarrollo del pensamiento computacional en estudiantes STEM. Particularmente, se utilizarán espacios de divulgación académica promovidos por la Universidad Nacional de Rafaela y otras universidades del país con vínculos estratégicos. Se organizarán talleres y jornadas técnicas abiertas a la comunidad educativa, a los actores del sistema productivo regional y a la sociedad civil donde estos eventos permitirán presentar directamente los resultados del proyecto, incluyendo las metodologías, materiales didácticos generados y buenas prácticas implementadas en las instituciones participantes. Se elaborarán guías didácticas y recursos digitales específicos que serán puestos a disposición de docentes e instituciones educativas, tanto de nivel medio como universitario.

4. FORMACIÓN DE RECURSOS HUMANOS

La estrategia de formación y capacitación de recursos humanos vinculados directamente al proyecto, particularmente estudiantes avanzados y becarios involucrados, dado su carácter académico y científico-tecnológico. Las acciones específicas previstas para la formación de recursos humanos son incorporación activa de estudiantes avanzados y becarios para participar en forma activa en tareas como revisión bibliográfica, relevamiento de campo, entrevistas, encuestas, análisis de datos cuantitativos y cualitativos, elaboración de informes técnicos, y presentación de resultados parciales. En este caso ya se están trabajando con tres becarios, uno de cada carrera, con el fin de articular las experiencias logradas y poder comparar resultados, no solo de las carreras, sino entre otros ámbitos donde desarrollen metodologías de gamificación y en particular sea STEM.

5. BIBLIOGRAFÍA

Bers, M. U. (2018). Coding and computational thinking in early childhood: The impact of ScratchJr in Europe. European Journal of STEM Education, 3(3), Article 08.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. Educational Technology & Society, 18(3), 75–88. Grover, S., & Pea, R. (2013). Computational Thinking in K–12: A Review of the State of the Field. Educational Researcher, 42(1), 38-43.

Kalelioğlu, F., & Gülbahar, Y. (2020). The effects of teaching programming via Scratch on problem-solving skills. Informatics in Education, 19(2), 231-254.

Kalelioğlu, F., & Gülbahar, Y. (2014). The effect of instructional techniques on critical thinking and computational thinking. Eurasia Journal of Mathematics,

Science & Technology Education, 10(6), 571-582.

Kazimoglu, C., Kiernan, M., Bacon, L., & MacKinnon, L. (2012). Learning Programming at the Computational Thinking Level via Digital Game-Play. Procedia Computer Science, 9, 522-531.

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2017). The effectiveness of computer and tablet assisted intervention in early childhood students' understanding of numbers. An empirical study conducted in Greece. Education and Information Technologies, 22(5), 1847-1871.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation. American Psychologist, 55(1), 68-78. Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.

Gestión del Inteligente del Conocimiento en la Educación Superior

Gustavo Illescas[1], Elías Todorovich[1], Claudio Aciti[1], Guillermo Rodriguez[2], Pia Silvestrini[3]

[1]Instituto de Investigación en Tecnología Informática Avanzada (INTIA-Centro Asociado CIC, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA).

[2]Instituto Superior de Ingeniería de Software Tandil (ISISTAN-CONICET), Facultad de Ciencias Exactas, UNCPBA.

[3]Instituto de Investigación en Ciencias de la Salud (IICS), Facultad de Ciencias de la Salud, UNCPBA.

{illescas, etodorov, caciti, grodri} @exa.unicen.edu.ar, pia.silvestrini@salud.unicen.edu.ar

RESUMEN

El objetivo principal de esta investigación es profundizar en el análisis de los datos académicos, demográficos y censales de los estudiantes para identificar los factores que influyen en el rendimiento académico, la graduación y la tasa de deserción, dando así continuidad al proyecto 03/C314 (2021 – 2024). Por otro lado, buscamos generar información valiosa para la toma de decisiones institucionales, analizando los datos disponibles para identificar oportunidades de mejora en los procesos académicos y administrativos.

Analizar los datos de los estudiantes nos ayudará a comprender mejor su experiencia universitaria, identificar sus necesidades y expectativas, y aportar ideas para diseñar estrategias de personalización del aprendizaje que permitan mejorar su satisfacción y rendimiento académico.

A través de técnicas de análisis de datos, machine learning e inteligencia artificial, se busca construir modelos predictivos que permitan identificar a los estudiantes en riesgo y diseñar estrategias de acompañamiento académico personalizadas.

Esta línea de investigación ha formado ya a seis tesistas de grado y 3 becarios EVC-CIN en años anteriores y continúa con instancias de formación en este nuevo proyecto en el que se enmarca la investigación.

Palabras clave: Analítica de Datos, Educación Superior, Indicadores, Optimización

CONTEXTO

Esta investigación se enmarca dentro del proyecto de incentivos 03/C349 titulado "Integración Tecnológica para la Gestión Inteligente del Conocimiento en los Procesos Organizacionales", en particular la Línea 1: Gestión Inteligente del Conocimiento en la Educación Superior: Impulsando la Innovación Académica. Se desarrolla en el Instituto de Tecnología Informática Avanzada (INTIA) de la Facultad de Ciencias Exactas (EXA), Universidad Nacional del Centro de la provincia de Buenos Aires (UNCPBA).

1. INTRODUCCIÓN

Las Bases de Datos de Educación Superior, particularmente la de EXA-UNCPBA (ver Figura 1), integran un vasto volumen de datos de sistemas de gestión universitarios, como SIU-Guaraní (alumnos), Kune (evaluación docente), Majen (planta funcional), SIU-Kolla (encuestas), el sistema de indicadores y el Aula Virtual Moodle (Illescas, et. al. 2022).

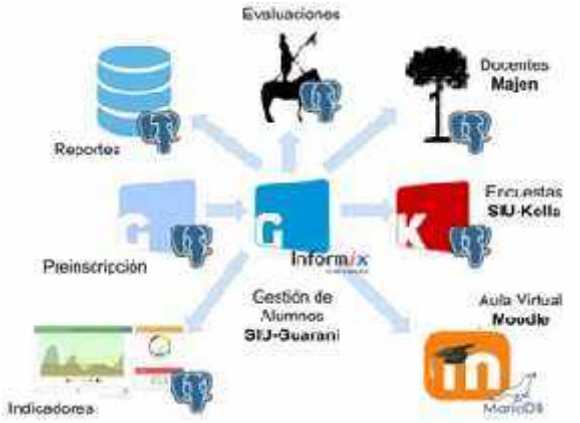


Figura 1. Esquema de Sistemas y Bases de datos disponibles en EXA. Fuente: área TICs EXA

Más allá de los reportes estándar de estos sistemas, las técnicas de Análisis de Datos en Educación Superior utilizadas en el proyecto anterior (03/C314, 2021-2024) demostraron su potencial para extraer conocimiento novedoso en seguimiento estudiantil y toma de decisiones institucionales. En esta etapa, ampliamos el enfoque con métodos no explorados previamente, incorporando el auge de la Inteligencia Artificial (IA), especialmente la IA Generativa, para potenciar el análisis predictivo, junto con chatbots interactivos (Jiménez-García 2025), cuyos avances iniciales ya hemos desarrollado de forma exploratoria.

Adicionalmente, integramos herramientas emergentes de estadística e investigación de operaciones: Redes Bayesianas para modelado probabilístico (Chen et. al. 2024), Cadenas de Markov para trayectorias estudiantiles (Yekeen, 2024), Procesos Jerárquicos Analíticos (AHP) para priorización de intervenciones (Saaty, 2008; Maral et. al. 2025), análisis de Camino Crítico en procesos académicos (Mendieta Lucas et. al. 2025) y Árboles de Decisión para clasificación de riesgos (Villarreal-Torres et. al. 2024). Esta combinación busca generar modelos más robustos y accionables, alineados con las necesidades de la educación superior argentina.

Metodología

Para abordar las preguntas de investigación que motivan este proyecto, se propone el método de trabajo en equipo *Design Thinking* (Brown, 2009). Esta metodología, basada en la colaboración y la iteración, facilita la generación de ideas y la entrega frecuente de resultados, lo que permite ajustar y mejorar las soluciones a medida que avanza la investigación.

Los proyectos interdisciplinarios, por su naturaleza compleja, requieren de un enfoque que permita aprovechar al máximo el potencial de cada integrante. Design Thinking ofrece un marco de trabajo sólido para abordar este desafío, ya que facilita la colaboración, la organización y la generación de soluciones centradas en el usuario a través de sus principios fundamentales: definición de roles, fases de trabajo, pautas de comportamiento, diseño centrado en el usuario y desarrollo iterativo e incremental.

A pesar de haber sido concebido originalmente para el diseño industrial, Design Thinking ha demostrado su eficacia en la generación de ideas e innovaciones en diversos campos, incluyendo la investigación y la educación. Los resultados positivos observados en la industria han impulsado su adopción como marco de trabajo en estos ámbitos

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Estado de Avance: el estado de avance está relacionado con el "Estado actual del conocimiento sobre el tema" donde los integrantes y colaboradores han realizado experiencias en la temática que han sido reportadas en diferentes comunicaciones.

Cronograma

El plan de trabajo propuesto se estructura en tres años, cada una con objetivos específicos:

Año 1: Exploración y definición. Se realizará una actualización bibliográfica continua, se analizará el estado del arte, se recopilará información relevante y se estudiarán los requerimientos del proyecto. Se explorarán técnicas para el diseño de indicadores y se difundirán los avances del proyecto a través de workshops.

Año 2: Experimentación y análisis. Se depurarán los datos, se definirán casos de estudio, se diseñarán experimentos y se realizarán ensayos con los datos disponibles.

Los resultados intermedios se presentarán en congresos.

Año 3: Validación y difusión. Se validarán los resultados obtenidos, se analizarán los ensayos, se elaborarán conclusiones y se propondrán futuras líneas de investigación. Los resultados finales se publicarán en revistas científicas.

Los aportes esperados en cuanto a difusión del conocimiento es la realización de publicaciones en workshops, congresos y revistas.

Este proyecto se fundamenta en la sólida trayectoria de los integrantes y colaboradores del grupo, quienes cuentan con investigaciones previas en áreas estrechamente vinculadas a la temática propuesta, fundamentalmente en el desarrollo del proyecto previo (2021-2024) en el cual han participado todos los integrantes que figuran en esta oportunidad.

-El Dr. Gustavo Illescas ha participado en actividades de EXA en el área de Coordinación de Asuntos Estudiantiles y Graduados (CAEG), en la Sec. de Coordinación General y en la Sec. de Extensión. Rol: coordinador del proyecto de implantación de un sistema de indicadores (2018-actualidad)

-El Dr. Todorovich ha participado como director en el Proyecto de Mejoramiento de la Enseñanza en Carreras de Informática (PROMINF), donde uno de los ejes de trabajo fue el mejoramiento de indicadores de ingreso, permanencia y promoción para la carrera de Ingeniería de Sistemas (EXA). Para ello, se puso en marcha un programa con tutores pares, que luego fue evolucionando y se relevaron los resultados para construir esos indicadores que marcaron el avance y resultado del proyecto. 2013-2017.

-El Dr. Claudio Aciti es el actual decano de EXA donde una de las áreas que coordina en forma directa es la CAEG (Coordinación de Asuntos Estudiantiles y Graduados) que se encarga del seguimiento del ingreso, permanencia y egreso. (2017-actualidad)

-El Dr. Guillermo Rodriguez ha participado en numerosos artículos relacionados a las técnicas de *Data Mining*, así mismo ha presentado diversos trabajos en el área de investigación junto al Dr. Gustavo Illescas.

-La Médica Veterinaria María Pía Silvestrini se desempeñó como co-directora del proyecto de investigación "Desarrollo de un sistema de indicadores para el monitoreo del desempeño académico en las carreras de Lic. en Enfermería y Medicina. ESCS-UNICEN" y ha participado activamente en la formación de estudiantes a través de la dirección de proyectos de iniciación a la investigación científica, promoviendo el aprendizaje de herramientas metodológicas y el pensamiento crítico en el ámbito académico.

En cuanto a los colaboradores:

-La Ing. Norma Montero ha participado en trabajos de investigación sobre experiencias desarrolladas en la Dirección de Formación Continua de la Provincia de Buenos Aires.

-La Mg. Rosana Cepeda se desempeñó como subsecretaria académica de EXA (2017-2024), es Licenciada en Ciencias Matemáticas, Magister en Matemáticas (Estadística).

-El Ing. Mariano Martinez y la Ing. Iris Figini pertenecen al área TICs de EXA donde desarrollan actividades en los sistemas SIU-Guarani, de Concursos (Kune), de gestión Docente (Majen), de encuestas SIU-Kolla y han participado en el sistema CONEAU para la acreditación de la carrera de Ingeniería de Sistemas.

-El T.G.U. Gabriel Montaruli (nodocente Rectorado-UNICEN) es Director del Centro de Información y Estadística Universitaria de la Secretaría Académica de la UNICEN, También es el responsable institucional del sistema SIU-Araucano ante el Ministerio de Educación, y es asesor en el área de Control de Gestión de la Secretaría General del Rectorado.

3. RESULTADOS OBTENIDOS/ESPERADOS

Aunque el foco principal del análisis se mantiene en los procesos de ingreso, permanencia y egreso estudiantil (Gutiérrez Pizarro, 2014), ampliamos el alcance para incorporar matices complementarios que enriquezcan la comprensión de la trayectoria universitaria y apoyen intervenciones preventivas. Estas extensiones metodológicas, alineadas con las tendencias en computación aplicada a la educación, permiten una visión

más holística de los desafíos en la educación superior argentina.

Aplicaciones Específicas Propuestas

AHP para orientación vocacional: Aplicamos Procesos Jerárquicos Analíticos (AHP) para guiar a estudiantes de secundaria en la elección de carreras, priorizando criterios como intereses, aptitudes y demanda laboral regional, reduciendo así la deserción temprana por inadecuación vocacional.

AHP en detección de habilidades no técnicas: Identificamos competencias blandas (liderazgo, trabajo en equipo, resiliencia) mediante AHP, integrando datos cualitativos con evaluaciones psicométricas, para diseñar programas de fortalecimiento que mejoren la empleabilidad.

Cadenas de Markov para deserción: Modelamos trayectorias de abandono como procesos estocásticos, estimando probabilidades de transición entre estados y prediciendo puntos críticos basados en variables socioeconómicas y académicas.

Movilidad estudiantil: Sistematizamos patrones de movilidad interfacultad como factor atenuante de deserción, analizando flujos con redes y métricas de centralidad para optimizar políticas de flexibilidad curricular.

Minería de texto en procedencia: Extraemos datos geográficos y socioeconómicos de campos textuales en SIU-Guaraní, clasificando orígenes para segmentar intervenciones regionales (Miranda 2017; Traini, et. al. 2021; Triventi et al., 2021).

Encuestas a desertores: Diseñamos y analizamos encuestas estructuradas para identificar causas raíz, como brechas económicas o falta de apoyo (Lorenz, 2024).

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo se muestra en la siguiente tabla:

Apellido y nombre	Título	Cargo	Funciones
Aciti, Claudio	Dr.	Prof. UNCPBA	Integrante
Illescas, Gustavo	Dr.	Prof. UNCPBA	Director
Rodríguez, Guillermo	Dr.	Prof. UNCPBA	Integrante
Silvestrini, Pía	Med. Vet.	Prof. UNCPBA	Integrante
Todorovich, Elías	Dr.	Prof. UNCPBA	Integrante

Colaboradores: Norma Montero (Docente EXA-UNICEN), Rosana Cepeda (Docente EXA-UNICEN), Mariano Martinez (Docente y no docente EXA-UNICEN), Iris Figini (no docente EXA-UNICEN), Gabriel Montaruli (no docente Rectorado-UNICEN).

Colaboradores externos: María Isabel Sanchez-Segura (UC3M, España), Fuensanta Medina-Dominguez (UC3M, España), José Arturo Mora-Soto (Univ. de Celaya, México), José Roberto Cantú-González (Universidad Autónoma de Coahuila, México), Antonio Hernandez (Univ. de Celaya, México)

Becarios

Sistematización de la movilidad estudiantil como atenuante de la deserción. Alumno: Thomas Ojeda (Becario EVC-CIN 2024). Director: Illescas, Gustavo. Codirector: Todorovich Elias

Tesis de grado

- Análisis del rendimiento académico en la carrera de Ing. de Sistemas utilizando clasificadores (*Random forest*, K-nn vecinos más cercanos, *Support vector machine* y Bayes). Dirección: Illescas G., Mora-Soto A (Univ. Celaya MEX.). Tesistas: Maylen Dell Oso, Florencia Paglione.
- Detección de habilidades no técnicas utilizando AHP. Dirección: Illescas G., Cantu-Gonzalez Roberto (Univ. Coahuila MEX.). Tesista: Jorge Berclaz

Los siguientes estudiantes de Ingeniería de Sistemas EXA se encuentran realizando trabajos finales de asignaturas relacionados al proyecto, bajo la dirección de los docentes integrantes del proyecto que se presenta en esta oportunidad.

Hernan Horquin, Matias Fino, Tomás Ceschini, Maximiliano Lujan, Rocio Lopez, Ezequiel Mapelli, Matías Pozzi, Facundo Martinez, Gonzalo Clementi, Leandro Salias.

5. BIBLIOGRAFÍA

- Illescas Gustavo, Todorovich Elías, Aciti Claudio, Rodriguez Guillermo, Silvestrini Pía (2022). Aplicación de Analítica de Datos en espacios de Educación Superior. XXIV Workshop de Investigadores en Ciencias de la Computación Universidad Champagnat. Mendoza, Argentina.

- Jiménez-García, E. (2025). Inteligencia Artificial y chatbots para una educación superior sostenible. Revista Iberoamericana de Educación a Distancia (RIED) E-ISSN: 1390-3306. Disponible en: educaia.com.ar/wp-content/uploads/2025/06/IA-y-chatbots-para-educacion-superior-sostenible.pdf

- Chen Xiaoyong, Zong Xuanyi, Zhang Duo, Cheng Shi (2024). "Evaluation of Teaching Effectiveness Based on Bayesian Network Algorithm in Teaching and Learning Process in Higher Education Institutions". Journal of Electrical Systems Vol. 20 No. 6s. https://doi.org/10.52783/jes.3097.

- Yekeen, W. (2024). "Modelling Students Academic Performance Progression Using Discrete-Time Markov Chain". Journal of Sciences. Journal of Sciences (FJS). ISSN: 2616-1370
Vol. 8 No. 6, December, (Special Issue) 2024, pp 444 - 449

- Saaty, T. (2008). "Decision Making with the Analytic Hierarchy Process", Int. J. Services Sciences, Vol. 1, Nº. 1, pp.83-98

- Maral, M. Ali Ö (2025). Una revisión sistemática sobre métodos de toma de decisiones multicriterio en investigación educativa. British Educational Research Journal.

- Mendieta Lucas, L., Cabrera Lituma, E., Ramírez Rivera, R., & Ramos Echevarría, M. (2025). Aplicación de la Metodología de Ruta Crítica en Proyectos Educativos de Educación Superior: Una Estrategia para Mejorar Tiempos y Recursos en la Ejecución Académica. Reincisol., 4(7), 4961–4985. https://doi.org/10.59282/reincisol.V4(7)4961-4985

- Villarreal-Torres, H., Ángeles-Morales, J., Marín-Rodriguez, W., & Cano-Mejía, J. (2024). Modelo de clasificación para la deserción estudiantil en las universidades públicas del Perú. Revista De Ciencias Sociales, 30(1), 452-469. https://doi.org/10.31876/rcs.v30i1.41667

- Brown, T. "Change by Design, How Design Thinking Transforms Organizations and Inspires Innovation" (2009), Ed. Harper Collins New York ISBN 978-0-06-193774-3.

- Gutiérrez Pizarro C. (2014). Estudio de la permanencia y deserción de los estudiantes de pregrado en una Facultad de Ingeniería, utilizando una metodología mixta de Investigación y árboles de decisión en su predicción. Tesis doctoral Universidad de Alcalá.

- Miranda M y Guzmán J. (2017). Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos-Form. Univ. vol.10 no.3 La Serena. Chile

- Traini, C., Kleinert, C., & Bittmann, F. (2021). "How does exposure to a different school track influence learning progress? Explaining scissor effects by track in Germany". Research in Social Stratification and Mobility, Volume 76, 100625

- Triventi, M., Barone, C., & Facchini, M. (2021). "Upper secondary tracks and student competencies: A selection or a causal effect? Evidence from the Italian case". Research in Social Stratification and Mobility, Volume 76, 100626.

- Lorenz T., Schneebaum, A. (2024). "Does Early Educational Tracking Contribute to Gender Gaps in Test Achievement? A Cross-Country Assessment". Jahrbücher für Nationalökonomie und Statistik, Vol. 244, no. 1-2, pp. 5-36.

Innovación y Desafíos en la Educación Digital

Claudia Russo[1], Tamara Ahmad[1], Gustavo Gnazzo[1], Paula Lencina[1], María Mercedes Guasch[1], Mariana Adó[1], María Rosa Piergallini[1], Hugo Ramón[1], Marina Rodríguez[1], Karina Marenzana, Fernando López Gil[1], Sola Tatiana Verónica[1], Carla Decoud[2], María Elena Chan Nuñez[3], Sarita Shenriksen[4], Renato Abreu[5] , Martha Leticia Quintanilla Acosta[6]

[1] Instituto de Investigación y Transferencia en Tecnología (ITT), UNNOBA-CIC.
Sarmiento N° 1169, 2do Piso, Junín (Bs. As.), Argentina – Te: +54 236 4477-050 INT 11610.

[2] Facultad Politécnica (FP), Universidad de Asunción (UNA).
Acceso peatonal de Mcal. Estigarribia a la UNA, San Lorenzo, Paraguay – Te: +595 21 588 7000.

[3] Centro Universitario de Guadalajara, Universidad de Guadalajara México

[4] Faculdade de Ciências da Linguagem, Comunicação e Artes (Fclca)

[5] Escola Superior de Tecnologia de Lisboa

6 Universidad Nacional Autónoma de Honduras, UNAH

{claudia.russo, tamara.ahmad, gustavo.gnazzo, paula.lencina, mercedes.guasch, mariana.ado, rosana.piergallini, hugo.ramon, marina.rodriguez, fernando.lopezgil}@itt.unnoba.edu.ar, karinamarenzana@unnoba.edu.ar, tvsola@comunidad.unnoba.edu.ar, carladecoud@gmail.com, elena.chan@virtual.udg.mx, hmdsarita@gmail.com, renato.abreu@estesl.ipl.pt, marthaquintanilla@unah.edu.hn

Resumen

La transformación digital en educación superior exige superar la mera incorporación de herramientas tecnológicas para avanzar hacia modelos basados en evidencia, automatización inteligente y sostenibilidad digital. El presente proyecto SIB 2025, enmarcado en el área de Tecnología Informática Aplicada en Educación (TIAE), propone investigar, diseñar y desarrollar soluciones informáticas innovadoras basadas en inteligencia artificial, analítica de datos y tecnologías emergentes con el objetivo de optimizar la calidad del aprendizaje, fortalecer la personalización educativa y promover prácticas sostenibles.

El proyecto articula cuatro líneas de investigación: desarrollo de plataformas y aplicaciones inteligentes, redefinición de instrumentos de evaluación, formación docente en tecnologías emergentes y sostenibilidad digital, y diseño de tableros de control orientados al seguimiento y análisis de procesos educativos digitales. Asimismo, se alinea con los Objetivos de Desarrollo Sostenible (ODS 9, 10, 12, 13 y 17), integrando criterios de equidad digital, responsabilidad ambiental y generación de información confiable para la toma de decisiones basada en evidencia.

Palabras claves: Educación digital; Inteligencia Artificial; Analítica del aprendizaje; Sostenibilidad digital; Dashboard educativo; Evaluación de impacto.

Contexto

La acelerada evolución de tecnologías (particularmente inteligencia artificial (IA), analítica de datos, realidad virtual (RV) y realidad aumentada (RA)) ha modificado profundamente la manera en que se digitalizan, procesan y analizan los datos educativos. Sin embargo, la incorporación de estas tecnologías en la educación digital no siempre ha estado acompañada por marcos sistemáticos de análisis, métricas compartidas

ni instrumentos que permitan comprender su impacto en los procesos formativos.

La literatura destaca que la analítica del aprendizaje constituye un campo en expansión que permite identificar patrones de desempeño y apoyar decisiones pedagógicas basadas en datos [1]. No obstante, se advierte sobre los desafíos éticos asociados al uso de grandes volúmenes de información educativa, especialmente en relación con la privacidad, el sesgo algorítmico y la transparencia [2].

En este escenario, muchas instituciones enfrentan tres desafíos estructurales:

- Fragmentación de datos académicos en múltiples plataformas.
- Escasa sistematización de indicadores para analizar procesos educativos digitales.
- Ausencia de herramientas integradas que permitan monitoreo y visualización en tiempo real.

En este marco, la implementación de un sistema centralizado —como un dashboard o tablero de control— orientado al seguimiento y análisis de información académica resulta una estrategia relevante para favorecer procesos de mejora y toma de decisiones basada en evidencia [3].

Paralelamente, la creciente digitalización conlleva un aumento del consumo energético y de la huella de carbono asociada a infraestructuras tecnológicas. La formación en sostenibilidad digital emerge así como un componente necesario en proyectos educativos contemporáneos, promoviendo prácticas responsables y un uso más eficiente de los recursos tecnológicos [4,5].

1. Introducción

La acelerada incorporación de tecnologías emergentes en educación superior ha generado un nuevo escenario en el que la digitalización de procesos ya no constituye un diferencial, sino una condición estructural del sistema educativo. En los últimos años, la inteligencia artificial (IA), la analítica de datos, las tecnologías inmersivas y los entornos virtuales han ampliado significativamente las posibilidades de personalización del aprendizaje, automatización de tareas administrativas y generación de información en tiempo real para la toma de decisiones.

No obstante, la expansión de la analítica del aprendizaje no siempre ha estado acompañada por instrumentos compartidos que permitan analizar de manera sistemática el aporte de estas tecnologías a los procesos formativos [1]. Asimismo, el uso intensivo de datos educativos requiere marcos éticos que regulen su tratamiento, considerando aspectos vinculados a privacidad, transparencia y posibles sesgos algorítmicos [2].

En este contexto, el proyecto propone superar la fragmentación de datos académicos y la dispersión de herramientas digitales mediante el diseño de soluciones informáticas integradas que articulen:

- Inteligencia artificial aplicada a la personalización y automatización.
- Analítica de datos para monitoreo académico.
- Tecnologías emergentes para experiencias de aprendizaje innovadoras.
- Instrumentos estandarizados de evaluación.
- Estrategias de sostenibilidad digital.

En este escenario, también se considera la dimensión de sostenibilidad digital como criterio transversal en el uso de tecnologías educativas [4] [5].

El objetivo general de este proyecto es investigar, diseñar y desarrollar soluciones informáticas innovadoras aplicadas a la educación digital, basadas en inteligencia artificial, analítica de datos y tecnologías emergentes, con el propósito de contribuir a la mejora y análisis de los procesos educativos en sus distintas modalidades, promoviendo la personalización del aprendizaje y la

incorporación de prácticas tecnológicas sostenibles.

Como objetivos específicos, se definen:

- Diseñar y validar herramientas tecnológicas que automaticen procesos administrativos y pedagógicos.
- Desarrollar modelos de personalización educativa basados en analítica de datos.
- Definir métricas e indicadores que permitan analizar el aporte de tecnologías emergentes en procesos educativos digitales.
- Diseñar un dashboard institucional que integre datos académicos y métricas pedagógicas.
- Incorporar formación en sostenibilidad digital en el uso de tecnologías educativas.
- Fortalecer la formación de recursos humanos y la articulación con redes de I+D+i.

Como hipótesis se plantean:

H1: La integración sistemática de inteligencia artificial, analítica de datos y tecnologías emergentes, acompañada por métricas de seguimiento, favorece el análisis y la mejora progresiva de los procesos educativos digitales.

H2: La implementación de un dashboard institucional integrado contribuye a facilitar la visualización de información académica y apoyar la toma de decisiones basada en datos.

H3: La incorporación de criterios de sostenibilidad digital en el uso de tecnologías educativas promueve prácticas tecnológicas más responsables en el ámbito universitario.

2. Líneas de investigación y desarrollo

El proyecto se estructura en cuatro líneas interrelacionadas que articulan desarrollo tecnológico, evaluación pedagógica, sostenibilidad digital y formación docente. Estas líneas permiten abordar de manera sistémica la transformación de los procesos educativos mediante tecnologías emergentes:

L1. Desarrollo de soluciones informáticas basadas en tecnologías emergentes e inteligencia artificial

Se desarrollarán plataformas, aplicaciones y prototipos que integren inteligencia artificial, analítica de datos y tecnologías inmersivas (RV y RA) con el propósito de apoyar la automatización de tareas administrativas y pedagógicas, así como la personalización del aprendizaje.

Se investigarán modelos de recomendación adaptativa y mecanismos de análisis basados en datos académicos. La IA aplicada a la educación ofrece oportunidades para fortalecer procesos de retroalimentación, detección de dificultades y adaptación de contenidos [6]. Asimismo, se explorará el potencial de tecnologías inmersivas para generar experiencias interactivas, en línea con estudios que evidencian su contribución a la motivación y comprensión conceptual [7].

L2. Diseño de instrumentos estandarizados e indicadores de calidad educativa digital

Uno de los desafíos centrales identificados es la ausencia de métricas integradas que permitan evaluar la eficacia pedagógica de tecnologías emergentes en educación digital. Esta línea se enfoca en la redefinición y diseño de instrumentos estandarizados que midan:

- Rendimiento académico.
- Retención del conocimiento.
- Satisfacción estudiantil.
- Eficiencia en procesos administrativos.
- Impacto en sostenibilidad digital.

Se desarrollarán modelos de evaluación basados en analítica del aprendizaje, considerando la necesidad de marcos éticos y transparencia en el uso de datos (Williamson, 2018). La construcción de indicadores confiables permitirá fortalecer la toma de decisiones institucionales basada en evidencia. [1], [8].

L3. Formación docente, alfabetización digital y sostenibilidad tecnológica

La transformación educativa requiere no sólo infraestructura tecnológica, sino también competencias digitales críticas y éticas por parte de los actores involucrados. Esta línea contempla el diseño de programas de capacitación orientados a la adopción pedagógica, ética y sostenible de tecnologías emergentes.

La alfabetización digital se concibe como una competencia integral que incluye habilidades críticas, creativas y reflexivas para el uso responsable de tecnologías (Iskandar et al., 2022). En este marco, la formación en sostenibilidad digital adquiere un rol estratégico, promoviendo prácticas que minimicen el impacto ambiental y reduzcan la huella de carbono asociada al uso intensivo de plataformas educativas. [4] [5].

Esta línea también contribuye a la reducción de brechas digitales, alineándose con los principios de equidad y accesibilidad en contextos diversos.

L4. Diseño e implementación de un dashboard institucional para educación digital

Esta línea constituye un eje estructurante del proyecto y articula resultados de las líneas anteriores. Se propone el diseño de un dashboard específico para la educación digital que integre:

- Datos académicos.
- Métricas de calidad pedagógica.
- Indicadores de uso de tecnologías emergentes.
- Variables vinculadas a sostenibilidad digital.

El dashboard permitirá visualizar información en tiempo real, identificar áreas críticas y facilitar la personalización de estrategias pedagógicas y administrativas. Desde una perspectiva técnica, se considerarán principios de modularidad, escalabilidad y reutilización en su diseño [3].

3. Resultados obtenidos/esperados

Si bien el proyecto SIB 2025 constituye una nueva etapa de investigación, se apoya en antecedentes consolidados del equipo en el ámbito de la educación digital, lo que ha permitido iniciar avances preliminares en varias de sus líneas estratégicas. Este proyecto está articulado con el área de Educación Digital de la universidad de manera estratégica en apoyo a las necesidades del SIEDD (Sistema Institucional de Educación a Distancia y Digital) de UNNOBA.

Entre los principales resultados obtenidos hasta el momento se destacan:

a) Sistematización inicial de indicadores para educación digital

Se ha avanzado en la identificación y organización preliminar de indicadores académicos relevantes para entornos virtuales e híbridos, incluyendo variables vinculadas a participación, desempeño, permanencia y uso de plataformas digitales. Este trabajo constituye la base conceptual para el desarrollo del dashboard institucional planteado en la línea L4.

b) Experiencias piloto con tecnologías emergentes

Se implementaron instancias de capacitación y experimentación con herramientas de inteligencia artificial, diseño de contenidos interactivos y tecnologías inmersivas en propuestas formativas institucionales. Estas experiencias permitieron relevar percepciones docentes, detectar oportunidades de automatización y ajustar criterios de integración pedagógica.

c) **Consolidación del Observatorio en Tecnologías Innovadoras Educativas (OTIE)**

El equipo ha fortalecido la producción y análisis de datos institucionales vinculados a educación digital, generando información estratégica para la mejora continua de

prácticas tecnopedagógicas.

d) Formación y desarrollo académico

Se registran avances en tesis de grado y posgrado vinculadas a la temática del proyecto, incluyendo investigaciones relacionadas con seguimiento académico, visualización de datos y aplicación de inteligencia artificial en contextos educativos digitales.

Asimismo, se espera que el proyecto contribuya al alineamiento institucional con los Objetivos de Desarrollo Sostenible (ODS 9, 10, 12, 13 y 17), promoviendo innovación tecnológica responsable, sostenibilidad digital y generación de datos educativos de calidad.

4. Metodología de investigación

El proyecto adopta un enfoque metodológico mixto, integrando técnicas cuantitativas y cualitativas para el diseño, desarrollo y validación de soluciones tecnológicas aplicadas a la educación digital.

En una primera etapa se realiza una revisión del estado del arte sobre inteligencia artificial en educación, analítica del aprendizaje, tecnologías emergentes y sostenibilidad digital, con el objetivo de fundamentar conceptualmente el diseño de herramientas e indicadores. [1], [6], [7].

Posteriormente, se lleva a cabo un diagnóstico institucional mediante encuestas, entrevistas y análisis de datos académicos disponibles, permitiendo identificar necesidades pedagógicas, procesos susceptibles de automatización y oportunidades de mejora en la gestión de información.

A partir de estos resultados, se diseñan y prototipan soluciones informáticas que integran modelos de analítica de datos, automatización y arquitecturas de visualización orientadas al desarrollo de un dashboard institucional. El diseño contempla criterios de modularidad, escalabilidad e interoperabilidad [3].

Los prototipos son sometidos a validación técnica (funcionalidad, seguridad y protección de datos) y validación pedagógica mediante implementaciones piloto en contextos reales. El impacto se evalúa a través de indicadores previamente definidos, considerando desempeño académico, participación, satisfacción estudiantil y criterios de sostenibilidad digital.

De manera transversal, el proyecto incorpora principios éticos en el uso de datos educativos, garantizando privacidad, transparencia y equidad (Williamson, 2018), así como un enfoque colaborativo e interdisciplinario en todas sus etapas. [2]

5. Formación de recursos humanos

Si bien la mayoría de los integrantes del proyecto ya cuentan con títulos de grado y algunos posgrados, la formación de recursos humanos constituye un eje estratégico del proyecto, orientado al fortalecimiento de capacidades en investigación aplicada.

El equipo cuenta con una beca de investigación BIAT vinculada al desarrollo de soluciones tecnológicas en educación digital, lo que refuerza la dimensión formativa del proyecto y permite la participación activa de jóvenes investigadores en tareas de diseño, análisis y validación de herramientas.

Asimismo, el proyecto enmarca el desarrollo de dos tesis doctorales: una que se encuentra en definición y análisis y la otra, centrada en el diseño y validación de un modelo de seguimiento académico y dashboard institucional basado en analítica de datos e inteligencia artificial. Esta investigación se articula directamente con la línea de desarrollo del sistema de medición integral propuesto, contribuyendo tanto al avance científico como a la transferencia institucional.

Se promueve además la realización y finalización de trabajos finales de grado y posgrado vinculados a las líneas del proyecto, así como la participación en congresos y publicaciones científicas nacionales e internacionales. De este modo, el proyecto no sólo impulsa desarrollos tecnológicos

innovadores, sino que consolida una masa crítica de investigadores en educación digital con proyección académica sostenida.

5. Bibliografía y referencias

[1] Domínguez Figaredo, D., Reich, J., & Ruipérez, J. A. (2020). Analítica del aprendizaje y educación basada en datos: Un campo en expansión. *RIED. Revista Iberoamericana de Educación a Distancia, 23*(2), 25–38.

[2] Williamson, B. (2018). *Big data in education: The digital future of learning, policy and practice*. Sage.

[3] Vázquez-Ingelmo, A., García-Peñalvo, F. J., & Therón, R. (2021). Beneficios de la aplicación del paradigma de líneas de productos software para generar dashboards en contextos educativos. *Education in the Knowledge Society, 22*, 1–15.

[4] Ayuso, J. (2023). La formación en sostenibilidad es un reto cultural y económico clave para transformar el futuro.

[5] Uniminuto. (2023). Educación virtual, un pilar para la sostenibilidad.

[6] Holmes, W., Bialik, M., & Fadel, C. (2023). *Artificial Intelligence in Education: Ethics and Possibilities for Human Flourishing*. UNESCO.

[7] Radianti, J., Majchrzak, A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*.

[8] OCDE. (2021). *Perspectivas de la educación digital*.

[9] Iskandar, N., Sumarni, T., Dewanti, R., & Asnur, L. (2022). Infusing digital literacy in authentic academic digital practices of English language teaching at universities. *International Journal of Language Education, 6*(1), 75–90.

INTEGRACIÓN DE INTELIGENCIA ARTIFICIAL GENERATIVA EN LA EDUCACIÓN MÉDICA: UN FRAMEWORK PARA EL ENTRENAMIENTO DE COMPETENCIAS BLANDAS Y JUICIO CLINICO

Ariel Adrián Blajos

Universidad Abierta Interamericana. Facultad de Medicina y Ciencias de la Salud. Centro de Altos Estudios en Ciencias Humanas y de la Salud (CAECIHS). Rosario. Santa Fe. Argentina
ArielAdrian.Blajos@uai.edu.ar

RESUMEN

La educación médica contemporánea enfrenta una disonancia estructural entre los modelos de acreditación por competencias y una infraestructura de simulación tradicional que suele limitarse a destrezas procedimentales. Este trabajo presenta el desarrollo de una arquitectura de software diseñada para la simulación clínica de "Alta Fidelidad Cognitiva" en escenarios ambulatorios. A diferencia de los simuladores lineales, la plataforma utiliza una orquestación de agentes de Inteligencia Artificial Generativa bajo el paradigma del "Narrador No Fiable", lo que obliga al estudiante a gestionar la incertidumbre y validar la información clínica. Técnicamente, el sistema opera mediante una arquitectura de tres capas: una capa probabilística de simulación basada en Large Language Models (LLMs); una capa determinista de supervisión ("Tutor-Centinela") que audita la conversación en tiempo real para mitigar alucinaciones ; y una capa de analítica de aprendizaje (Teaching Analytics) para detectar puntos ciegos curriculares. La propuesta incluye un protocolo de validación experimental para medir la progresión del razonamiento clínico, garantizando equidad formativa y escalabilidad institucional.

Palabras clave: Simulación Clínica; Inteligencia Artificial Generativa; Práctica Deliberada; Educación Médica; Habilidades Comunicacionales.

CONTEXTO

La presente línea de investigación se enmarca en el Proyecto de Investigación Subsidiado (2026-2027) titulado: "Análisis del Aporte Pedagógico de la Simulación Conversacional con IA como Estrategia de Andamiaje para el Desarrollo de Competencias Clínicas en Estudiantes de Medicina". El mismo se encuentra radicado en el Centro de Altos Estudios en Ciencias Humanas y de la Salud (CAECIHS), Sede Rosario, de la Facultad de Medicina y Ciencias de la Salud de la UAI.

1. INTRODUCCIÓN

La educación médica del siglo XXI transita hacia un modelo basado en la adquisición de competencias (CBME), donde la seguridad del paciente exige que los estudiantes certifiquen sus habilidades en entornos controlados antes de la práctica real. Organismos de acreditación y aseguramiento de la calidad (como la CONEAU en Argentina) han formalizado estas exigencias normativas [1]. Sin embargo, existe una disonancia estructural frente a la capacidad instalada universitaria. La simulación tradicional de "Alta Fidelidad Física", aunque eficaz para emergencias procedimentales, presenta barreras logísticas y económicas insalvables para el entrenamiento masivo del juicio clínico cotidiano, subordinando el

aprendizaje a la aleatoriedad de la casuística hospitalaria.

1.1 La Brecha Epistémica: Del "Saber Cómo" al "Mostrar Cómo"

Desde la perspectiva de la evaluación pedagógica estructurada en la Pirámide de Miller [2], los métodos tradicionales certifican los niveles inferiores, pero el nivel "Mostrar Cómo" en competencias complejas —como el razonamiento diagnóstico y la comunicación empática— queda relegado a costosos exámenes tipo ECOE. Esto genera un vacío curricular donde los estudiantes carecen de espacios de "práctica deliberada" [3]. Dado que el feedback docente inmediato suele ser diferido o inexistente en la práctica clínica saturada, se impide la consolidación del aprendizaje experiencial y la corrección iterativa de errores.

1.2 Estado del Arte en Simulación Digital: La Limitación del Determinismo

Históricamente, los simuladores de software han intentado suplir estas carencias mediante plataformas virtuales [4] basadas en arquitecturas de árboles de decisión rígidos. Esta tecnología presenta una limitación epistémica severa: el determinismo. Al entregar información estática y secuencial ante preguntas pre-guionadas, el modelo falla en replicar la realidad clínica. En la consulta real, el paciente es un "Narrador No Fiable" que olvida datos, miente o desconoce la relevancia de sus síntomas. Al carecer de Procesamiento de Lenguaje Natural (NLP) dinámico, los simuladores tradicionales no entrenan la astucia clínica para gestionar la incertidumbre.

1.3 La Irrupción de la IA Generativa: Oportunidades y Riesgos

La reciente emergencia de los Grandes Modelos de Lenguaje (LLMs) con aplicaciones en medicina [5] ha transformado este escenario, permitiendo crear agentes conversacionales con una precisión diagnóstica comparable a la de médicos de atención primaria [6] y un realismo percibido por los usuarios superior al 90% [7]. No obstante, su adopción acrítica conlleva riesgos sustanciales, como la "descarga cognitiva" o dependencia tecnológica del estudiante [8], y el grave fenómeno de las "alucinaciones". Un paciente virtual que inventa signos vitales introduce un ruido pedagógico inaceptable que puede fijar conocimientos erróneos.

1.4. Propuesta: Hacia la Alta Fidelidad Cognitiva

Frente a la dicotomía entre simuladores deterministas seguros y simuladores generativos alucinados, proponemos la Simulación de Alta Fidelidad Cognitiva Supervisada. Esta arquitectura de orquestación neuro-simbólica no conecta al estudiante directamente con el LLM, sino que media la interacción con capas de control determinista. Mediante un sistema de andamiaje automatizado ("Tutor-Centinela") fundamentado en la "Zona de Desarrollo Próximo" [9], el estudiante se enfrenta a la complejidad de un paciente "no fiable" dentro de un marco de seguridad académica que garantiza la adherencia a las Guías de Práctica Clínica.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se estructura sobre tres ejes interconectados orientados a superar la disonancia estructural en la educación médica mediante una solución tecnológica escalable.

2.1. Arquitectura de Orquestación Neuro-Simbólica y Seguridad Académica

Para abordar la mitigación de alucinaciones en la IA generativa, proponemos una arquitectura híbrida con una topología de tres capas. La capa probabilística implementa agentes conversacionales basados en LLMs multimodales, ajustados para emular el lenguaje natural y no verbal de un paciente. Por su parte, la capa determinista incorpora un "Agente Tutor-Centinela" que audita la interacción en tiempo real contra una matriz de verdad clínica, bloqueando cualquier invención fisiopatológica para garantizar la seguridad. Finalmente, una capa de analítica (*Teaching Analytics*) transforma los registros no estructurados de la simulación en métricas cuantitativas de desempeño institucional.

2.2. Ingeniería Semántica y Modelado del "Narrador No Fiable"

Desde la perspectiva pedagógica, exploramos la transición hacia la "Alta Fidelidad Cognitiva" mediante el diseño de *System Prompts* complejos. A diferencia de los modelos deterministas tradicionales, programamos agentes que omiten información crítica, presentan inconsistencias cronológicas y manifiestan barreras comunicacionales o emocionales. Esto obliga al estudiante a triangular datos y gestionar la incertidumbre epistémica, competencias centrales que corresponden al nivel "Mostrar Cómo" de la Pirámide de Miller [2].

2.3. Validación Experimental de la Transferencia de Aprendizaje

Para trascender la mera evaluación de usabilidad, establecemos un diseño cuasiexperimental prospectivo (pre-test/post-test con grupo control) que busca aportar evidencia sobre la eficacia clínica del dispositivo. Las variables dependientes principales incluyen la precisión diagnóstica y la competencia comunicacional, medida a través de la *Jefferson Scale of Empathy* [10] y la adherencia a protocolos como la escala CICAA [11]. Asimismo, investigamos el fenómeno de "desvanecimiento del andamiaje" (*scaffolding fading*), analizando cómo la reducción progresiva de la ayuda del tutor virtual se correlaciona con el aumento de la autonomía cognitiva del alumno.

3. RESULTADOS OBTENIDOS/ESPERADOS

Dada la naturaleza y proyección del estudio para el período 2026-2027, los resultados se estructuran tanto en hitos de desarrollo técnico ya consolidados como en metas experimentales proyectadas para la fase piloto.

3.1. Resultados Tecnológicos y Arquitectónicos (Obtenidos)

Se ha logrado definir y prototipar exitosamente una arquitectura de software SaaS basada en la orquestación de servicios, superando así las limitaciones del modelo monolítico de interacción directa con LLMs. Esta topología de tres capas interpone un middleware de control determinista que permite al "Agente Tutor" auditar las respuestas del paciente virtual en tiempo real, bloqueando eficazmente las alucinaciones fisiopatológicas antes de que lleguen a la interfaz del usuario. Las pruebas preliminares demuestran que esta orquestación de múltiples agentes opera con latencias compatibles con una conversación clínica fluida, validando la viabilidad técnica en entornos web estándar.

3.2. Resultados Pedagógicos Esperados (Validación Experimental)

Para la fase de campo en una cohorte piloto de 40 a 60 estudiantes del ciclo clínico, se han establecido rigurosos indicadores de éxito. En cuanto a la progresión del razonamiento clínico, se proyecta que al menos el 70% del grupo experimental exhiba una mejora estadísticamente significativa en la precisión diagnóstica. Respecto a la toma de decisiones críticas, se espera alcanzar una tasa de detección superior al 90% de las "Banderas Rojas" en escenarios de riesgo vital simulado. Asimismo, se anticipa un impacto positivo en el desarrollo de competencias blandas, evidenciado por mejoras en los puntajes de la

Jefferson Scale of Empathy [10] y una adherencia superior al 80% a la escala CICAA [11], superando ampliamente al grupo control.

3.3. Resultados Institucionales (Impacto y Transferencia)

A nivel macro, la implementación del dispositivo busca instaurar un modelo robusto de "Equidad Formativa", garantizando que la totalidad de la cohorte acceda a una casuística clínica estandarizada y diversa. Complementariamente, se prevé la consolidación de un dashboard de *Teaching Analytics* que transforme el rol docente, permitiendo identificar puntos ciegos curriculares recurrentes y ajustar los contenidos de la asignatura en tiempo real a partir de evidencia masiva de desempeño.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación se constituye como un espacio interdisciplinario de formación, fomentando la convergencia estratégica entre las Ciencias de la Salud y las Ciencias de la Computación en el ámbito *HealthTech*. El equipo está dirigido por un docente-investigador con doble perfil clínico y experto en IA Aplicada, garantizando la dirección académica y técnica integral del proyecto. La estructura operativa integra activamente a estudiantes de grado de la carrera de Medicina enfocados en el diseño narrativo de los casos clínicos y el testeo de usabilidad. Asimismo, el proyecto establece una sólida vinculación con el posgrado y grado mediante líneas específicas para Tesis de Maestría y Trabajos Finales de Ingeniería, abordando el desarrollo de algoritmos de Procesamiento de Lenguaje Natural (NLP) y validación de instrumentos de evaluación automática. Complementariamente, se proyecta incorporar un becario informático para optimizar los flujos de orquestación en plataformas de automatización y la integración de APIs, promoviendo una alfabetización digital cruzada que forme médicos que comprenden la lógica algorítmica e ingenieros que asimilan la complejidad del razonamiento clínico.

5. BIBLIOGRAFÍA

[1] Comisión Nacional de Evaluación y Acreditación Universitaria [CONEAU]. (2024). *Estándares de acreditación para las carreras de medicina*. Resolución Ministerial vigente. República Argentina.

[2] Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, *65*(9), S63-S67.

[3] Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10), S70-S81.

[4] Cook, D. A., & Ellaway, R. H. (2015). Virtual patients: A critical literature review and proposed next steps. *Medical Education*, *49*(1), 6-18.

[5] Thirunavukarasu, A. J., et al. (2023). Large language models in medicine. *Nature Medicine*, *29*, 1930–1940.

[6] Tu, T., et al. (2024). Towards conversational diagnostic AI. *arXiv*. Preprint. https://arxiv.org/abs/2401.05654

[7] Feigerlova, E., et al. (2025). Clinical Simulation with ChatGPT: A revolution in medical education? *Journal of CME*, *14*(1), 2525615.

[8] Masters, K. (2023). Ethical use of Artificial Intelligence in Health Professions Education: AMEE Guide No. 158. *Medical Teacher*, *45*(6), 574-584.

[9] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

[10] Hojat, M. (2016). *Empathy in health professions, education and patient care*. Springer.

[11] Ruiz-Moral, R., et al. (2015). *Comunicación clínica: Principios y habilidades para la práctica*. Editorial Médica Panamericana.

INTELIGENCIA ARTIFICIAL GENERATIVA (IAG): RELEVAMIENTO AL ESTADO DEL ARTE DE SOBRE LAS TRANSFORMACIONES PEDAGÓGICAS ACTUALES Y LAS PLATAFORMAS EDUCATIVAS

Soledad Ayala[1], Alejandro Sartorio[1], Alejandro Hernández[1], Juan Gaseli[1], Mateo Dip[1], Mateo Durán[1]

[1] Centro de Altos Estudios de Tecnología Informática (CAETI) Rosario / Facultad de Tecnología Informática / Universidad Abierta Interamericana (UAI) Rosario.

soledad.ayala@uai.edu.ar, alejandro.sartorio@uai.edu.ar, alejandromario.hernandez@uai.edu.ar, juanamado.gaseli@alumnos.uai.edu.ar, mateo.dip@alumnos.uai.edu.ar, mateo.duran@alumnos.uai.edu.ar

RESUMEN

Este artículo resume los primeros acercamientos a la construcción del objeto de estudio del proyecto de investigación Sitios web educativos e inteligencia artificial (IA): un análisis de los estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG)", que tiene sede en el CAETI Rosario. En la actualidad, se está llevando adelante la etapa de relevamiento de estudios previos y tomando decisiones ciertas decisiones teóricas. Una de ellas, por ejemplo, fue la de reemplazar sitios web por plataformas educativas; otra, la de incorporar estudios vinculados a las transformaciones pedagógicas y; por último, darle un lugar a la inteligencia artificial generativa (IAG), teniendo en cuenta el contexto actual. De esta manera, se presenta de manera breve un análisis al primer relevamiento sobre los tres ejes antes mencionados. Por otro lado, se recupera y se reafirma la posición epistemológica desde la cual se investiga: el relativismo tecnológico y una idea sobre la tecnología en tanto dispositivo socio-tecnológico-educativo. Próximamente, se empezarán a relevar estudios sobre los estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG) y de elegirán determinadas plataformas educativas para tomarlas como caso de estudio.

Palabras clave: estado del arte – plataformas educativas – transformaciones pedagógicas – inteligencia artificial generativa.

CONTEXTO DE LA INVESTIGACIÓN

Cabe aclarar, en primer lugar, que la investigación parte de un punto de vista constructivista y relativista de la tecnología, entendiéndola como un dispositivo socio-tecnológico y no una "mera herramienta". De esta manera, se aleja de miradas deterministas, lineales e instrumentalistas de la tecnología.

El trabajo que se presenta a continuación sintetiza las primeras aproximaciones al objeto de estudio que se está construyendo en el proyecto de investigación interdisciplinario iniciado en 2025 en el CAETI-UAI, orientado al análisis de "Sitios web educativos e inteligencia artificial (IA): un análisis de los

estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG)" y del impacto de la Inteligencia Artificial Generativa (IAGen). A fin de organizar mejor la investigación, se decidió abordarla en tres ejes: transformaciones pedagógicas, plataformización educativa e IAGen como tecnología socio-técnica. Esto nos permite generar una mirada amplia y una reflexión crítica sobre la temática a investigar. De esta manera, se partió de una revisión sistemática del estado del arte, con el objetivo de identificar debates actuales, enfoques teóricos predominantes y líneas de investigación emergentes en los tres ejes mencionados que son fundamentales. El primer eje se focaliza en relevar estudios previos que analicen las transformaciones que están ocurriendo en las prácticas de los docentes universitarios. En este sentido, nos preguntamos hasta qué punto y en qué aspectos el uso de la IAGen y de ciertas plataformas educativas modifican las prácticas docentes en la actualidad y, de ser así, cuáles son esos cambios y, si estos cambios, están contemplando o no los estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG). El segundo eje de análisis está puesto en las plataformas educativas y su uso en la educación superior. Si bien algunas plataformas son de software libre -como Moodle- y otras de software privativo -como Teams-, lo cierto es que desde 2020 las plataformas se han tornado en un actor fundamental en los procesos de enseñanza y de aprendizaje en la educación superior. Así, se entiende que las plataformas educativas deben ser entendidas en tanto dispositivos socio-tecnológicos, es decir, en tanto tecnologías digitales específicas, producto de un proceso complejo de interrelación entre factores sociales, técnicos, culturales, políticos, de intereses económicos y comunicacionales. (Ayala, 2020; Thomas, 2008; Zukerfeld, 2015). En otras palabras, las plataformas educativos -en reemplazo de la idea de sitios web educativos- están modificando gradualmente las prácticas educativas, generando no solo transformaciones pedagógicas; sino también modificaciones importantes en la referidas al modo en que se concibe el tiempo, el espacio, la presencia y la realización de actividades grupales. El tercer y último eje, está vinculado a la aparición y uso de la Inteligencia Artificial Generativa (en adelante IAGen). Aunque el inicio del desarrollo de este tipo de inteligencia se remonta a mediados del siglo pasado, desde la irrupción en 2023 de la herramientas OpenIA conocida como ChatGPT, ha "revolucionado" y puesto sobre la mesa debates tales como: si una máquina piensa y escribir por nosotros, entonces ¿cómo serán los rasgos de construcción del conocimiento reflexivo propio de la educación superior?, ¿de qué manera identificar y enseñar los sesgos y las ilusiones existentes en las respuestas de la IAGen?, ¿cómo pensar la accesibilidad a estas herramientas?, ¿de qué manera re configura las condiciones de acceso material y simbólico?

1. OBJETIVOS TEÓRICOS Y EXPERIMENTALES

Objetivo general:
Analizar críticamente el impacto de la IAGen y las plataformas educativas en las prácticas pedagógicas universitarias.

Objetivos específicos:
Identificar transformaciones reales en las prácticas docentes. Analizar los efectos de la plataformización en la organización de la enseñanza y la evaluación. Caracterizar la IAGen desde una perspectiva socio-técnica no determinista.
Construir un marco teórico crítico para futuras investigaciones empíricas.

El proyecto se basa en un análisis de "Sitios web educativos e inteligencia artificial (IA): un análisis de los estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG)". Actualmente se encuentra en su primera etapa, la cual está

focalizada en la revisión bibliográfica sistemática y análisis documental y en un primer relevamiento de sitios educativos web que tengan esas características.

2. FORMACIÓN DE RECURSOS RECURSOS HUMANOS

El proyecto contribuye a la formación de investigadores jóvenes y estudiantes universitarios mediante su participación en tareas de relevamiento bibliográfico, análisis crítico de fuentes, sistematización teórica y producción académica. Se promueve una alfabetización digital crítica orientada al uso reflexivo de tecnologías emergentes en educación. De hecho, el proyecto responde, a los intereses e inquietudes de los estudiantes co-autores del presente trabajo, a fin de que el espacio del mismo no sea solo un espacio de formación, sino de interés y conocimiento por la investigación.

3. RESULTADOS ESPERADOS

En esta primera etapa se espera la construcción de un mapa conceptual del estado del arte sobre las plataformas educativas, las transformaciones pedagógicas, la IAGen y los estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG). La elaboración de este estado del arte ayudará a identificar líneas teóricas dominantes; elaborar criterios para la integración pedagógica crítica de tecnologías; reconocer las condiciones reales de usabilidad, y generar bases teóricas para futuras investigaciones empíricas en diversos contextos. Si bien este es un primer acercamiento a los estudios sobre los ejes antes mencionados, la investigación evita la reproducción de discursos y miradas determinista; lo cual implica un análisis riguroso de los estudios encontrados. Ante esto, consideramos que uno de los principales desafíos actuales es recuperar y revalorizar los análisis de corte constructivistas, situados, coyunturales, que incorporen preguntas que inviten a un pensamiento crítico-reflexivo, de manera tal de evitar caer en una posición en la que predomine un adjetivo calificativo del tipo "bueno-malo" o miradas futuras basadas en catástrofes o meras opiniones carentes de base empírica. Es decir, es necesario empezar a hacer inteligible la complejidad que IAGen ha generado en relación a las transformaciones pedagógicas, el uso de plataformas y su propia configuración socio-tecnológicas.

Si bien cada uno de estos ejes tiene sus particularidades, podemos decir que forman parte de un "paraguas mayor" que los abarca y que responden a la necesidad de saber cómo funcionan las tecnologías digitales, si es que se están respetando los estándares de usabilidad establecidos en las Web Content Accessibility Guidelines (WCAG), y, por lo tanto, a la necesidad de construir procesos de alfabetización digital para docentes y estudiantes. Es decir, procesos de alfabetización digital que interrelacionen conocimientos sobre cómo funcionan los sistemas algorítmicos, los patrones de datos, el modo en que los mismos son recolectados, procesados y los efectos que esos análisis -en su mayoría cuantitativos- tienen en la toma de decisiones educativas, tanto en las políticas públicas y regulaciones como en las estrategias pedagógicas que se llevan a cabo en el aula. Sin estos procesos de alfabetización digital, se corre el riesgo de que las transformaciones pedagógicas que se realicen con la incorporación de la IAGen adquieran características instrumentales, que el conocimiento construido con el uso de las plataformas predomine la idea de "cajas negras" y que la IAGen sea vista como una "simple herramienta". Lo cual, en un futuro, inferimos que también afectaría a los estándares de usabilidad.

El abordaje del estado del arte sobre estos tres ejes, da cuenta de la importancia de continuar elaborando un mapa que permita identificar autores, ideas y miradas concretas, a fin de construir conocimiento crítico, amplio, pero, especialmente, para recordar el rol de la

educación en la construcción de ciudadanía y ciudadanía digital en lo que respecta a los cambios que se co-construyen con la aparición de nuevas tecnologías digitales y su convivencia con otras tecnologías.

4. REFERENCIAS BIBLIOGRÁFICAS

[1] Ayala, S. (2020). El reinado del papel: prácticas de lectura universitarias, un análisis desde la construcción social de la tecnología. 1a ed. Bernal: Universidad Nacional de Quilmes. Disponible online en: http://unidaddepublicaciones.web.unq.edu.ar/libros/el-reinado-del-papel-practicas-de-lectura-universitarias-un-analisis-desde-la-construccion-social-de-la-tecnologia/

[2] Morozov, E. (2011). El desengaño de internet, Los mitos de la libertad en la red. Ed. Destino, Barcelona, 2011, trad. Eduardo G. Murillo.

[3] Morozov, E., & Piñeiro, N. V. (2016). La locura del solucionismo tecnológico (1a. ed). Capital intelectual.

[4] Selwyn, N. (2019). ¿Deberían los robots sustituir al profesorado? La Inteligencia Artificial y el future de la educación. Ediciones Morata, S. L.

[5] Selwyn, N. (2022). Education and technology: Key issues and debates. Bloomsbury Academic.

[6] Srnicek, N. (2017). Platform capitalism (1st ed). Polity.

[7] UNESCO (2023). Guía para el uso de IA generativa en educación e investigación. Disponible online en: https://unesdoc.unesco.org/ark:/48223/pf0000389227

[8] Van Dijck, J.; Poell, T. & de Waal, M. (2018). The platform society. Public values in a connective world. New York: Oxford University Press. 240 páginas. ISBN: 9780190889777

[9] Williamson,B.; Trotta, L.; Atairo, D. (2023). Entrevista a Williamson. Implicancias de la transformación digital en la educación superior: privatización, comercialización y pérdida de soberanía. Política universitaria (9), 22-30. Disponible online en: https://www.memoria.fahce.unlp.edu.ar/art_revistas/pr.18829/pr.18829.pdf

[10] Winkler, R., y Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. En Academy of Management Annual Meeting (AOM). https://www.alexandria.unisg.ch/publications/254848

[11] Thomas, H., Lalouf, A., Garrido S. (2014) Carpeta de Trabajo. Estudios sociales de la ciencia y la tecnología. 1a ed. Bernal: Universidad Virtual de Quilmes. E-Book. ISBN 978-987-3706-25-7

[12] UNESCO (2023). Guía para el uso de IA generativa en educación e investigación.

[13] Zukerfeld, M. (2015). La tecnología en general, las digitales en particular. Vida, milagros y familia de la "Ley de Moore", en Revista Hipertextos, Vol. 2, N° 4, Buenos Aires, Julio/Diciembre de 2015.

Investigaciones emergentes sobre tecnologías digitales en escenarios educativos

Sanz Cecilia[1,2] [0000-0002-9471-0008], Artola Verónica[1] [0000-0003-0833-1077], Gonzalez Alejandro[1] [0000-0002-9854-8618], Zangara Alejandra[1] [0000-0002-9854-8618], Iglesias Luciano[1] [10000-0002-3284-6997], Ibáñez Eduardo[1] [0000-0002-1211-9370], Violini Lucía[1,3] [0000-0002-0537-5528], Fachal Adriana[1] [0000-0002-6379-2856], Del Gener Aldana[1] [0009-0008-0024-3703], Abásolo María José[1,2] [0000-0003-4441-3264], Dieser Paula[4] [0000-0002-2066-0183], Coma-Roselló Teresa[5] [0000-0002-7439-8459], Handula Lorenzo[1], Lombardo Sabrina[1], Segovia Maite[1], Pesado Patricia[1] [0000-0003-0000-3482]

[1] Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC. Facultad de Informática, Universidad Nacional de La Plata
[2] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
[3] Ha contado con beca UNLP
[4] Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa
[5] Departamento de Ciencias de la Educación, Universidad de Zaragoza, España

{csanz, vartola, agonzalez, eibanez, lviolini, aldanamdg, mjabasolo, ppesado}@lidi.info.unlp.edu.ar, li@info.unlp.edu.ar, {alejandra.zangara, paula.dieser, lorenzohandula, maitesegovia.a}@gmail.com, afachal@hotmail.com, tcoma@unizar.es, sabrina.lombardo@alu.ing.unlp.edu.ar

RESUMEN

La transformación digital se ha convertido en un tema crítico para la sociedad, ya que se entrama con todos los sectores, y en particular con el ámbito educativo. Este trabajo presenta las líneas de investigación, desarrollo e innovación del subproyecto "Creación de tecnologías digitales para el escenario educativo", que forma parte de uno de los proyectos acreditados del Instituto de Investigación en Informática LIDI (III-LIDI). Este subproyecto aborda temas críticos para la transformación digital en el escenario educativo. En el III-LIDI se cuenta con una amplia trayectoria de investigación en el área de tecnologías digitales y educación; en este caso, se detallan los resultados alcanzados en el proyecto durante el año 2025.

Palabras clave: trabajo colaborativo con tecnologías digitales, juegos serios, autorregulación del aprendizaje en entornos digitales, inteligencia artificial, materiales educativos digitales

CONTEXTO

El contexto de las líneas de I+D+i aquí presentadas es el subproyecto SP3: "Creación de tecnologías digitales para el escenario educativo". Este subproyecto forma parte de un proyecto acreditado del Instituto de Investigación en Informática LIDI (III-LIDI) de la Facultad de Informática de la UNLP, titulado: "Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital". En este contexto se abordan temas críticos relacionados con el área de tecnología educativa y transformación digital que van desde creación de metodologías para la mediación de diferentes procesos en el ámbito de la educación hasta el desarrollo de tecnologías digitales concretas y el análisis de su impacto.

1. INTRODUCCIÓN

Las tecnologías digitales forman parte del ecosistema que caracteriza a los escenarios educativos actuales. Desde la pandemia a causa del COVID-19, se ha visto un esfuerzo por parte de las instituciones educativas, especialmente las instituciones de educación superior, por iniciar/profundizar en procesos de transformación digital (Sandoval-Benavides y López-Ornelas, 2025).
En los estudios realizados en el marco de este proyecto se visualizan temas de innovación y tendencia, que abonan a procesos de transformación digital, tales como: los juegos serios (Ferreyra et al., 2026; Sanz y Artola, 2025) y las estrategias de gamificación (Sipone et al., 2025; Vera et al., 2025), la adopción incremental de modalidades educativas híbridas con mediación a través de entornos digitales (Sandoval-Benavides y López-Ornelas, 2025), el uso de herramientas de IAGen (García, Sanz y Artola, 2025; Chiu, 2024; Sanz, 2024), el trabajo colaborativo y su seguimiento en entornos digitales (Zheng, Shi y Gao, 2026; Vazquez Bermúdez, Sanz y Zangara, 2025a), los entornos personales de aprendizaje y redes sociales (Anantasopon y Chaveesuk, 2022) y el desarrollo de competencias digitales en docentes y estudiantes (Cosgrove y Cachia, 2025; Sandí Delgado y Sanz, 2020), entre otros temas. En el estudio bibliográfico y estado del arte que se desarrolla en el subproyecto se encuentran análisis teóricos y empíricos con

experiencias concretas en cada una de estas líneas. Avanzar sobre la investigación y abordar estas líneas de manera crítica resulta fundamental para quienes trabajan en la creación de tecnologías digitales. Además, en el III-LIDI se abordan metodologías de investigación-acción, de manera interdisciplinaria, con vinculación con diferentes ámbitos educativos, docentes y estudiantes.

Las líneas de investigación aquí presentadas guardan relación directa con las temáticas de la Maestría y Especialización en Tecnología Informática Aplicada en Educación de la misma Facultad, por lo que se desarrollan tesis y trabajos finales en estos temas. Se cuenta además con una agenda de cooperación a nivel nacional e internacional, que permite profundizar el trabajo que se lleva adelante en el III-LIDI.

2. LÍNEAS DE INVESTIGACIÓN / DESARROLLO

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- **Entornos digitales** para la mediación de procesos educativos: entornos virtuales de enseñanza y aprendizaje, entornos 3D, redes sociales y PLE, entornos virtuales gamificados. Funcionalidades y formas de acceso a estos espacios, trazabilidad de las actividades, estándares. Estrategias para el diseño e implementación de estos tipos de entornos. Potencialidades y barreras en la Integración de Inteligencia Artificial en estos entornos.
- **Materiales educativos digitales (MEDs) y actividades interactivas**. Metodologías para su diseño y producción. Objetos de aprendizaje. Multimedia e hipermedia en escenarios educativos. Nuevos entramados de medios, soportes y lenguajes. Estándares.
- **Integración de tecnologías digitales en procesos educativos**. Hibridación de las modalidades educativas. Diseño de MOOC. E-actividades. Actividades transmedia.
- **Inteligencia Artificial (IA) en Educación**, herramienta de IA generativa (IAGen).
- **Gamificación y Juegos Serios Educativos** con diferentes paradigmas de interacción. Metodologías para su creación. Objetos activos.
- **Trabajo colaborativo mediado por tecnologías digitales y su seguimiento**. *Mirroring*. Herramientas para promover la autorregulación y el desarrollo de capacidades metacognitivas. Conceptualización, análisis y desarrollo de software y metodologías. Actividades colaborativas aprovechando dispositivos móviles.
- **Computación afectiva en entornos digitales** para el escenario educativo.
- **Salas de escape educativas.**
- **Formación de recursos humanos** en el área de Informática. Alfabetización digital, pensamiento computacional. Desarrollo de la empatía, y la resiliencia académica en estudiantes de carreras en Informática.

Para llevar adelante estas líneas de I+D+i se cuenta con un equipo de docentes-investigadores, tesistas, becarios, pasantes y estudiantes que aportan al proyecto. Al mismo tiempo, se establecen vinculaciones con otros grupos de profesionales y docentes y estudiantes de diferentes niveles con los que se realizan estrategias de co-diseño y diseño participativo.

3. RESULTADOS OBTENIDOS

En esta sección se describen los resultados alcanzados durante el año 2025 y los avances en las líneas de I+D+i del subproyecto mencionado. A continuación se presentan los resultados organizados por temáticas.

Según Moschini y Sindoni (2022) la mediación se entiende como un proceso de comunicación durante el cual los participantes utilizan recursos semióticos basados en medios digitales para comunicar, crear y compartir conocimiento, representaciones y prácticas, generando cambios en los entornos sociales y culturales. Bajo esta concepción las tecnologías digitales configuran parte de las prácticas educativas en un diálogo con quienes las utilizan, y donde también subyace la visión y creencias de quiénes las crean. En este subproyecto se abordan estudios teóricos, diseños y desarrollos de entornos digitales que consideran las necesidades de la comunidad que los utiliza y aportan a la mediación de diferentes procesos educativos.

En 2025 se avanzó en el desarrollo y mantenimiento del entorno virtual de enseñanza y aprendizaje (EVEA), IDEAS. Se profundizó en los estudios para integrar un chatbot basado en IAGen (Lombardo & Sanz, 2024). En la actualidad se está finalizando una versión que permite responder preguntas a partir del entrenamiento con los propios materiales de lectura de un curso en IDEAS. También se está iniciando el trabajo de cambiar la tecnología subyacente de este entorno a las versiones más actuales de Symfony y php.

En relación a los EVEA, en 2025 se culminaron dos tesis doctorales dirigidas por miembros del subproyecto. Una relacionada con gamificación en EVEAs, con la propuesta de un modelo tecno-pedagógico para guiar procesos de gamificación. En el marco de esta tesis se desarrollaron dos casos de estudio siguiendo el modelo propuesto en la tesis, y utilizando el entorno Moodle (Vera Mora et al., 2025). La segunda tesis doctoral se enfocó en el desarrollo de una matriz de indicadores de seguimiento en procesos colaborativos mediados por herramientas y entornos digitales, una metodología de diseño de actividades colaborativas, y un dashboard que posibilita llevar adelante estrategias de *mirroring* (Vazquez Bermudez, Sanz y Zangara, 2025a; Vazquez Bermudez et al. 2025b). El *mirroring* refleja la actividad de los miembros de un grupo que trabaja en colaboración,

mediante la presentación de comentarios textuales y visualizaciones gráficas de las acciones individuales y grupales en tiempo real. Su propósito incluye aumentar la visibilidad de las contribuciones, promover la conciencia sobre el progreso del grupo y los roles individuales, y fomentar la responsabilidad en los esfuerzos colectivos (Zangara y Sanz, 2018). Es de destacar que se continúan implementando estrategias de *mirroring* en el marco de un curso de postgrado de la Maestría en Tecnología Informática Aplicada en Educación. Estas metodologías en conjunto con el diseño de actividades colaborativas conforman resultados concretos y transferidos de este subproyecto (Dieser, Sanz y Zangara, 2023).

Se finalizó y aprobó un trabajo de especialización que indaga sobre la resiliencia académica en EVEAs y su potencial impacto en el rendimiento académico. Cabe aclarar que la resiliencia académica es definida como la habilidad multidimensional de adaptarse y superar obstáculos manteniendo la motivación académica (Ragone, Rexachs del Rosario, y Sanz, 2025). Este trabajo amplía el abanico de estudios que se realizan en torno a los impactos de la mediación de procesos educativos con diferentes tecnologías digitales.

También se entregó y está en evaluación un trabajo de especialización que investiga sobre la formación de docentes en modalidades híbridas y con utilización de e-actividades (Ruiz Díaz et al., 2025).

En vinculación con la línea de integración de tecnologías digitales en entornos educativos y modalidades híbridas, se finalizó y está en evaluación una tesis de maestría que estudia la importancia del componente de diálogo en propuestas educativas mediadas por EVEAs (Spandre, Dieser y Sanz, 2022). También se avanzó en una tesis que estudia modalidad híbridas en disciplinas proyectuales (Díaz Reinoso y Sanz, 2023) y se complementa con otra tesis de maestría en la que se analizan la conformación de entornos personales de aprendizaje, y en particular el uso de redes sociales en este tipo disciplinas proyectuales (Jimenez, Artola, y Sanz, 2025).

Una de las líneas que ha cobrado relevancia en el proyecto se vincula con el desarrollo e integración de juegos serios en procesos educativos (Gubaro, 2025), y también el estudio de estrategias de gamificación. Se trabaja fuertemente en la investigación y aplicación, en diferentes contextos educativos, de este tipo de juegos y estrategias. Se crearon juegos serios orientados a temas de aprendizaje para la salud (Figura 1), tales como carrera de sellos[1] y guerra de alimentos. Este ecosistema de juegos y su metodología de diseño, creación e integración fue premiado en el marco de los *Best Practices Awards* del Programa Unita Geminae[2], en el año 2025, en el contexto de la Conferencia Geminae desarrollada en Zaragoza, España. Los juegos creados se encuentran en el Centro de Innovación y Transferencia Tecnológica (CIyTT) de la Facultad de Informática. En este contexto se reciben visitas y se llevan adelante talleres donde se utilizan estas tecnologías y se reflexiona con ellas en los contenidos que median. Se crean también entornos interactivos con experiencias de realidad virtual, realidad aumentada e interacción tangible. Es de destacar que se llevan a cabo estudios sobre las preferencias de los estudiantes por estos paradigmas de interacción y sus aportes en el aprendizaje (Sanz y Artola, 2025). Además, se estudian otros potenciales usos de los juegos serios como el caso de su aplicación para el desarrollo de competencias digitales (Sandí Delgado y Sanz, 2020), su rol en procesos de evaluación sigilosa analizando rasgos de personalidad (Ferreyra et al., 2026) y sus posibilidades como *exergames* para promover la actividad física y mejorar situaciones de sedentarismo tecnológico (Del Gener y Sanz, 2025; Del Gener, Sanz e Iglesias, 2023).

En el CIyTT funcionan también dos salas de escape creadas con fines educativos, y que son resultados de la investigación en este subproyecto. Se utilizan para promover metodologías de aprender haciendo, y constituyen actividades lúdico-educativas de gran interés. Al mismo tiempo, una de las salas ha sido utilizada para investigaciones en relación a temas de computación afectiva (Zeballos y Sanz, 2025).
Se destaca la importancia de la formación de recursos humanos en el marco de este subproyecto, que derrama en acciones de formación a través de juegos serios (Sanz, Artola y Baldassarri, 2024), entornos interactivos, y el trabajo que se realiza para el desarrollo del pensamiento computacional. La actividad del "Semillero de innovadores" (Sanz et al., 2024, los talleres con el juego serio Astrocódigo[3] para abordar habilidades fundamentales del pensamiento computacional y el desarrollo de un modelo para el trabajo de la empatía en estudiantes de carreras en Informática (Sanz et al., 2023) dan cuenta de estas acciones realizadas en el marco del subproyecto.
En temas vinculados a materiales educativos digitales, en diciembre de 2025, se entregó una tesis doctoral que aborda el desarrollo de un *framework* para la creación de objetos de aprendizaje (Violini, Sanz y Pesado, 2023).
Finalmente, se mencionan algunos resultados de los estudios sobre el uso y el impacto de la Inteligencia

[1] Video de presentación de Carrera de Sellos: http://163.10.22.82/OAS/Videos/carrerasellos.mp4

[2] https://univ-unita.eu/these-are-the-five-geminae-universities-recognised-for-their-best-practices-social-inclusion-innovation-and-sustainable-development-among-the-values-highlighted/

[3] Astrocódigo es un juego serio creado en el marco del proyecto en un trabajo de fin de carrera

Artificial en procesos educativos. Una propuesta de trabajo final de la Especialización en TIAE sobre las herramientas de IAGen en la formación docente (Zangara y Gonzalez, 2026). En 2025 se avanzó con una tesis de maestría en la que se investiga el uso de IAGen en la enseñanza de la matemática y en general, en el ámbito de la educación superior (García, Sanz y Artola, 2025).



Figura 1 – Ecosistema de juegos para el aprendizaje para la salud, en particular sobre alimentación saludable y nutrición

Cabe aclarar que todos estos resultados se desarrollaron en vinculación con las líneas aquí presentadas.

En cuanto a los proyectos vinculados y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se cuenta con un acuerdo de colaboración con la Universidad de Zaragoza y la Universidad de Islas Baleares.
- Se ha obtenido financiamiento en el marco de una convocatoria K171 de Erasmus+ para movilidades entre la Universidad de Zaragoza y la Universidad Nacional de La Plata en 2024. Se ha participado del Congreso de Unita Geminae 2025 en Zaragoza con una exposición.
- Se participa de la iniciativa CIN-UNIUEAR. Alianza Universitaria Argentina Europea para la Transformación Digital.
- Se participa del Proyecto "*Construction and validation of a methodology for digital transformation of localities in developing countries to promote their sustainable development*", aceptado en el marco de la convocatoria de proyectos de "Perez Guerrero Trust Fund (PGTF) 2023" para países miembro del grupo de 77. Participan de este proyecto Ecuador, Argentina y Cuba.
- Se participa en la RedAUTI: Red temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva.
- Algunos miembros son colaboradores del proyecto TEMOR: Ecosistema tecnológico para el reconocimiento del estado de ánimo de los pacientes durante el proceso de rehabilitación cardíaca. Proyecto de la Universidad de Zaragoza, aprobado en 2022.
- Se participa del proyecto: Experiencias Lúdicas con Agentes Sociales Interactivos y Robots: Aprendizaje Social y Comunicación Intergeneracional (PLEISAR) y su continuación. Dirigido por la Dra. Cerezo de la Universidad de Zaragoza.
- Se cuenta con los siguientes acuerdos de cooperación 2025: 1-con el Museo de Ciencias Naturales de La Plata; 2- Instituto Privado De Neuromedicina S.A. (IPNM) con quienes se trabaja en aplicaciones interactivas para adultos mayores.

4. FORMACION DE RRHH

Se trabaja en la formación de recursos humanos a través de la dirección de tesis de doctorado, maestría, trabajos de especialización y tesinas de grado; también con becas de diferentes organismos de ciencia y técnica. En 2025, se han finalizado 2 tesis doctorales, y otra entregada para su evaluación, dos tesis de maestría entregadas y en evaluación, 2 trabajos de especialización finalizados y otro evaluación.

5. BIBLIOGRAFÍA

- Anantasopon, P., & Chaveesuk, S. (2022). Uso de aplicaciones de redes sociales en el aprendizaje colaborativo activo para mejorar el rendimiento de los estudiantes en el aprendizaje. IEEE Access. https://ieeexplore.ieee.org/document/9850242
- Chiu, T. K. F. (2024). The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. Interactive Learning Environments, 32(10), 6187–6203. https://doi.org/10.1080/10494820.2023.2253861
- Zheng, L., Shi, Z., Gao, L. (2026). A generative artificial intelligence-enhanced multiagent approach to empowering collaborative problem solving across different learning domains, Computers & Education, Volume 241,105489, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2025.105489.
- Centro de Innovación y Transferencia Tecnológica, FI, UNLP, https:/ciytt.info.unlp.edu.ar/
- Cosgrove, J. & Cachia, R., DigComp 3.0: European Digital Competence Framework - Fifth Edition, Publications Office of the European Union, Luxembourg, 2025, https://data.europa.eu/doi/10.2760/0001149, JRC144121.
- Díaz Reinoso, V. C., & Sanz, C. V. (2023). Revisión Sistemática de Literatura de estudios sobre modalidades híbridas educativas en Disciplinas Proyectuales. In XVIII Congreso Nacional de Tecnología en Educación y Educación en Tecnología-TE&ET 2023 (Hurlingham, 15 y 16 de junio de 2023).
- Dieser, M. P., Sanz, C. V., Zangara, A. (2023), Autorregulación del aprendizaje, estrategias de participación y rendimiento académico. Un estudio de su relación en comunidades virtuales de indagación. Tesis de Maestría en TIAE. Disponible en: https://sedici.unlp.edu.ar/handle/10915/158434

- Ferreyra, P., Dieser, P., Sanz, C., Coma-Roselló, T. (2026). Evaluación del rasgo responsabilidad a través de un videojuego serio de química y matemática. Revista Iberoamericana De Tecnología En Educación Y Educación En Tecnología, 42, e3. https://doi.org/10.24215/18509959.42.e3
- Del Gener A., Sanz C., & Iglesias L. (2023). Propuesta de un Gamepad para Sensar Movimientos del Jugador y su Integración a un Exergame. Revista AIPO, 4(2), 68-77.
- Del Gener, A. M., & Sanz, C. V. (2025). *Development of a Hand Motion Sensing Glove for Exergames: Design Evolution and Future Perspectives*. In XIII Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 24 al 26 de junio de 2025).
- García, P. J., Sanz, C. V., & Artola, V. (2025). Revisión sistemática sobre inteligencia artificial generativa y sus posibilidades de uso en educación universitaria. In XX Congreso de Tecnología en Educación y Educación en Tecnología (TE & ET)(UNDAV, 12 y 13 de junio de 2025).
- Gubaro, M. (2025). Juegos Serios Educativos digitales orientados a la enseñanza de la Matemática en educación secundaria.
- Jimenez, S., Artola, V. & Sanz, C. (2025). 3300-3860/21-001. Propuesta de Tesis de Maestría en Tecnología Informática Aplicada en Educación titulada "Usos y posibilidades que ofrecen las redes sociales en el aprendizaje de estudiantes de Educación superior. El caso de alumnos de una asignatura Proyectual".
- Lombardo, S, & Sanz, C. (2024). “Aportes al desarrollo y actualización del entorno virtual de enseñanza y aprendizaje Ideas. Integración de un chatbot para estudiantes y docentes”. Informe final de Práctica profesional supervisada de la carrera Ing. en Computación. Fac. Informática. UNLP
- Moschini, I. & Sindoni, M. G. (Eds.). (2022). Mediation and Multimodal Meaning Making in Digital Environments. Routledge.
- Ragone, M. I., Rexachs del Rosario, D. I., & Sanz, C. V. (2025). Análisis de la utilización de entornos virtuales de enseñanza y aprendizaje y su relación con la resiliencia de estudiantes universitarios. In XX Congreso de Tecnología en Educación y Educación en Tecnología (TE & ET)(UNDAV, 12 y 13 de junio de 2025).
- Ruiz Díaz, S., Sanz, C., Artola, V. y Violini, L. (2025). El desarrollo de las competencias digitales de los estudiantes del profesorado a partir de propuestas de aula extendida, mediante el uso de eactividades” Informe final del trabajo de Especialización en Tecnología Informática Aplicada en Educación. FI. UNLP
- Sandí Delgado, J. C. S., & Sanz, C. V. (2020). Juegos serios para potenciar la adquisición de competencias digitales en la formación del profesorado/Serious Games to Enhance Digital Competencies Acquisition for Training Faculty. Educación, 44(1)
- Sanz, C. V., & Artola, V. (2025). Experiencias con juegos y aplicaciones educativas usando diferentes paradigmas de IPO con escuelas secundarias. In XX Congreso de Tecnología en Educación y Educación en Tecnología (TE & ET)(UNDAV, 12 y 13 de junio de 2025).
- Sanz, C. (2024) La Inteligencia Artificial viste a la moda. Reflexiones sobre sus posibilidades para el escenario educativo. Revista Iberoamericana De Tecnología En Educación Y Educación En Tecnología, 37, e1. https://doi.org/10.24215/18509959.37.e1
- Sanz, C., Coma-Rosello, T., Aguelo, A., Álvarez, P., & Baldassarri, S. (2023). Model and methodology for developing empathy: an experience in computer science engineering. IEEE Transactions on Education, 66(3), 287-298.
- Sipone S., Abella, V., Rojo, M. & Moura, J.L. (2025). Gamifcation and player type: Relationships of the HEXAD model with the learning experience Journal of New Approaches in Educational Research (2025) 14:1 https://doi.org/10.1007/s44322-024-00021-w
- Spandre, O., Dieser, P., & Sanz, C. (2022). Analyzing Attendance Indicators in Digital Technologies-Mediated Courses in Tertiary Technician Education. In Argentine Congress of Computer Science (pp. 47-58). Cham: Springer Nature Switzerland.
- Sandoval-Benavides, V. L., & López-Ornelas, M. (2025). Transformación digital en la educación superior desde la perspectiva internacional: mapeo sistemático de la literatura. Texto Livre, 18, e51996.
- Sanz, C. V., Artola, V., & Baldassarri, S. (2024). Developing educational serious games as a proposal for university training in the area of Child Computer Interaction. International Journal of Child-Computer Interaction, 41, 100672.
- Sanz, C., Artola, V., Medina, S., Zeballos, M., & Lombardo, S. (2024). Crear con tecnologías digitales para el desarrollo del pensamiento computacional. El caso del proyecto de semillero de innovadores.
- Vazquez Bermúdez, M., Sanz, C., y Zangara, A. (2025a). Desarrollo de un modelo de mirroring con indicadores y herramientas para el seguimiento y visualización de procesos educativos colaborativos mediados por tecnología digital en educación superior. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, (42), e24-e24.
- Vásquez-Bermúdez, M., Sanz, C., Zangara, M.A., Hidalgo, J. (2025b). Evaluation of the Tool Visualization of Interaction Tracking in Collaborative Work in a Forum Activity in Moodle. In: Valencia-García, R., Borodulina, T., Del Cioppo-Morstadt, J., Moran-Castro, C.E., Vera-Lucio, N. (eds) Technologies and Innovation. CITI 2024. Communications in Computer and

Information Science, vol 2276. Springer, Cham. https://doi.org/10.1007/978-3-031-75702-0_15

- Vera Mora, G. R., Sanz, C., Coma-Roselló, T & Baldassarri, S. (2025). Propuesta de un modelo de gamificación para entornos virtuales de enseñanza y aprendizaje en Instituciones de Educación Superior y su impacto en la presencia social, cognitiva y docente en dichos entornos (Tesis doctoral, Universidad Nacional de La Plata). https://doi.org/10.3390/educsci14080907
- Violini, M. L., Sanz, C. V., & Pesado, P. M. (2023). Proceso de diseño y desarrollo del framework MarCOA para la creación de Objetos de Aprendizaje. In XVIII Congreso Nacional de Tecnología en Educación y Educación en Tecnología-TE&ET 2023 (Hurlingham, 15 y 16 de junio de 2023).
- Zangara, A., & Sanz, C. (2018). Interacción e interactividad en el trabajo colaborativo mediado por tecnología informática (Doctoral dissertation, Universidad Nacional de La Plata).
- Zangara, A. & Gonzalez A. (2026). Propuesta de trabajo final de la Especialización en TIAE sobre las herramientas de IAGen en la formación docente.
- Zeballos, M., & Sanz, C. (2025). Informe final Práctica Profesional Supervisada: “Prototipo funcional de un firmware de registro de sensores fisiológicos para el análisis de emociones. El caso de una sala de escape inmersiva y educativa

La Tecnología Informática como mediadora del lenguaje oral en el nivel inicial: una experiencia de robótica educativa con Blue-Bot

Autor: Licenciado Roberto Nelson Rolón. Correo Electrónico: rnrolon@gmail.com

Carrera Licenciatura en Tecnología Educativa. Facultad Regional Resistencia. Universidad Tecnológica Nacional.

RESUMEN:

La investigación se desarrolló en el nivel inicial y se inscribió en el campo de la tecnología informática aplicada a la educación, a partir del análisis de una experiencia de robótica educativa orientada al desarrollo del lenguaje oral. Desde un enfoque cualitativo, se controlará el uso de un dispositivo programable como mediador pedagógico, priorizando las interacciones, la verbalización y la participación de los niños en contextos lúdicos mediados por la tecnología. Los resultados evidenciaron que la robótica educativa favoreció la expresión oral y la construcción colectiva de significados, abriendo nuevas líneas de reflexión sobre su potencial en las prácticas educativas.

PALABRAS CLAVES:

Tecnología informática aplicada a la educación, Robótica educativa, Lenguaje oral, Nivel inicial, Pensamiento computacional temprano

CONTEXTO:

La investigación se desarrolló en el nivel inicial y se inscribió en el campo de la tecnología informática aplicada a la educación, a partir del análisis de una experiencia de robótica educativa orientada al desarrollo del lenguaje oral. Desde un enfoque cualitativo, se controlará el uso de un dispositivo programable como mediador pedagógico, atendiendo a los procesos de verbalización, interacción entre pares y participación activa que emergieron en contextos lúdicos. La propuesta fue implementada por el docente del grupo, mientras que el investigador asumió un rol de observador, lo que permitió registrar las prácticas y respuestas infantiles sin intervención directa. Los resultados evidenciaron que la robótica educativa generó condiciones favorables para la expresión oral, abriendo proyecciones sobre su potencial pedagógico en el nivel inicial.

INTRODUCCIÓN:

Objetivos teóricos

1. Analizar el desarrollo del lenguaje oral en el nivel inicial desde una perspectiva pedagógica y sociocultural, considerando el aporte de la tecnología informática aplicada a la educación como marco conceptual para la incorporación de propuestas innovadoras.

2. Fundamentar la robótica educativa como una derivación pedagógica de la tecnología informática, reconociendo su potencial para favorecer la interacción, la comunicación oral y la construcción colectiva de significados en contextos educativos tempranos.

Objetivos experimentales

1. Implementar una propuesta didáctica mediada por tecnología informática aplicada a la educación, mediante el uso del robot Blue-Bot, orientada a promover la expresión oral y la interacción verbal de niños de cinco años del nivel inicial.

2. Observar y analizar las interacciones verbales producidas durante las actividades de robótica educativa, identificando el rol mediador de la docente y los efectos de la participación activa en el enriquecimiento del lenguaje oral.

El desarrollo del lenguaje oral en los niños del nivel inicial fue reconocido como un aspecto central para su crecimiento integral, en tanto posibilitó la comunicación, fortaleció las habilidades sociales y sentó las bases para futuros aprendizajes escolares. Diversos estudios coincidieron en la importancia de estimular esta capacidad desde las primeras etapas de la vida, considerando que los niños se encontraron predispuestos a adquirir el lenguaje desde el nacimiento, tal como señalaron Rodríguez Valeriano (2020) y Lybolt y Gottfred (2003).

En este marco, la literatura especializada destacó el papel fundamental que desempeñaron los adultos en el desarrollo lingüístico infantil. Investigaciones como las de Vernon Alvarado (2014) evidenciaron que la interacción temprana con padres y docentes influyó de manera directa en la adquisición del lenguaje del oral y en el desarrollo de habilidades sociales y cognitivas, especialmente en contextos educativos iniciales.

No obstante, estas oportunidades de estimulación no siempre estuvieron garantizadas. En contextos rurales, como el Jardín N.º 44 de la provincia de Corrientes, las condiciones socioeconómicas de las familias limitan, en muchos casos, las interacciones verbales enriquecedoras en el hogar. Esta realidad planteó desafíos significativos para el desarrollo del lenguaje oral de los niños, reforzando la necesidad de intervenciones pedagógicas situadas y contextualizadas.

Frente a este escenario, la institución incorporó propuestas innovadoras vinculadas a la tecnología informática aplicada a la educación, mediante la implementación de robótica educativa con el uso del robot Blue-Bot. Esta estrategia permitió integrar el juego, la programación y la interacción grupal como medios para favorecer la expresión oral. Estudios como los de Méndez Porras (**2021**) y colaboradores señalan que la robótica educativa incrementó la motivación, el rendimiento académico y el desarrollo de habilidades comunicativas, mientras que investigaciones de Palma (2017) y Bernabel León (2019) alertaron sobre la alta presencia de dificultades en el lenguaje oral en niños preescolares.

En consonancia con las políticas educativas nacionales y provinciales que promovieron la integración de tecnologías digitales, programación y robótica en el nivel inicial, la experiencia desarrollada se propuso analizar de qué manera el uso del Blue-Bot contribuyó a enriquecer el lenguaje oral de los niños de cinco años. A partir de observaciones no participantes, se evidencia que la participación activa en actividades mediadas por tecnología informática, junto con el acompañamiento pedagógico de la docente, generó condiciones favorables para el diálogo, la verbalización y la construcción compartida de significados, abriendo nuevas líneas de reflexión sobre el

potencial de estas propuestas en contextos educativos diversos.

LÍNEAS de INVESTIGACIÓN y DESARROLLO:

El análisis del lenguaje oral en el nivel inicial abrió diversas líneas de investigación y desarrollo que articularon los aportes teóricos clásicos con enfoques contemporáneos vinculados a la tecnología informática aplicada a la educación. Desde una perspectiva lingüística integral, el lenguaje oral fue comprendido como un sistema complejo que involucró dimensiones fonológicas, morfosintácticas, semánticas y pragmáticas, tal como lo definió Bernabel León (2019), lo que permitió abordar su desarrollo desde una mirada multidimensional y contextualizada.

Una primera línea de investigación se vinculó con las teorías del desarrollo del lenguaje y su relación con los procesos cognitivos y sociales. Los aportes de Jean Piaget, recuperados por Pérez F. (2009), permitieron comprender la transición del lenguaje egocéntrico al lenguaje socializado como un proceso estrechamente ligado a la interacción con el entorno. En este sentido, el lenguaje fue concebido no solo como una manifestación del desarrollo cognitivo, sino también como una herramienta fundamental para la construcción del conocimiento colectivo.

Desde un enfoque sociocultural, los planteos de Lev Vygotsky, analizados por Yépez (2021), refuerzan la idea de que el lenguaje se desarrolló a partir de la interacción social y el contexto cultural. Esta perspectiva habilitó una línea de investigación centrada en el valor de las experiencias comunicativas compartidas y en el rol del entorno educativo como mediador del desarrollo lingüístico, especialmente en contextos con limitadas oportunidades de estimulación verbal.

Otra línea relevante se relacionó con los enfoques innatistas del lenguaje, a partir de la teoría de Noam Chomsky, también retomada por Yépez (2021). La noción de un dispositivo innato de adquisición del lenguaje permitió comprender la predisposición biológica de los niños para estructurar el lenguaje, destacando la necesidad de entornos ricos en estímulos que potencien dicha capacidad. Esta perspectiva dialogó con los enfoques pedagógicos que promovieron intervenciones tempranas y significativas.

En este entramado teórico, surgió una línea de desarrollo vinculada a la robótica educativa como recurso pedagógico innovador. Los aportes de Cebrián de la Serna (2021) y Quiroga (2017) permitieron comprender la robótica como un entorno de aprendizaje interdisciplinar que promovió el pensamiento lógico, el trabajo colaborativo y el pensamiento computacional. En el nivel inicial, estas propuestas se articularon con el desarrollo del lenguaje oral al generar situaciones de diálogo, explicación, anticipación y verbalización de procesos.

Finalmente, se consolidó una línea de investigación orientada a la integración curricular de la tecnología digital y la robótica en el nivel inicial, en consonancia con los lineamientos normativos y curriculares vigentes. El uso de kits de robótica educativa, como el Blue-Bot, se presentó como una oportunidad para enriquecer el lenguaje oral mediante actividades lúdicas y significativas, favoreciendo la ampliación del vocabulario, la interacción social y la expresión verbal. Estas líneas abrieron nuevas posibilidades para futuras investigaciones que profundicen en la articulación entre lenguaje, tecnología y educación desde edades tempranas, dejando planteados nuevos interrogantes y horizontes de desarrollo en el campo educativo.

Resultados obtenidos

Los resultados obtenidos evidenciaron que la integración de la tecnología informática, a través de un dispositivo de robótica educativa programable, generó escenarios de aprendizaje que favorecieron la participación activa de los niños y la emergencia de múltiples situaciones comunicativas en el aula. Durante el desarrollo de las propuestas, se observaron instancias en las que los niños verbalizaron acciones, anticiparon recorridos y compartieron decisiones, poniendo en juego el lenguaje oral de manera espontánea y significativa. La interacción con el dispositivo tecnológico promovió dinámicamente el intercambio entre pares, en las que la palabra tomó un lugar central para explicar, acordar y resolver situaciones planteadas. Estas prácticas estuvieron atravesadas por nociones propias del pensamiento computacional temprano, como la secuenciación y la organización de acciones, que se integraron de forma natural a las actividades lúdicas sin desplazar el carácter expresivo del nivel inicial. Asimismo, se identificó que la mediación docente resultó clave para orientar el uso pedagógico de la tecnología informática, favoreciendo espacios de diálogo y reflexión compartida. La observación permitió reconocer que el dispositivo tecnológico actuó como un catalizador de la comunicación oral, abriendo posibilidades para repensar el lugar de la informática educativa en las prácticas del nivel inicial y su potencial para enriquecer los procesos de enseñanza y aprendizaje desde una perspectiva integral.

Formación de recursos humanos

La experiencia desarrollada aportó elementos significativos para reflexionar sobre la formación de recursos humanos en el campo de la tecnología informática aplicada a la educación, especialmente en el nivel inicial. La implementación de propuestas mediadas por robótica educativa puso de manifiesto la importancia de contar con docentes capacitados no solo en el uso técnico de los dispositivos tecnológicos, sino también en su integración pedagógica, atendiendo a los procesos comunicativos y expresivos propios de la infancia. Durante la experiencia, el rol docente se configuró como un factor clave en la mediación entre la tecnología informática y las dinámicas del aula. La planificación de actividades, la formulación de consignas y el acompañamiento de las interacciones evidenciaron la necesidad de una formación que articule saberes pedagógicos, didácticos y tecnológicos, permitiendo resignificar el uso de dispositivos programables en contextos educativos tempranos. Asimismo, el estudio permitió visibilizar la relevancia de generar espacios de formación continua y de reflexión profesional orientados a la informática educativa, que favorecen el intercambio de experiencias y la construcción colectiva de conocimientos. En este sentido,la incorporación de la tecnología informática en el nivel inicial se presenta como una oportunidad para fortalecer el desarrollo profesional docente y para abrir nuevas líneas de formación que acompañan los desafíos emergentes de la educación contemporánea.

BIBLIOGRAFÍA:

Bernabel León, L. (2019). *El lenguaje oral en niños de cinco años de una institución educativa inicial pública del Callao.* Universidad San Ignacio de Loyola.

Cebrián de la Serna, M., Huerta Gaytán, P., Flores Almendárez, J. M., González Fernández, M. O., Gómez Rodríguez, H., Ruíz Rey, F. J., ... & Cebrián Robles, V. (2021). *Robótica Educativa. Una perspectiva didáctica en el aula.*

Lybolt, J., & Gottfred, C. H. (2003). Cómo fomentar el lenguaje en el nivel preescolar. *Traducido y editado por: UNESCO (2006). Suiza*.

Méndez-Porras, A., Alfaro-Velasco, J., & Rojas-Guzmán, R. (2021). Videojuegos educativos para niñas y niños en educación preescolar utilizando robótica y realidad aumentada. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E42), 482-495.

Palma Mallma, C. (2017). *Lenguaje oral y comprensión lectora en los niños y niñas de 5 años de la Institución Educativa Inicial Nº 107 Los Rojitos Huancavelica, 2016*.

Pérez, F. (01 de agosto de 2009). ETAPAS DEL DESARROLLO DEL LENGUAJE SEGÚN PIAGET. Recuperado el noviembre de 2018, de https://es.scribd.com/doc/17922603/ETAPAS-DEL-DESARROLLO-DEL-LENGUAJE-SEGUN-PIAGET

Quiroga, L. P. (2017). La robótica educativa y la Educación en Preescolar. *Revista educación y pensamiento*, *24*(24).

Rodríguez Valeriano, M. B. (2020). *La valoración del lenguaje oral en los niños de educación inicial.*

Vernon, S. A., y Alvarado, M. (2014). Aprender a escuchar, aprender a hablar. La lengua oral en los primeros años de escolaridad. *Materiales para Apoyar la Práctica Educativa*. México: INEE. https://www.inee.edu.mx/wp-content/uploads/2019/01/P1D417.pdf

Yépez, E. R., & Álvarez, G. C. P. (2021). La oralidad y las dimensiones del lenguaje en los infantes. *Revista Cognosis. ISSN 2588-0578*, *6*(EE-I-), 01-22.

Modelado y Automatización de Procesos para el Aseguramiento de la Calidad Universitaria bajo ISO 21001:2018

Fernanda CARMONA[1], Emanuel Agustín IRRAZABAL[2], Mara Lía Inés ROVERO[1], Ana Carolina MANRIQUE [1], Paula María RIVAS [2], Nelson Javier CORDOBA VILTE[1], Amadeo Jesús CLARAMONTE [1]

[1]Departamento de Ciencias Básicas y Tecnológicas
Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina; [2] Grupo de Investigación en Calidad de Software, Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, 25 de Mayo 868, Corrientes, Argentina

{fbcarmona, mrovero, amanrique, privas }@undec.edu.ar, {emanuelirrazabal, javiercordoba0899, claramontejesus}@gmail.com

RESUMEN

Gestionar la Calidad implica un conjunto de acciones y herramientas coordinadas para dirigir y controlar una organización de acuerdo a ciertos requisitos y estándares. Respecto a la calidad en la educación superior existen varias acepciones, suele definirse como el grado en el que un conjunto de rasgos diferenciadores inherentes a la educación superior cumple con una necesidad o expectativa establecida, también como la propiedad de una institución o programa que cumple los estándares previamente establecidos por una agencia u organismo de acreditación. La norma ISO 21001:2018 alineada con la norma ISO 9001 que promueve la adopción de un enfoque a procesos al desarrollar, implementar y mejorar la eficacia de un Sistema de Gestión de Organizaciones Educativas, para aumentar la satisfacción del estudiante y de otros beneficiarios mediante el cumplimiento de los requisitos de estos. El crecimiento y desarrollo de la UNdeC, requiere de un proceso de identificación de las actividades desarrolladas y de los procesos relacionados con el objeto de implementar acciones preventivas, correctivas y de análisis para la satisfacción de los destinatarios finales de los servicios que la institución brinda. Esta línea de investigación tiene por objeto el análisis y estudio de los requisitos definidos por la norma ISO 21001:2018 para la definición y desarrollo de estrategias que permitan implementar la mejora continua del Sistema de Aseguramiento de la Calidad de la UNdeC.

Palabras clave: Gestión de Calidad, Enfoque a procesos, Educación Superior, Normas ISO.

CONTEXTO

La Ley N° 24.521 de Educación Superior establece que la autoevaluación institucional y la evaluación externa apuntan a la mejora y el aumento en la calidad del funcionamiento de las universidades y del cumplimiento de sus propósitos. Dispone que las instituciones universitarias deben asegurar el funcionamiento de instancias internas de evaluación institucional, que tendrán por objeto analizar los logros y dificultades en el cumplimiento de sus funciones, así como sugerir medidas para su mejoramiento; instaura la tutela del bien público en actividades que puedan ser afectadas por intervenciones profesionales, estableciendo para ello mecanismos de acreditación periódicas; entre otras funciones propias de las instituciones que conforman la Educación Superior, establece la de garantizar crecientes niveles de calidad y

excelencia en todas las opciones institucionales del sistema.

En este contexto desde el año 2005 la UNdeC viene transitando diversos procesos de acreditación de carreras y en el 2021 firma el acuerdo con CONEAU para implementar el proceso de evaluación institucional actualmente finalizado.

Con el objeto de capitalizar la experiencia adquirida en estos procesos y de sistematizar los procedimientos y aprendizajes de futuros procesos, en abril del año 2018 la UNdeC crea la Dirección de Evaluación y Acreditación, RR Nº 203/18, con la responsabilidad primaria de asegurar el adecuado desarrollo de las instancias de evaluación institucional y de acreditación de carreras, con el fin de garantizar la calidad de la oferta académica y de la gestión institucional y su mejora continua.

La Resolución 2597/2023 del ME establece que las instituciones universitarias deberán garantizar, en cada una de ellas y en el marco de sus pautas organizativas y de gobierno, el funcionamiento de sistemas institucionales de aseguramiento de la calidad (SIAC) a cargo de las unidades de gestión correspondientes. Los SIAC serán, sin perjuicio de otras funciones que la institución les asigne, responsables de promover y/o fortalecer la cultura de evaluación de calidad; administrar y desarrollar las estrategias, instrumentos y procesos de evaluación institucional; realizar la certificación de carreras; promover planes de mejora acordes con las evaluaciones realizadas y en el marco del respectivo plan de desarrollo institucional.

El logro de estos requerimientos implica perfeccionar mecanismos y dispositivos de autorregulación de las instituciones universitarias en sus procesos de mejora de la calidad. Es así que resulta recomendable la institucionalización de procesos, dispositivos y mecanismos estables, orgánicos, regulares y sustentables para la evaluación, planeamiento y coordinación de los procesos de mejora en el cumplimiento de las funciones institucionales, conforme a especificaciones y estándares de calidad nacional e internacional.

En Argentina, la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU) es el organismo descentralizado que funciona con la finalidad de contribuir al mejoramiento de la educación universitaria.

La norma ISO 21001:2018 es la primera norma internacional orientada a la especificación de los requisitos de un Sistemas de Gestión de Organizaciones Educativas (SGOE).

La UNdeC como Institución de Educación Superior se encuentra con el desafío de responder a las nuevas exigencias y competencias relacionadas con la sociedad del conocimiento y con ello a las nuevas actividades pedagógicas y a los requerimientos de acciones más eficientes en la gestión institucional.

El mejoramiento continuo requiere de estrategias, herramientas, planificación de acciones y de despliegue de recursos para el logro eficiente de sus objetivos institucionales. El aseguramiento de la calidad, es la acción de garantizar la calidad de los procesos y resultados en función de criterios y estándares preestablecidos, siendo la base para asegurar la estabilidad del proceso educativo. [7]

La Resolución Rectoral UNdeC N° 777/24 aprueba el financiamiento y ejecución del proyecto “Especificación de procesos bajo la norma ISO 21001:2018 para el fortalecimiento del sistema de aseguramiento de la calidad de la UNdeC.” presentado en la convocatoria UNdeC - Financiamiento para estímulo y desarrollo de la Investigación Científica y Tecnológica, período de ejecución 2024-2025. Este proyecto tiene por objeto el análisis y estudio de los requisitos definidos por la norma ISO 21001:2018 para especificar los procesos

del sistema de aseguramiento de la calidad de la UNdeC con el fin de desarrollar, implementar y mejorar su eficacia y eficiencia.

INTRODUCCIÓN

La ISO 9000 define la calidad como todas las características de un producto o servicio que satisfacen los requisitos del cliente y cumplen con todos los reglamentos y políticas aplicables. [5]

Las organizaciones buscan gestionar la calidad por medio de la implementación de estándares de calidad que aseguren su rentabilidad y la satisfacción del cliente involucrando los esfuerzos de todo el personal y utilizando metodologías que permiten conocer las necesidades particulares del cliente y determinar su percepción de la calidad para traducirlas en especificaciones de las características que debe cumplir el producto o el servicio que se ofrece. En este sentido, la familia de estándares de calidad más utilizada en las administraciones públicas y privadas es la de las normas ISO 9000 que incluye las siguientes: ISO 9000:2015, Sistemas de gestión de la calidad, Fundamentos y vocabulario; ISO 9001:2015, Sistemas de gestión de la calidad, Requisitos; ISO 9002:2016, Sistemas de gestión de la calidad, Directrices para la aplicación de la Norma ISO 9001:2015; ISO 9004:2018, Gestión de la calidad, Calidad de una organización, Orientación para lograr el éxito sostenido.

Gestionar la Calidad implica la definición de un conjunto de acciones y herramientas coordinadas para dirigir y controlar una organización con respecto a la calidad. Según la Norma ISO 9000: 2015 "la Gestión de Calidad puede incluir el establecimiento de las políticas y objetivos de calidad junto con los procesos para lograr estos objetivos de la calidad a través de la planificación de la calidad, el aseguramiento de la calidad, el control de la calidad y la mejora de la calidad"[4]. La Gestión de Calidad incluye:

- Planificación de la calidad: orientada a establecer los objetivos de la calidad y la especificación de los procesos operativos y los recursos relacionados necesarios para lograr los objetivos de la calidad.
- Aseguramiento de la calidad: orientada a proporcionar confianza en que se cumplirán los requisitos de calidad.
- Control de calidad: orientada al cumplimiento de los requisitos de la calidad.
- Mejora de la calidad: orientada a aumentar la capacidad de cumplir con los requisitos de la calidad.

Por su parte, un Sistema de Gestión de Calidad (SGC) comprende actividades mediante las cuales la organización identifica sus políticas, objetivos y determina los procesos y recursos requeridos para lograr sus objetivos. A su vez, gestiona los procesos posibilitando a la alta dirección optimizar el uso de los recursos considerando sus decisiones a largo y corto plazo y abordar las consecuencias previstas y no previstas en la provisión de productos y servicios.[4]

Respecto de la calidad en la educación superior existen varias acepciones, suele definirse como el grado en el que un conjunto de rasgos diferenciadores inherentes a la educación superior cumple con una necesidad o expectativa establecida, se define además como la propiedad de una institución o programa que cumple los estándares previamente establecidos por una agencia u organismo de acreditación. [10]

La evaluación de la calidad en la educación universitaria constituye uno de los temas claves en la agenda de reformas que se comienzan a implementar en este sector hacia mediados de la década del '80 en casi todas las regiones del mundo. En este contexto, en casi todos los países latinoamericanos han surgido agencias públicas o privadas de evaluación cuyos objetivos generales son asegurar y mejorar la calidad de la formación universitaria. En el caso de Argentina, la CONEAU cumple la misión institucional de asegurar y mejorar la calidad de

las carreras e instituciones universitarias que operan en el sistema universitario argentino por medio de actividades de evaluación y acreditación de la calidad de la educación universitaria.

Para lograr una mejora continua, las instituciones de educación deben reconocer a los estudiantes, como la razón principal de su existencia y aplicar políticas que mantengan relaciones positivas con estos, el personal académico, administrativo y todos los miembros que conforman la universidad. [11]

La norma ISO 21001:2018 es una norma de sistema de gestión independiente, alineada con la Norma ISO 9001 que promueve la adopción de un enfoque a procesos al desarrollar, implementar y mejorar la eficacia de un Sistema de Gestión de Organizaciones Educativas (SGOE), para aumentar la satisfacción del estudiante y de otros beneficiarios mediante el cumplimiento de los requisitos de estos.

La comprensión y gestión de los procesos interrelacionados como un sistema contribuye a la eficacia y eficiencia de la organización en el logro de sus resultados previstos. El enfoque a procesos implica la definición y gestión sistemática de los procesos y sus interacciones, con el fin de alcanzar los resultados previstos de acuerdo con la política, los objetivos y el plan estratégico de la organización. La aplicación del enfoque a procesos en un SGOE permite la comprensión y coherencia en el cumplimiento de objetivos; la consideración de los procesos en términos de valor agregado; el logro del desempeño eficaz del proceso; y la mejora de procesos con base en la evaluación de los datos y la información. Una organización necesita planificar e implementar acciones para abordar los riesgos y las oportunidades, esto establece una base para aumentar la eficacia del sistema de gestión de calidad, alcanzar mejores resultados y prevenir los efectos negativos. [6]

LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Gobernanza y políticas institucionales.
- Calidad académica y administrativa.
- Diseño curricular y planificación académica.
- Indicadores de desempeño alineados con normas de calidad.
- Gestión de Procesos y Evaluación de Desempeño.
- Automatización y digitalización de procesos para el aseguramiento de la calidad.
- Interoperabilidad.
- Preservación digital.
- Acceso y recuperación de la información.

OBJETIVOS

El desarrollo de esta línea de investigación está permitiendo cumplir con los siguientes objetivos:

- Analizar en detalle los requisitos de la Norma ISO 21001:2018 y su aplicación en Instituciones Educativas.
- Realizar un diagnóstico de la situación actual de los procesos de evaluación y acreditación implementados para determinar el grado de cumplimiento de los requisitos establecidos por la Norma ISO 21001:2018.
- Establecer un plan de trabajo para el diseño, especificación y actualización de los procesos del Sistema de Aseguramiento de la Calidad que permita dar cumplimiento a los requisitos de la norma ISO 21001:2019.
- Diseñar el mapa de procesos y especificar los procesos identificados en la norma ISO 21001:2018.
- Definir los indicadores que permitan evaluar la mejora continua en los procesos especificados.

- Analizar la factibilidad de automatización de los procesos y la sistematización de la información generada.

RESULTADOS OBTENIDOS/ESPERADOS

- Identificación de los requisitos establecidos en la norma para la especificación de procesos.
- Diagnóstico de los procesos de evaluación y acreditación implementados para determinar el grado de cumplimiento de los requisitos establecidos por la norma.
- Especificación y actualización de los procesos del sistema de aseguramiento de la calidad.
- Especificación, actualización de procedimientos, instructivos y registros vinculados al proceso de acreditación de carreras de grado y posgrado.
- Desarrollo del Manual de Procedimientos del proceso de acreditación de carreras de grado y posgrado, aprobado por Ord. HCS N° 011-25 Manual de Procedimientos Carreras de Grado-Posgrado.
- Especificación de requerimientos para la automatización de los procedimientos.
- Implementación del sistema Trello, herramienta de gestión visual de proyectos, para la planificación, monitoreo y colaboración entre los actores institucionales durante los procesos de acreditación de carreras.
- Modelado de flujogramas y automatización de tareas incluidos en los procedimientos de gestión vinculados al proceso de acreditación de carreras grado y posgrado, desarrollados en Bonita BPM como plataforma de Gestión de Procesos de Negocio.
- Se avanza en la integración de la planificación de tareas desarrollada en el sistema Trello con los flujogramas modelados en Bonita BPM, con el propósito de lograr una mayor coherencia operativa entre la organización de tareas, la automatización de procesos y el seguimiento en tiempo real.

FORMACIÓN DE RECURSOS HUMANOS

El equipo está conformado por tres docentes investigadoras del Departamento de Básicas y Tecnológicas, una docente investigadora del Departamento de Ciencias Sociales, Jurídicas y Económicas de la UNdeC, un docente investigador de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional de Nordeste. Dos estudiantes avanzados de la UNdeC.

De los integrantes del equipo: una docente investigadora en el área de Ingeniería de Software y Tecnologías en Educación, categoría IV del programa de incentivos, ha ocupado cargos de gestión en diferentes áreas de la UNdeC relacionadas con la gestión académica e institucional; un docente investigador de la UNNE, categoría III del programa de incentivos, Director del Grupo de Investigación en Calidad del Software y de proyectos de investigación relacionados con calidad del software, acreditado como Auditor Jefe Norma ISO 15504 por AENOR y Auditor de personal I+D por la Agencia Española de Normalización. Dos de las docentes investigadoras integrantes del equipo están vinculadas a la Dirección de Evaluación y Acreditación de la UNdeC como soporte a los procesos de evaluación institucional y acreditación de carreras, una de ellas es docente investigadora en el área de Auditoría y Calidad de Software, integrante de comisiones de autoevaluación en los procesos de acreditación de las carreras de la UNdeC. Los estudiantes que forman parte del equipo, son estudiantes avanzados de la carrera Ingeniería en Sistemas de la UNdeC, los dos desarrollando sus trabajos finales sobre temáticas vinculadas a esta línea de investigación.

La formación y capacidades de cada uno de los miembros del equipo de trabajo y en especial la colaboración de experiencias de docentes investigadores de dos instituciones universitarias fomenta la reflexión, el debate y contraste de diferentes puntos de vistas, fortaleciendo el conocimiento, capacidades, habilidades y destrezas en la utilización de herramientas, métodos y técnicas que forman parte del proceso de investigación.

BIBLIOGRAFÍA

[1]. ARAICA ZEPEDA, Ricardo. Diseño de un Sistema de Gestión y Aseguramiento de la Calidad de las Instituciones de Educación Superior de Nicaragua: Caso de una Universidad pública. 2023. Tesis Doctoral. Universidad Nacional Autónoma de Nicaragua.

[2]. CONEAU. Guía para la implementación y consolidación de sistemas internos de aseguramiento de la calidad universitaria. 2023

[3]. CORDOBA MORENO, Romina Alejandra. Sistema de gestión de la calidad en la Universidad Nacional de Rosario. Programa de implementación y certificación de un sistema de gestión de la calidad según los requerimientos de la Norma IRAM-ISO 9001. 2023. Tesis de Maestría..

[4]. DE NORMALIZACIÓN, Asociación Española. UNE-ISO 9000 Sistema de gestión de calidad. Fundamentos y vocabulario. 2015.

[5]. GOETSCH, David L.; DAVIS, Stanley B. Quality management for organizational excellence: Introduction to total quality. pearson, 2016.

[6]. ISO 21001:2018. Norma Internacional ISO 21001:2018 traducción oficial ISO, primera edición. 2018.

[7]. PÉREZ, Ivonne Rodríguez. Sistema de gestión de la calidad y su influencia en la calidad educativa. Revista Iberoamericana de Producción Académica y Gestión Educativa, 2015, vol. 2, no 3.

[8]. SALGADO CHILON, Kelly Gabriela. Sistemas de gestión de calidad bajo la norma ISO 21001: 2018, para procesos académicos de la Escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca. 2024.

[9]. SOLORZANO, Diana; CALIUSCO, María Laura; ARAUJO, Pedro. Identificación de stakeholders para la autoevaluación de calidad en una universidad pública: caso de estudio UTN-FRT. En XV Simposio de Informática en el Estado (SIE 2021)-JAIIO 50 (Modalidad virtual). 2021.

[10]. STRAH, Martín. Los sistemas nacionales de aseguramiento de la calidad de la Educación Superior en Iberoamérica. Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU) y Organización de Estados Iberoamericanos (OEI), 2020.

[11]. SUNDER M, Vijaya. Constructos de calidad en los servicios de educación superior. Revista internacional de gestión de la productividad y el rendimiento, 2016, vol. 65, núm. 8, pág. 1091-1111.

[12]. CARMONA, Fernanda Beatriz, et al. Estrategias para la mejora continua del sistema de aseguramiento de la calidad de la UNdeC. En XXVII Workshop de Investigadores en Ciencias de la Computación (Mendoza, 10 y 11 de abril de 2025). 2025. http://sedici.unlp.edu.ar/handle/10915/184412

[13]. CARMONA, Fernanda Beatriz, et al. Integración de la Norma ISO 21001:2018 en la Mejora Continua de los Procesos de Aseguramiento de la Calidad de la UNdeC. Libro de actas-XXXI Congreso Argentino de Ciencias de la Computación-CACIC 2025. http://sedici.unlp.edu.ar/handle/10915/189846

Objetos de aprendizaje y metodologías activas: Innovación pedagógica para la retención en carreras de ingeniería con planes de estudio por competencias

Minnaard Claudia, Torres Zulma, Ramos Florencia

Instituto de Investigaciones de Tecnología y Educación (IITE)

Facultad de Ingeniería – Universidad Nacional de Lomas de Zamora

Buenos Aires, Argentina

minnaardclaudia@gmail.com, iite@ingenieria.unlz.edu.ar

RESUMEN

Bajo la órbita del Instituto de Investigaciones en Tecnología y Educación (IIT&E) de la Facultad de Ingeniería de la UNLZ, la presente investigación se alinea con los ejes estratégicos institucionales orientados a la optimización de los procesos de enseñanza-aprendizaje. El estudio se centra en el fortalecimiento de las competencias de egreso en disciplinas científico-tecnológicas, mediante la integración sistemática de las TIC y el desarrollo de objetos de aprendizaje. A continuación, se exponen los objetivos fundamentales del proyecto y la hoja de ruta de las acciones ejecutadas."

CONTEXTO

La permanencia estudiantil en las facultades de ingeniería representa un desafío crítico para la gestión universitaria contemporánea. Al respecto, Brito (2011) recupera el modelo pionero de Vincent Tinto, subrayando que la retención es un fenómeno multifactorial donde convergen el capital cultural previo, el respaldo del entorno familiar y la capacidad de integración social y académica del individuo. No obstante, si bien la literatura científica ha profundizado en el diagnóstico de la deserción, la implementación de dispositivos institucionales de retención sigue siendo una tarea pendiente. En este sentido, Mastache y Goggi (2017) postulan que la formación técnica requiere de una enseñanza centrada en la autonomía del estudiante, donde la resolución de problemas complejos se logre mediante la sinergia entre el ingenio divergente y el rigor lógico-matemático. Bajo esta visión, la retención no solo depende de factores externos, sino de la capacidad de la universidad para ofrecer experiencias formativas que promuevan la reflexión y la articulación interdisciplinar de contenidos.

Los objetivos del proyecto son los siguientes:

- Analizar la interrelación entre la adquisición de competencias profesionales en el estudiantado de ingeniería, las metodologías de enseñanza-aprendizaje, los sistemas de evaluación y sus efectos sobre los índices de retención académica.
- Desarrollar objetos de aprendizaje (OA) fundamentados en metodologías activas, diseñados específicamente para fortalecer la persistencia estudiantil y la autonomía en el proceso formativo.
- Determinar las variables determinantes y los criterios de diseño en la creación de recursos educativos digitales que impacten positivamente en la disminución del abandono escolar.

- Implementar Clases Online Complementarias (COCOs) como dispositivos de apoyo pedagógico para áreas críticas y programas de transición, específicamente en:

Matemática, Probabilidad y Estadística, Medios de Representación I y II, y el Programa de Acompañamiento y Nivelación (PANI).

1. INTRODUCCIÓN

El aprendizaje por competencias se orienta hacia el fortalecimiento de las aptitudes y facultades críticas para una integración social eficiente. Para consolidar tales capacidades, resulta imperativo que los educadores cuenten con instrumentos didácticos pertinentes que logren adaptarse a las diversas demandas de aprendizaje de los estudiantes (Comoglio y Minnaard, 2020; Minnaard et al., 2020, 2021; Minnaard y Minnaard, 2022; Morgado et al., 2015). Este modelo formativo exige una visión integral y sistémica de los componentes de la labor universitaria, lo que conlleva una interacción constante y una participación proactiva de todos los actores del proceso instructivo.

Desde esta óptica, la gestión curricular se fundamenta en "identificar con claridad las prácticas implícitas o explícitas que se tienen respecto a la formación, con el fin de tomar conciencia de ellas, modificarlas (si es necesario) y buscar generar las condiciones que lleven a tener personas con alto compromiso ético, autorrealización, emprendimiento e idoneidad para afrontar los diferentes retos del contexto" (Tobón Tobón, 2011)

Al respecto, Caraballo Carmona et al. (2019) sostienen que el éxito de estos propósitos reside en la competencia misma, lo que demanda una planificación, enseñanza y evaluación bajo este enfoque. La competencia no es solo una capacidad de ejecución procedimental; por el contrario, integra necesariamente dimensiones teóricas y actitudinales.

Por otra parte, Amaya Chávez (2021) afirma que el desarrollo de las competencias en el alumnado requiere el uso de metodologías activas. Entre las más relevantes se encuentran el aprendizaje colaborativo, el estudio de casos y el aprendizaje basado en problemas o proyectos. Recientemente, se han sumado el aula invertida, la gamificación, el visual thinking, el design thinking y el studio based learning (Palma et al., 2017).

Estas metodologías se sustentan en que el estudiante, bajo la tutoría docente, asuma con autonomía su aprendizaje, mejorando sus desempeños conceptuales, procedimentales y axiológicos. Resulta esencial, entonces, un cambio metodológico que sitúe al alumno como centro de la construcción de significados mediante un proceso activo y constructivo, tanto individual como colaborativo.

En el panorama educativo actual, el desarrollo de competencias se fundamenta en una integración sistémica de metodologías activas y tecnologías emergentes. Investigaciones recientes (Dávila Morán y López Gómez, 2025) demuestran que el Aprendizaje Basado en Problemas (ABP) continúa siendo la estrategia con mayor impacto en la adquisición de competencias profesionales en la educación superior. Asimismo, autores como Ramos Caballero et al. (2025) enfatizan la relevancia de la neurodidáctica para potenciar competencias científicas mediante la participación dinámica del discente. Por otro lado, la implementación de metodologías basadas en retos y el uso de inteligencia artificial se consolidan como ejes transformadores que fomentan el pensamiento crítico y la autonomía (Guillén-Mendoza et al., 2025). Estos enfoques no solo optimizan el desempeño conceptual y procedimental, sino que preparan al estudiante para afrontar desafíos ecosociales complejos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El aprendizaje eficaz requiere que los alumnos operen activamente en la manipulación de la información a ser aprendida, pensando y actuando sobre ello para revisar, expandir y

asimilarlo. Por lo tanto y desde una perspectiva didáctica, estamos construyendo competencias.

Para alcanzar los resultados esperados en el presente proyecto de investigación, y en función de los objetivos planteados, se requiere trabajar con datos provenientes de distintas fuentes, los que serán sometidos a distintos tipos de análisis.

A continuación, se presentan los objetivos planteados y se desarrolla la explicación solicitada

-Analizar la interrelación entre la adquisición de competencias profesionales en el estudiantado de ingeniería, las metodologías de enseñanza-aprendizaje, los sistemas de evaluación y sus efectos sobre los índices de retención académica.

Encuestas, entrevistas y focus group a los alumnos ingresantes. Interpretación y análisis de discurso social de los resultados de las entrevistas y focus group. Análisis estadístico (descriptivo y correlacional). Relevamiento de estrategias innovadoras que los docentes integran a sus prácticas y sistematización de casos describiendo: objetivos, metodología desarrollada, resultados esperados y alcanzados.

-Desarrollar objetos de aprendizaje (OA) fundamentados en metodologías activas, diseñados específicamente para fortalecer la persistencia estudiantil y la autonomía en el proceso formativo.

Diseñar e implementar propuestas pedagógicas adecuadas. Diseño de rúbricas ajustadas a las propuestas.

-Determinar las variables determinantes y los criterios de diseño en la creación de recursos educativos digitales que impacten positivamente en la disminución del abandono escolar.

Relevamiento de estrategias innovadoras que los docentes integran a sus prácticas y sistematización de casos describiendo: objetivos, metodología desarrollada, resultados esperados y alcanzados.

- *Implementar Clases Online Complementarias (COCOs) como dispositivos de apoyo pedagógico para áreas críticas y programas de transición, específicamente en: Matemática, Probabilidad y Estadística, Medios de Representación I y II, y el Programa de Acompañamiento y Nivelación (PANI).*

Desarrollar e implementar propuestas pedagógicas adecuadas. Diseño de rúbricas ajustadas a las propuestas

3. RESULTADOS OBTENIDOS Y ESPERADOS

Los destinatarios directos de los resultados del presente programa son estudiantes y docentes de las carreras de Ingeniería Mecánica, Industrial, Ferroviaria y Mecatrónica de la FI UNLZ , a través de las aplicaciones y conocimientos que se generen en el ámbito de las cátedras y espacios institucionales en los que se desempeñan los miembros del grupo de investigación.

4. FORMACIÓN DE RECURSOS HUMANOS

La formación de Recursos Humanos incluye la formación de 1 becario del programa EVC-CIN para el período 2025-2026

5. BIBLIOGRAFÍA

Amaya Chávez, D. (2021). Diseño, análisis, y evaluación de un método de aprendizaje basado en problemas (abp) mediado por tecnologías interactivas desarrollado con estudiantes de ingeniería.

Caraballo Carmona, C., Meléndez Ruiz, R., & Iglesias Triana, L. (2019). Reflexiones acerca del concepto competencias y aprendizaje por competencias en las instituciones de educación superior y su incidencia en el aprendizaje de las matemáticas. Opuntia Brava, 11(1), 297-307.

Comoglio, M. & Minnaard, C. (2020) Desafíos y adaptaciones en el proceso de enseñanza en las distintas cátedras de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora en tiempos de pandemia. V Encuentro Centros CIC, Buenos Aires, Argentina.

Dávila Morán, R. C., y López Gómez, H. E. (2025). Impacto de las metodologías activas en la adquisición de competencias de estudiantes universitarios. Conrado, 21(102)

Guillén-Mendoza, S. V., et al. (2025). Metodologías Activas en la Educación Superior: Estrategias para el Aprendizaje Significativo. Editorial Erevna Ciencia

Mastache y Goggi (2017) El Comité de Mejora una experiencia colaborativa en Ingeniería. Revista Argentina de Educación Superior N° 14.

Minnaard, C. & Minnaard, V. (2014). Evaluación por competencias en Entornos Virtuales de Aprendizaje. XVI Workshop de Investigadores en Ciencias de la Computación. En: http://sedici.unlp.edu.ar/bitstream/handle/10915/43712/Documento_completo.pdf?sequence=1&isAllowed=y

Minnaard,A.; Iravedra, C.; Garrido, G.& Gonzalez, B.(2020) Enseñanza por competencias: la experiencia de los COCOs. VII Jornadas Nacionales y III Latinoamericanas de Ingreso y Permanencia en Carreras Científico – Tecnológicas

Minnaard, C.;Comoglio, M.; Iravedra, C.;Garrido, G. & Gonzalez, B.(2020) Enseñanza por competencias: la experiencia de los COCOs. 18th LACCEI International Multi-Conference for Engineering, Education, and Technology: "Engineering, Integration, and Alliances for a Sustainable Development" "Hemispheric Cooperation for Competitiveness and Prosperity on a Knowledge-Based Economy", 29-31 July 2020, Buenos Aires, Argentina.

Minnaard,C.; Minnaard, V. & Torres, Z. (2021) Adaptaciones en el proceso de enseñanza en la cátedra de Probabilidad y Estadística en tiempos de pandemia. XVI Congreso Internacional de Ciencias Sociales Interdisciplinares, Oxford, Gran Bretaña.

Minnaard, V. & Minnaard, C. (2022) Design and Validation of an Instrument to Analyze Mobile Applications that Favor mLearning: Mobile Applications Subject to Analysis. XVIII International Conference on Technology , Knwoledge and Society. Universidad Changhua, Taiwan.

Morgado, E. M., Peñalvo, F. G., Ortuño, R. C., & Hidalgo, C. A. (2015). Desarrollo de competencias a través de objetos de aprendizaje. Revista de Educación a Distancia, (36).

Palma Luengo, M. et al (2017) Metodologías activas para la formación universitaria. Universidad de Bio Bio.

Ramos Caballero, E. J., et al. (2025). Metodologías Activas: Desarrollo de las Competencias Científicas en Estudiantes de la Media Académica. Unaciencia, 18(34), 4–26.

Tobón, S. (2011). Formación basada en competencias. Pensamiento complejo, diseño curricular y didáctica. 2 ed. Bogotá: ECOE Ediciones

OPTIMIZACIÓN DE UN MODELO DE RIESGO DE ABANDONO ESTUDIANTIL MEDIANTE MINERÍA DE DATOS EN LA UTN LA PLATA: ANÁLISIS E INTEGRACIÓN DE FACTORES PSICOLÓGICOS Y SOCIOLÓGICOS

ISTVAN, Romina; LASAGNA, Valeria; PLAQUIN, Lorena; DIAZ LAPERGOLA, Ayelén; FERRARA, Damian

Universidad Tecnológica Nacional, Facultad Regional La Plata
Laboratorio de Ingeniería en Sistemas de Información, LINES UTN FRLP
Av. 60 s/n° esquina 124, CP 1900, La Plata, Buenos Aires, Argentina.
{ristvan; valerial; lplaquin; madlapergola; dferrara}@frlp.utn.edu.ar

RESUMEN

El presente proyecto tiene como objetivo desarrollar una versión optimizada del modelo de riesgo de abandono estudiantil para la UTN La Plata, integrando a las actuales variables demográficas y académicas, factores psicológicos y sociológicos que permitan mejorar tanto la comprensión como la predicción de las trayectorias educativas, mediante técnicas de minería de datos. Para ello, se prevé el diseño, entrenamiento y evaluación de modelos predictivos utilizando algoritmos de aprendizaje automático supervisado.

En este marco, se espera consolidar una herramienta tecnológica que permita detectar tempranamente estudiantes en situación de riesgo y fortalecer las estrategias institucionales de permanencia y egreso, con posibilidad de transferencia a otras Facultades Regionales de la UTN.

El proyecto se sustenta en antecedentes nacionales e internacionales vinculados al estudio del abandono universitario, y en particular en investigaciones que aplican técnicas de minería de datos a este campo. Asimismo, se apoya en la trayectoria del equipo de investigación, responsable del desarrollo del sistema de gestión tutorial ESDEU y de modelos predictivos previos orientados al análisis de las trayectorias académicas (Stirz et al., 2020; Istvan et al., 2023).

Palabras clave: Abandono estudiantil; Modelo de riesgo de abandono estudiantil; Sistemas de Gestión Tutorial; Minería de datos

CONTEXTO

Con el objetivo de abordar la permanencia estudiantil, la UTN La Plata diseñó y desarrolló el Sistema ESDEU en el marco del Proyecto de Investigación y Desarrollo (PID) «Estudio Sistemático de Deserción Estudiantil Universitaria», homologado por el Rectorado de la UTN bajo el código TEINNLP0003786.

ESDEU es un sistema software de gestión tutorial que integra un modelo de riesgo de deserción generado con variables demográficas y académicas, con tres ejes de análisis: perfil socioeconómico, inasistencias y rendimiento académico. A partir del analisis de datos, el sistema visualiza el nivel de riesgo de cada estudiante mediante un semaforo de colores con cuatro categorías: verde (sin riesgo), amarillo (riesgo leve), naranja (riesgo moderado) y rojo (riesgo crítico), generando alertas tempranas de abandono.

Desde el año 2020, esta línea de trabajo se

profundizó mediante la ejecución de nuevos proyectos de tecnología con impacto en la deserción estudiantil. En continuidad con estos desarrollos, en el año 2026 inicia el PID «Optimización de un modelo de riesgo de abandono estudiantil mediante minería de datos en la UTN La Plata: análisis e integración de factores psicológicos y sociológicos», homologado bajo el código LPTEEC898, cuyo propósito es ampliar el enfoque analítico del modelo existente mediante la incorporación de dimensiones psicológicas y sociológicas, fortaleciendo su capacidad predictiva y su aplicación a nivel institucional.

1. INTRODUCCIÓN

El abandono en la educación superior representa una preocupación de alcance global, con manifestaciones que varían según los contextos sociales, económicos y educativos. En muchas regiones del mundo, el acceso a la universidad aumentó significativamente en las últimas décadas, impulsado por políticas orientadas a la democratización del conocimiento y a la expansión de la oferta académica.

Según datos de la Secretaría de Políticas Universitarias (SPU, 2022), en Argentina la Tasa Bruta de Escolarización Superior alcanzó el 67,3% en 2021, un valor comparable al de países como Dinamarca y los Países Bajos, y superior al de Austria, Bélgica, Portugal, Reino Unido y Suecia (UNESCO, 2023). No obstante, esta alta tasa de acceso contrasta con una de las tasas de graduación más bajas a nivel mundial, que apenas alcanza el 19,7% (UNESCO, 2023). Este desfasaje pone en evidencia una brecha significativa entre el ingreso a la educación superior y la finalización efectiva de los estudios.

Frente a esta realidad, las universidades enfrentan el desafío de disponer de herramientas analíticas capaces de identificar de manera temprana situaciones de riesgo y de diseñar estrategias de intervención oportunas. Este requerimiento se enmarca en un contexto de creciente digitalización de los procesos académicos y administrativos; donde la mera disponibilidad de datos no garantiza un aprovechamiento efectivo, pero adecuadamente analizados, ofrecen un potencial significativo para el abordaje del abandono.

En este marco, la minería de datos emerge como una herramienta clave al permitir transformar grandes volúmenes de datos en conocimiento útil para la toma de decisiones (Rodelo Mendoza & Gómez Cuello, 2022; Romero & Ventura, 2020). A diferencia de los enfoques tradicionales, facilita la identificación de patrones ocultos en conjuntos de datos complejos mediante diversas técnicas.

Entre los métodos más utilizados en la predicción del abandono estudiantil se encuentran los árboles de decisión, Random Forest y la regresión logística, todos ellos pertenecientes a los métodos supervisados. Estos se seleccionan frente a otros enfoques debido a que los datos disponibles contienen información etiquetada sobre el estado de los estudiantes (si abandonaron o continuaron sus estudios), lo que permite entrenar modelos que predicen resultados concretos. Además, posibilitan clasificar a los estudiantes en distintos niveles de riesgo, estimando la probabilidad de abandono de manera cuantificable (Han et al., 2011; Márquez Vera et al., 2013).

Dentro de estos enfoques, los árboles de decisión se destacan por su alta interpretabilidad. Utilizan

una estructura jerárquica en forma de diagrama de flujo, donde cada nodo representa una condición sobre una variable específica, guiando hacia diferentes clases o resultados (Quinlan, 1996). Esta característica los convierte en herramientas especialmente valiosas para el ámbito educativo, ya que permiten comprender fácilmente qué factores contribuyen al riesgo de abandono y facilitan la creación de reglas claras para intervenciones institucionales (Romero & Ventura, 2020).

Por su parte, la regresión logística y Random Forest ofrecen una mayor capacidad de generalización y precisión predictiva. Ambos modelos permiten estimar la probabilidad de abandono, lo que resulta útil para asignar niveles de riesgo y priorizar acciones de seguimiento. Si bien su interpretación puede ser menos directa que la de los árboles simples, su utilizacion mejora la calidad del análisis y contribuye a decisiones institucionales más informadas (Delen, 2010).

La combinación de modelos altamente interpretables con modelos de mayor precisión permite aprovechar lo mejor de ambos enfoques, generando predicciones confiables y comprensibles que sirven de base para el diseño de políticas de acompañamiento y prevención del abandono estudiantil.

En los últimos años, diversas universidades argentinas desarrollaron modelos predictivos basados en técnicas de minería de datos. Sin embargo, una limitación recurrente de estos desarrollos es su énfasis casi exclusivo en variables académicas y demográficas, tales como calificaciones, asistencia y antecedentes educativos, dejando de lado otros factores que también influyen en la permanencia estudiantil. Esta limitación evidencia la necesidad de complementar los modelos cuantitativos con enfoques que aborden los procesos subjetivos y contextuales del estudiante. Entre ellos, los enfoques psicológico y sociológico ocupan un lugar central, ya que permiten analizar de manera directa los procesos que inciden en la permanencia estudiantil. En este marco, investigaciones recientes resaltan la importancia de variables como la motivación, el autoconcepto académico, la integración social y, especialmente, la resiliencia, entendida como la capacidad de afrontar el estrés académico, regular las emociones y preservar la salud mental frente a situaciones adversas (Eisenberg et al., 2016; Brewer et al., 2019; Ang et al., 2022).

Estos factores adquieren especial relevancia en el caso de los estudiantes de primera generación, quienes suelen enfrentar presiones adicionales asociadas a la adaptación cultural al ámbito universitario, las expectativas del entorno familiar y la ausencia de redes de apoyo académico e institucional consolidadas. Esta combinación de demandas incrementa su vulnerabilidad frente al abandono y refuerza la necesidad de priorizar variables psicológicas y sociológicas en el análisis del problema (Lipson et al., 2023).

De este modo, la integración de estas dimensiones se consolida como una línea de trabajo necesaria para avanzar hacia modelos explicativos más integrales.

Bajo este contexto, el presente trabajo propone desarrollar una versión optimizada del modelo de riesgo de abandono estudiantil para la UTN La Plata incorporando a las variables demográficas y académicas, actualmente

utilizadas, factores psicológicos y sociológicos, mediante técnicas de Mineria de Datos.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La línea de investigación comprende el desarrollo y la aplicación de Tecnología Informática Aplicada en Educación, con énfasis en el diseño de modelos predictivos basados en técnicas de Minería de Datos para el análisis de las trayectorias estudiantiles.

Desde un enfoque interdisciplinario, se incorporan aportes de la Psicología Cognitiva y la Sociología Educativa con el fin de enriquecer los modelos existentes mediante la integración de factores subjetivos y contextuales.

3. RESULTADOS OBTENIDOS / ESPERADOS

La UTN La Plata desarrolló e implementó el sistema de gestión tutorial ESDEU, junto con un modelo predictivo inicial de abandono estudiantil basado en variables demográficas y académicas, estructurado en tres ejes de análisis: perfil socioeconómico, inasistencias y rendimiento académico. Esto permitió generar alertas tempranas y gestionar de manera eficiente el seguimiento de los estudiantes antes de que pierdan contacto con la institución. En este marco, el presente proyecto se propone optimizar el modelo de riesgo mediante técnicas de minería de datos. Para ello, se ampliarán las dimensiones de análisis mediante la incorporación de variables psicológicas y sociológicas. El propósito final es mejorar la capacidad de anticipación y detección del abandono estudiantil.

De este modo, se busca consolidar una base tecnológica más robusta, que facilite el diseño de estrategias institucionales más efectivas, con potencial de transferencia a otras Facultades Regionales de la UTN.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por un Director, un Codirector, tres Docentes Investigadores de la Carrera de Ingeniería en Sistemas de Información (DISI), dos Docentes del Equipo Pedagógico de la regional y tres estudiantes becarios de investigación.

Cuenta en curso con una tesis de Doctorado, tres tesis de Maestría en Docencia Universitaria y una de Especialización en la carrera de Tecnología Informática Aplicada en Educación, asimismo se prevé el desarrollo de Prácticas Supervisadas (PS) de estudiantes, necesarias para la obtención del título de grado de la carrera de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

Ang, W. H. D., Lau, S. T., Cheng, L. J., Chew, H. S. J., Tan, J. H., Shorey, S., & Lau, Y. (2022). Effectiveness of resilience interventions for higher education students: A meta-analysis and metaregression. Journal of Educational Psychology, 114(7), 1670.

Brewer, M. L., Van Kessel, G., Sanderson, B., Naumann, F., Lane, M., Reubenson, A., & Carter, A. (2019). Resilience in higher education students: A scoping review. Higher Education Research & Development, 38(6), 1105-1120.

Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. Decision Support Systems, 49(4), 498-506.

Eisenberg, D., Lipson, S. K., & Posselt, J.

(2016). Promoting resilience, retention, and mental health. En M. B. Paulsen (Ed.), Higher education: Handbook of theory and research (Vol. 31, pp. 281–335).

Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

Istvan, R.; Plaquin, L.; Stirz, A.; Mauriño, A.; (2023). Impulso Ingeniería: Potenciando el Éxito Estudiantil a través de Tutorías Motivacionales Pares en la UTN La Plata; Jornada de Enseñanza de la Ingeniería (JEIN 2023); Universidad Tecnológica Nacional - Facultad Regional Paraná. Paraná, Entre Ríos; 7 y 8 de septiembre de 2023

Lipson, S. K., Diaz, Y., Davis, J., & Eisenberg, D. (2023). Mental health among first-generation college students: Findings from the national Healthy Minds Study, 2018-2021. Cogent mental health, 2(1), 2220358. https://doi.org/10.1080/28324765.2023.2220358

Márquez-Vera, C., Cano, A., Romero, C., y Ventura, S. (2013). Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Applied Intelligence, 38(3), 315-330.

Quinlan, J.R. (1996). Improved use of continuous attributes in C4.5. Journal of Artificial Intelligence Research, 4, 77-90.

Rodelo Mendoza, M., y Gómez Cuello, N. (2022). Aplicación de técnicas de minería de datos para identificar factores determinantes en la deserción estudiantil del programa de Ingeniería de Sistemas de la Universidad de La Guajira. Revista Ingenio, 19(2), 141–158. https://dialnet.unirioja.es/descarga/articulo/9585948.pdf

Romero, C., y Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3). https://doi.org/10.1002/widm.1355

SPU, Secretaría de Políticas Universitarias, Ministerio de Educación Argentina (2022). Síntesis de Información Estadísticas Universitarias 2021-2022.

https://www.argentina.gob.ar/sites/default/files/sintesis_2021-2022_sistema_universitario_argentino_-_ok.pdf

Stirz, A.; Lyde, E.; Istvan, R.; Lasagna, V. (2020). Herramienta para la identificación temprana de riesgo de deserción estudiantil y gestión tutorial, 8° Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI); Trabajo seleccionado para Exposición, UTN San Francisco, San Francisco, Provincia de Córdoba; 5 y 6 de noviembre de 2020; http://conaiisi2020.frsfco.utn.edu.ar/actas.html

UNESCO, Instituto Internacional de Planeamiento de la Educación, IIPE (2023). Indicadores estadísticos. https://siteal.iiep.unesco.org/indicadores

Pensar con datos: desarrollo de competencias en Ciencia de Datos para el nivel secundario

Sofía Martin[1], Claudia Banchoff[1], Paula Venosa[1], Liliana Hurtado[2]
[1]LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata
Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528
{smartin, cbanchoff, pvenosa}@info.unlp.edu.ar}
[2]Facultad de Ciencias Económicas, Jurídicas y Sociales. Universidad Nacional de Salta
lilianahur@hotmail.com

RESUMEN

Este trabajo presenta el estado de avance de una línea de investigación que aborda la enseñanza de aspectos básicos de la Ciencia de Datos en la escuela secundaria. Se describen los resultados obtenidos en experiencias con docentes y estudiantes y los nuevos desafíos que surgen de las experiencias realizadas, destacando el uso de herramientas de software libre y datos abiertos. También se describe la articulación entre extensión, investigación y docencia, que ha permitido desarrollar herramientas y materiales que benefician estudiantes universitarios y de nivel secundario.

Palabras clave: Ciencia de Datos, educación secundaria, Python, datos abiertos, Ciencia de la Comutación.

CONTEXTO

Durante los últimos años, el equipo de investigación del LINTI[1] ha desarrollado experiencias educativas para introducir la Ciencia de Datos en la escuela secundaria, apoyándose en una trayectoria previa en pensamiento computacional y robótica[1,2]. La Ciencia de Datos, como campo interdisciplinario que integra la Ciencia de la Computación, la Estadística y el conocimiento del dominio específico, ofrece un marco potente para el análisis significativo de problemas del mundo real [3,4]. Pese a los avances, persisten desafíos como la adaptación curricular, la adecuación tecnológica y la implementación efectiva en el aula [5]. Esta línea se enmarca en el proyecto **Tecnologías digitales para la Inclusión, la Equidad y la Sostenibilidad** (Programa de Incentivos) y se nutre de dos proyectos de extensión de la SPU que permitieron implementar y evaluar propuestas en escuelas de la región. Además, un proyecto I+D de alumnos de la Facultad de Informática contribuyó a la mejora de herramientas tecnológicas para la enseñanza de la Ciencia de Datos. Esta articulación entre investigación, extensión y docencia ha sido clave para consolidar la línea.

INTRODUCCIÓN

La creciente disponibilidad de datos ha cobrado una relevancia creciente en los últimos años debido al aumento exponencial de datos disponibles y su aporte a diversas áreas [6]. En consecuencia ha generado una demanda de alfabetización en datos desde edades tempranas [5,7,8]. La Ciencia de Datos, por su carácter interdisciplinario[4], permite enmarcar el aprendizaje en contextos significativos y fomentar el pensamiento crítico y ético [9,10]. Su enseñanza en secundaria enfrenta desafíos: formación docente, equidad en el acceso y disponibilidad de herramientas adecuadas [11, 12]. A partir de experiencias previas, se identificaron dos aspectos fundamentales: (a) uso de herramientas accesibles que permitan trabajar sin barreras técnicas [13], y (b) acceso a datos significativos que potencien la interdisciplinariedad.

En cuanto a las herramientas, se adopta un enfoque progresivo: se comienza con entornos de alto andamiaje como Pandalyze (basada en bloques y con código Python visible), se

[1] Sitio del Laboratorio LINTI: https://www.linti.unlp.edu.ar/

continúa con Jupyter Notebook[2][14] y se finaliza con Streamlit[3] para crear aplicaciones web interactivas. Esta progresión reduce la carga cognitiva inicial [15] y facilita la transición hacia entornos profesionales. Respecto a los datos, el uso de datos abiertos [16, 17] permite trabajar problemas locales y globales, pero enfrenta dificultades de búsqueda y adecuación para docentes no especializados [18, 19]. Como estrategia pedagógica, la gamificación ha demostrado ser efectiva para motivar y facilitar la apropiación de conceptos abstractos [5, 12, 20], promoviendo además habilidades transversales como el trabajo en equipo y la evaluación entre pares.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea central es analizar, desarrollar y evaluar estrategias para enseñar análisis de datos en secundaria, con foco en habilidades de estudiantes y docentes. Los temas abordados incluyen: relevamiento de experiencias nacionales e internacionales; análisis de indicadores de evaluación; identificación de repositorios de datos abiertos de nuestro país; diseño de materiales didácticos interdisciplinarios; y evaluación de actividades con escuelas.

3. RESULTADOS Y OBJETIVOS

El presente trabajo tiene como objetivo principal describir los principales resultados surgidos en las experiencias realizadas y la investigación:

Sobre los conjuntos de datos: se observó la necesidad de generar guías que orienten a docentes en la exploración de repositorios, y se fomentó la generación de datos propios (encuestas) para abordar problemáticas comunitarias.

Sobre el tratamiento de los datos y visualización de resultados: en los talleres surgió la necesidad de trabajar la limpieza de datos de forma guiada, con niveles de complejidad controlados. También se reflexionó sobre aspectos éticos (sesgos, verificación) y se analizaron distintas formas de comunicar resultados.

Respecto a las herramientas informáticas utilizadas: la evaluación de diversas herramientas [21, 22] confirmó que la programación en bloques es propicia para principiantes. Se desarrolló Pandalyze [23], una aplicación web de código abierto que ofrece bloques focalizados en el análisis de datos, utilizando la librería pandas[4]. En el año 2025 en el marco del proyecto de la Facultad de Informática **Herramienta Pandalyze con nuevas funcionalidades**, se incorporó a esta herramienta la visualización de mapas con Folium[5] (Fig. 1) y funcionalidades como guardar programas, nuevos ejemplos y tutoriales en video[6] (Fig. 2). Pandalyze se utilizó en talleres con buena aceptación por parte de estudiantes y docentes.

Respecto a las propuestas didácticas: las guías didácticas están en constante adecuación. de acuerdo a las devoluciones de los docentes y las observaciones de las experiencias. Se elaboraron actividades incluyendo aspectos de gamificación, incrementando la motivación y dinamismo en los distintos encuentros con los estudiantes. En el marco de estas propuestas, una de las actividades que generó mayor motivación fue la "Maratón de Ciencia de Datos", desarrollada con un formato lúdico y colaborativo que resultó especialmente efectivo para afianzar los contenidos trabajados. La dinámica consistió en una serie de postas con niveles de dificultad creciente, cada una compuesta por una instancia de precalentamiento —donde los estudiantes se familiarizaban con las operaciones necesarias— y desafíos grupales que debían resolverse simultáneamente. En una etapa

[2] Sitio oficial de Jupyter Notebook: https://jupyter.org/
[3] Sitio oficial de Streamlit: https://streamlit.io/
[4] Sitio oficial de pandas: https://pandas.pydata.org/
[5] Sitio oficial de Folium: https://python-visualization.github.io/folium/latest/
[6] Canal de Youtube con tutoriales: https://www.youtube.com/@CienciadeDatosenescuelas

posterior, cada grupo elaboró sus propias preguntas desafiantes sobre otros conjuntos de datos, para que otro equipo las resolviera. Esta propuesta promovió habilidades fundamentales como la claridad en la formulación de consignas, la resolución de problemas y, especialmente, la crítica constructiva entre pares, al generar un espacio de retroalimentación y mejora colectiva.

Es de destacar que la incorporación de la herramienta **Pandalyze** fue un punto de inflexión en la investigación. Esta herramienta surge como el desarrollo de una tesina de grado en el año 2024.



Figura 1. Bloques específicos de Pandalyze asociadas a pandas y Folium.



Figura 2. Nuevas funcionalidades incorporadas a Pandalyze.

En el marco del trabajo con estudiantes de nivel medio de la EET 9 de La Plata que se encuentran desarrollando sus prácticas profesionales supervisadas (PPS) en el LINTI, se desarrollaron actividades avanzadas como el uso de Streamlit que habilitó la creación de una página web interactiva donde podían integrar y comunicar los resultados de los análisis realizados durante el taller.

Como se mencionó anteriormente, el **objetivo principa**l de esta línea de investigación es: analizar, formular y evaluar propuestas didácticas que permitan incorporar aspectos básicos de la Ciencia de Datos en las escuelas secundarias, conforme a los lineamientos dados por la Provincia de Buenos Aires. Estas propuestas involucran la evaluación de los datasets a utilizar y el énfasis en el uso de herramientas de licencia libre y de datos abiertos.

Para alcanzar este objetivo, se proponen los siguientes **objetivos específicos**:

- Revisar y actualizar el relevamiento de las experiencias pedagógicas en la región para la enseñanza de contenidos relacionados tanto con la Ciencia de Datos como con las Ciencias de la Computación en escuelas secundarias. Especialmente, respecto a la revisión de propuestas relacionadas a la programación de computadoras, debido a la estrecha vinculación con la Ciencia de Datos.
- Revisar y evaluar las metodologías y herramientas informáticas utilizadas en la enseñanza de la Ciencia de Datos.
- Evaluar librerías para el trabajo con Python, que permitan generar en forma simple visualizaciones de los análisis realizados, como ser la librería Folium.
- Continuar con la extensión y actualización de la herramienta Pandalyze ampliando tanto las ayudas y ejemplos, como también las funcionalidad de pandas y de visualización provista.
- Revisar las propuestas pedagógicas existentes y elaborar nuevas propuestas para la inclusión de aspectos básicos de la Ciencia de Datos en la escuela secundaria que incluyan el uso de Pandalyze, como también guías que ayuden a los docentes a incorporar nuevos datasets.
- Diseñar, implementar y evaluar las intervenciones con docentes y estudiantes de nivel secundario.
- Actualizar el sitio web público[7] de referencia con el que se cuenta, incorporando ejemplos de análisis de datos

[7] Sitio público con recursos elaborados: https://ciencia_datos_escuela.gitlab.io/

que utilicen la visualización de mapas interactivos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación está formado por docentes investigadores categorizados del LINTI, docentes de la Universidad de Salta y estudiantes de la Facultad de Informática, con trayectoria en el trabajo con escuelas: más de 10 años en temas relacionados a enseñanza de programación y robótica educativa y, más recientemente, incluyendo la temática de la Ciencia de Datos.

La herramienta Pandalyze, mencionada en el artículo, fue desarrollada en el marco de una tesina de grado y actualizada en el marco del proyecto **Herramienta Pandalyze con nuevas funcionalidades** donde participaron tanto docentes como estudiantes durante el año 2025. Así mismo, también se encuentra en la etapa final de desarrollo la tesis de maestría titulada **Espacio de aprendizaje de la Ciencia de Datos para establecimientos educativos de la Provincia de Buenos Aires**, en la que las autoras de este trabajo están profundizando la investigación y desarrollo de estrategias para la introducción de aspectos básicos de la Ciencia de Datos en el contexto educativo provincial de nivel secundario.

Al igual que en años anteriores, esta línea de trabajo también se aborda en la asignatura de Seminario de Lenguajes, opción Python de las carreras de Licenciatura en Informática y Sistemas, y, desde el año 2025, el Taller de Lenguajes de la carrera Ciencia de Datos en las Organizaciones. También se continúa trabajando con estudiantes de nivel medio de la EET 9 de La Plata, a través del espacio de prácticas profesionales supervisadas (PPS). Estas experiencias permiten obtener evidencias empíricas para continuar la línea de investigación.

5. BIBLIOGRAFÍA

[1] Banchoff Tzancoff, C. M., Martin, E. S., Gómez, N. S., & López, F. E. M. (2019) Experiencias en robótica educativa: Diez años trabajando con escuelas argentinas. XIV TE&ET. ISBN 978-987-733-196-7, pág. 84-93.

[2] Queiruga C., Banchoff Tzancoff C, Martin S. y Kimura I. (2024) Propuestas de acercamiento a la Informática para la escuela secundaria: contenidos, estrategias didácticas y materiales 53 JAIIO - SAEI - ISSN: 2451-7496. pp. 116-134.

[3] National Academies of Sciences, Engineering, and Medicine and others. (2023). Foundations of Data Science for Students in Grades K 12 A Workshop | National Academies. https://www.nationalacademies.org/our-work/foundations-of-data-science-for-students-in-grades-k-12-a-workshop

[2]Danyluk, A., & Leidig, P. Computing Competencies for Undergraduate Data Science Curricula: Vol. ACM Data Science Task Force. Association for Computing Machinery.

[4]Danyluk, A., & Leidig, P. Computing Competencies for Undergraduate Data Science Curricula: Vol. ACM Data Science Task Force. Association for Computing Machinery.

[5] Avila-Garzon, C., & Bacca-Acosta, J. (2025). Curriculum, Pedagogy, and Teaching/Learning Strategies in Data Science Education. Education Sciences, 15(2), 186. https://doi.org/10.3390/educsci15020186

[6]Hicks, S. C., & Irizarry, R. A. (2018). A Guide to Teaching Data Science. The American Statistician, 72(4), Article 4. https://doi.org/10.1080/00031305.2017.1356747

[7] Berikan, B., & Özdemir, S. (2020). Investigating "Problem-Solving With Datasets" as an Implementation of Computational Thinking: A Literature Review. Journal of Educational Computing Research, 58(2), 502-534. https://doi.org/10.1177/0735633119845694

[8]Borchardt, Mara, Klinkovich, Vanina, & Iocca, Natalia (with Llambí, Cecilia, Martinez, Cecilia, & Scasso, Martín). (s. f.). Aprendizajes y desafíos para la enseñanza de las Ciencias de la Computación en las escuelas. La iniciativa Program. AR de Argentina. Banco de Desarrollo de América Latina y el Ca- ribe (CAF, publicaciones@caf.com) y

Fundación Sadosky (info@fundacionsadosky.org.ar). Recuperado 29 de agosto de 2023, de https://program.ar/libro-programar-caf-diezanios/
[9]Finzer, W. (2013). The Data Science Education Dilemma. Technology Innovations in Statistics Education, 7(2), Article 2. https://doi.org/10.5070/T572013891
[10]Heinemann, B., Opel, S., Budde, L., Schulte, C., Frischemeier, D., Biehler, R., Podworny, S., & Wassong, T. (2018). Drafting a Data Science Curriculum for Secondary Schools. Proceedings of the 18th Koli Calling International Conference on Computing Education Research, Koli Calling '18, 1-5. https://doi.org/10.1145/3279720.3279737
[11]Zhang, S. (2023, junio 1). An Analysis of Data Analytics Curriculum Development through an NSF Research Experience for Teachers (RET) Program in Arkansas. https://doi.org/10.18260/1-2--42601
[12]Aponte, C. G., Malallah, S., & Shamir, L. (2025, junio 22). A Review of K-12 Data Science Education in the United States: Trends, Tools, and Gaps. 2025 ASEE Annual Conference & Exposition. https://peer.asee.org/a-review-of-k-12-data-science-education-in-the-united-states-trends-tools-and-gaps
[13]Banchoff Tzancoff, C., Martin, S., Marcón, P., & Duggan, T. (2024). Herramientas Informáticas utilizadas para Introducir Aspectos Básicos de Ciencia de Datos en la Escuela Secundaria. 30° Congreso Argentino de Ciencias de la Computación - CACIC 2024, 350-359. https://doi.org/10.35537/10915/172755
[14]Randles, B. M., Pasquetto, I. V., Golshan, M. S., & Borgman, C. L. (2017). Using the Jupyter Notebook as a Tool for Open Science: An Empirical Study. 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL), 1-2. https://doi.org/10.1109/JCDL.2017.7991618
[15]Luxton-Reilly, A., Simon, Albluwi, I., Becker, B. A., Giannakos, M., Kumar, A. N., Ott, L., Paterson, J., Scott, M. J., Sheard, J., & Szabo, C. (2018). Introductory programming: A systematic literature review. Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education, 55-106. https://doi.org/10.1145/3293881.3295779
[16]Atenas, J., & Havemann, L. (2015). Open Data as Open Educational Resources: Case Studies of Emerging Practice. 2561242 Bytes. https://doi.org/10.6084/M9.FIGSHARE.1590031
[17]Saddiqa, M., Rasmussen, L., Magnussen, R., Larsen, B., & Pedersen, J. M. (2019). Bringing open data into danish schools and its potential impact on school pupils. Proceedings of the 15th International Symposium on Open Collaboration, OpenSym '19, 1-10. https://doi.org/10.1145/3306446.3340821
[18]Martin, S. S., Banchoff Tzancoff, C. M., Venosa, P., & Hurtado, L. (2025). Avances en la elaboración de estrategias para la inclusión de la ciencia de datos en la escuela secundaria. XXVII Workshop de Investigadores en Ciencias de la Computación (Mendoza, 10 y 11 de abril de 2025). http://sedici.unlp.edu.ar/handle/10915/184394
[19]Saddiqa, M., Magnussen, R., Larsen, B., & Pedersen, J. M. (2021). Open Data Interface (ODI) for secondary school education. Computers & Education, 174, 104294.
[20]Apiola, M. (with López-Pernas, S., & Saqr, M.). (2023). Past, present and future of computing education research: A global perspective (1st ed). Springer International Publishing AG.
[21] Martin S., Banchoff Tzancoff C., Marcón P. y Duggan T. (2014) Herramientas Informáticas utilizadas para Introducir Aspectos Básicos de Ciencia de Datos en la Escuela Secundaria. XXX CACIC. ISBN 978-950-34-2428-5. pp 350-359.
[22] Martin S., Banchoff Tzancoff C., Venosa P. y Hurtado L. (2024) Estrategias para incorporar análisis de datos en la escuela secundaria. XXVI WICC. ISBN 978-987-8352-57-2. pp 846-851
[23]Marcón P., Duggan T. (2024) Pandalyze: aprender Ciencia de Datos con programación en bloques. http://sedici.unlp.edu.ar/handle/10915/173859

Predicción temprana de la deserción universitaria en primer año: Un modelo multivariable basado en aprendizaje automático.

Pablo Vilaboa, Roxana Martínez

Universidad Abierta Interamericana (UAI)
Ciudad Autónoma de buenos Aires, Argentina
{Pablo.Vilaboa, Roxana.Martinez}@uai.edu.ar

RESUMEN

La deserción estudiantil en los primeros años de la formación universitaria constituye un desafío significativo tanto para las instituciones educativas como para los propios estudiantes, dado su impacto directo en el aprovechamiento de los recursos institucionales y en la planificación de estrategias académicas orientadas a cada cohorte.

En este contexto, la disponibilidad de herramientas que permitan identificar de manera temprana a los estudiantes en riesgo resulta fundamental para optimizar la asignación de recursos y fortalecer las políticas de acompañamiento académico. La incorporación de modelos predictivos basados en múltiples variables posibilita una detección más precisa y oportuna de situaciones de riesgo, superando las limitaciones de los enfoques tradicionales.

Asimismo, el desarrollo de un modelo predictivo multivariable complementado con una metodología que oriente su implementación en el ámbito institucional representa un enfoque innovador, al integrar información proveniente de distintos dominios y contribuir tanto al avance del conocimiento académico como a la mejora de los sistemas de alerta temprana.

La investigación se fundamenta en la necesidad de desarrollar y evaluar un enfoque metodológico integral que articule modelos predictivos y criterios sistemáticos de análisis para la detección temprana de abandono estudiantil. Con el propósito de disponer de manera holística variables académicas, conductuales, psicoemocionales e institucionales, superando así las limitaciones de los enfoques tradicionales y aprovechando las oportunidades que ofrecen las nuevas fuentes de recolección de datos educativas.

Palabras clave: *deserción universitaria; modelo multivariable; aprendizaje automático*

CONTEXTO

Este proyecto se suscribe en el marco de las iniciativas de investigación aplicada e innovación tecnológica radicada en la UAI, orientado a fortalecer la calidad académica y la gestión basada en datos en la educación superior.

La propuesta articula los avances en ciencia de datos, aprendizaje automático y analítica educativa con las necesidades institucionales vinculadas a la retención y permanencia estudiantil.

Con el propósito de integrar de manera holística variables académicas, conductuales, psicoemocionales e institucionales, superando así las limitaciones de los enfoques tradicionales y aprovechando las oportunidades que ofrecen las nuevas fuentes de recolección de datos educativas.

En este sentido, se orienta al estudio de las diferentes propuestas con el propósito de construir un marco estructurado que permita integrar dominios de datos contextuales (sociales, académicos, psicológicos e institucionales) del estudiante para establecer un mecanismo replicable de análisis de estos datos y traducir los resultados en lineamientos operativos para la toma de decisiones.

INTRODUCCIÓN

La deserción universitaria en primer año de carreras del área tecnológica, tales como ingeniería en sistema y licenciatura en videojuegos, constituye una problemática crítica de múltiples dimensiones que afecta tanto la trayectoria individual del estudiante como la sostenibilidad institucional. En el contexto latinoamericano, aproximadamente entre el 40%

de los estudiantes no completan sus estudios superiores. (Hermida Bravo, 2025)

Uno de los aspectos importantes, es determinar y comprender el termino deserción en el ámbito universitario. Este concepto, se concibe como el abandono de la formación académica, independiente de las condiciones y modalidades, y ajeno a cualquier expulsión forzada. (Lorenzi, 2024; Seminaria, 2018)

A su vez, Tinto (1975) señala que existe una amplia gama de comportamientos denominados deserción, pero agregó que este término no debe incluir todos los abandonos de estudio, por lo que, podemos acotar el marco del concepto considerando que no todas las acciones en este ámbito pueden ser consideradas como deserción universitaria. Ante estos argumentos, se entiende como abandono a aquellos estudiantes que, por razones sociológicas, psicológicas, académicas o institucionales, optan formalmente por presentar la documentación para solicitar la baja de la carrera.

El fenómeno de la deserción universitaria ha sido tradicionalmente analizado desde diversas perspectivas en relación con el conjunto de datos utilizados, aunque estos han generado propuestas muy valiosas, ocasionalmente se integran en un único conjunto de datos. Respeto a esto, podemos subclasificar las variables en un grupo referido al perfil del estudiante y otro según su rendimiento académico y su relación con el entorno institucional. En relación con esto último, por un lado, se expone la perspectiva psicológica que destaca variables relacionadas con la motivación, la autoeficacia y la adaptación (Zabala et al, 2018; Villegas, 2024). Por otro lado, se identifican los datos orientados desde una perspectiva sociodemográfico y económico la cual analiza factores contextuales como el capital cultural, las condiciones económicas familiares y las redes de apoyo social (Bordieu & Passeron, 1979), que han sido incorporados y enmarcado en propuestas como el modelo de integración de Tinto (Tinto, 1975) y adaptados a contextos específicos como el modelo académico (Saldaña Villa & Barriga, 2010). Sin embargo, estas propuestas teóricas, pese a su solidez conceptual, presentan limitaciones significativas de dominio de variables y su consolidación en un esquema de datos integral y aplicable para la modelización predictiva en tiempo real.

En la actualidad, la revolución digital en educación superior, caracterizada por el marcado aumento de las plataformas educativas (learning management system - LMS) y la digitalización de los procesos administrativos, ha generado volúmenes masivos de datos estudiantiles. Esta nueva realidad tecnológica posibilita un abordaje innovador: la aplicación de técnicas de aprendizaje automático (Machine learning - ML) para la identificación temprana de patrones de riesgo. Sin embargo, existe una brecha significativa entre la riqueza teórica multidimensional sobre la deserción y la aplicación de un modelo predictivo, que suele limitarse a variables de fácil medición como calificaciones académicas o la asistencia, ignorando aspectos fundamentales vinculados a las variables de dominios psicológicos, emocionales y sociales. (Valencia-Arias, 2025)

Factores Predictores de deserción

Como hemos comentado en párrafos anteriores, el impacto de la deserción en la educación superior es centro de interés de las autoridades. Este aspecto, ha sido muy estudiado y debatido desde distintos aspectos y disciplinas. Desde una perspectiva técnica, los avances actuales en ciencia de datos y aprendizaje automático ofrecen la posibilidad de abordar el fenómeno mediante la construcción de patrones de riesgo predictivos. No obstante, la calidad y utilidad de estos modelos dependen directamente de un diseño riguroso del conjunto de los datos que los alimenta, el cual debe contemplar criterios de relevancia teórica, validez de medición, y tratamiento de valores faltantes o calidad de datos. Cada variable debe estar respaldada por literatura previa y justificar su pertenencia a un dominio especifico trabajando con escalas homogéneas.

El enfoque multivariable fue perfectamente abordado por diversos autores, identificando la deserción universitaria como un fenómeno multidimensional. En esta propuesta, se propone relevar y organizar dichas variables con el fin de aportar un enfoque clasificatorio que pudiera determinar los dominios relevantes para el

formato del conjunto de datos. (Santana & Cortez, 2021; Porras, 2023)

En este escenario multivariable, se incorporan variables vinculadas a los aspectos psicológicos del alumno, poniendo foco en el individuo. Considerando factores como, la motivación, y las habilidades de adaptación a la vida universitaria. En este sentido, es necesario identificar factores como el correcto sentimiento vocacional del alumno, sus motivaciones, y su grado de autoestima. Estos últimos son elementos importantes para relevar puesto que pueden condicionar el deseo de continuar o abandonar la carrera elegida. (Donoso et al, 2010)

Otro de los dominios apropiados para comprender el comportamiento de alumnos, es la sociología. En este enfoque, se agrega el análisis del grupo familiar y social, reconociendo que las condiciones que constituyen su hábitat influyen de manera significativa en cada uno. Respecto a esto, los autores lo clasifican en dos grupos: los microsistémicos, que refieren a cambios en la constitución del grupo familiar o enfermedades. Y los meso-sistémicos, que abonan a relaciones de carácter personal como pérdida de empleos o falta de soporte social. (Canales & De los rios, 2009). En este sentido, el nivel educativo de los padres y las condiciones económicas de la familia pueden ser relevantes en la toma de decisión respecto a la continuidad de sus estudios. Aunque, son variables que no se consideraran como un indicador de carácter absoluto, estas variables permiten estimar el capital cultural y las barreras de acceso al mundo universitario. (Santana & Cortez, 2021, Hermida Bravo, 2025)

Por otro lado, tenemos las variables institucionales y organizacionales. Estas variables, muy variables en las instituciones aseguran cierta contención a los alumnos de los primeros años. En grupo, se consideran cuestiones como horarios y compatibilidad con otras obligaciones. Percepción de calidad docente e infraestructura y apoyos académicos o económicos. (Suárez & Díaz, 2015)

Diferentes autores señalan que, sentirse parte de la institución es relevante para promover la retención. Haciendo foco en esto último, la no participación en actividades extracurriculares o tutorías, así como la poca participación puede ser un indicador que evidencia la ausencia de este sentido de pertenencia. (Canales & De los Rios, 2009)

Por último, se encuentran las variables académicas, de todas son las más claras para analizar. Puesto que un mal rendimiento del estudiante puede impactar en los aspectos psicológicos y con ello, promover el abandono. A continuación, enumeramos aquellas más representativas, tales como, calificaciones en asignaturas, la cantidad de materias aprobadas, cantidad de inasistencias, y por último el promedio de horas de estudio. (Eckert & Suénaga, 2015; Navarrete & Puchi, 2013)

En términos metodológicos, no se considera alguna estandarización respecto a las escalas de medición (Cuantitativa: discreta u ordinal, o cualitativa: Nominal u ordinal), o alguna normalización respecto a variables temporales (semestres, año, etc.) como cualquier indicador derivado (ratios entre tareas, complimiento, o participación en el LMS). Siendo esto un aspecto un aspecto a evaluar en un enfoque metodológico futuro. Así mismo en el marco de dicha metodología, se evaluará el impacto de variables que pudieran introducir sesgos o cuya información no pueda obtenerse de forma ética o consistente.

La literatura analizada, que se resume en la *Figura 1*, se muestra el encuadre para definir el conjunto de variables que serán insumos para el estudio.

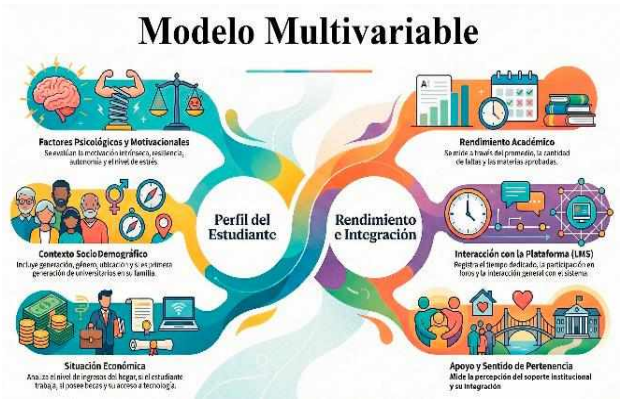


Figura 1. *Distribución y clasificación de variables.*

En esta Figura, se considera la multidimensionalidad del fenómeno, integrando variables psicológicas, sociológicas, económicas, institucionales, académicas y de integración en un único conjunto de dominios

interrelacionados. Del material relevado, se clasificaron las variables encontradas con el propósito de filtrar aquellas con menor impacto o con un impacto de difícil medición. Luego de un análisis holístico se podrá construir un modelo predictivo útil que nos permita detectar de manera temprana alumnos que cuadren con el patrón de abandono. Permitiendo a las instituciones implementar políticas personalizadas de retención y acompañamiento institucional.

LINEAS DE INVESTIGACIÓN y DESARROLLO

La investigación adopta un enfoque cuantitativo, de tipo aplicado y descriptivo–predictivo, dado que busca identificar relaciones causales entre variables, al mismo tiempo, desarrollar un modelo con aplicaciones prácticas en contextos académicos universitarios. Este enfoque se sustenta en el análisis estadístico y técnicas de aprendizaje automático, con el fin de construir y evaluar un modelo predictivo capaz de anticipar la deserción en el primer año. El diseño de la metodología es no experimental, transversal y predictivo, basado en el análisis de datos históricos de estudiantes de primer año, mientras que la muestra se seleccionará a partir de registros académicos de una o más cohortes, considerando criterios de disponibilidad y calidad de los datos.

RESULTADOS OBTENIDOS / ESPERADOS

Los objetivos de la presente investigación esperan fundamentar la necesidad de desarrollar y evaluar un enfoque metodológico integral que articule modelos predictivos que permita identificar patrones de riesgo asociados a la deserción en el nivel universitario durante los primeros años de la carrera. Para alcanzar esta meta, el modelo propuesto integrará dominios multivariables, con el fin no solo de apoyar a la detección temprana de estudiantes con mayor probabilidad de abandonar la educación superior, sino también facilitar la generación de evidencia formal para la formulación de modelos sostenibles de alerta temprana.

En función de este propósito, el estudio plantea, en primer lugar, realizar un mapeo sistemático de la literatura científica sobre los modelos de deserción estudiantil, poniendo especial atención en los aportes provenientes de diferentes dominios y la interrelacione entre ellos. A partir de esta revisión, se analizarán y compararán los distintos modelos teóricos y predictivos reportados, con el objetivo de identificar aquellos que demuestren mayor solidez y respaldo empírico.

Con base en lo anterior, se diseñará y construirá un dataset unificado que integre las variables que han sido consideradas relevantes en investigaciones previas, siguiendo criterios teóricos y metodológicos respaldados por la literatura. Posteriormente, se seleccionarán e implementarán algoritmos de aprendizaje automático acordes con el problema de estudio, y su desempeño se evaluará mediante métricas estandarizadas y pruebas empíricas controladas.

Asimismo, se examinará la capacidad predictiva del modelo desarrollado y su utilidad práctica como herramienta de apoyo para la toma de decisiones académicas. De manera complementaria, se explorarán técnicas de visualización de datos que permitan interpretar y comunicar los resultados de forma clara dentro de contextos institucionales.

Finalmente, a partir de los hallazgos obtenidos, se elaborará un conjunto estructurado de recomendaciones orientadas al diseño e implementación de sistemas de alerta temprana, enfocados en la identificación oportuna de estudiantes con riesgo de deserción.

FORMACIÓN DE RECURSOS HUMANOS

El presente trabajo se desarrolla en el marco de una tesis de maestría en tecnología informática de la UAI y es llevado adelante bajo la dirección académica de una directora de tesis con perfil alineado en calidad y ciencia de datos.

BIBLIOGRAFÍA

Bordieu, P., & Passeron, J. (1979). La Reproducción. Barcelona: Editorial Laia, S.A. Barcelona: Laia, S.A.

Canales, A., & De los rios, D. (2009). Retención de estudiantes vulnerables en la educación universitaria.

Revista Calidad en la Educación, 30, 50-83. doi:https://doi.org/10.31619/caledu.n30.173

Donoso, S., Donoso, G., & Arias, Ó. (2010). Análisis de deserción-permanencia de estudiantes universitarios utilizando técnica de clasificación en minería de datos. Calidad En La Educación, (33), 15-61. doi:https://doi.org/10.31619/caledu.n33.138

Eckert, K. B., & Suénaga, R. (2015). Scielo. Obtido de Analysis of Attrition-Retention of College Students Using Classification Technique in Data Mining.: https://dx.doi.org/10.4067/S0718-50062015000500002

Hermida Bravo, K. R. (2025). Factores psicosociales que inciden en la deserción universitaria. Prohominum, 7(3), 174–185. doi:https://doi.org/10.47606/ACVEN/PH0364

Lorenzi, G. (2024). Deserción en la educación superior Argentina. Las competencias digitales como nuevo desafío para la inclusión. *Revista de Ciencia, Tecnología e Innovación*, *22*(30), 269-278.

Navarrete, S. C., & Puchi, R. (2013). Scielo. Obtido de Factors associated with the dropout/ retention of Mapuche students of Universidad de la Frontera and the impact of academic support programs: https://dx.doi.org/10.4067/S0718-45652013000200003

Porras, J., Lara, J., Romero, C., & Ventura, S. (2023). A Case-Study Comparison of Machine Learning Approaches for Predicting Student's Dropout from Multiple Online Educational Entities. . Algorithms, 16. 554. 10.3390/a16120554.

Saldaña Villa, M., & Barriga, O. A. (2010). Adaptación del modelo de deserción universitaria de Tinto a la Universidad Católica de la Santísima Concepción, Chile. Obtido em 19 de 1 de 2026, de http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1315-95182010000400005&lng=es&nrm=iso

Santana, A., & Cortez, M. (2021). Deserción estudiantil en la educación superior: reflexiones sobre la gestión enfocada en la retención o la permanencia. Revista Educación, vol. 45, núm. 1, pp. 519-533. doi:https://doi.org/10.15517/revedu.v45i1.40602

Suárez, N., & Díaz, L. (2015). Estrés académico, deserción y estrategias de retención de estudiantes en la educación superior. Revista de Salud Pública, 17(2), 300-313. doi:https://doi.org/10.15446/rsap.v17n2.52891

Tinto, V. (1975). Dropout from Higher Education: A theoretical synthesis of recent Research. Review of educational Research, 45(1), 89-195. doi:https://doi.org/10.3102/00346543045001089

Valencia-Arias A, Valera Aredo JC, Valencia J, Cardona-Acevedo S, Patiño-Vanegas JC and Uribe Bedoya H (2026) Key determinants of university dropout: academic, economic, and psychosocial factors in student retention. Front. Educ. 10:1701644. doi: 10.3389/feduc.2025.1701644

Villegas, B., & Muñez Lira, L. (2024). Factores asociados a la deserción estudiantil en el ámbito universitario. Una revisión sistemática 2018-2023. Ride: Revista iberoamericana para la investigacion y el desarrollo educativo, Vol. 14, Núm. 28. doi:https://doi.org/10.23913/ride.v14i28.1923

Zabala, M., Vázquez, M., Gonzalez, I., & Bazán, A. (2018). Factores Internos, Externos y Bilaterales asociados con la deserción en estudiantes universitarios. Interacciones, 4(1), 59-69. doi:https://doi.org/10.24016/2018.v4n1.103

Proyecto para la Gestión de la Información Académica en las Carreras de Ingeniería de la FacultadRegional Córdoba de UTN

Roberto Miguel Muñoz, Franco Mana, Luis Esteban Damiano, LorenaBeatriz Genoud, Clarisa Liliana Stefanich,

{*rmunoz; fmana; ldamiano; lgenoud; cstefanich*}*@frc.utn.edu.ar*

GIEIMeT Grupo de Investigación en Enseñanza de la Ingeniería Mediada por Tecnologías
Universidad Tecnológica Nacional, Facultad Regional Córdoba.

RESUMEN

El proyecto, que se presenta, permitirá analizar el gran volumen de datos almacenados durante másde 30 años en la Facultad Regional Córdoba de la Universidad Tecnológica Nacional. Actualmente el registro de la información académica se hacea través del sistema SysAcad, el cual ofrece posibilidades de múltiples consultas, tales como las inscripciones a cursado o exámenes, docentes por curso, expedientes, etc. Sin embargo, estos datos no siempre se transforman en información útil para la toma de decisiones. En este contexto, el proyecto propone el diseño e implementación de una estrategia integral, basada en el análisis sistemático de datos académicos, orientado a mejorar la gestión académica en la institución y a fortalecer la toma de decisiones, basadas en evidencias de las ocho carreras de grado de la Facultad. La propuesta se fundamenta en la experiencia previa del equipo, en el diseño e implementación del Sistema de Monitoreo Académico Estudiantil (SMAE) y articuladicha trayectoria con la definición sistemática de indicadores académicos y el desarrollo de una plataforma analítica digital. Esta integración permite transformar datos institucionales en información estratégica, visualizada mediante tableros dinámicos e interactivos, orientados a apoyar la toma de decisiones en los distintos niveles de gestión académica.

Palabras clave: *gestión académica, analítica de datos, ingeniería, indicadores académicos.*

CONTEXTO

Lo que sucede habitualmente es que las instituciones universitarias generan y almacenan grandes volúmenes de datos académicos como resultado de los procesos de enseñanza, aprendizaje y gestión [1, 2, 3, 4]. Esto es así también en la Facultad Regional Córdoba (FRC) de la Universidad Tecnológica Nacional (UTN) y cuando se desea determinada información, por ejemplo, que resuma situaciones o comportamientos, se deben iniciar solicitudes no programadas y significa tiempo de desarrollo. Por tal motivo el equipo orienta sus esfuerzos en este proyecto. Una temática ampliamente abordada e investigada, es que las carreras de ingeniería atraviesan, a nivel nacional e internacional, problemáticas vinculadas con la deserción estudiantil, la prolongación de las trayectorias académicas y las dificultades para el egreso oportuno [5, 6]. La FRC no es ajena a esas situaciones, con una matrícula numerosa, en las ocho especialidades de Ingeniería que se ofrecen actualmente, y un escenario atravesado por diversas variables, muchas de ellas externas a la unidad académica, como problemas sociales y económicos que afectan a la comunidad estudiantil. Estas situaciones se ven potenciadas por la magnitud y complejidad de las estructuras

Académicas, que abarcan desde planes de estudios recientemente actualizados hasta un elevado número de cátedras y las comisiones en las que se brindan, las cuales generan datos de manera continua en los sistemas académicos institucionales. La FRC dispone de una base de datos académica robusta, administrada y alimentada por un sistema propio de la UTN (SysAcad), empleada principalmente con fines administrativos académicos, con más de 30 años de registros históricos y adoptada por 28 facultades regionales, de las 30 que conforman a la UTN. La información allí contenida es accedida a través de múltiples consultas y vistas de usuarios, pero no se presenta en formatos que faciliten el análisis integral y la toma de decisiones estratégicas por parte de directores de carrera, jefes de cátedra, docentes y autoridades de gestión. Este proyecto de investigación también se encuentra en concordancia con la unidad de gestión "Unidad de Visibilidad y Transparencia de la Universidad Tecnológica Nacional", que es un espacio institucional dedicado a asegurar el acceso a la información pública, la transparencia de la gestión y la difusión de políticas, programas y actividades desarrolladas por la universidad [7]. Esta unidad forma parte de las políticas institucionales que promueven la rendición de cuentas y el ejercicio de una gestión abierta, lo cual es coherente con los objetivos del proyecto de investigación, al buscar que la información académica institucional se transforme en conocimiento útil para la toma de decisiones y el fortalecimiento de la gestión educativa. En este sentido, la estrategia de gestión de información académica propuesta no solo se alinea con la mejora interna de procesos, sino también con los principios de transparencia y gobernanza institucional promovidos por la UTN.

1. INTRODUCCIÓN

La toma de decisiones basada en evidencias se configura como un eje estratégico para la mejora continua de la calidad educativa en el ámbito universitario [3, 5]. En el ámbito universitario, contar con información académica confiable, actualizada y pertinente resulta clave, para diseñar estrategias de acompañamiento estudiantil, optimizar prácticas pedagógicas y fortalecer los procesos de gestión, favoreciendo así intervenciones más eficaces y contextualizadas. El presente trabajo se inscribe en el campo de la Tecnología Educativa y la Enseñanza de la Ingeniería y se vincula al Proyecto de Investigación y Desarrollo (PID) "Estrategia para la gestión de la información académica en las carreras de Ingeniería de la Facultad Regional Córdoba de UTN", identificado con el código COTEEC523 y homologado por la Secretaría de Investigación, Desarrollo y Posgrado de la UTN. El mismo se desarrolla en el Grupo de Investigación en Enseñanza de la Ingeniería Mediada por Tecnologías (GIEIMeT), de la FRC. El período de ejecución se extiende desde el 1 de abril de 2025 hasta el 1 de abril de 2027. Asimismo, recupera la experiencia previa de parte del equipo de investigación en el desarrollo del Sistema de Monitoreo Académico Estudiantil (SMAE), iniciado en 2016 en el Departamento de Ingeniería en Sistemas de Información de la FRC[8].

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación está orientada a la gestión de la información académica mediante plataformas analíticas digitales, combinando enfoques de análisis cuantitativo y cualitativo [2, 4]. Es objetivo colaborar en fortalecer los procesos de toma de decisiones académicas y de gestión en las carreras de ingeniería, a partir del aprovechamiento sistemático de los datos institucionales.

2.1 METODOLOGÍA DE ENFOQUE Y DISEÑO DE LA INVESTIGACIÓN

Desde una perspectiva teórica, la investigación se apoya en los aportes de la analítica de datos aplicada a la educación superior, la gestión académica basada en evidencias y la enseñanza de la ingeniería mediada por tecnologías. En este sentido, se concibe a los sistemas de información académica no sólo como herramientas administrativas, sino como fuentes estratégicas de conocimiento para el análisis de trayectorias estudiantiles, la evaluación de políticas académicas y el diseño de acciones de mejora continua.

En el plano metodológico y tecnológico, el proyecto articula los siguientes ejes:

- Definición y validación de indicadores académicos relevantes para la gestión de carreras, cátedras y trayectorias estudiantiles, considerando dimensiones como ingreso, permanencia, regularidad, rendimiento, recursado y egreso.
- Procesos de extracción, transformación y carga de datos (ETL) desde el sistema académico institucional, orientados a la construcción de repositorios analíticos que integren información histórica y actual, garantizando criterios de seguridad, confidencialidad y calidad de los datos.
- Análisis exploratorio y descriptivo de datos, con el objetivo de identificar patrones, tendencias y comportamientos significativos asociados a las trayectorias académicas y al desempeño deestudiantes y cátedras.
- Desarrollo de una plataforma analítica digital con tableros dinámicos e interactivos, adaptados a distintos perfiles y niveles de gestión.

Desde el desarrollo tecnológico, se priorizan herramientas compatibles con la infraestructura institucional, acceso web y perfiles según roles de gestión. El PID prevé incorporar alertas tempranas y analítica predictiva para detectar riesgos académicos, bajo un uso ético y responsable de la información.

3. RESULTADOS OBTENIDOS Y ESPERADOS

3.1 RESULTADOS OBTENIDOS Y AVANCESCONSOLIDADOS DEL PROYECTO

A partir del trabajo desarrollado por el equipo de investigación, en el marco del PID, y de los aportes sistematizados en el artículo presentado en CoNaIISI 2025 [8], se han alcanzado avances sustantivos que permiten dar cuenta no sólo de la viabilidad técnica de la propuesta, sino también de su pertinencia metodológica e importancia en los espacios de gestión y decisión de la FRC, en el contexto institucional en el que se inscribe.

Entre los principales resultados obtenidos, hasta el momento, se destacan:

- Definición estratégica del proyecto, con objetivos y alcances claros, basada en el diálogo institucional y el análisis de necesidades de gestión académica.
- El acceso controlado y la gobernanza inicial de los datos académicos, en articulación con las autoridades de la Facultad y el Centro de Cómputos, garantizando criterios de confidencialidad, uso responsable y acuerdos de actualización diaria.
- La comprensión y modelado de la arquitectura de datos del sistema académico SysAcad, abordando inconsistencias estructurales y semánticas, y estableciendo criterios de normalización que facilitan el análisis posterior.
- La implementación de procesos ETL, que permiten integrar información histórica y actual en repositorios analíticos independientes de los sistemas operativos.
- La definición y validación de un conjunto base de indicadores académicos, alineados con las necesidades de gestión de carreras y cátedras, que contemplan variables tales como ingreso, inscripción, estados académicos, trayectorias, rendimiento y

egreso.

- El desarrollo de tableros de control experimentales, que posibilitan la visualización de la información y la identificación de patrones relevantes.
- La validación inicial de los resultados con actores institucionales, lo que permitió ajustar requerimientos según las necesidades de los usuarios.

3.2 RESULTADOS ESPERADOS Y PROYECCIÓN DE LA LÍNEA DE INVESTIGACIÓN

En coherencia con la línea de investigación definida y tomando como referencia los avances y proyecciones planteadas en el trabajo presentado en CoNaIISI 2025 [7], los resultados esperados del proyecto se orientan a profundizar y consolidar la estrategia desarrollada, ampliando su impacto académico, pedagógico y de gestión.

En este sentido, se espera avanzar en:

- Consolidación de una plataforma analítica web institucional que integre información académica de las carreras de ingeniería con acceso por perfiles diferenciados.
- Ampliación y refinamiento de indicadores académicos, incorporando dimensiones de docentes, cátedras y asignaturas para habilitar análisis comparativos y longitudinales.
- Profundización del análisis de trayectorias académicas para identificar cuellos de botella y factores asociados al recursado y la deserción.
- Desarrollo de alertas tempranas para detectar riesgos académicos y diseñar estrategias de acompañamiento basadas en evidencia.
- Avance hacia analítica predictiva para anticipar escenarios académicos, bajo criterios éticos, de transparencia y protección de datos.
- El fortalecimiento de la gobernanza y calidad de los datos, mediante políticas de acceso, con- troles de calidad y mecanismos de trazabilidad de la información.
- La integración de la plataforma analítica con otros sistemas institucionales, favoreciendo una visión más integral de los procesos académicos.
- El diseño e implementación de estrategias multinivel de capacitación y transferencia, orientadas a promover el uso efectivo de la información generada.

4. Formación de recursos humanos

El proyecto contempla la formación de recursos humanos en el área de analítica de datos y gestión académica, articulando la participación de docentes investigadores, investigadores en formación, personal de apoyo y estudiantes avanzados de ingeniería, en el marco del PID. El equipo está conformado por la dirección, investigadores, personal no docente y estudiantes becarios; cuatro deestos últimos se encargan de la programación yel soporte tecnológico para la implementación y evolución de la plataforma analítica digital.

5. Bibliografía

[1] J. Arias-Montes, “Aplicación de las herramientas de calidad al rendimiento académico,” AA- CINI - Revista Internacional de Ingeniería Industrial, no. 1, pp. 21–32, 2023.

[2] A. Brown and B. Miller, “Unlocking the potential of educational data: A conceptual framework,” Educational Researcher, vol. 48, no. 8, pp. 499–513, 2019.

[3] C. García and M. López, “Utilización de datos académicos para la mejora continua en la educación universitaria,” Revista Iberoamericana de Educación Superior, vol. 11, no. 31, pp. 42–61, 2020.

[4] R. Gómez and S. Pérez, “Análisis de datos académicos en la universidad tecnológica nacional: retos y oportunidades,” Revista

Latinoamericana de Educación Superior, vol. 7, no. 2, pp. 98–115, 2021.
[5] R. Martínez and N. Góngora, Evaluación de la gestión universitaria. República Argentina: CONEAU, 2000.
[6] UNESCO, Proyecto Regional de Indicadores Educativos para América Latina y el Caribe. Oficina Regional de Educación para América Latina y el Caribe, 2007.
[7] Universidad Tecnológica Nacional. (2026) Unidad de visibilidad y transparencia. [Online]. Available: https://utn.edu.ar/es/la-universidad/ unidad-de-visibilidad-y-transparencia
[8] R. M. Muñoz, F. Mana, C. L. Stefanich, L. Genoud, A. R. Kabusch, and L. E. Damiano, "Perspectivas para la construcción del sistema de monitoreo académico estudiantil," Congreso Nacional de Ingeniería en Informática/Sistemas de Información (CoNaIISI), Argentina, 2025.

REINGENIERÍA Y EXPANSIÓN DE UNA INFRAESTRUCTURA CAVE DE REALIDAD VIRTUAL PARA SIMULACIÓN INMERSIVA MULTIUSUARIO EN INGENIERÍA

Javier J. Rosenstein[1,2], Carlos A. Tomba[2], Rodrigo Gonzalez[1], Verónica Míguez[1], Mario Mirallas[1], Martín Wieckowski[2], Tomás Fiorentini[2]
[1]Instituto de Investigaciones – Facultad de Informática y Diseño – Universidad Champagnat – Mendoza - Argentina
[2]Facultad de Ingeniería – Universidad Nacional de Cuyo – Mendoza - Argentina

[1]rosensteinjavier@uch.edu.ar, [2]javier.rosenstein@ingenieria.uncuyo.edu.ar, [2]alejandro.tomba@ingenieria.uncuyo.edu.ar, [1]gonzalezrodrigo@uch.edu.ar, [1]miguezveronica@uch.edu.ar, [1]mirallasmario@uch.edu.ar, [2]tincho92wk@gmail.com, [2]tomasabelfiorentini@gmail.com

RESUMEN

El presente trabajo describe la reingeniería integral de una infraestructura CAVE (Cave Automatic Virtual Environment) instalada en la Facultad de Ingeniería de la Universidad Nacional de Cuyo y destinada al aprendizaje de estudiantes de la carrera de Ingeniería en Petróleo mediante la simulación de procesos de perforación en una planta de extracción de Petróleo. El sistema original, desarrollado en Unity bajo una arquitectura distribuida multi-PC y sin acceso al código fuente, presenta limitaciones críticas de mantenimiento y escalabilidad, dado el alto grado de utilidad que tiene actualmente para el aprendizaje de los estudiantes es que se decide rediseñar la misma mejorando sus capacidades y buscando mayor inmersividad en su utilización.

El presente trabajo se enmarca en la continuidad de la línea de investigación en Realidad Virtual iniciada en la Universidad Champagnat y actualmente extendida, mediante articulación interinstitucional, a la Facultad de Ingeniería de la Universidad Nacional de Cuyo, particularmente en la asignatura Realidad Virtual de la carrera Ingeniería Mecatrónica.

Se propone una nueva arquitectura unificada basada en una estación de trabajo de alto rendimiento con GPU multi-salida capaz de controlar los cuatro proyectores del entorno inmersivo, complementada con un modelo de interacción multiusuario mediante dispositivos móviles conectados inalámbricamente.

Actualmente se está terminando de desarrollar un MVP funcional que valida la viabilidad técnica del enfoque.

El proyecto es financiado conjuntamente por la Universidad Champagnat y la Universidad Nacional de Cuyo, articulando continuidad académica con infraestructura estratégica institucional, ejecutado operativamente con participación activa de docentes y estudiantes de la Universidad Nacional de Cuyo, institución beneficiaria directa por ser poseedora de la infraestructura CAVE.

Palabras clave: Realidad Virtual, CAVE, simulación inmersiva, reingeniería de sistemas, interacción multiusuario, educación en ingeniería.

CONTEXTO

La línea de investigación en Realidad Virtual dirigida por el Mgter. Javier J. Rosenstein se inició en la Universidad Champagnat, orientada al diseño de entornos inmersivos interactivos aplicados a contextos académicos y tecnológicos.

En su etapa actual, el proyecto se articula con la Facultad de Ingeniería de la Universidad Nacional de Cuyo, específicamente con la asignatura Realidad Virtual de la carrera de Ingeniería Mecatrónica, ampliando su alcance hacia la simulación técnica aplicada.

La Facultad de Ingeniería posee una infraestructura CAVE utilizada en Ingeniería en Petróleo para simulaciones de perforación. Una CAVE, según Cruz-Neira et al. [1], constituye un entorno inmersivo multi-proyección que permite interacción tridimensional envolvente mediante superficies sincronizadas.

La CAVE existente a mejorar está compuesta por cuatro proyectores con ópticas gran angular que cubren piso y tres paredes, operando originalmente con cuatro computadoras independientes y una estación adicional para HMI y control.

El sistema existente opera sobre una arquitectura distribuida de múltiples computadoras y fue desarrollado en Unity. La falta de acceso al código fuente y la obsolescencia tecnológica limitan su uso y expansión, generando la necesidad de una reingeniería completa.

1. INTRODUCCIÓN

El objetivo principal del proyecto consiste en la reingeniería y refuncionalización de la Cave, para lo cual es establecen los siguientes objetivos teóricos y experimentales:

Objetivos teóricos

•Diseñar un modelo arquitectónico unificado para entornos CAVE académicos.

•Formalizar un esquema de interacción multiusuario distribuida.

•Proponer un modelo replicable de laboratorio inmersivo universitario.

•Consolidar una línea de investigación interinstitucional sostenible.

Objetivos experimentales

• Implementar un prototipo funcional en Unity con renderizado multi-superficie sincronizado.

•Validar el control de cuatro proyectores desde una única estación de trabajo.

•Desarrollar interfaces móviles en Processing Java y Python para interacción inalámbrica.

•Relevar requerimientos con docentes de Perforaciones I y II para reconstrucción del simulador.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

De este trabajo se desprenden 4 lineas claras de trabajajo que están fuertemente ligadas, desde la reingeniería del modelo actual, necesario para la continuidad y evolución de la Cave, pasando por el diseño de una nueva Arquitectura gráfica para comandar los 4 proyectores en sincronía, la nueva modalidad pensada de interacción multiusuario de manera inalámbrica y el planteo interinstitucional que permite que ambas instituciones sumen recursos para llevar adelante la financiación del proyecto, estas lineas se definen de la siguiente manera:

2.1 Reingeniería de infraestructura inmersiva

Se aborda la sustitución de un sistema legacy sin código fuente mediante rediseño completo, aplicando principios de arquitectura de sistemas interactivos distribuidos.

2.2 Arquitectura gráfica multi-proyección

La migración hacia una única estación con GPU multi-salida reduce complejidad y mejora sincronización, alineándose con tendencias

actuales en sistemas inmersivos de alto rendimiento [2].

2.3 Interacción multiusuario con dispositivos móviles

Se adopta un modelo cliente-servidor en red local donde smartphones actúan como controladores inalámbricos. Este enfoque incrementa la inmersión y participación activa, coherente con estudios sobre RV colaborativa en educación en ingeniería [13][15].

2.4 Transferencia tecnológica interinstitucional

El proyecto se desarrolla en forma colaborativa desde la Universidad Champagnat como continuidad de su línea de investigación, mientras que la implementación práctica y el despliegue físico se realizan en la Universidad Nacional de Cuyo, generando un modelo colaborativo de beneficio mutuo.

La UNCuyo resulta directamente fortalecida al recuperar operatividad plena de su CAVE y ampliar su potencial como plataforma transversal de simulación académica.

3. RESULTADOS

Resultados obtenidos

•Diseño arquitectónico consolidado.

•MVP funcional con proyección multi-superficie.

•Comunicación inalámbrica preliminar con dispositivos móviles.

•Integración operativa de estudiantes becarios de UNCuyo en el desarrollo.

Resultados esperados

•Simulador completo de perforación actualizado.

•Plataforma CAVE multiusuario estable.

•Apertura del laboratorio a otras carreras.

•Generación de tesis de grado, Proyectos finales integradores y publicaciones conjuntas.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto es dirigido por el Mgter. Javier J. Rosenstein (UCH–UNCuyo) y co-dirigido por el Ing. Carlos A. Tomba (UNCuyo).

La financiación de la línea de investigación proviene de la Universidad Champagnat. La ejecución técnica del MVP se desarrolla principalmente en la Universidad Nacional de Cuyo, con participación activa de los estudiantes Martín Wieckowski y Tomás Fiorentini.

Este esquema fortalece:

•Formación en desarrollo inmersivo avanzado.

•Integración investigación-docencia.

•Transferencia tecnológica intrauniversitaria.

•Consolidación de una línea interinstitucional sostenida.

Todos los avances logrados y las implementaciones realizadas relacionadas con el desarrollo del presente proyecto desde su primer etapa, será utilizado para el estudio de perforaciones de la carrera de ingeniería en Petróleo asi como también lo diseñado y desarrollado se utilizará como recurso para el dictado de la materia Realidad Virtual y serán capitalizados para brindar bases de desarrollo sobre las temáticas de Realidad Virtual de la carrera de Mecatrónica en particular y de ambas instituciones en general.

5. BIBLIOGRAFÍA

Realidad Virtual y entornos inmersivos

[1] C. Cruz-Neira, D. J. Sandin y T. A. DeFanti, “Surround-Screen Projection-Based Virtual Reality: The Design and

Implementation of the CAVE," *SIGGRAPH*, 1993.

[2] J. Jerald, *The VR Book: Human-Centered Design for Virtual Reality*, Morgan & Claypool Publishers, 2016.

[3] M. Slater y A. Sanchez-Vives, "Enhancing Our Lives with Immersive Virtual Reality," *Frontiers in Robotics and AI*, 2016.

[4] R. Pallavicini, F. Ferrari y M. Mantovani, "Applications of Virtual Reality in Education: A Meta-Review," *Computers & Education*, vol. 147, pp. 103778, 2020.

Unity y desarrollo de entornos virtuales

[5] A. Watt y M. Policarpo, *3D Games: Real-Time Rendering and Software Technology*, Addison-Wesley, Edición 2, 2016. *(contexto general sobre gráficos)*

[6] J. Gregory, *Game Engine Architecture*, 3ª Ed., CRC Press, 2018. *(orientado a motores gráficos como Unity)*

[7] S. H. Lee, H. Ko y J. Park, "Development of Immersive Virtual Environments Using Unity 3D for Interactive Simulation," *IEEE Access*, vol. 8, pp. 128045-128057, 2020.

[8] Y. Liu et al., "A Unity-based Framework for Large-Scale Virtual Environment Rendering," *Journal of Computer Graphics Techniques*, vol. 9, no. 4, pp. 55-78, 2020.

Sincronización multi-pantalla y arquitecturas de proyección

[9] M. Cordeil et al., "ImAxes: Immersive Axes as Embodied Affordances for Interactive Multidimensional Data Visualisation," *CHI Conference*, 2019.

[10] K. Kaufmann y D. Schmalstieg, "Mathematics and Geometry Education with Collaborative Augmented Reality," *Computers & Graphics*, vol. 27, pp. 339-345, 2003. *(uso de multi-proyección)*

[11] P. Milgram y F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE Trans. Information Systems*, vol. E77-D, no. 12, pp. 1321-1329, 1994.

Educación mediante simulaciones y tecnologías inmersivas

[12] K. J. Bower et al., "Blended Reality Simulations for Engineering Education: Benefits and Challenges," *Computer Applications in Engineering Education*, vol. 27, no. 4, pp. 840-852, 2019.

[13] P. Radianti et al., "A Systematic Review of Immersive Virtual Reality Applications for Higher Education: Design Elements, Lessons Learned, and Research Agenda," *Computers & Education*, vol. 147, 2020.

[14] N. F. N. A. Rahman et al., "Virtual Reality in STEM Education: A Review on Technology and Developer Tools," *Applied Sciences*, vol. 10, no. 18, 2020.

[15] V. D. Nguyen et al., "Impact of Immersive Virtual Reality on Teaching and Learning in Higher Engineering Education," *Education and Information Technologies*, vol. 25, pp. 1359-1381, 2020.

[16] Rosenstein, Javier J. (2020). *Uso de VRPN para la implementación de una interfaz cerebro-computadora* [Tesis de Maestría en Teleinformática, Universidad de Mendoza].

RESULTADOS PROYECTO DEMOCRATIZANDO EDUCACIÓN Y TIC , IDICE UNSJ EXPANDIDA CON CURRICULUM INTEGRADO E INTERNACIONALIZADO CON AI

Emilse Elizabeth Carmona[1], Luciana Buteler2, Carolina González Zibarelli[3].
1,2,3 Instituto de Investigaciones en Ciencias de la Educación,
Facultad de Filosofía, Humanidades y Artes, UNSJ
3, Instituto Preuniversitario Escuela de Comercio Libertador General San Martín
emilsecoaching@gmail.com, butelerluciana.unsj.@gmail.com, gzibarellic@gmail.com

RESUMEN

El proyecto CICITCA, dirigido por la Mgter. Emilse Carmona y Mgter. Luciana Buteler, se sustenta en una investigación aplicada con enfoque mixto (cuali-cuantitativo) que analiza prácticas de enseñanza innovadoras. El eje central es el metamodelo de aprendizaje inverso (Flipped Learning) y el Neuroaprendizaje, los cuales se articulan con la gamificación y el paradigma de "aprender a hacer". Estas metodologías se entrelazaron para transformar el aula tradicional en un espacio de aprendizaje ubicuo, colaborativo e inclusivo, donde las tecnologías de la información se convierten en Tecnologías Accesibles al Conocimiento (TAC) con el uso de tecnologías emergentes como la Inteligencia Artificial aplicada a la educación. Una de las acciones de mayor impacto fue la creación de "puentes educativos" entre la Universidad Nacional de San Juan (UNSJ) y Ministerio de Educación Provincial, junto a las escuelas rurales y urbano-periféricas y urbanas de la provincia en sus distintos niveles educativos, acercando la universidad a la escuela y la escuela a la universidad a través de un curriculum integrado e internacionalizado para favorecer y potenciar el derecho a la educación superior.

CONTEXTO

El trabajo forma parte del proyecto CICITCA denominado "Democratizando Educación y Tic". En octubre de 2019 crearon la red Educación, Subjetividades y Tecnología (EDUSyT-FFHA), integrada por un equipo interdisciplinario de docentes y estudiantes, realizando las siguientes acciones detalladas dentro del proyecto:

Creación de espacios de reflexión: Propicia instancias de análisis y diálogos interdisciplinarios para reflexionar sobre los sentidos y la significación de las prácticas de enseñanza innovadoras.

Gestión de políticas educativas: Desarrolla procesos de elaboración de políticas y articulación que favorecen la creación de curriculum integrados con un enfoque tecno-pedagógico-didáctico.

Acompañamiento académico: Trabaja en el fortalecimiento de dispositivos destinados al acompañamiento de las trayectorias académicas de los estudiantes de la UNSJ.

Vinculación territorial con escuelas: Ejecuta acciones para acercar la universidad a estudiantes de escuelas secundarias rurales y urbano-marginales de San Juan, con el fin de garantizar el derecho a la educación superior.

Prevención de la deserción: Colabora en estrategias de articulación que buscan disminuir la deserción temprana en las carreras universitarias.

Análisis de metodologías activas: Investiga y analiza el uso de estrategias como el Neuroaprendizaje, el aprendizaje basado en retos, problemas, casos y la gamificación, evaluando su impacto en la colaboración, la

ubicuidad y la inclusión educativa.

Recopilación de experiencias previas: La red considera y amplía el análisis de relatos de experiencias obtenidos en cursos y jornadas dictados durante la pandemia y postpandemia, realizados en articulación con el Sistema Institucional de Educación a Distancia, el área de Educación a Distancia de la Universidad Nacional de San Juan.

1. INTRODUCCIÓN

Entre 2023 y 2024, la red EDUSyT ha pasado de la respuesta a la emergencia (postpandemia) a una planificación estratégica de la democratización tecnológica, donde la inclusión de la Inteligencia Artificial (IA) aplicada a la Educación y el fortalecimiento de los vínculos con las escuelas rurales de San Juan han sido los pilares fundamentales.

1.1. ACTIVIDADES Y RESULTADOS LOGRADOS EN EL MARCO DE LA RED EDUSyT y EL PROYECTO CICITCA

Investigación aplicada: Participa en la implementación de una estrategia metodológica de enfoque mixto (cuali-cuantitativo) para analizar las políticas y prácticas de enseñanza actuales.

Estas actividades se realizan en articulación constante con el Área de Educación a Distancia (EAD SIED UNSJ) y la Secretaría de Posgrado de FACSO UNSJ.

Preparar informe detallado sobre las metodologías activas

Este proyecto profundiza en la implementación y el impacto del aprendizaje basado en retos, problemas y casos, así como en la gamificación dentro de las prácticas de enseñanzas innovadoras. Se analiza su articulación con la Inteligencia Artificial en distintos niveles, con aprendizaje inverso y aula inversa, y el Neuroaprendizaje, enfocándose en cómo estos modelos, metamodelos pedagógicos y estrategias promueven la inclusión educativa, ubicua, colaborativa en las escuelas de San Juan.

En cuanto al análisis de las acciones de la red Educación, Subjetividades y Tecnología (EDUSyT-FFHA) durante el periodo 2023-2024 se revela una intensa actividad centrada en la innovación pedagógica, la articulación territorial y la inclusión educativa a través de las TIC y TAC como tecnologías educativas emergentes e Inteligencia Artificial.

La red es coordinada por Mg. Emilse Elizabeth Carmona, quien es directora del proyecto CICITCA "Democratizando Educación y TIC", logrando consolidar puentes entre la educación superior y contextos de escuelas urbanas, urbanas-periféricas y rurales, en situación de vulnerabilidad.

A continuación se detallan las principales acciones y logros de la red en este periodo:

1. Articulación Territorial y Vínculo con Escuelas Rurales

Un eje central ha sido la extensión hacia sectores rurales y urbano-marginales de San Juan para garantizar el derecho a la educación superior y disminuir la deserción.

IV Jornadas Taller (2024): Se denominaron “Democratizando el conocimiento con Educación y TIC: Desafíos creando redes en aulas híbridas”.

Acción en Territorio: En mayo de 2024, se realizó un taller presencial en la Escuela República Argentina (Bermejo) sobre puentes entre la UNSJ y escuelas rurales, utilizando Neuroaprendizaje y gamificación.

Debate Ético: En noviembre de 2024, se coordinó un proceso de indagación con estudiantes de la escuela secundaria Alas Argentinas (Rawson) sobre los retos de incluir la Inteligencia Artificial en la educación.

Tercer Taller de Reconstrucción y Diálogos (2023): Enfocado en construir lazos educativos entre la universidad y diversos contextos

escolares, con la participación de 130 asistentes.

2. Innovación Pedagógica y Metodologías Activas

La red ha promovido el uso de estrategias como el metamodelo de aprendizaje inverso (flipped learning), el Neuroaprendizaje y las aulas híbridas.

Clases en Espejo e Internacionalización: En junio de 2024, se llevaron a cabo clases en espejo en modalidad de aula híbrida, articulando con la Universidad Católica de Santiago de Guayaquil (Ecuador) para aplicar el metamodelo de aula inversa.

Gamificación: Se implementaron experiencias de gamificación con juegos de mesa modernos para desarrollar habilidades sociales y neurocognitivas en jóvenes.

3. Foco Emergente: Inteligencia Artificial (IA)

Durante 2024, la red desplazó gran parte de sus esfuerzos hacia la capacitación en IA aplicada a la educación.

Entrenamiento en IA: En octubre de 2024, se dictó el curso "Entrenamiento intensivo en Inteligencia Artificial aplicada a la Educación", destinado a docentes de universidades nacionales.

Debate de Vanguardia: Se organizaron paneles y conversatorios sobre el sentido ético de la IA en las Ciencias Humanas y Sociales, integrando a equipos de investigación de la FFHA y FACSO.

4. Formación y Capacitación de Posgrado

La red ha sido fundamental en el diseño e implementación de trayectos formativos especializados.

Diplomaturas: Continuó el soporte y gestión de la Diplomatura en Gestión Educativa Estratégica (enfocada en Neuroaprendizaje y tutorías) y la Diplomatura en Tecnologías Accesibles al Conocimiento (TAC), diseñada para la inclusión de estudiantes con discapacidad.

Tutorías Virtuales: Se fortalecieron los dispositivos de acompañamiento de las trayectorias académicas mediante la figura del "Tutor Coach" en entornos virtuales.

5. Producción Científica y Divulgación

La red ha sistematizado sus experiencias en diversas publicaciones y eventos científicos.

Publicaciones Destacadas: En 2023 se publicó el artículo "Articulaciones de propuestas de enseñanzas innovadoras..." en e-dialnet y capítulos de libros sobre mediaciones tecnológicas en pandemia y postpandemia, tesis de grado y posgrado en cuanto al grado de Magister.

Eventos: Participación en las XI Jornadas Nacionales sobre Discapacidad y Derechos (2023) y en el I Encuentro Bienal de Divulgación Científica en Educación (2024).

12. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación y desarrollo de se centra Se investiga la eficacia de modelos que desplazan el centro del proceso hacia el estudiante.

Línea 1: Aprendizaje Inverso (Flipped Learning): Aplicación del metamodelo de aula inversa en entornos virtuales, presenciales e híbridos.

Línea 2: Gamificación y "Aprender a Hacer": Uso de estrategias basadas en juegos y retos para desarrollar habilidades neurocognitivas y sociales.

Línea 3. Neuroaprendizaje: Integración de la neuropsicología en las prácticas de enseñanza y en el acompañamiento de las trayectorias académicas.

Línea 4. Tecnologías emergentes en

articulación con Inteligencia Artificial Generativa aplicada a la Educación y la Inclusión educativa.

La investigación busca transformar las TIC en herramientas de acceso real al conocimiento.

Tecnologías Accesibles al Conocimiento (TAC): Diseño e implementación de propuestas para estudiantes con discapacidad visual, auditiva, motora e intelectual y TEA.

Tiflotecnología: Desarrollo de ayudas informáticas específicas para personas con dificultades visuales totales o parciales.

Línea 5. Educación a Distancia y Escenarios Híbridos

Investigación sobre la mediación pedagógica en la bimodalidad.

Aulas Híbridas y Clases en Espejo: Estudio de la hibridación pedagógica-didáctica en redes interuniversitarias nacionales e internacionales (Ecuador).

Tutoría Virtual (Tutor Coach): Desarrollo de dispositivos de acompañamiento tutorial mediante técnicas de PNL3 y coaching para prevenir la deserción.

Línea 5. Democratización y Vínculo con el Territorio

Escuelas Rurales y Urbano-Marginales: Investigación aplicada para acercar la universidad a contextos de vulnerabilidad y garantizar el derecho a la educación superior.

Políticas de Gestión Educativa: Análisis de currículums integrados (tecno-pedagógico-didáctico) y la planificación estratégica de la educación superior.

Estas líneas se articulan bajo el programa de investigación del IDICE-FFHA denominado "Desarrollo y Aplicación de Nuevas Tecnologías Educativo Informáticas en Educación Especial", vigente desde 1996en 3.

RESULTADOS

El proyecto "Democratizando Educación y TIC" (CICITCA 2023-2025), ha alcanzado resultados significativos en diversas áreas que van desde la producción académica hasta la intervención territorial e institucional.

A continuación se detallan los principales logros organizados por sus ejes estratégicos:

2. Producción Científica y Académica

El equipo ha generado un volumen importante de conocimiento sistematizado:

Publicaciones: Artículos en revistas especializadas como e-dialnet y capítulos en libros colectivos sobre mediación tecnológica en pandemia y postpandemia.

Libros: La directora ha publicado obras clave como "Desafíos Innovadores en Educación: Estrategias con PNL3 y Quantum Coaching en Aulas Virtuales" y "Tutor coach en EVEA".

.4. FORMACIÓN DE RECURSOS HUMANOS

1. Formación de Recursos Humanos

Uno de los resultados más destacados es la formación de un total de 60 integrantes, distribuidos de la siguiente manera:

Tesis dirigidas: Se contabiliza la dirección de 15 tesis de grado, 3 de maestría y 1 de doctorado.

Becarios y pasantes: El proyecto cuenta con 3 becarios de investigación y un total de 22 pasantes de grado y maestría enfocados en áreas como gamificación, neuropsicología y tecnologías con IA.

Adscriptos: Se han integrado numerosos graduados y estudiantes como adscriptos a las cátedras de Informática Educativa y Estadística Aplicada.

Eventos: Se registran 36 acciones de comunicación pública de la ciencia y la

tecnología, incluyendo conferencias y ponencias en congresos nacionales e internacionales.

3. Impacto en el Territorio y Extensión

El proyecto ha logrado una vinculación efectiva con la comunidad a través de:

Escuelas Rurales y Urbano-Marginales: Realización de las IV Jornadas Taller presenciales en la Escuela República Argentina (Bermejo) y la Escuela Alas Argentinas (Rawson), enfocadas en Neuroaprendizaje y gamificación.

Contextos de Encierro: Implementación del proyecto "Música para la Libertad", que utiliza la educación musical como herramienta de transformación social en el Servicio Penitenciario de San Juan.

Servicios Comunitarios: Diseño de programas de beneficios socioeconómicos para estudiantes y asesorías educativas en territorio rural.

4. Internacionalización

Clases en Espejo: Se consolidó el vínculo con la Universidad Católica de Santiago de Guayaquil (Ecuador) mediante la realización de clases sincrónicas híbridas que integran el currículum de ambos países.

Redes: Coordinación de la red EDUSyT (Educación, Subjetividades y Tecnología), facilitando el intercambio académico con otras universidades latinoamericanas.

5. Logros Institucionales y Tecnológicos

Nuevas Carreras: Creación de la "Especialización en Tecnología Educativa accesible al conocimiento con mención en discapacidad" (Ord. 03/25-CD-FFHA).

Diplomaturas: Implementación de la Diplomatura en Tecnologías Accesibles al Conocimiento (TAC) y la Diplomatura en Gestión Educativa Estratégica, ambas en modalidad a distancia.

Espacios Inclusivos: Diseño e implementación de un Tiflocubículo en la biblioteca de la FFHA, un espacio equipado con tecno-ayudas informáticas para personas con discapacidad visual.

Alcance Digital: El video de formación sobre IA aplicada a la educación superó las 8,500 visitas, demostrando un alto impacto en la capacitación docente en servicio. De manera presencial en el año 2025 se formaron en Inteligencia Artificial aplicada a la educación ¡AI está la cuestión! 11.500 docentes de distintos niveles de la provincia de San Juan.

5. REFERENCIAS

Artículos en Revistas Científicas

- **Carmona, E. E.**, Gutiérrez, M. L., Pósito, R. M., Buteler, L., Mercado, P., & Guzmán, S. (2023). Articulaciones de propuestas de enseñanzas innovadoras entre área de educación a distancia en red educación subjetividades y tecnología, y articulaciones de posgrado de Universidad Nacional de San Juan. ***e-dialnet***, *2*, 1-12. ISSN 2953-4860.
- **Carmona, E. E.** (2021). Taller Reconstrucción y Diálogos, Derecho a la Educación: construyendo lazos educativos entre la escuela en contexto rural y Facultad de Filosofía, Humanidades y Artes - UNSJ. ***Contribuciones Científicas***, (2), 110-115. eissn 2684-0596.

Libros y Capítulos de Libros

- **Carmona, E. E.** (2019). ***Desafíos innovadores en educación: Estrategias con PNL3 y Quantum Coaching en aulas virtuales***. Editorial Fundación Universidad Nacional de San Juan.
- **Carmona, E. E.** (2019). ***Desafíos innovadores en educación: Tutor coach en EVEA*** (1.ª ed.). Facultad de Filosofía, Humanidades y Artes, Universidad Nacional de San Juan.
- **Carmona, E. E.**, Peralta, S., Harraca, M., Mateos, P. H., & Videla, A. N. (2025). Innovación educativa y modelos pedagógicos didácticos en la era digital. En S. Peralta & M. Harraca (Comps.), ***Educación en la frontera: Tramas***

situadas en la educación pública (1.ª ed., pp. 1-484). Universidad Nacional de San Juan. ISBN 978-950-605-964-4.

- Pósito, R. M., **Carmona, E. E.**, Ghiraldi, L., Dominguez, A., Morel, M., Diana Bunge, P., & Díaz Reinoso, V. (2023). Capítulo II: Ciencias Humanas. En ***Primeras Jornadas de buenas prácticas de docencia en los entornos educativos virtuales en tiempos de pandemia en la UNSJ*** (pp. 65-103). Editorial de la Universidad Nacional de San Juan. ISBN 978-950-605-934-7.

Informes Técnicos y Planes de Trabajo

- **Carmona, E. E.**, & **Buteler, L.** (2024). ***Informe de avance 2024: Proyecto CICITCA Democratizando Educación y TIC***. Secretaría de Ciencia y Técnica, Universidad Nacional de San Juan.
- **Carmona, E. E.**, & **Buteler, L.** (2023). ***Plan de trabajo: Democratizando Educación y TIC*** (Código 80020220100058SJ). Instituto de Investigaciones en Ciencias de la Educación (IDICE), Universidad Nacional de San Juan.

Resoluciones e Instrumentos Legales

- Facultad de Filosofía, Humanidades y Artes, Universidad Nacional de San Juan. (2025). ***Ordenanza N° 03/25-CD-FFHA: Carrera de Posgrado Especialización en Tecnología Educativa accesible al conocimiento con mención en discapacidad***.
- Facultad de Filosofía, Humanidades y Artes, Universidad Nacional de San Juan. (2025). ***Resolución N° 036/25-CD-FFHA: V Jornadas Taller Democratización del conocimiento con Educación y TIC***.
- Universidad Nacional de San Juan. (2023). ***Resolución N° 1005/23-CS: Aprobación del Proyecto Democratizando Educación y TIC***.

Documentos de Extensión

- **Carmona, E. E.**, Gutierrez, M. L., & Bonilla Morán, L. A. (2025). ***Proyecto de Extensión: V Jornadas Taller Democratización del conocimiento con Educación y TIC***. Universidad Nacional de San Juan; Universidad Católica de Santiago de Guayaquil.

SISTEMA INTELIGENTE DE SOPORTE A LA RETENCIÓN ESTUDIANTIL BASADO EN MINERÍA DE DATOS EDUCACIONAL EN UNCAUS

Stella Maris Gerzel,
Brollo, Magdalena Elisabet, Carlos Emanuel Aguirre, Rubén Andrés Cáceres, Paszco, Gustavo Ariel

Universidad Nacional del Chaco Austral. Departamento de Ciencias Básicas y Aplicadas. Carrera de Ingeniería en Sistemas de Información. Chaco. Argentina

licgerzelstella@gmail.com, magdalenabrollo@uncaus.edu.ar, aguirreemanuel@uncaus.edu.ar, rubencaceres@uncaus.edu.ar, paszcogustavo@uncaus.edu.ar

RESUMEN

La retención estudiantil en los primeros años universitarios constituye un problema estructural en carreras del área informática. Estudios recientes demuestran que el rendimiento académico puede explicarse por variables motivacionales, emocionales y sociodemográficas [1], así como por factores metodológicos e institucionales [2].

En la Universidad Nacional del Chaco Austral (UNCAUS), la carrera Ingeniería en Sistemas de Información presenta indicadores persistentes de bajo rendimiento en asignaturas troncales como “Algoritmos y Estructuras de Datos” y “Sistemas y Organizaciones”. Las intervenciones pedagógicas tradicionales no han revertido estructuralmente esta tendencia, evidenciando la necesidad de incorporar sistemas de soporte basados en análisis predictivo.

La literatura reciente en Minería de Datos Educacional y Learning Analytics demuestra que los modelos predictivos permiten identificar riesgo de fracaso académico con suficiente anticipación para implementar estrategias de intervención [3] [4]. En este contexto, la UNCAUS desarrolla un proyecto de investigación en ejecución orientado al diseño de una arquitectura tecnológica institucional para la detección temprana de perfiles académicos de riesgo.

Palabras clave: rendimiento académico, minería de datos educativos, prevención del abandono escolar

CONTEXTO

El proyecto se desarrolla en el ámbito de la carrera Ingeniería en Sistemas de Información de la UNCAUS, donde se han observado históricamente bajos porcentajes de regularización y elevados niveles de abandono en asignaturas troncales del primer año, particularmente “Algoritmos y Estructuras de Datos” y “Sistemas y Organizaciones”. Estas asignaturas son muy importantes estructuralmente dentro del plan de estudios, ya que condicionan la continuidad académica mediante un régimen de correlatividades que impacta directamente en la trayectoria estudiantil. La problemática se enmarca en un escenario más amplio de desgranamiento temprano en educación superior, especialmente en carreras tecnológicas, lo que exige el diseño de mecanismos institucionales que permitan analizar sistemáticamente las variables incidentes en el rendimiento académico y anticipar situaciones de riesgo antes de que se traduzcan en abandono efectivo.

1. INTRODUCCIÓN

El proyecto se fundamenta en el paradigma de Knowledge Discovery in Databases [5], en la integración de Data Warehouse orientado a soporte de decisiones [6] [7] y en técnicas de modelado predictivo [8].

La literatura internacional enfatiza la importancia de Learning Analytics institucionales para apoyar decisiones pedagógicas basadas en evidencia [3]. Este proyecto adapta dichas tendencias al contexto regional de UNCAUS.

Las técnicas de minería de datos aplicadas al ámbito educativo incluyen modelado predictivo supervisado y segmentación por clustering. Los modelos supervisados permiten inducir relaciones entre variables independientes y rendimiento académico [9]. Asimismo, investigaciones recientes destacan la importancia de modelos de retención estudiantil basados en aprendizaje automático y análisis de equidad [10] [11].

El proyecto de investigación tiene por finalidad determinar las variables que inciden en el rendimiento académico de los estudiantes de primer año, particularmente en las asignaturas "Algoritmos y Estructuras de Datos" y "Sistemas y Organizaciones". A partir de la construcción de un Data Warehouse institucional y la aplicación de técnicas de Minería de Datos Educacional, incluyendo modelos descriptivos, clustering y algoritmos predictivos supervisados con el propósito de identificar perfiles académicos asociados a distintos niveles de desempeño. El proyecto integra variables socioeconómicas, educativas y actitudinales, con el objetivo de desarrollar un sistema de detección temprana de estudiantes en riesgo de bajo rendimiento o deserción, permitiendo así la implementación de estrategias pedagógicas proactivas basadas en evidencia empírica y fortaleciendo la retención estudiantil en los primeros años de la carrera.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El tema de investigación del proyecto se compone de los siguientes ejes:

a. **Rendimiento académico en el primer año universitario:** análisis del rendimiento académico en asignaturas troncales de primer año de Ingeniería en Sistemas de Información, considerando su incidencia en la regularización, aprobación y deserción temprana.
b. **Identificación de variables incidentes:** estudio sistemático de variables socioeconómicas, educativas y actitudinales (situación laboral, nivel educativo familiar, dedicación al estudio, uso de TIC, entre otras) como factores explicativos del desempeño académico.
c. **Arquitectura de gestión de datos educativos:** diseño e implementación de un Data Warehouse institucional y análisis multidimensional (OLAP) para consolidar, estructurar y analizar información académica histórica por cohortes.
d. **Minería de datos educacional y modelado predictivo:** aplicación de técnicas de minería de datos descriptiva y supervisada (regresión, árboles de decisión, clustering) para la construcción de modelos predictivos de riesgo académico.
e. **Detección temprana y retención estudiantil:** desarrollo de un sistema de soporte a decisiones que permita identificar perfiles de estudiantes en riesgo y orientar intervenciones pedagógicas proactivas destinadas a mejorar la retención.
f. **Innovación tecnológica aplicada a la gestión académica:** integración de tecnologías de análisis de datos en el ámbito institucional como mecanismo de transformación de la gestión educativa desde un enfoque reactivo hacia uno predictivo y basado en evidencia.

Metodología

- Se adopta un enfoque cuantitativo-computacional. Las unidades de análisis son los estudiantes de las cohortes 2024, 2025, 2026 y 2027.
- Las variables independientes incluyen factores socioeconómicos, antecedentes académicos y uso de TIC. La variable dependiente corresponde al rendimiento académico y condición final del estudiante.
- Los modelos son evaluados mediante métricas de precisión, sensibilidad y capacidad de discriminación. La

validación cruzada garantiza confiabilidad y robustez estadística.

Arquitectura Tecnológica

En la Figura 1, se muestra la arquitectura que se propone, la cual se estructura en cuatro capas:

1. Captura de datos: la capa de captura integra datos académicos, sociodemográficos y actitudinales.
2. Almacenamiento en Data Warehouse: el Data Warehouse consolida información histórica por cohortes (2024 a 2027), permite un análisis multidimensional mediante cubos OLAP [12] y responde al modelo orientado a decisiones.
3. Análisis predictivo: en la capa analítica se aplican modelos supervisados y técnicas de clustering. La validación cruzada y el ajuste iterativo anual permiten mejorar la precisión predictiva y la estabilidad del modelo.
4. Retroalimentación académica: los resultados se traducen en reportes institucionales que clasifican estudiantes según nivel de riesgo, habilitando intervenciones personalizadas. La detección temprana permite optimizar recursos docentes y diseñar estrategias específicas de acompañamiento con intervención personalizada y mejora de indicadores de regularización. La evidencia empírica indica que intervenciones basadas en predicción reducen riesgo de abandono.



Figura 1: Arquitectura propuesta de cuatro capas

3. RESULTADOS ESPERADOS

Se espera que el proyecto permita identificar, modelizar y validar un conjunto de variables incidentes en el rendimiento académico de los estudiantes de primer año de la carrera Ingeniería en Sistemas de Información de la UNCAUS, particularmente en las asignaturas "Algoritmos y Estructuras de Datos" y "Sistemas y Organizaciones".

En el plano científico y técnico, se prevé:

- El diseño e implementación de un Data Warehouse institucional que consolide información académica y sociodemográfica por cohortes.
- La construcción de modelos descriptivos y predictivos basados en técnicas de Minería de Datos Educacional (clustering, regresión logística, árboles de decisión).
- La validación empírica de dichos modelos mediante métricas estadísticas de desempeño, garantizando su confiabilidad y estabilidad.
- La determinación de perfiles académicos diferenciados asociados a distintos niveles de rendimiento y riesgo de deserción.

En el plano institucional y pedagógico, se espera:

- Desarrollar un sistema operativo de detección temprana de estudiantes en riesgo académico.
- Proporcionar a las cátedras involucradas información estructurada para orientar intervenciones pedagógicas específicas.
- Contribuir a la mejora de los indicadores de regularización y aprobación en las asignaturas troncales.
- Reducir el desgranamiento temprano y fortalecer la retención estudiantil en los primeros años de la carrera.

4. RESULTADOS OBTENIDOS

Hasta el momento se ha definido el aplicativo para la recolección de información, el data warehouse para almacenar la información del proyecto, los hipercubos OLAP para el análisis multidimensional y los modelos de minería de datos descriptivos y predictivos que se utilizarán.

Se ha realizado la carga de los datos correspondiente al año 2024 y 2025 que serán procesados y analizados.

Conclusion

La arquitectura implementada en UNCAUS busca transformar la gestión académica reactiva en un modelo proactivo basado en evidencia cuantitativa, alineado con tendencias internacionales en Learning Analytics.

La innovación del proyecto radica en la implementación de una arquitectura tecnológica institucional que integra Data Warehouse, análisis multidimensional (OLAP) y técnicas de Minería de Datos Educacional para la detección temprana de perfiles académicos asociados al riesgo de bajo rendimiento. A diferencia de intervenciones pedagógicas tradicionales basadas en análisis retrospectivos o criterios subjetivos.

5. FORMACIÓN DE RECURSOS HUMANOS

- El proyecto consolida la línea iniciada en [13] [14] [15] [16] [17] sobre perfiles académicos y minería de datos educacional, fortaleciendo la capacidad técnica institucional.
- Fortalece al Grupo de investigación, promoviendo formación en Ciencia de Datos Educacional, integración docente-investigador y producción científica nacional e internacional.

6. BIBLIOGRAFÍA

[1] Grimaldo, M., & Manzanares-Medina, E. (2023). Variables intervinientes en el rendimiento académico en ingresantes universitarios. *Revista Electrónica Educare*, 27(1), 1–14. https://doi.org/10.15359/ree.27-1.14860

[2] García Jiménez, M., Alvarado Izquierdo, J., & Jiménez Blanco, A. (2000). La predicción del rendimiento académico: Regresión lineal versus regresión logística. *Psicothema*, 12(2), 248–252.

[3] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1355. https://doi.org/10.1002/widm.1355

[4] Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student retention using educational data mining and predictive analytics: A systematic literature review. *IEEE Access*, 10, 72480–72503. https://doi.org/10.1109/ACCESS.2022.3184567

[5] Fayyad, U., Grinstein, G., & Wierse, A. (2001). Information visualization in data mining and knowledge discovery. Morgan Kaufmann.

[6] Inmon, W. H. (1996). Building the data warehouse. John Wiley & Sons.

[7] Simon, A. (1997). Data warehouse, data mining and OLAP. John Wiley & Sons.

[8] Han, J., & Kamber, M. (2001). Data mining: Concepts and techniques. Morgan Kaufmann.

[9] Wilson, R. L., & Hardgrave, B. C. (1995). Predicting graduate student success in an MBA program: Regression versus classification. *Educational and*

Psychological Measurement, 55(2), 186–195. https://doi.org/10.1177/0013164495055002003

[10] Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2023). Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice*, 25(1), 51–75. https://doi.org/10.1177/15210251211024740

[11] Quy, T. L., Friege, G., & Ntoutsi, E. (2023). A review of clustering models in educational data science towards fairness-aware learning. En A. Peña-Ayala (Ed.), *Educational data science: Essentials, approaches, and tendencies*. Springer. https://doi.org/10.1007/978-3-031-12345-6

[12] Ballard, C., Rollins, J., Ramos, J., Perkins, A., Hale, R., Dorneich, A., Cas Milner, E., & Chodagam, J. (2007). *Dynamic warehousing: Data mining made easy*. IBM Press.

[13] La Red Martínez, D. L., Karanik, M., Giovannini, M., & Pinto, N. (2015a). Academic performance profiles: A descriptive model based on data mining. *European Scientific Journal*, 11(9), 17–38.

[14] La Red Martínez, D. L., Karanik, M., Giovannini, M., & Pinto, N. (2015b). Perfiles de rendimiento académico: Un modelo basado en minería de datos. *Revista Científica Iberoamericana de Tecnología Educativa*, 4(1), 12–30.

[15] La Red Martínez, D. L., Karanik, M., Giovannini, M., & Scappini, R. (2016). Towards a predictive model of academic performance using data mining in the UTN-FRRe. *Journal of Systemics, Cybernetics and Informatics*, 14(2), 36–41.

[16] La Red Martínez, D. L., Gerzel, S. M., Aguirre, C. E., Cáceres, R. A., & Berecoechea Galarza, J. A. (2020a). Uso de minería de datos para la determinación de perfiles socioeconómicos y sanitarios en la UNCAus. *European Scientific Journal*, 16(33), 88–104. https://doi.org/10.19044/esj.2020.v16n33p88

[17] La Red Martínez, D. L., Gerzel, S. M., Aguirre, E., Cáceres, R. A., Berecoechea Galarza, J. A., & Vallejo, S. (2022). Determinación de perfiles socioeconómicos y sanitarios en actividades curriculares universitarias. *En XXIV Workshop de Investigadores en Ciencias de la Computación* (WICC 2022) (pp. 133–137).

SmartLab: Innovación con tecnología para Espacios Educativos

Sarobe Mónica[1], Cicerchia Lucas Benjamín[1], Ramón Hugo[2], Giordano Walter[1], Pablo Luengo[1], Jaszczyszyn, Adrián[1], Cascardo Ernesto[1], Alvarez, Eduardo[1]. Commisso, Sergio[1], Useglio, Gustavo[1], Skeppstedt, Carola[1] , Spreafichi, Ignacio[1], Puppio, Nehuén[1], Gambarte, Camila[1], Laiun, Francisco[1], Di Carlo, José[1], Moyano, María Emilia[1]

Instituto de Investigación y Transferencia en Tecnología (ITT)
Centro asociado a la Comisión de Investigaciones Científicas (CIC)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

{monica.sarobe, lucas.cicerchia, hugo.ramon, pablo.luengo, adrian.jaszczyszyn, eduardo.alvarez, sergio.commisso}@itt.unnoba.edu.ar

{wgiordano, ghuseglio, ispreafichi, napuppio, cgambarte, flaiun, mdicarlo, memoyano411}@comunidad.unnoba.edu.ar

{cskeppstedt, acascardo}@unnoba.edu.ar

Resumen

Este proyecto de investigación interdisciplinario, tiene por objetivo transformar los espacios educativos tradicionales en entornos de aprendizaje inteligentes y colaborativos mediante la integración de tecnologías innovadoras y su impacto en los resultados educativos. El mismo se enfoca en el diseño y equipamiento de aulas y laboratorios que permitan a estudiantes y docentes interactuar de manera dinámica con los contenidos educativos y en el análisis cuantitativo y cualitativo de su impacto en los participantes. De esta manera, docentes y estudiantes pueden tener acceso a plataformas y laboratorios para experimentar con nuevas tecnologías, prototipos y metodologías educativas. Buscando crear soluciones tecnológicas personalizadas combinando infraestructura con tecnologías emergentes para mejorar procesos y crear espacios de innovación que permitan beneficiar tanto a la comunidad educativa como a otras áreas de la sociedad. Se centra en el diseño, desarrollo, implementación y evaluación de impacto de laboratorios remotos, configurados para su integración en entornos educativos inteligentes.

En este sentido, se trabaja en la remotización de Equipos de manera de transformar dispositivos y sistemas físicos en recursos accesibles a distancia al implementar soluciones que permitan el acceso remoto a laboratorios con capacidad para monitorear, operar y realizar simulaciones en tiempo real.

Palabras claves: Tecnologías en educación; Educación; Laboratorios remotos; Automatización, Robotización.

Contexto

Las líneas de investigación a describir se enmarcan en el proyecto de investigación: ***SmartLab: Innovación con tecnología para Espacios Educativos***, con lugar de trabajo en el Instituto de Investigación y Transferencia en Tecnología (ITT), presentado en la convocatoria de Subsidios a la Investigación Bianuales (SIB) 2025, aprobado y financiado por la Secretaría de Investigación de la Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA). Su objetivo es Desarrollar y poner en marcha laboratorios remotos que transformen los espacios educativos tradicionales en entornos inteligentes, colaborativos e innovadores mediante la integración de tecnologías emergentes, evaluando su impacto en la calidad educativa, en el desarrollo de competencias tecnológicas en los estudiantes y las posibilidades de aplicación en diferentes contextos.

1. Introducción

Desde hace más de una década, la UNNOBA trabaja en la consecución de mejoras en la calidad de los métodos y las técnicas de enseñanza-aprendizaje.

Este proyecto busca abordar el diseño y la implementación de espacios educativos inteligentes, con el objetivo de transformar el modelo de enseñanza-aprendizaje tradicional en uno más flexible y adaptado a las necesidades del

[1] Docente Investigador ITT-UNNOBA. [2] Docente Investigador ITT. Inv. Asociado Adjunto sin director CIC.

entorno digital y colaborativo actual. La falta de acceso a laboratorios de investigación remotos y la escasa interacción entre las herramientas tecnológicas y los procesos educativos son problemas adicionales que se intentarán resolver mediante este enfoque innovador.

Buscando ser un aporte a la necesidad de adaptar los espacios educativos a las demandas actuales del mundo digital. Integrar tecnologías avanzadas en los procesos educativos no solo mejora la calidad de la enseñanza, sino que también permite una mayor personalización del aprendizaje, fomenta la creatividad de los estudiantes y promueve una mejor preparación para el futuro laboral. La implementación de laboratorios remotos ofrece una oportunidad para que los estudiantes interactúen con equipos reales, lo que enriquecerá su formación práctica.

La Escuela de Tecnología cuenta con un Laboratorio de Desarrollo de Ingeniería Mecánica e Industrial (LaDIMI) en el que se desarrollan actividades académicas de vinculación con las asignaturas como actividad principal, consolidándose como un puente entre la teoría y la práctica para docentes y estudiantes fomentando el aprendizaje aplicado y significativo. En este entorno se desarrollan también actividades que impulsan la investigación aplicada de tecnologías emergentes aplicadas en contextos educativos, fortaleciendo también la extensión y vinculación con la comunidad.

La colaboración entre el Instituto de Investigación y Transferencia de Tecnología (ITT) y el LaDIMI, donde se implantarán algunos de los laboratorios remotos desarrollados, facilitará la transferencia de conocimientos y el desarrollo de soluciones tecnológicas que se pueden aplicar tanto en la educación como en otros sectores de la sociedad. Este enfoque interdisciplinario contribuye a formar un entorno colaborativo, inclusivo, accesible y tecnológicamente avanzado. Donde, además, se espera trabajar de manera colaborativa con otros dos proyectos presentados por ITT como son: “Innovación y Desafíos en la Educación Digital” y “Técnicas y algoritmos avanzados para la optimización de procesos: seguridad, sostenibilidad e innovación”.

2. Líneas de investigación, alcances, impacto y antecedentes

Habiendo planteado la hipótesis de que la implementación de espacios educativos, equipados con tecnologías de laboratorios remotos, mejora significativamente el aprendizaje de los estudiantes, promoviendo una mayor interacción, creatividad y desarrollo de competencias tecnológicas.

El equipo de trabajo se centra en:

- Analizar los requisitos, diseñar y equipar aulas y laboratorios inteligentes que faciliten el aprendizaje interactivo y el uso de tecnologías emergentes.
- Implementar plataformas educativas que permitan el acceso remoto a equipos y sistemas reales para realizar experimentos y actividades de aprendizaje colaborativo.
- Evaluar, mediante métodos cuantitativos y cualitativos, el impacto de los nuevos espacios educativos.
- Desarrollar soluciones tecnológicas personalizadas que optimicen procesos y que puedan ser aplicadas en diferentes contextos.
- Fomentar la colaboración entre los usuarios y expertos tecnológicos para generar propuestas innovadoras y adaptativas a los cambios tecnológicos.
- Desarrollar materiales educativos en función de los laboratorios remotizados.
- Propiciar la colaboración con otros grupos de I+D+i que hayan presentado proyectos a la convocatoria SIB en la temática de este proyecto.

En este sentido, se trabaja en la **Remotización de Equipos** de manera de transformar dispositivos y sistemas físicos en recursos accesibles a distancia al **implementar soluciones que permitan el acceso remoto** a laboratorios con capacidad para monitorear, operar y realizar simulaciones en tiempo real.

Para ello se trabaja en equipar laboratorios con tecnologías emergentes como placas SBC (Single Board Computers), cámaras IP, software de comunicación y gestión remota, que permitan una experiencia inmersiva y colaborativa; diseñar un entorno integrado para la interacción remota que sea escalable y adaptable a diferentes disciplinas y modalidades educativas.

Durante el desarrollo del presente se encuentran en etapa de definición **Indicadores de Impacto** de manera de determinar métricas para evaluar el impacto educativo y tecnológico de los laboratorios, estas podrían incluir:

1. **Indicadores de acceso**: número de estudiantes y docentes que utilizan los laboratorios remotizados.
2. **Indicadores de aprendizaje**: mejora en competencias técnicas y habilidades prácticas medida mediante evaluaciones pre y post uso.
3. **Indicadores de eficiencia**: reducción en tiempos de uso y aumento en la disponibilidad de equipos para prácticas simultáneas.

Impacto

El impacto tecnológico y educativo de este proyecto se manifiesta en múltiples dimensiones:

1. **Tecnológico**:

- Consolidación de un sistema de laboratorios remotos que integren tecnologías como IoT (Internet of Things), automatización industrial y plataformas en la nube.

- Promoción de avances en la remotización de equipos, creando entornos interactivos accesibles desde cualquier lugar.

2. **Investigativo y Metodológico**:

- Generación de nuevos enfoques para la medición de la efectividad de los entornos educativos inteligentes, con análisis de datos cuantitativos y cualitativos basados en los indicadores definidos.

El concepto de **laboratorios remotos** se refiere a entornos de aprendizaje donde los estudiantes interactúan con equipos físicos o simulados a través de interfaces digitales, eliminando la necesidad de presencia física. Estos laboratorios han evolucionado en las últimas dos décadas debido a avances en conectividad, automatización, y tecnologías emergentes como el Internet de las Cosas (IoT) y la inteligencia artificial (IA).

Entre los casos de éxito que podemos nombrar se encuentran **MIT iLab Project:** Proyecto que desarrolló laboratorios remotos para diversas disciplinas, demostrando su impacto positivo en la educación de ingeniería a nivel global. **Labshare (Australia):** Iniciativa que permitió compartir laboratorios entre universidades, maximizando el uso de recursos y democratizando el acceso a equipos costosos.

La educación universitaria ha comenzado a explorar el acceso remoto a prácticas de formación experimental en laboratorios mucho antes de la pandemia e inspirados por experiencias internacionales como el proyecto llamado Virtual Instruments Systems In Reality (VISIR) (Visir, s.f.), que es un proyecto con socios de Argentina, Austria, Brasil, Portugal, España y Suecia, financiado por el programa Erasmus+ de la Unión Europea (CONFEDI, 2016) (Castro et al, 2019).

Para el año 2021, basándose de la experiencia VISIR, las Unidades Académicas con Carreras de Ingeniería socias plenarias del Consejo Federal de Decanos de Ingeniería de la República Argentina (CONFEDI), plantearon la conformación de una red de Universidades Argentinas que permita el acceso remoto a la utilización de experiencias de laboratorios de cada institución en beneficio de otras instituciones miembros. Constituyeron la Red Argentina Colaborativa de Laboratorios de Acceso Remoto cuya sigla es CONFEDI R-LAB, y además, acordaron un programa junto con la Secretaría de Políticas Universitarias (SPU) del Ministerio de Educación para promover esta red (CONFEDI R-LAB,2021).

La Universidad del Noroeste de Buenos Aires (UNNOBA) en 2022, conformó un equipo multidisciplinario de alumnos, docentes de la Escuela de Tecnología e investigadores del Instituto de Investigación y Transferencia (ITT) y del Laboratorio de Desarrollo de Ingeniería Mecánica e Industrial (LaDIMI) para ser parte miembro de la red. El equipo, remotizó un laboratorio donde el alumnado puede acceder a través de una página web, a la experiencia de programar una secuencia lógica en PLC, para representar un proceso productivo, simularlo y reproducir los movimiento de una máquina industrial a través de actuadores neumáticos. Donde dichos avances se presentaron los avances en el II Congreso Multidisciplinario de la UNNOBA en la sesión de posters de ingeniería, recibiendo el primer premio de ingeniería (Multidisciplinario, 2022)(LA OPINIONLINE, 2022).

En cuanto a la evaluación de su impacto en lo referente a resultados académicos se han realizado en diferentes instituciones, estudios comparando laboratorios remotos, virtuales y tradicionales donde se analizaron su impacto en el aprendizaje práctico y teórico en la mejora de la retención de conceptos complejos y habilidades técnicas.

Este trabajo toma como antecedentes los proyectos de tecnología aplicada a la educación presentados en convocatorias SIB anteriores enmarcados al igual que el presente en el ITT. Por otro lado, el LaDIMI, dependiente de la Escuela de Tecnología de la UNNOBA será el espacio donde se desarrolle el equipamiento necesario y donde se realicen las prácticas de laboratorio.

3. Resultados obtenidos/esperados

A continuación, se mencionan los proyectos de laboratorios remotos que se encuentran en marcha y que cuentan con financiamiento externo:

1- Programación de secuencia lógica para la simulación del movimiento de actuadores neumáticos de máquinas industriales:

El LaDIMI, contaba con un equipamiento didáctico (totalmente desarrollado y construido por personal propio) que utilizaba el escritorio remoto de una computadora "esclava" donde se habían realizado pruebas piloto. Estaba operativo para las prácticas realizadas por los estudiantes la simulación de un proceso mediante el movimiento de actuadores. En el marco del presente, se logró que el equipo funcione en forma remota bajo el sistema interno y se encuentra disponible para conectarse a la Red R-Lab cuando esta esté disponible.

2- Determinación de la curva característica de la bomba, curva característica de la instalación, y punto óptimo de funcionamiento:

Se cuenta con un equipamiento didáctico totalmente desarrollado y construido por personal propio llamado Parral de Cañerías, donde los estudiantes realizan prácticas de laboratorio en las asignaturas Máquinas Hidráulicas y Neumáticas. Tecnología de los Servicios Industriales. Mecánica de los fluidos y Mecatrónica. Además se están realizando mejoramientos que requiere de cambios de componentes, sensores, placas, computadoras y cámaras.

3- Banco de prueba de inyectores:

En términos técnicos, un banco de inyectores es un equipo que cuenta con la capacidad para proporcionar alimentación eléctrica a cuatro inyectores de manera individual o grupal, los cuales pueden suministrar una cantidad de fluido precisa en diversas condiciones de funcionamiento. Con el uso de este banco de prueba es posible detectar el funcionamiento de cada inyector como así también realizar el lavado de los mismos cuando esto sea requerido. Se proyectan ensayos consistentes en observar el funcionamiento del banco de inyectores y sus componentes principales; rampa de inyectores, bomba de combustible y controlador lógico programable. Esta tecnología se presentó y su financiamiento fue aprobado en el mes de noviembre de 2025 por Fundación YPF por un total de $ 3.032.508. Al momento de la presentación su estado es preliminar dado que se están comprando las componentes necesarias.

4- Viscosímetro de Stokes:

El viscosímetro de Stokes permite determinar la viscosidad dinámica de fluidos mediante la caída libre de una esfera, aplicando la ley de Stokes. Cuando la esfera se mueve a velocidad constante en un fluido viscoso, se equilibran la fuerza de arrastre viscosa, la fuerza gravitatoria y la fuerza de empuje hidrostático.

Este equipamiento fue aprobado para ser financiado por la fundación YPF quien se encuentra enviando el equipamiento necesario para la remotización por un monto de $2.557.541.

5- Intercambiador de Calor y Torre de enfriamiento:

Este desarrollo está en la misma situación que el anterior, se obtuvo el financiamiento por Fundación YPF a fines de 2025 y está siendo desarrollada su remotización. El conjunto experimental consta de dos módulos: un intercambiador de calor de tubos concéntricos y una torre de enfriamiento.

1. Intercambiador de calor de tubos concéntricos: Montado sobre una estructura metálica que aloja el sistema eléctrico y el circuito hidráulico de agua caliente. Este circuito cerrado comienza en una bacha, desde donde una bomba periférica de 1/4 HP impulsa el fluido hacia un calefón trifásico de 12 kW y luego al intercambiador, circulando por el tubo interior; el agua de enfriamiento pasa por el espacio anular, en régimen paralelo a contracorriente. Tras el intercambio térmico, el fluido retorna a la bacha. El sistema eléctrico se divide en dos subcircuitos: trifásico (calefón) y monofásico (bomba). Cuatro sensores analógicos permiten medir temperaturas de entrada y salida en ambos flujos, facilitando el monitoreo térmico.

2. Torre de enfriamiento En fase de optimización, presenta un diseño de flujo paralelo a contracorriente y mide aproximadamente 800 × 800 × 2000 mm. Incluye un ventilador axial (IP66) en la parte superior. Se compone de:

- Estructura: chasis, bastidor y carcasa.
- Eléctrico: tablero, cableado y controles.
- Hidráulico: bacha, bomba, distribuidor y relleno.
- Succión: rejillas, separador de gotas y ventilador.

El circuito hidráulico es cerrado, con reposición por evaporación. El agua se impulsa al distribuidor, desciende por el relleno, intercambia calor con el aire ascendente y vuelve a la bacha.

3. Integración: Ambos módulos conforman un sistema didáctico-industrial para simular procesos de enfriamiento y transferencia de calor. El agua de procesos transfiere calor al agua de enfriamiento, que luego se refrigera en la torre, cerrando el ciclo.

Además de los resultados obtenidos a nivel de financiamiento y desarrollo de tecnología se realizaron las siguientes presentaciones y publicaciones:

1- Presentación de dos prácticas de laboratorio al **Concurso Docente de Guías de Trabajos Prácticos RLAB 2025,** convocatoria realizada por CONFEDI. Las mismas fueron aprobadas para publicación de la RADI (Revista Argentina de Ingeniería).

2- "Transformación digital en educación: aplicaciones base para R-LAB". Publicación presentada y aprobada para publicación en el 1° Congreso Nacional de Innovación Universitaria: Enseñanza, Investigación, Gestión y Territorio. Consejo Interuniversitario Nacional (CIN) / Universidad Nacional de Rosario.

3- "Control remoto del AutoClean en máquinas I3D". Jornada de Jóvenes Investigadores UNNOBA.

4- "Trabajo práctico en laboratorio de acceso remoto". VI Workshop de Innovación y Transformación Educativa UNNOBA.

5- "SmartLab UNNOBA: Modernización de Prácticas de Laboratorio mediante la Integración de PLC y Herramientas IoT." V Congreso virtual de microcontroladores y sus aplicaciones. UTN.

4. Bibliografía y referencias

1. Villalobos, M., & Romero, R. (2023). Implementación y desafíos de laboratorios remotos en la educación superior en Argentina. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (RIITET)*, *12*(23), 45–60. https://sedici.unlp.edu.ar

2. Facultad de Ingeniería de la Universidad de Buenos Aires. (2023). *Laboratorios remotos: 2da etapa. Programa del Consejo Federal de Decanos de Ingeniería (CONFEDI).* https://ingenieria.uba.ar/noticias/laboratorios-remotos-2da-etapa

3. Capuya, F. G., Montero-Miranda, E., Arguedas-Matarrita, C., & Idoyaga, I. (2023). Laboratorios remotos: Un recurso para el aprendizaje de la temática de gases en cursos universitarios masivos en Argentina durante la pandemia de la COVID-19. *Revista Innovaciones Educativas*, *25*(38), 246–262.

4. Castro, M., Nilsson, K., Pozzo, M. I., García-Loro, F., & Fernández, R. M. (2019). Workshop: Teaching practices with VISIR remote lab: Technical, educational and research fundamentals from the PILAR Project. En *2019 IEEE World Conference on Engineering Education (EDUNINE)* (pp. 1–3). IEEE.

5. CONFEDI. (2016). *Convocatoria participación ERASMUS VISIR CONFEDI 2016.* https://confedi.org.ar/wp-content/uploads/2016/04/Convocatoria-Participaci%C3%B3n-ERASMUS-VISIR-CONFEDI-2016.pdf

6. CONFEDI R-LAB. (2021). *Acta fundacional R-LAB.* https://confedi.org.ar/wp-content/uploads/2021/04/Acta-Fundacional-R-LAB.pdf

7. La Opinionline. (2022). Entregaron los premios del Congreso Multidisciplinario de la UNNOBA. https://laopinionpergamino.com.ar/nota/9827/2022/09/entregaron-los-premios-del-congreso-multidisciplinario-de-la-unnoba

8. Congreso Multidisciplinario UNNOBA. (2022). *Jueves 22.* https://congresomultidisciplinario.unnoba.edu.ar/jueves-22/

9. VISIR. (s.f.). *VISIR remote laboratory.* https://www2.isep.ipp.pt/visir/index.php?page=inicio

10. Russo, C., Sarobe, M., & Ahmad, T. (2021). Definición de indicadores: Calidad en cursos virtuales. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET)*, (28).

Tecno pedagogía como Motor de la Transformación Educativa

Mg. Roberto Bertone[1], Mg. José Luis Filippi[2], Lic. Guillermo Lafuente[3], Mg. Carlos Ballesteros[4], Lic. Gustavo Lafuente[5], Pablo Etcheverry[6], Valentín Filippi[7]
LIAU[9] - Facultad de Ingeniería – UNLPam.
pbertone@ada.info.unlp.edu.ar[1]
{filippij[2], lafuente[3], balleste[4], gustavo[5], etcheverry[6]}@ing.unlpam.edu.ar
valefilippi08@gmail.com[7]
[9]Laboratorio de Investigación de Ambientes Ubicuos

RESUMEN

La tecno pedagogía es la tecnología aplicada a la educación. A través de ella se busca dar respuesta pedagógica al impacto de las herramientas digitales en el proceso formativo.

La tecno pedagogía utiliza estrategias didácticas como el aprendizaje significativo y colaborativo, conjuntamente incorpora recursos tecnológicos permitiendo el desarrollo de actividades de enseñanza virtuales que permitirán a los estudiantes accionar de forma autónoma.

Los centros educativos del siglo XXI se han de enfrentar a las nuevas demandas de una sociedad cambiante y con una clara vocación tecnológica. Aparecen nuevas formas de abordar lo tecnológico: aprendizaje adaptativo y móvil, inteligencia artificial, realidad virtual y aumentada y surgen las interfaces de uso natural (Johnson et al., 2016; Becker et al., 2017). [1]

El propósito del proyecto, de naturaleza teórico-práctico, consiste en indagar las posibilidades que ofrecen los dispositivos tecnológicos de última generación, como instrumentos aplicados a diferentes situaciones de aprendizaje, considerando los aspectos pedagógicos y las teorías del aprendizaje como marcos flexibles de la organización pedagógica. Las actividades que se proponen tienen lugar en el contexto de la UNLPam.

Palabras claves: Tecno pedagogía. Aprendizaje Móvil. Aprendizaje Colaborativo.

CONTEXTO

Tipo de Investigación: Aplicada
Campo de Aplicación Principal: 7 1802 Computación, 7 1803 Comunicaciones.
Campos de Aplicación posibles: 13 1040 Ciencia y Tecnología, 7 4399 Otras – Educación – Tecnología Aplicada a la Educación
Institución que Coordina el Proyecto: Facultad de Ingeniería, Universidad Nacional de La Pampa.

1. INTRODUCCION

Con la aparición del coronavirus a comienzos del 2020, los cambios tecno pedagógicos que se realizaban de manera natural en los centros educativos, se precipitaron como resultado de la situación de encierro ciudadano.

A partir del 2022 con la vuelta a la formación presencial, se visualiza un modelo educativo donde lo presencial y lo virtual van a convivir de forma permanente, y en este punto el accionar docente es vital para desplegar su actividad profesional según los modelos que la sociedad moderna exige.

Aparecen nuevas estrategias formativas como el microlearning **[2]** y el aprendizaje inversivo. Surge el desafío educativo, definir metodologías y procesos para desplegar una formación pedagógica eficaz. Por ello se requiere un diseño tecno pedagógico instruccional adaptado al contexto en el que

se va a desplegar la tarea educativa, con docentes capacitados en desempeñar su profesionalidad de forma hibrida. **[3].**

A partir del estrepitoso cambio de paradigma educativo que se articuló en marzo del 2020, donde se determina el paso de la formación presencial a la virtual, numerosos actores se pronuncian en priorizar el desarrollo de habilidades integrales además de las cognitivas. Por ello es necesario "…construir otras formas de hacer educación, y que los cambios curriculares den cuenta de la posibilidad, la necesidad y la urgencia de vincular los contenidos disciplinares con habilidades de aprendizaje socioemocionales en las aulas…". **[4]**

Aquí radica el objetivo primario del proyecto: indagar las posibilidades que ofrecen los dispositivos tecnológicos de última generación, como instrumentos aplicados a diferentes situaciones de aprendizaje, considerando los aspectos pedagógicos y las teorías del aprendizaje como marcos flexibles de la organización pedagógica.

Situación Actual del Problema.

La educación influenciada por ésta realidad, acepta la necesidad de reconfigurar el proceso educativo, partiendo de algunas premisas:

a. Redefinir nuevas acciones para docentes y estudiantes.

El creciente número de dispositivos móviles y la conectividad a internet de forma permanente, permite a los estudiantes adquirir nuevos contenidos en línea y configurar sus conocimientos a partir de intereses propios. El estudiante es artífice de su futuro, asume un papel activo, decide que aprender y cómo.

El rol tradicional del docente es superado por nuevas exigencias, requiere nuevas capacidades y conocimientos para desempeñar su actividad en entornos virtuales de aprendizaje.

Frente a la realidad del mundo digital que rodea a los estudiantes de hoy en día, es justo que los docentes replanteen su rol en la construcción del conocimiento sobre ellos. De allí que la competencia exigida a un profesor del Siglo XXI es precisamente que se enfoquen "en las necesidades de los alumnos, supervisando su búsqueda de información e intentando facilitar la búsqueda de información individual de los alumnos ya que el papel de suministrador de conocimiento ha sido superado por las TIC" **[5]**.

Mucho se habla de cuáles serían las competencias digitales para un docente, un acercamiento a ello lo propone Núñez-Torrón Stock **[6]**, quien plantea cinco competencias básicas para el docente en la educación, y entre ellas se pueden destacar dos:

1. Usar herramientas de trabajo en línea.
2. Utilizar dispositivos móviles en el aula.

En la mayoría de las universidades de todo el mundo se observa la incorporación de los sistemas de Gestión de Aprendizaje (LMS), para implementar entornos virtuales que permiten la gestión del aprendizaje. **[7]** Sin embargo algunas investigaciones indican que las motivaciones de los estudiantes son vistas como una obligación al participar en entornos de aprendizaje institucionales cerrados, mientras que la motivación es mayor cuando el proceso formativo es abierto a diferentes tecnologías en contextos informales. Estos espacios son conocidos como PLE (Personal Learning Environments) donde los estudiantes configuran su propio entorno. **[8]**

Algunos autores ya incorporan un nuevo concepto, el de mPLE (Mobile Personal Learning Environments) como una nueva estructura de aprendizaje para la generación que hace uso de los dispositivos móviles. **[9]**

La agregación del modelo tecno pedagógico, que integra las tecnologías en todas las áreas del conocimiento, va a permitir desplegar las competencias digitales de los docentes en el aula, influenciando de manera positiva el desempeño de los estudiantes.

La tecno pedagogía utilizada como herramienta de planificación pedagógica

expone diferentes actividades formativas que conectan lo que se enseña en la práctica con lo que aprende el estudiante.

Es un modelo que viabiliza la construcción del conocimiento de forma colaborativa, donde los docentes efectúan su aporte según su área del conocimiento específico, tecnología, pedagogía y/o contenidos. Así, los docentes aprenden unos de otros, lo que puede derivar en el desarrollo colaborativo del currículo. **[10]**

La organización de la tecno pedagogía en el aula, integra las tecnologías con la pedagogía y el conocimiento científico. Para que la integración sea efectiva, los docentes no solo deben poseer conocimientos en tecnología y pedagogía, se debe poseer además el conocimiento propio de la disciplina, y desde allí integrar el conocimiento tecnológico y el conocimiento pedagógico. **[11]**.

b. Facilitar la continuidad del aprendizaje.

El aprendizaje móvil permitirá a los estudiantes dar continuidad a su formación a lo largo de toda su vida.

Los estudiantes universitarios no solo deben dominar las disciplinas correspondientes a su carrera, además deben incorporar habilidades transversales como el pensamiento crítico, resolución de problemas, persistencia y trabajo colaborativo. **[12]**

En la actualidad se está transitando tiempos de pos pandemia, y si algo bueno se puede rescatar de esta situación, es la consolidación del uso de los dispositivos tecnológicos de última generación durante los últimos dos años, escenario que alcanzo a todos los sectores sociales incluidas las instituciones educativas en sus diferentes niveles. Entre los instrumentos tecnológicos que sobresalieron en las actividades formativas virtuales que se desplegaron durante la pandemia y que van a prevalecer pos pandemia, podemos indicar **[13]**, aulas colaborativas, educación personalizada e integrar en las actividades pedagógicas las tecnologías emergentes como nuevos ambientes de enseñanza y aprendizaje. Se trata no solo de afianzar los conocimientos adquiridos de los estudiantes, sino el de desarrollar en ellos las habilidades cognitivas y meta cognitivas para la realización de tareas en su quehacer diario, objetivo primario del diseño tecno pedagógico. **[14]**

c. Grandes volúmenes de datos. Big data.

Los investigadores que estudian el aprendizaje en línea, los sistemas de tutoría inteligente, los laboratorios virtuales, las simulaciones y los sistemas de gestión del aprendizaje están explorando maneras de entender y utilizar mejor la analítica del aprendizaje, a fin de mejorar la actividad del docente y seguir avanzando en la educación para todos. **[15]**

En éste contexto el punto de partida debe ser siempre tecno pedagógico. No se trata de innovar por innovar, hay que explotar el potencial que ofrecen las nuevas herramientas tecnológicas que hacen posible el aprendizaje móvil, a partir de una planificación docente previa dentro de un marco teórico formalizado.

El BigData trae beneficios de cara a la personalización de la educación en el alumno/a de modo que estudiantes como profesores sepan en qué punto del aprendizaje se encuentran, cuáles son sus debilidades y fortalezas. **[16]**

2. LINEA DE INVESTIGACION y DESARROLLO

El plan de actividades corresponde al proceso de investigación aplicada, con objetivos de corto, mediano y largo plazo, y una duración prevista de cuatro años.

Primer año.

- Identificar el porcentaje de inserción de dispositivos móviles en la institución (teléfonos y/o tablets).

- Analizar la inserción del aprendizaje móvil en las universidades de todo el mundo.
- Estudiar como el aprendizaje móvil puede cerrar la brecha entre el aprendizaje formal y el informal.
- Identificar los escenarios educativos que muestren aspectos móviles en las actividades formativas del ámbito académico propio.
- Examinar herramientas tecnológicas móviles orientadas al proceso educativo de acceso libre y gratuito.

Segundo y Tercer año.

- Definir el enfoque tecno pedagógico que posibilite un aprendizaje móvil de calidad.
- Estudiar las tecnologías disponibles para implementar acciones de aprendizaje móvil en el contexto educativo universitario.
- Definir prioridades según necesidades de inmediatez en la virtualización de las disciplinas que se han de impartir.
- Gestionar el uso de plataformas de formación virtual que posibilite una doble modalidad, presencial y mediada por las nuevas tecnologías móviles.
- Incorporar el uso de redes sociales que favorezcan la práctica educativa.
- Confeccionar objetos de aprendizaje acorde a las herramientas disponibles en la nube y a las características de los dispositivos móviles que van a operar el producto final.
- Diseñar y/o desarrollar aplicaciones móviles a partir de las necesidades que se presenten durante el transcurso del proceso educativo virtual.
- Capacitar a la comunidad educativa en general en el desarrollo de objetos de aprendizaje a partir de los requerimientos de carácter institucional.
- Difundir los avances a toda la comunidad universitaria los progresos a medida que van trascurriendo, a través de jornadas, congresos y/o revistas científicas.
- Propiciar el intercambio de información permanente con grupos de investigación que den valor agregado a nuestra actividad profesional.

Cuarto año.

- Instituir las aplicaciones desarrolladas para dispositivos móviles que hayan alcanzado buen nivel de aceptación.
- Conformar un repositorio de objetos de aprendizajes de acceso libre.
- Registrar y difundir los resultados alcanzados con la finalidad de que se puedan utilizar, ampliar y mejorar a través de trabajos futuros.
- Disponer de un catálogo de trabajos realizados en el sitio web del grupo de investigación GIAU (Grupo de Investigación de Ambientes Ubicuos) a partir del cual los interesados podrán acceder al material requerido.
- Presentar en jornadas, congresos y/o revistas de todo el mundo los resultados alcanzados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Durante el 2025 se llevaron a cabo varios trabajos destacados, que se presentaron en diferentes reuniones científicas; a modo de ejemplo se pueden citar: **REVISTA CABB 2025,** La revolución de los datos en el baloncesto: Desarrollo de una plataforma de análisis de datos para entrenadores. **RUEDA 2025**, "Integrando la Tecnología en la Educación Física: El Caso de CoachAssist y el Baloncesto escolar". **WICC 2025**, Aprendizaje Personalizado y Colaborativo. El Potencial de la Tecno pedagogía en la Formación Universitaria. **WITE 2025**, Análisis del rendimiento y la toma de decisiones en el baloncesto escolar: El caso de una herramienta multiplataforma (móvil y web).

4. FORMACION DE RECURSOS HUMANOS

Director de Proyecto

Co-Director de Proyecto

6 Investigadores

5. BIBLIOGRAFIA

[1] Integración de Tecnologías de la Información y la Comunicación a la Docencia - 2017 https://nuestroscursos.net/mod/url/view.php?id=1857

[2] Que es un Nanodegree. Webinar. 9 de febrero de 2016 https://www.itmadrid.com/que-es-un-nanodegree/

[3] Tendencias eLearning 2022. https://www.itmadrid.com/wp-content/uploads/2021/12/itmadrid-tendencias-elearning-2022.pdf

[4] [UNESCO (2021), Educación pos pandemia: salud mental y prevención de la violencia con la comunidad educativa, Memoria, UNESCO, en: UNESCO (2021), Educación pos pandemia: salud mental y prevención de la violencia con la comunidad educativa, Memoria, UNESCO, en: https://es.unesco.org/sites/default/files/doc-memoria-foro-educacion-post-pandemia.pdf.

[5] webdelmaestrocmf.com, "Estas son las capacidades digitales de un docente moderno." [Online]. Available: http://webdelmaestrocmf.com/portal/5-competencias-digitales-que-deben-tener-los-profesores-actuales/?utm_source=blogsterapp&utm_medium=Facebook. [Accessed: 27-Oct-2017].

[6] A. Núñez-Torrón Stock, "5 competencias digitales que deben tener los profesores actuales." [Online]. Available: http://www.ticbeat.com/educacion/5-competencias-digitales-profesores/. [Accessed: 27-Oct-2017].

[7] M. P. Prendes, "Plataformas de campus virtual con herramientas de software libre: Análisis comparativo de la situación actual en las universidades españolas," Murcia, Project Report EA-2008-0257, Jun. 2009. [Online]. Available: http://www.um.es/campusvirtuales/informe_final_CVSL_SF.pdf

[8] F. J. García-Peñalvo, "Docencia," in Libro blanco univ. digital 2010, J. Laviña Orueta and L. Mengual Pavón, Eds. Barcelona, Spain: Ariel, 2010, pp. 29-61.

[9] G. Attwell, J. Cook, and A. Ravenscroft, "Appropriating Technologies for Contextual Knowledge: Mobile Personal Learning Environments," in Best Practices for the Knowledge Society. Knowledge, Learning, Development and Technology for All, M. D. Lytras, P. O. de Pablos, E. Damiani, D. Avison, A. Naeve, and D. G. Horner, Eds. Springer Berlin Heidelberg, 2009, pp. 15-25.

[10] Technological Pedagogical Content Knowledge (TPACK) in Action. https://www.researchgate.net/publication/271937971_Technological_Pedagogical_Content_Knowledge_TPACK_in_Action

[11] [Kiray, S. A., Çelik, I., & Çolakoğlu, M. H. (2018).TPACK Self-efficacy perceptions of science teachers:A structural equation modeling study. Egitim ve Bi-lim, 43(195), 253-268. http://egitimvebilim.ted.org.tr/index.php/EB/article/view/7538]

[12] World Economic Forum, 2015.

[13] Tendencias educativas que marcaran el 2022 y el futuro de las aulas. https://www.entreestudiantes.com/2022/01/tendencias-educativas-marcaran-2022-futuro-aulas/.

[14] Pedroza, O., & Crespo, M. (2017). Importancia del diseño tecno pedagógico basado en el enfoque de la acción, para reforzar el dominio del idioma inglés como segunda lengua. Revista Colombiana de Computación, 18(2), 7-21, e-ISSN: 2539-2115. Recuperado de: https://doi.org/10.29375/25392115.3214

[15] Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief.US Department of Education, Office of Educational Technology, 1-57.

[16] Otero, A. (2020). El Big Data. Como se usaría en la educación, vermislab. Recuperado de https://www.vermislab.com/big-data-que-es-como-nos-afecta-y-su-uso-en-educacion/

TRANSFORMACIÓN DIGITAL DOCENTE. IMPACTO Y REALIDADES EN EL NIVEL PRIMARIO DE LA CIUDAD DE ORÁN

Lorena E. Del Moral Sachetti, Angel R. Barberis, Viviana Encalada, Virginia Rosa
Facultad Regional Orán & Facultad de Ciencias Exactas – Universidad Nacional de Salta
Orán, Salta - Argentina
lorena.dms.7@gmail.com, barberis@cidia.unsa.edu.ar

RESUMEN

La alfabetización digital refiere al conjunto de competencias que permiten utilizar herramientas tecnológicas con diversos fines, desde tareas cotidianas hasta actividades laborales. En la actualidad, muchos niños ya dominan estas herramientas; sin embargo, su integración en los procesos educativos resulta aún limitada o inexistente. Esto evidencia una desconexión entre las prácticas pedagógicas y el entorno digital en el que se desarrollan los estudiantes, especialmente en el nivel primario.

Como resultado, los contenidos escolares suelen alejarse de la realidad de los alumnos y se pierde la oportunidad de fomentar un uso crítico de la tecnología. En este contexto surge un proyecto de investigación sobre alfabetización digital en docentes de nivel medio de Orán (Salta), orientado a abordar esta problemática mediante formación y acompañamiento docente para fortalecer la integración pedagógica de herramientas digitales.

Iniciado a mediados de 2025, el proyecto detectó limitaciones culturales y de recursos tecnológicos, e impulsa capacitaciones continuas en escuelas de Orán para incorporar estas herramientas en las prácticas educativas, sosteniendo que la alfabetización docente es clave para la inclusión digital del alumnado.

Palabras claves: *Alfabetización Digital, Competencias Digitales, Integración Pedagógica, Formación Docente, Inclusión Digital.*

CONTEXTO

El proyecto se propone trabajar con tres instituciones educativas de gestión pública ubicadas en la ciudad de San Ramón de la Nueva Orán, situada al norte de la provincia de Salta. Para la localización de las instituciones educativas, se recurrió a datos estadísticos disponibles en el sitio web oficial Argentina.gob.ar, el cual ofrece un mapa cartográfico de los establecimientos educativos del país.

Según esta fuente, en la ciudad de San Ramón de la Nueva Orán existen 25 escuelas de nivel primario de gestión pública, distribuidas entre contextos urbanos y rurales: 16 de ellas se encuentran en zonas urbanas, mientras que las 9 restantes pertenecen a zonas rurales.

En este marco, el proyecto prevé la selección de entre tres y cuatro instituciones, en función de la disponibilidad temporal, con el objetivo de implementar instancias de capacitación personalizada orientadas al uso pedagógico de herramientas digitales en las prácticas docentes. No obstante, y aprovechando las posibilidades que brinda la virtualidad, se ofrecerán también capacitaciones remotas abiertas y gratuitas para todos los docentes interesados, ampliando así el alcance de la propuesta.

El foco de la investigación estará centrado en identificar y analizar posibles transformaciones en las prácticas pedagógicas a partir de la intervención, estableciendo un contraste entre la situación previa y posterior al desarrollo del proyecto. En este sentido, se propone analizar el grado de competencias digitales que poseen los docentes de nivel primario, así como realizar un diagnóstico a

través de instrumentos de autoevaluación que permitan identificar sus niveles de dominio tecnológico.

De manera complementaria, el proyecto contempla la implementación de talleres de alfabetización digital orientados a fortalecer las habilidades tecnológicas de los docentes, tanto en el uso de hardware como de software con fines educativos. A su vez, se propone fomentar la integración de las TIC en las prácticas pedagógicas, promoviendo su utilización no solo en actividades de enseñanza, sino también en tareas de gestión y organización institucional.

Igualmente, se prevé relevar información en el conjunto de las 25 instituciones respecto a la disponibilidad y estado de los denominados "carritos tecnológicos". Estos dispositivos, junto con netbooks, cámaras fotográficas, pizarras digitales e impresoras, fueron distribuidos en el marco del Plan Nacional Conectar Igualdad. En este sentido, se busca indagar si dichos recursos se encuentran actualmente en uso, si cuentan con soporte técnico activo y, en aquellos casos en que no lo tengan, si los equipos se encuentran bloqueados y qué posibilidades existen para su reactivación y aprovechamiento pedagógico.

En términos generales, este contexto evidencia la persistencia de desigualdades en el acceso, uso y aprovechamiento pedagógico de las tecnologías digitales, no solo vinculadas a la disponibilidad de recursos, sino también a las condiciones institucionales y formativas de los docentes. En este sentido, el proyecto se inscribe en la necesidad de fortalecer las capacidades docentes para una integración significativa de las tecnologías, promoviendo prácticas educativas más inclusivas y contextualizadas que respondan a las demandas de la sociedad digital contemporánea.

1. INTRODUCCION

Las Nuevas Tecnologías de la Información y la Comunicación (NTIC) se han incorporado de manera significativa en múltiples ámbitos de la vida social. El campo educativo no ha permanecido ajeno a estos avances, dado que dichas tecnologías se encuentran presentes en todos los niveles del sistema educativo a través de computadoras, software educativo, programas multimedia, plataformas virtuales y redes de información disponibles en Internet. En este sentido, su integración en los planes de estudio y en los procesos de enseñanza y aprendizaje se ha ido consolidando progresivamente (Area Moreira, 2007; Cabero, 2006).

Teniendo en cuenta, este contexto, es posible afirmar que los estudiantes suelen desarrollar tempranamente habilidades en el manejo de tecnologías digitales, al haber nacido en entornos altamente digitalizados, lo que ha sido conceptualizado como "nativos digitales" (Prensky, 2001). En contraste, muchos docentes presentan un vínculo más limitado con estas herramientas, lo que evidencia una brecha generacional en el ámbito educativo respecto al uso de las tecnologías de la información y la comunicación (TIC) (Cabero & Llorente, 2008).

Con el objetivo de atender estas transformaciones, el gobierno argentino ha implementado diversas políticas públicas orientadas a promover la inclusión digital y fortalecer la formación docente desde el año 2003. Entre ellas se destaca el programa Conectar Igualdad, creado en 2010, cuyo propósito fue democratizar el acceso a tecnologías digitales mediante la distribución de equipamiento tecnológico en instituciones educativas (Ministerio de Educación de la Nación, 2010). A lo largo de los años, este programa atravesó diferentes etapas de continuidad, reformulación y suspensión, reflejando la complejidad de sostener políticas públicas de inclusión digital en el tiempo.

En el marco de estas iniciativas, muchas escuelas públicas fueron provistas de "carritos tecnológicos" que incluyen netbooks, impresoras, pizarras digitales y otros dispositivos, con el objetivo de favorecer su integración en los procesos de enseñanza y aprendizaje. Sin embargo, a pesar de las capacitaciones ofrecidas, en numerosos casos estos recursos permanecen subutilizados, lo que pone en evidencia la necesidad de fortalecer las competencias digitales docentes y de generar

estrategias de acompañamiento sostenido (Dussel, 2012; Lugo & Kelly, 2011).

En este sentido, la alfabetización digital docente se presenta como un elemento clave para la incorporación significativa de las TIC en el ámbito educativo. Tal como señalan diversos autores, esta no solo implica el desarrollo de habilidades técnicas, sino también la capacidad de analizar, diseñar, evaluar y utilizar críticamente las tecnologías como herramientas pedagógicas (Gutiérrez Martín, 2003; Area Moreira, 2010).

El presente proyecto tiene como propósito investigar el nivel de alfabetización digital de los docentes de nivel primario en la ciudad de San Ramón de la Nueva Orán, así como identificar las principales limitaciones que dificultan la integración de las TIC en sus prácticas pedagógicas. Asimismo, se propone analizar el grado de implementación y el impacto de las políticas públicas de alfabetización digital en el contexto local.

De manera complementaria, el proyecto contempla el diseño e implementación de instancias de capacitación orientadas a fortalecer las competencias digitales docentes, organizadas en tres ejes principales: el uso y manejo del equipamiento tecnológico (hardware), la utilización de herramientas digitales para tareas administrativas y académicas, y el empleo de software educativo específico para la enseñanza de contenidos.

2. LINEAS DE INVESTIGACION y DESARROLLO

El proyecto se inscribe en el campo de estudio de la tecnología educativa y, específicamente, en las líneas de investigación vinculadas con la alfabetización digital, la integración pedagógica de las tecnologías de la información y la comunicación (TIC) y la formación docente en entornos digitales.

En primer lugar, una de las líneas centrales de investigación está orientada al análisis de las competencias digitales docentes. En este sentido, se busca diagnosticar el nivel de alfabetización digital de los docentes de nivel primario, identificando sus conocimientos, habilidades y actitudes frente al uso de tecnologías en el ámbito educativo. Esta línea se articula con marcos teóricos contemporáneos que destacan la necesidad de desarrollar competencias digitales integrales para una enseñanza efectiva en contextos mediados por tecnologías.

En segundo lugar, el proyecto aborda la problemática de la integración pedagógica de las TIC, entendida no solo como la disponibilidad de recursos tecnológicos, sino como su incorporación significativa en las prácticas de enseñanza. En esta línea, se investigan las principales barreras que dificultan dicha integración, tales como factores culturales, institucionales y formativos, así como las resistencias docentes frente al uso de tecnologías digitales.

Una tercera línea de investigación se centra en el análisis de políticas públicas de inclusión digital, particularmente aquellas implementadas en el sistema educativo argentino, como el programa Conectar Igualdad. Se busca comprender el alcance, la implementación y el impacto de estas políticas en el contexto local, así como las condiciones que favorecen o limitan su sostenibilidad en las instituciones educativas.

En cuanto a las líneas de desarrollo, el proyecto propone el diseño e implementación de estrategias de formación docente orientadas a la alfabetización digital. Estas acciones se materializan en talleres de capacitación que abordan tanto el uso técnico de los recursos (hardware), como la aplicación de herramientas digitales en tareas administrativas, académicas y pedagógicas, así como el uso de software educativo específico.

Finalmente, el proyecto incorpora una línea de desarrollo vinculada a la evaluación del impacto de las intervenciones formativas. A través de un enfoque comparativo, se pretende analizar las transformaciones en las prácticas docentes antes y después de la implementación de las capacitaciones, con el objetivo de generar conocimiento aplicable que contribuya a mejorar los procesos de integración de las TIC en el ámbito educativo.

En conjunto, estas líneas de investigación y desarrollo buscan no solo comprender la problemática de la alfabetización digital en el contexto local, sino también generar propuestas concretas de intervención que promuevan una educación más inclusiva, pertinente y acorde a las demandas de la sociedad digital contemporánea.

3. RESULTADOS OBTENIDOS / ESPERADOS

En el marco de la primera etapa del proyecto, se llevó a cabo un acercamiento inicial a una institución educativa de la ciudad de San Ramón de la Nueva Orán. Como resultado de este relevamiento, se constató que la escuela dispone de un carrito tecnológico equipado con aproximadamente 30 netbooks en buen estado de conservación, debido a su escaso uso. Asimismo, se identificaron otros recursos disponibles, tales como una impresora, una cámara digital y un proyector, los cuales también se encuentran en condiciones operativas.

No obstante, se evidenció que dichos recursos no son utilizados de manera sistemática en las prácticas pedagógicas. Esta situación se vincula principalmente con la falta de soporte técnico desde la discontinuidad del programa Conectar Igualdad, lo que ha derivado en el bloqueo de los dispositivos y, en consecuencia, en la imposibilidad de su uso por parte de los docentes. Actualmente, se están realizando gestiones para establecer contacto con autoridades del sistema educativo que puedan facilitar las herramientas necesarias para el desbloqueo de los equipos, priorizando soluciones institucionales antes que intervenciones técnicas directas sobre el hardware.

Por otra parte, se relevó que aproximadamente el 40 % de los docentes incorpora, en alguna medida, recursos digitales en sus propuestas de enseñanza. Sin embargo, se observa una marcada predisposición e interés por ampliar sus conocimientos y fortalecer sus competencias en el uso pedagógico de las tecnologías.

En cuanto a los resultados esperados, se prevé continuar con el relevamiento en las restantes instituciones, aunque se reconoce como limitación potencial la disponibilidad de tiempo para abarcar la totalidad de los casos en profundidad. No obstante, gran parte de la información ya ha sido recolectada y actualmente se encuentra en proceso de sistematización y análisis.

Se espera que este análisis identifique con mayor precisión las principales restricciones en la integración de herramientas digitales, para diseñar capacitaciones más pertinentes y contextualizadas.

4. FORMACION DE RECURSOS HUMANOS

La estructura del equipo de investigación es de 4 (cuatro) miembros incluidos el Director.

Uno de los miembros desarrolla la tesis de Maestría en Ingeniería de Sistema de Información (UTN-FRC).

Otro de los miembros, desarrolla su tesis de la Maestría en Docencia Universitaria de Disciplinas Tecnológicas (UNCA) con tema en el marco del presente proyecto.

De manera complementaria, el proyecto contempla la participación ocasional de estudiantes en el marco de una perspectiva de gestión a escala micro, mediante la implementación de proyectos de extensión con participación estudiantil. En estas instancias intervienen alumnos de carreras vinculadas a la informática, tanto de la Facultad Regional Orán como de la Facultad de Ciencias Exactas de la ciudad de Salta.

A través de estos proyectos, los estudiantes realizan un primer acercamiento al campo laboral, enfrentándose a situaciones reales en contextos educativos. Entre sus principales actividades se encuentran el relevamiento de los equipos tecnológicos disponibles en las instituciones, el análisis de su estado de funcionamiento y, en muchos casos, la explicación de su uso a los docentes de nivel medio.

Asimismo, se les ofrece participar en capacitaciones para desarrollar y aplicar

habilidades comunicativas, como el uso adecuado del lenguaje técnico y la transmisión clara de conocimientos, fortaleciendo competencias clave para su futura inserción profesional en análisis de sistemas y gestión tecnológica.

5. BIBLIOGRAFIA

Area Moreira, M. (2007). La integración de las tecnologías de la información y la comunicación en los sistemas educativos. Madrid: Fundación Telefónica.

Area Moreira, M. (2010). El proceso de integración y uso pedagógico de las TIC en los centros educativos. Revista de Educación, (352), 77–97.

Area Moreira, M., & Pessoa, T. (2012). De lo sólido a lo líquido: Las nuevas alfabetizaciones ante los cambios culturales de la Web 2.0. Comunicar, 19(38), 13–20.

Cabero, J. (2006). Bases pedagógicas del e-learning. Revista de Universidad y Sociedad del Conocimiento, 3(1), 1–10.

Cabero, J., & Llorente, M. C. (2008). La alfabetización digital de los alumnos. Competencias digitales para el siglo XXI. Revista Portuguesa de Pedagogía, 42(2), 7–28.

Coll, C., & Monereo, C. (2008). Psicología de la educación virtual: Aprender y enseñar con las tecnologías de la información y la comunicación. Madrid: Morata.

Dussel, I. (2012). Más allá del mito de los "nativos digitales": Jóvenes, escuela y saberes en la cultura digital. Buenos Aires: Santillana.

Gisbert, M., & Esteve, F. (2011). Digital learners: La competencia digital de los estudiantes universitarios. La Cuestión Universitaria, (7), 48–59.

Gutiérrez Martín, A. (2003). Alfabetización digital: Algo más que ratones y teclas. Barcelona: Gedisa.

Kenton, J., & Blummer, B. (2010). Promoting digital literacy skills: Examples from the literature and implications for academic librarians. Community & Junior College Libraries, 16(2), 84–99.

Lugo, M. T., & Kelly, V. (2011). La gestión de las TIC en las escuelas: El desafío de gestionar la innovación. Buenos Aires: IIPE-UNESCO.

Marqués Graells, P. (2000). Impacto de las TIC en la educación: Funciones y limitaciones. Barcelona: Universidad Autónoma de Barcelona.

Mirete Ruiz, A. B. (2016). La alfabetización digital: Una competencia clave para el profesorado del siglo XXI. Revista Electrónica Interuniversitaria de Formación del Profesorado, 19(2), 25–34.

Pérez Escoda, A. (2015). Alfabetización digital y competencias digitales en el marco de la educación superior. Pixel-Bit. Revista de Medios y Educación, (47), 85–98.

Pérez Tornero, J. M. (2004). La alfabetización digital y mediática: Nuevos desafíos educativos. Barcelona: UOC.

Prensky, M. (2001). Digital natives, digital immigrants. On the Horizon, 9(5), 1–6.

Un abordaje de la enseñanza de algoritmos evolutivos en la materia inteligencia artificial en la carrera de ISI UTN FRC desde los ambientes de aprendizaje colaborativos

Olariaga, Sandra[1]; Paéz, Nancy del Valle[2]; Stefanich, Clarisa[3];
Hass, Silvina[4]; De la Quintana Segundo Javier[5]
Departamento de Ingeniería en Sistemas de Información
Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN-FRC)
Maestro Marcelo López y Cruz Roja Argentina – Ciudad Universitaria – Córdoba, Argentina
(1)solariaga@yahoo.com;(2)nancydelv@gmail.com;(3)clarystefanich@gmail.com
(4)silhass13@gmail.com; (5)segudlq@gmail.com

Eje Temático del Trabajo: TIAE Tecnología Informática Aplicada enEducación.

RESUMEN

El uso de la tecnología como apoyo al proceso de enseñanza y aprendizaje, se ha convertido en necesario e imprescindible. Las estrategias didácticas mediadas por tecnologías permiten presentar conceptos e ideas a los estudiantes de manera multimedial, ágil e interactiva, situándolos en un rol activo y participativo, lo cual les permite ejercitar y extender sus conocimientos. En el caso concreto de la asignatura Inteligencia Artificial, ya se han incorporando, herramientas como soporte del estudio de los contenidos curriculares; en esta ocasión, el enfoque está puesto en la temática de algoritmos evolutivos. Con el nombre de algoritmos evolutivos se identifica a un grupo de procedimientos que se distinguen de otras técnicas porque están inspirados en las teorías que sintetizan la evolución Darwiniana y la herencia genética descripta por Mendel, también se basan en la inteligencia colectiva de grupos de organismos, como bandadas de aves o cardúmenes de peces, para resolver problemas de optimización.

A través de este proyecto se pretende focalizar en el uso de estas tecnologías como medio para que el estudiante construya su propio conocimiento, pero en una dinámica grupal y colaborativa, lo que actualmente se denomina ambiente colaborativo de aprendizaje

Palabras clave: algoritmos evolutivos, inteligencia artificial, aprendizaje colaborativo.

CONTEXTO

El presente proyecto está siendo llevado a cabo en el Departamento de Ingeniería en Sistemas de la Universidad Tecnológica Nacional Facultad Regional Córdoba. (U.T.N.-F.R.C.).

El Código del proyecto es TETEUCO0008805 dentro del área deAplicaciones tecnológicas en la educación superior, radicado dentro de la Secretaría de Ciencia y Tecnología de la Universidad

Tecnológica Nacional Facultad Regional Córdoba. Fecha de inicio: 01/04/2023 Fecha de finalización: 31/03/2026. Con pedido de prorroga.

1. IN TRODUCCIÓN

La aplicación de las tecnologías en la educación, ha generado importantes cambios en la interacción y comunicación entre los estudiantes, aplicándose de esta manera el aprendizaje colaborativo.

Los ambientes colaborativos de aprendizaje brindan espacios en los que se pueda explorar los conceptos o las situaciones problemáticas que se desean tratar, apoyados con tecnología, así como de estrategias pedagógicas como la investigación y pensamiento crítico. De esta manera se favorece el desarrollo de habilidades mixtas (aprendizaje y desarrollo personal y social), donde cada miembro del grupo es responsable tanto de su aprendizaje como del aprendizaje de los restantes miembros del grupo, lo que implica un cambio de roles asociados a profesores y estudiantes.

Este modelo permite aprovechar las ventajas del trabajo grupal y la colaboración entre los mismos, ya que cada estudiante asume su aprendizaje a su propio ritmo, potencialidades, alentando la actividad autónoma, pero de manera colaborativa donde cada uno comprende la necesidad de aportar lo mejor de si al grupo, para lograr un resultado sinérgico; favoreciendo de esta manera los procesos de aprendizaje individuales, las relaciones interpersonales y la productividad. Además, favorece la capacidad de resolver problemas de forma creativa, a partir de estrategias de mediación y la búsqueda grupal de alternativas, fomentando la motivación y una actitud positiva hacia la materia de estudio.

Con el nombre de algoritmos evolutivos se identifica a un grupo de procedimientos inspirados en las teorías que sintetizan la evolución Darwiniana y la herencia genética descripta por Mendel, también se basan en la inteligencia colectiva de grupos de organismos, como bandadas de aves o cardúmenes de peces, para resolver problemas de optimización. Estos algoritmos se destacan por combinar exitosamente dos mecanismos de búsqueda: exploración y explotación. La exploración permite identificar áreas prometedoras del espacio de búsqueda y la explotación lleva a cabo una búsqueda especializada dentro de estas áreas para encontrar soluciones cercanas al óptimo. Su aplicación se extiende en áreas de la actividad humana como por ejemplo en salud, educación, ciencias, industria, economía y defensa. Se analizan dentro de los Algoritmos Evolutivos los dos más importantes los algortimos genéticos y el de optimización por

enjambre de partículas, ambos están orientados a encontrar mínimos o máximos globales. Los algoritmos genéticos se basan en hacer evolucionar poblaciones de soluciones hacia valores óptimos del problema a través de mecanismos de combinación y mutación.

En el caso de los algoritmos de enjambre de partículas se parte de la creación de un enjambre inicial (posibles soluciones), las cuales se evalúan y se procede a la actualización de los parámetros de esta solución hasta llegar a la óptima.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se inscribe dentro de los lineamientos de investigación en innovaciones curriculares en Educación Superior Universitaria dentro del área de ambientes colaborativos de aprendizaje.

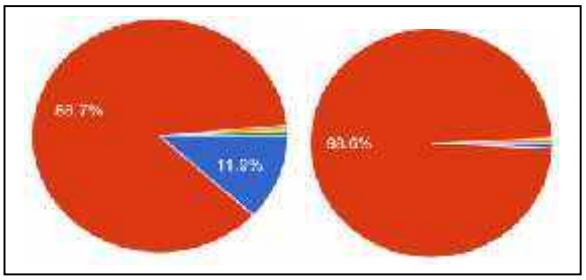
Para el desarrollo de este trabajo de investigación se aplicará el método empírico-analítico, que se basa en la experimentación y en la lógica empírica, junto a la aplicación y evaluación de la herramienta elaborada.

3. RESULTADOS OBTENIDOS / ESPERADOS

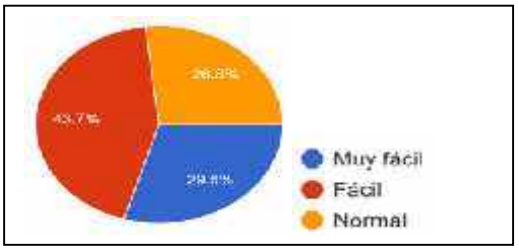
El eje sobre el que gira nuestro trabajo es la elaboración e implementación de una herramienta de software que nos permite evaluar la enseñanza de la temática de algoritmos evolutivos.

Una vez formulada la encuesta, ésta se les presentó a los estudiantes que ya habían utilizado la herramienta desarrollada. Se realizó un análisis de las respuestas obtenidas respecto de la utilización de la herramienta desarrollada el que se presenta a continuación.

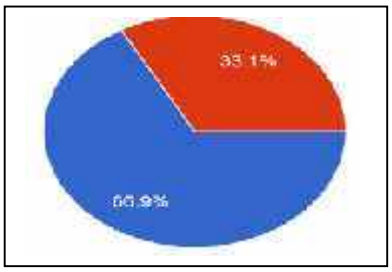
En lo sucesivo nos referiremos a algoritmos genéticos como (AG) y a enjambre de partículas como (EP). Cabe aclarar que la herramienta presenta el mismo problema a ser resuelto tanto por el AG como por el EP, a fin de que se comparen la forma de resolución de cada uno. A continuación los resultados generales:Un 86.7% expresó no tener conocimiento de los AG, y un 98.6 % de EP.



Al 43.7 % le fue sencillo la comprensión de la herramienta. Hubo algunos problemas de entendimiento de los datos presentados en la interfaz gráfica.



Para el 66,9 % los resultados fueron coherentes y comprensibles.



Entre las sugerencias generales para mejorar la herramienta se destacan las siguientes:

* Requerimiento de más explicaciones en cada una de las etapas de cada uno de los algoritmos.

* Sugieren cambios en la forma en que se ingresan los diferentes parámetros del problema.

Analizaremos estas respuestas a fin de perfeccionar la herramienta la cual una vez modificada será puesta a disposición de los estudiantes de este año.

Una vez que los estudiantes hayan utilizado la herramienta perfeccionada aplicaremos nuevamente la encuesta a fin de analizar las respuestas de los estudiantes obtenidas en esta ocasión.

Los resultados finales obtenidos serán presentados ante el jefe de cátedra de la materia Inteligencia Artificial, para que previa evaluación de la herramienta ya perfeccionada pueda ser implementada en todos los cursos que integran dicha cátedra.

4. FORMACIÓN DE RECURSOS HUMANOS

El desarrollo del Proyecto que engloba este trabajo, contribuirá a ampliar el conocimiento de los algoritmos evaluativos particularmente AG y EP, dentro de los ambientes colaborativos de aprendizaje, de manera de que los resultados obtenidos puedan llegar a ser el punto de partida de actividades que favorezcan el mejoramiento de los procesosde enseñanza y de aprendizaje en el aula universitaria.

Los investigadores que recién inician forjarán sus primeras herramientas en I+D lo que les permitirá ampliar conocimientos en el área en cuestión además de incrementar sus antecedentes académicos. Este proyecto contribuirá a fortalecer la formación de un becario, estudiante de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

ADELL, J., Mengual-Andrés, S., & Roig-Vila, R. (2015). Monográfico. Webquest: 20 añosutilizando Internet como recurso para el aula. Edutec. Revista Electrónica de Tecnología Educativa, 52.http://www.edutec.es/revista/index.php/edutec-e/article/view/622

COELLO Carlos Artemio Coordinador Computación Evolutiva en colaboración con la AcademiaMexicana de Computación Primera edición: 2019 Academia Mexicana de Computación, A. C.

DÍAZ BARRIGA: Estrategias docentes para un aprendizaje significativo. Una interpretación constructivista.México, Mcgraw-Hill. (1999)

FERNANDEZ MUÑOZ, R. (2003). Competencias profesionales del docente en la sociedad del siglo XXI. Praxis, 1,4-8.

GUTIERREZ ESTEBAN, P., Yuste, R., Cubo, S., & Lucero, M. (2011). Buenas prácticas en el desarrollo de trabajocolaborativo en materias TIC aplicadas a la educación. Profesorado, Revista de Currículum y Formación delProfesorado, 15. Recuperado de http://www.ugr.es/~recfpro/rev151ART13.pdf

LAISEMA, S., & Wannapiroon, P. (2014). Design of Collaborative Learning with Creative Problem-solvingProcess Learning Activities in a Ubiquitous Learning Environment to Develop Creative Thinking Skills. Procedia- Social and Behavioral Sciences, 116, 3921-3926. doi:10.1016/j.sbspro.2014.01.867

LUCERO, M. M: Entre el trabajo colaborativo y el aprendizaje colaborativo Revista Iberoamericana de EducaciónISSN 1681-5653 (2003).

MORIN, E. Los siete saberes necesarios para la educación del futuro. Organización de las Naciones Unidas,para la Educación, la Ciencia y la Cultura.Traducción de Mercedes Vallejo-Gómez, Medellín. Colombia. (2000).

SALINAS, J., de Benito, B., & Lizana, A. (2014). Competencias docentes para los nuevos escenarios deaprendizaje. Revista Interuniversitaria de Formación del Profesorado, 28(1), 145-163. Recuperado de http://www.redalyc.org/articulo.oa?id=27431190010.

TUDGE: Vigotsky: la zona de desarrollo próximo y su colaboración en la práctica de aula. NuevaYork,Universidad de Cambridge. (1994).

VIGOTSKY, L. (1979). El desarrollo de los procesos psicológicos superiores.

ZHU, C. (2012). Student Satisfaction, Performance, and Knowledge Construction in Online CollaborativeLearning. Educational Technology & Society, 15(1), 127-136.

Videojuegos y analítica educativa en educación primaria: evaluación y desarrollo de propuestas pedagógicas

Rodrigo René Cura, Romina Stickar

Departamento de Informática, Facultad de Ingeniería,
Universidad Nacional de la Patagonia San Juan Bosco.
Puerto Madryn, Chubut, Argentina. +54 280-4472885 – Int. 116
rodrigo.renecura@gmail.com, romistickar@gmail.com

RESUMEN

Este trabajo presenta una línea de investigación y desarrollo orientada a la integración de videojuegos educativos en el nivel primario, incorporando un sistema de analítica educativa que permite recolectar y visualizar datos procesuales de las sesiones de juego. La propuesta busca validar el potencial de esta información para apoyar la toma de decisiones pedagógicas, tanto en contextos sincrónicos dentro del aula como en instancias asincrónicas.

La plataforma, desarrollada en colaboración con la asociación Formate Digital, registra la manera en que los estudiantes abordan y resuelven los desafíos propuestos, generando reportes que pueden ser utilizados por los y las docentes para ajustar sus prácticas didácticas. El proyecto se encuentra en etapa de implementación piloto y procura producir evidencia sobre la integración significativa entre videojuegos, análisis de datos y enseñanza en contextos escolares reales.

Palabras clave: VIDEOJUEGOS, ANALÍTICA EDUCATIVA, PEDAGOGÍA.

CONTEXTO

El proyecto se desarrolla en el marco de los programas de investigación avalados por la Secretaría de Ciencia y Técnica de la Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB). Se inscribe en las actividades del Grupo de Desarrollo de Videojuegos (GDV), perteneciente al Departamento de Informática de la Facultad de Ingeniería, sede Puerto Madryn.

Desde su conformación en 2015, el GDV ha orientado sus actividades al diseño y desarrollo de productos interactivos, combinando formación académica, producción tecnológica y vinculación con el medio. En los últimos años, el grupo ha consolidado una línea de trabajo específica en el desarrollo de videojuegos con fines educativos, integrando saberes de programación, diseño de interacción y didáctica disciplinar.

El presente trabajo es posible además gracias a la colaboración con la asociación Formate Digital, creadora de la plataforma homónima, integrada por las docentes e investigadoras Patricia Sadovsky, Mónica Becerril y Mónica Urquiza. Esta articulación fortalece el proyecto desde el punto de vista pedagógico y curricular, aportando experiencia en didáctica de la matemática y en el diseño de propuestas formativas para docentes de nivel primario. La convergencia entre el desarrollo tecnológico impulsado desde la Facultad de Ingeniería y el trabajo pedagógico sostenido por Formate Digital configura un espacio interdisciplinario que potencia la calidad y pertinencia de la propuesta.

En este marco, la plataforma Formate Digital constituye un desarrollo tecnológico en fase de consolidación que integra videojuegos educativos con herramientas de seguimiento y analítica educativa orientadas al acompañamiento docente. La iniciativa se apoya en la experiencia previa del grupo en proyectos

de vinculación con instituciones educativas de la ciudad y la región, lo que permite situar el desarrollo en problemáticas concretas del sistema escolar patagónico.

El contexto territorial presenta desafíos vinculados con la disponibilidad desigual de recursos tecnológicos y con la necesidad de propuestas pedagógicas contextualizadas. En este escenario, el proyecto busca contribuir al fortalecimiento de capacidades locales para el desarrollo de soluciones tecnológicas educativas, evitando la dependencia exclusiva de plataformas externas y promoviendo la producción de conocimiento situado.

Asimismo, la línea se proyecta como un espacio abierto a la cooperación interinstitucional, tanto en el plano académico como en el desarrollo de experiencias de validación compartidas. La integración de desarrollo tecnológico, analítica educativa y trabajo de campo en escuelas primarias configura una propuesta que articula investigación, innovación y transferencia, en consonancia con los objetivos de fortalecimiento del sistema universitario nacional.

1. INTRODUCCIÓN

La enseñanza de la matemática en la educación primaria enfrenta desafíos persistentes vinculados con la motivación del estudiantado, la diversidad de estilos de aprendizaje y la necesidad de recursos pedagógicos que favorezcan la comprensión de conceptos abstractos. En este contexto, el uso de videojuegos educativos ha demostrado potencial para enriquecer el proceso de enseñanza-aprendizaje, promoviendo entornos más interactivos y significativos [1, 2, 3].

La incorporación de videojuegos en el ámbito educativo se fundamenta en la posibilidad de generar experiencias lúdicas que favorezcan el aprendizaje significativo y el desarrollo de habilidades cognitivas [3], así como en su capacidad de actuar como dispositivos de acceso a la cultura digital contemporánea [4]. Diversas investigaciones han señalado su incidencia en la mejora de la resolución de problemas y en el estímulo de la creatividad [3], además de su contribución a la reducción de la brecha digital en contextos escolares con acceso limitado a recursos tecnológicos [5].

Sin embargo, gran parte de las experiencias con videojuegos educativos se han centrado en el potencial motivacional o en los resultados finales obtenidos por los estudiantes, sin profundizar en el análisis sistemático de los procesos de resolución que se despliegan durante la interacción con el entorno digital. En este sentido, emerge como campo relevante la **analítica educativa (learning analytics)**, entendida como la medición, recopilación, análisis y reporte de datos sobre los aprendices y sus contextos con el propósito de comprender y optimizar los procesos de aprendizaje y los entornos en los que estos ocurren [6, 7].

La analítica educativa se ha consolidado como una disciplina emergente que articula tecnología, pedagogía y análisis de datos, aportando herramientas para transformar información generada en entornos digitales en insumos significativos para la toma de decisiones docentes [6]. Asimismo, se ha señalado su potencial para mejorar la retroalimentación, personalizar intervenciones y fortalecer la comprensión de las dinámicas de aprendizaje [8].

En el marco de la convergencia entre informática educativa, didáctica de la matemática y desarrollo de tecnologías interactivas, resulta pertinente explorar modelos que integren videojuegos educativos con herramientas de seguimiento y análisis de datos. Esta articulación permite trascender la concepción de los videojuegos como meros recursos motivacionales o dispositivos de entretenimiento [9], situándolos como entornos capaces de generar evidencia procesual sobre las estrategias y dificultades del estudiantado.

En este contexto, la línea de investigación desarrollada por el Grupo de Desarrollo de Videojuegos (GDV) de la UNPSJB se orienta al diseño y validación de una plataforma que integra videojuegos educativos con herramientas de analítica educativa destinadas a docentes de

nivel primario. El propósito no se limita a evaluar el impacto de los juegos en términos de resultados, sino a analizar cómo la disponibilidad de datos sobre las sesiones de juego, tanto en instancias sincrónicas dentro del aula como en actividades asincrónicas, puede contribuir a la reflexión pedagógica y al ajuste de las prácticas de enseñanza.

De este modo, el proyecto se posiciona en la intersección entre innovación tecnológica y mejora pedagógica, proponiendo un modelo de retroalimentación continua en el cual la experiencia lúdica genera datos, los datos se transforman en información pedagógicamente significativa y dicha información orienta decisiones didácticas contextualizadas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto se inscribe en una línea de investigación y desarrollo orientada al diseño, implementación y validación de videojuegos educativos como mediadores pedagógicos en el área de matemática en el nivel primario. Su rasgo innovador radica en la integración de principios de analítica educativa dentro de la plataforma Formate Digital, desarrollada en el ámbito universitario con trayectoria en investigación y educación.

A diferencia de propuestas centradas exclusivamente en el uso lúdico, la plataforma incorpora un sistema de recolección, procesamiento y visualización de datos que registra la interacción del estudiantado con los desafíos matemáticos, tanto en instancias sincrónicas dentro del aula como en actividades asincrónicas. No se limita a reportar resultados finales, sino que recopila información procesual, como cantidad de intentos, tiempos de resolución, tipos de errores, secuencias de acciones y niveles de progresión, con el propósito de transformarla en información pedagógicamente significativa.

La innovación reside en articular experiencia lúdica y análisis de datos para generar un ciclo de retroalimentación en el que la actividad de juego produce evidencias que orientan decisiones didácticas posteriores. En este marco, el proyecto busca validar si la incorporación de analítica educativa en entornos digitales facilita la identificación de patrones de error, favorece ajustes en la planificación docente y contribuye a un seguimiento continuo del aprendizaje más allá del momento presencial.

Desde el punto de vista teórico, la línea analiza el aporte de la analítica educativa en contextos lúdicos aplicados a la enseñanza de la matemática, así como su incidencia en la reflexión pedagógica y en modelos de integración curricular que combinen mediación digital y análisis de datos, atendiendo además a criterios de uso ético y contextualizado de la información escolar. En el plano empírico, se prevé la implementación de experiencias piloto en escuelas primarias, el análisis de los registros generados por el sistema y la evaluación de la utilidad pedagógica de la información desde la perspectiva docente, incorporando ajustes iterativos tanto en el diseño de los videojuegos como en la interfaz de visualización.

La estrategia de desarrollo contempla acuerdos de colaboración con instituciones escolares, instancias de capacitación docente y la aplicación de instrumentos cualitativos y cuantitativos que permitan triangular registros digitales y testimonios, consolidando un modelo que integra desarrollo tecnológico, analítica educativa e innovación pedagógica orientado a la mejora continua de las prácticas de enseñanza.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto se encuentra en etapa inicial de implementación. Se han desarrollado los primeros prototipos funcionales de videojuegos integrados a la plataforma Formate Digital, junto con módulos de recolección y visualización de datos asociados a las sesiones de juego. Paralelamente, se avanza en la selección de escuelas primarias participantes y en la

planificación conjunta con los equipos docentes para integrar la herramienta en propuestas didácticas concretas.

En el corto plazo se prevé realizar experiencias piloto orientadas a validar tanto el funcionamiento técnico del sistema como la pertinencia pedagógica de la información generada mediante analítica educativa. Se evaluará la claridad de los reportes, la utilidad de los datos para identificar dificultades conceptuales y la adecuación de la visualización para su interpretación docente en contextos sincrónicos y asincrónicos.

En el mediano plazo se busca obtener evidencia sobre el modo en que la disponibilidad de datos procesuales derivados del juego contribuye a la toma de decisiones didácticas, favoreciendo ajustes en la planificación, intervenciones diferenciadas y seguimiento de trayectorias individuales y grupales. La estrategia metodológica contempla la triangulación entre registros digitales, entrevistas y observaciones docentes, con el fin de analizar el impacto pedagógico del sistema en contextos escolares reales.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto integra a docentes-investigadores de la Licenciatura en Informática del Departamento de Informática de la Facultad de Ingeniería (UNPSJB), quienes desarrollan sus trabajos finales de Maestría en Enseñanza en Escenarios Digitales en articulación con esta línea.

Asimismo, participan estudiantes de las carreras de Licenciatura en Informática y Analista Programador Universitario en tareas de desarrollo, implementación y análisis de datos de la plataforma, fortaleciendo su formación en desarrollo de software educativo y analítica educativa aplicada.

La línea promueve la integración entre docencia, investigación y desarrollo tecnológico, generando oportunidades de participación en proyectos de innovación con impacto en el sistema educativo regional.

5. BIBLIOGRAFÍA

[1] Jaramillo, L. M., & Puga, L. A. (2016). El pensamiento lógico-abstracto como sustento para potenciar los procesos cognitivos en la educación. Sophia, colección de Filosofía de la Educación, (21), 31-55. https://doi.org/10.17163/soph.n21.2016.01

[2] Martínez, J. (2019). Percepciones de estudiantes y profesores acerca de las competencias que desarrollan los videojuegos. Pensamiento Educativo, 56(2), 1-21. https://doi.org/10.7764/PEL.56.2.2019.3

[3] Roncancio-Ortiz, M. F., Ortiz-Carrera, H., Llano-Ruiz, M.J., Malpica-Lopéz, J. J., & Bocanegra García, J. (2017). El uso de los videojuegos como herramienta didáctica para mejorar la enseñanza-aprendizaje: una revisión del estado del tema. Revista Ingeniería Investigación y Desarrollo, 17(2), 36-46. https://doi.org/10.19053/1900771X.v17.n2.2017.7184

[4] Lion, C., & Perosi, V. (2018). Los videojuegos serios como escenarios para la construcción de experiencias. e-ducadores DEL MUNDO. Revista telecolaborativa internacional, mayo, 4-8. http://e-ducadores.org/wp-content/uploads/2018/05/E-ducadores-del-mundo-2.pdf

[5] Thai, K. P., Bang, H. J., y Li, L. (2022). Accelerating early math learning with research-based personalized learning games: A cluster randomized controlled trial. Journal of Research on Educational Effectiveness, 15(1), 28-51. http://dx.doi.org/10.1080/19345747.2021.1969710

[6] Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. American Behavioral Scientist, 57(10), 1380–1400. https://doi.org/10.1177/0002764213498851

[7] Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE Review, 46(5), 31–40.

[8] Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4(5/6), 304–317. https://doi.org/10.1504/IJTEL.2012.051816

[9] Moncada, J. & Chacón, Y. (2012). El efecto de los videojuegos en variables sociales, psicológicas y fisiológicas en niños y adolescentes. Retos. Nuevas Tendencias en Educación Física. Deporte y Recreación, 21, 43-49. https://www.redalyc.org/pdf/3457/345732287009.pdf

xSTEAM: un enfoque sistémico para el desarrollo de soluciones STEAM en contextos extremos

Gilda R. Romero [1, 2, 3 y 4 [0009-0002-3524-5967]], Daniela López De Luise[1, 2, 3, 4 y 5 [0000-0003-3130-873X]], Nicolás Rizzo [1, 2, 3 y 4], Pablo Pescio[1, 2, 3 y 4], Cecilia Salina[1, 2, 3 y 4], Gisela Giménez [1, 2 y 5]

[1] *Centro de Altos Estudios en Tecnología Informática (CAETI). Universidad Abierta Interamericana - Facultad de Tecnología Informática. Buenos Aires, Argentina*

[2] *CI2S Labs (Computational Intelligence & Information Systems Labs). Buenos Aires, Arg.*

[3] *Sociedad Científica Argentina. Buenos Aires, Arg.*

[4] *Universidad Abierta Interamericana - Facultad de Tecnología Informática. Buenos Aires, Arg.*

[5] *Universidad Autónoma de Entre Ríos - Facultad de Ciencia y Tecnología. Entre Ríos, Arg.*

gilda.romero@gmail.com; mdldl.CI2S@gmail.com; nicolasrizzo@gmail.com; pablo.d.pescio@gmail.com; ceva25@hotmail.com; gisela.gimenez@uader.edu.ar

RESUMEN

El enfoque STEAM ha demostrado ser una de las estrategias más eficaces para la integración interdisciplinaria en educación; sin embargo, su aplicación en contextos heterogéneos, tecnológicos y organizacionales complejos requiere modelos más flexibles y escalables. Este trabajo estudia específicamente las actividades xSTEAM (eXtreme STEAM), término que se define en el contexto actual como experiencias STEAM en situaciones de restricciones inusuales para el desempeño de actividades integradoras. Se propone una estructuración conceptual estratificada que integra contexto, modelo, protocolos, tecnologías y mejora continua, permitiendo gestionar la complejidad de los procesos STEAM en condiciones extremas. En el trabajo se describe cómo la propuesta conceptual xSTEAM podría articularse con la plataforma y la arquitectura de la plataforma tecnológica del proyecto LINCIEVIS, diseñada para la reutilización, adaptabilidad y transferencia del conocimiento a partir de experiencias STEAM. El trabajo posiciona a xSTEAM como un método sistemático que habilita la incorporación de comunidades hasta ahora impedidas de este tipo de actividades, disponiendo de una nueva forma de diseño, implementación y evaluación de experiencias educativas STEAM, como un marco extendido y abarcativo.

Palabras clave: *STEAM, xSTEAM, métrica educativa, Sistemas Inteligentes, innovación educativa, contextos extremos*

CONTEXTO

Este proyecto forma parte de los proyectos de la Sociedad Científica Argentina (SCA) en colaboración con el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). Además, forma parte del laboratorio denominado LINCIEVIS (Laboratorio de INvestigaciones CIEntíficas VIdeojuegos y STEAM) de la Sociedad Científica Argentina (SCA, Instituto de Comunicaciones digitales presidido por la doctora Daniela López De Luise), iniciado con expertos del Museo Sarmiento, IEEE UCSE DASS, Universidad Autónoma de Entre Ríos y que evolucionó hasta incorporar y radicarse en el centro CAETI de la UAI.

El grupo LINCIEVIS de la SCA, a cargo desde su comienzo por la Dra. Daniela López De Luise, estudia y aplica experiencias STEAM desde 2018, tanto en Argentina como en otros países y con la colaboración de instituciones prestigiosas como el centro CAETI de la UAI, Universidad Católica de Santiago del Estado sede San Salvador de Jujuy, Universidad Autónoma de Entre Ríos en Concepción del Uruguay, y el IEEE Computational Intelligence Society capítulo Argentina. Las investigaciones y experiencias desarrolladas han derivado en varias líneas de

investigación y publicaciones donde se identificaron diferentes variables y elementos de trabajo a ser considerados en este tipo de experiencias; algunos resultados se detallan en López De Luise & Tabares (2020), López De Luise (2021) y Salina et *al.* (2022). En ese recorrido, destaca una de las experiencias desarrolladas durante 2022 entre comunidades de Chaco y Jujuy (Romero et *al.*, 2023) durante la cual ha emergido una serie de elementos que acontecen sólo en situaciones especiales. El resultado sería una interesante alteración del desarrollo y de la ejecución final de la experiencia. El hito principal fue la situación tecnológica, socioeconómica, cultural y geográfica peculiar de ambas comunidades, que las condiciona fuertemente. Una evaluación posterior permitió determinar que en estos casos es menester un cambio de paradigma, dado que su contexto podría considerarse extremadamente árido para instalar una experiencia de innovación pedagógica como la planteada. Este paradigma STEAM, en contextos extremos (xSTEAM o eXtreme STEAM), es un punto de inflexión en la forma tradicional en la que se ha entendido el desarrollo de las experiencias STEAM y de allí su relevancia, pues requiere considerar variables que circunscriben la experiencia y también se presentan como barreras a las prácticas STEAM tradicionales, pues las condicionan.

El trabajo se desarrolla en el marco del proyecto de investigación denominado "*Encapsulado sistematizado de actividades educativas STEAM*" (CAETI, 2022).

1. INTRODUCCIÓN

En los últimos años, el enfoque STEAM ha cobrado relevancia como estrategia para promover aprendizajes interdisciplinarios integrando ciencia, tecnología, ingeniería, arte y matemática. Sin embargo, su implementación efectiva enfrenta importantes desafíos cuando se traslada a contextos educativos caracterizados por heterogeneidad de trayectorias, restricciones estructurales, tecnológicas, organizacionales o socioculturales, lo que impacta directamente en el diseño, la ejecución y la sostenibilidad de las experiencias educativas. La comunidad en general coincide en que la educación STEAM en contextos rurales y no metropolitanos (también llamados periurbanos) requiere enfoques sensibles al territorio, a la identidad docente y a la participación comunitaria (Starrett et *al.*, 2021; Du & Gan, 2025). También observan que estas condiciones no constituyen meras variaciones contextuales, sino que configuran escenarios de complejidad, donde los enfoques STEAM tradicionales resultan insuficientes (Flanagan et *al.*, 2024; Casado-Mansilla et *al.*, 2023). En este marco, surge el paradigma xSTEAM, que puede definirse como una respuesta posible a este tipo de desafíos. La estrategia propuesta consiste en conceptualizar las experiencias STEAM condicionadas por variables endógenas y exógenas que afectan su ciclo de vida (Romero et *al.* 2023). Por otra parte, es necesario que se definan modelos sistematizados que integren algún tipo de diagnóstico contextual temprano. De esta manera se habilita la toma de decisiones asistida por métricas objetivas incorporando las herramientas necesarias para la evaluación continua del proceso sujeto al contexto.

A diferencia de los enfoques STEM centrados principalmente en el diseño didáctico o en experiencias aisladas, xSTEAM se concibe como una evolución, parte de un sistema educativo-computacional que integra modelos, protocolos formales y tecnologías digitales, que han evolucionado desde antes del 2019, y que se han ido plasmando en una serie de publicaciones sucesivas desde los primeros resultados. Esta consolidación permite diseñar, implementar, evaluar y mejorar experiencias educativas STEAM de manera sistemática, escalable y transferible. Desde esta perspectiva, el enfoque sistémico permite lograr una propuesta única que integra: diagnóstico contextual temprano, adaptación metodológica, encapsulado de actividades reutilizables y soporte tecnológico inteligente.

El presente trabajo tiene como objetivo mostrar el estado actual de una propuesta enfocada en modelar el paradigma xSTEAM como una propuesta educativo-tecnológica; posicionándolo como evolución desde un

paradigma educativo hacia un enfoque sistémico socio-técnico específico en situaciones de bajo grado de aplicabilidad tecnológica.

En particular, el proyecto plantea un modelo específico que se aplicará en el futuro como extensión de la plataforma STEAM del proyecto y será base para otras extensiones que permitan incorporar un sistema inteligente y herramientas orientadas a asistir a los equipos ejecutores en la toma de decisiones y a mejorar la calidad de las experiencias educativas. De esta manera, se pretende alinear innovación educativa, investigación aplicada y desarrollo tecnológico.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El paradigma STEM fue inicialmente planteado por el genio matemático Seymour Papert en los 80 con un juego de programación para niños llamado Lego-Logo (Schulkin, 2018). Lo interesante de este paso fundacional es la concepción misma de la necesidad de desarrollar e incentivar estrategias de pensamiento en los jóvenes. Durante un espacio temporal de aproximadamente una década desarrollar vocaciones científico-tecnológicas se asoció a la educación STEM, con marcado énfasis en las matemáticas, ciencia y tecnología. En 2011, se incluyó el arte, dando lugar a lo que conocemos como STEAM (Agreda Montoro et *al.*, 2016). En 2017, de manera pionera para Argentina, la SCA y el MHS coinciden en iniciar proyectos STEAM en el propio Museo Histórico Sarmiento y en algunos pocos centros escolares de Jujuy, Provincia de Buenos Aires, Ciudad de Buenos Aires y Entre Ríos.

A la fecha, el paradigma inicial sufrió adaptaciones fruto de la experiencia adquirida en la comunidad y, en virtud de la colaboración del centro CAETI de la UAI con la SCA, este trabajo da otro paso en la formalización del paradigma xSTEAM como un sistema educativo-computacional, superando enfoques descriptivos o experiencias aisladas de implementación STEM.

Los resultados preliminares indican que las experiencias STEAM estarían condicionadas por factores hasta ahora no considerados y que se presentan en esta investigación como un conjunto curado de variables endógenas y exógenas. En la presente extensión de la clausura conceptual del modelo se incluyen factores asociados al contexto, los actores, la infraestructura, la organización y la cultura. El modelo y su ejecución se encuentran en proceso de evaluación para establecer la interrelación hallada previamente, como un tipo de repercusión sobre desarrollo, ejecución y ciclo de vida (Salina et *al.*, 2022), desvíos que este proyecto ya identifica como experiencias xSTEAM (ver figura 1). La fenomenología extra debe asimismo determinarse previamente a toda actividad a través de un preciso diagnóstico xSTEAM, esencial para ajustar de forma temprana las estrategias, recursos y ciclo de vida del proceso STEAM maximizando los alcances, reduciendo riesgos de fracaso y mejorando la trazabilidad de los efectos logrados en el entorno. La presente propuesta se enmarca en cuanto a espíritu y forma a la educación STEAM como un modelo interdisciplinar.

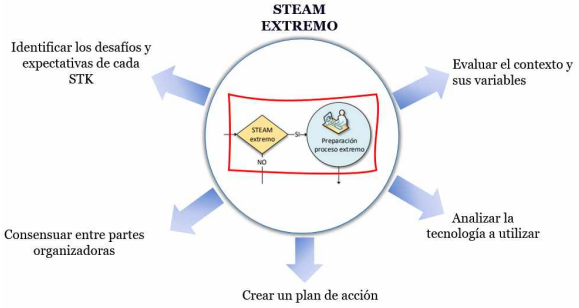


Fig. 1. xSTEAM: perspectivas en consideración.

3. RESULTADOS OBTENIDOS/ESPERADOS

Las experiencias xSTEAM se identifican, en términos generales, por variables endógenas y exógenas que afectan su ciclo de vida. A partir de esta premisa, es posible determinar ciertas métricas que, correctamente evaluadas, permiten definir en qué circunstancias el contexto tiene carácter de extremo riesgo dado el escaso acceso a ciertos recursos o a las características únicas socioculturales del entorno directamente involucrado. A partir de los datos compilados en experiencias STEAM

previas (López De Luise & Tabares, 2020), fue posible determinar las variables esenciales para modelar las características de la entidad, su ecosistema formal e informal y, por lo tanto, evaluar de manera sistemática la aplicabilidad y riesgo de una actividad. En ese camino, la identificación de indicadores apropiados permite determinar si corresponde a un caso xSTEAM, así como su estructura y su perfil. De ahí se puede derivar el diagnóstico objetivo y evaluar los Factores Críticos de Éxito (FCE). Con estos elementos de guía, se pretende maximizar la posibilidad de éxito en la implementación de experiencias STEAM en contextos extremos. Este enfoque posiciona a xSTEAM como una concepción educativa que no solo extiende la educación STEAM, sino que además mantiene las cualidades planteadas inicialmente en el proyecto LINCIEVIS: reproducibilidad, trazabilidad, mejorabilidad continua e inserción con STK científicos, culturales y empresariales solidarios con la formación académica.

En relación con los objetivos actualmente el proyecto ha logrado:

- La revisión y sistematización del estado del arte sobre STEAM y xSTEAM así como la identificación y categorización preliminar de variables endógenas y exógenas que condicionan las experiencias STEAM (Romero et *al.*, 2023),
- La definición del Modelo xSTEAM, estructurado en niveles que permite describir el ciclo de vida de la experiencia y,
- El diseño del protocolo xSTEAM (Romero & López De Luise, 2025).

A la fecha, la tarea se centra en la transformación de las variables detectadas en métricas específicas asociadas con indicadores y se espera poder publicar las estadísticas correspondientes, junto con las posibles mejoras que emerjan del proceso. Esto es parte de la consolidación de los resultados obtenidos mediante la validación empírica del modelo. A partir de allí se espera realizar más aplicaciones del protocolo xSTEAM en diferentes contextos educativos, a fin de evaluar las mejoras en la aplicabilidad, métricas obtenidas de impacto y posibles nuevas métricas para lograr medir la capacidad de adaptación a priori y post-mortem.

Este trabajo realiza las siguientes contribuciones al campo de la educación STEAM y las tecnologías aplicadas:

1) Conceptualiza el paradigma xSTEAM como una evolución del enfoque STEAM tradicional.
2) Propone un modelo formal xSTEAM para describir y analizar el ciclo de vida de experiencias xSTEAM.
3) Define un protocolo xSTEAM orientado a la adaptación de procesos STEAM y a la identificación de FCE, facilitando la repetición, el monitoreo y la mejora continua de las experiencias educativas.
4) Establece las bases para poder integrar a futuro la arquitectura tecnológica basada en la plataforma del proyecto LINCIEVIS, única por su tratamiento encapsulado de actividades, plataformas digitales y la proyección de un sistema inteligente.
5) Posiciona a xSTEAM como una propuesta educativa-tecnológica, superando enfoques descriptivos o experienciales, y habilitando su transferencia, escalabilidad y evaluación sistemática.



Fig. 2. xSTEAM: anatomía de la propuesta.

La figura 2 presenta la anatomía de la propuesta xSTEAM como una arquitectura estratificada. La estructura está diseñada para la escalabilidad y la reutilización, con base en la plataforma actual que ya existe en el proyecto LINCIEVIS, de acceso gratuito y universal desde hace más de 3 años.

4. FORMACIÓN DE RECURSOS

El equipo está conformado por investigadores formados, estudiantes de grado y posgrado de la UAI, UNLP y de la UADER, quienes están desarrollando sus respectivos trabajos finales y tesis. La primera autora de este trabajo se encuentra desarrollando el Doctorado en Informática en la Facultad de Tecnología Informática de la UAI.

5. BIBLIOGRAFÍA

Agreda Montoro, M., Ortiz Colón, A. M., Trujillo Torres, J. M. (2016). Adquisición de competencias STEAM: propuesta didáctica en el Grado de Educación Primaria de las Facultades de Ciencias de la Educación de Jaén y Granada. Tecnología, innovación e investigación en los procesos de enseñanza-aprendizaje. ISBN 978-84-9921-848-9, págs. 23-31.

Casado-Mansilla, D., García-Zubia, J., Cuadros, J., Serrano, V., Fadda, D., & Canivell, V. (2023). *Remote experiments for STEM education and engagement in rural schools: The case of project R3*. https://doi.org/10.1016/j.techsoc.2023.102404

Centro de Altos Estudios en Tecnología Informática (CAETI). (2022). *Proyecto TI/22/105 - Encapsulado sistematizado de actividades educativas STEAM*. https://caeti.uai.edu.ar/investigaci%C3%B3n/l%C3%ADneas-de-investigaci%C3%B3n/proyecto_detalle/?l=0000000017&r=BUE&id=0000002137

Du, J., & Gan, F. (2025). *Servant leadership and citizenship behavior via STEAM education: Community action research on a Chinese charity organization for rural children*. https://doi.org/10.1016/j.ijer.2025.102647

Flanagan, B., Hourigan, M., & Leavy, A. (2024). *Implementation of lesson study in the context of STEM education in a rural setting in Ireland: Challenges and opportunities.* https://doi.org/10.1108/IJLLS-11-2023-0162

Leiva, S. M. (2022). *Evaluación estadística y minería de rendimiento* [Tesis de grado, Universidad Abierta Interamericana]. https://repositorio.uai.edu.ar/items/a06a3772-4bb1-4637-884f-5b48250f74ec

López De Luise, D. (2021). *Museo inmersivo STEAM* [Tesis de grado, Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales].

López De Luise, D. & Ruiz Tabarez, E. A. (2020). *Factores relevantes en la educación STEAM: Desarrollo de métricas y modelos automatizados.* https://repositorio.uai.edu.ar/items/d79f758a-4639-452a-9019-21f7a9625319

López De Luise, D., González González, V., Draier, E., Salina, C., Leiva, S. M., & Ramírez, S. (2022). Experiencias educativas desde la comunidad del museo histórico Sarmiento. Dinámicas de una cultura científico-tecnológica inmersiva. *Revista Avatares de la Comunicación y la Cultura (23)*.

Romero, G. R., & López De Luise, D. (2025). *xSTEAM en acción.* JAIIO, Jornadas Argentinas de Informática, 11(8), 259-268. https://revistas.unlp.edu.ar/JAIIO/article/view/19962

Romero, G. R., Salina, C., Domen, M., López De Luise, D., & Bejarano, M. (2023). *Generación de STEAM en condiciones extremas con xSTEAM.* En Libro de actas – XXIX Congreso Argentino de Ciencias de la Computación (CACIC 2023) (pp. 662–671). Universidad Nacional de La Plata. https://sedici.unlp.edu.ar/handle/10915/165076

Salina, C. V. (2023). *Plataforma digital de cápsulas STEAM: ciclo de vida STEAM, integrando aplicaciones web progresivas*. [Tesis de grado, Universidad Abierta Interamericana]. https://repositorio.uai.edu.ar/handle/123456789/2272

Salina, C., Dome, M., López De Luise, M. D., Rizzo, N. & Romero, G. R. (2024). *PWA para potenciar una plataforma STEAM*. En el XXX Congreso Argentino de Ciencias de la Computación (CACIC) (La Plata, Argentina, 7–11 de octubre de 2024). https://sedici.unlp.edu.ar/handle/10915/176338

Salina, C., López De Luise, D., Leiva, S., Ramírez, S., & Romero, G. R. (2022). *A life-cycle to improve STEAM virtual museum activities exposition*. En 10th International Workshop on Soft Computing Applications (Arad, Romania, 21–23 November 2022).

Starrett, A., Yow, J., Lotter, C., Irvin, M. J., & Adams, P. (2021). *Teachers connecting with rural students and places: A mixed methods analysis*.https://doi.org/10.1016/j.tate.2020.103231

Schulkin, J. (2018). *Qué es la educación STEAM: 5 ideas para transmitir a los niños*. Otras Voces en Educación. https://otrasvoceseneducacion.org/archivos/278981

Temas Interdisciplinarios: Esports

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

Desarrollo del ecosistema de Esports en la universidad: experiencias de formación, investigación y extensión en la Universidad Abierta Interamericana

Pablo Vilaboa, David Curras, Juan Silva

Universidad Abierta Interamericana. Facultad de Tecnología Informática.
Centro de Altos Estudios en Tecnología Informática. Buenos Aires. Argentina

{pablo.vilaboa; david.curras; juanignacio.silva}@uai.edu.ar

RESUMEN

Este trabajo de investigación analiza la integración de los esports y la industria de los videojuegos en la educación superior como catalizadores de innovación educativa. El proyecto aborda el fenómeno desde una perspectiva integral que trasciende la formación teórica, articulando dimensiones tecnológicas, organizacionales y sociales para fortalecer la calidad académica y la identidad institucional. A través de un enfoque documental y aplicado, con un diseño metodológico no experimental y transversal, la investigación combina el análisis cualitativo de experiencias estudiantiles con métricas cuantitativas de participación y crecimiento. El estudio parte de una revisión de la literatura científica global para contrastarla con las iniciativas pioneras desarrolladas en la UAI, tales como la organización de competencias y la vinculación con actores clave del sector tecnológico. Los resultados esperados incluyen la sistematización de un modelo de gestión de ecosistemas de esports universitarios que favorezca la retención estudiantil y el desarrollo de competencias digitales. Como aporte final, la investigación propone la creación de un decálogo de lineamientos estratégicos replicable en otras instituciones de la región. Este marco conceptual busca reducir la brecha documental existente en América Latina, posicionando a los deportes electrónicos no solo como entretenimiento, sino como un espacio interdisciplinario de aprendizaje activo, profesionalización y experimentación científica en la Sociedad del Conocimiento.

Palabras claves: Esports universitarios, Innovación educativa, Ecosistemas digitales,

CONTEXTO

Este proyecto se suscribe en el marco de las iniciativas de innovación educativas y desarrollo tecnológico impulsadas por en la UAI, orientado a fortalecer la calidad académica mediante la integración de nuevas prácticas vinculadas a la industria de videojuegos y los deportes electrónicos. La propuesta articula los avances en videojuegos, gestión de comunidades y organización de competencias con las necesidades institucionales relacionadas con la formación de estudiantes en entornos tecnológicos emergentes.

Por otro lado, con el propósito de integrar de manera integral el fenómeno de los esports en el ámbito universitario, se consideran recursos tecnológicos, organizacionales, y sociales ampliando los enfoques tradicionales centrados exclusivamente en la formación teórica. En este sentido, se busca aprovechar las oportunidades que ofrecen los esports como espacio de aprendizaje activo, desarrollo de competencias, fortalecimiento social y la exigencia en competencias digitales [13].

1. INTRODUCCIÓN

Los esport han experimentado un crecimiento en los últimos años, impulsados por la expansión de internet de alta velocidad, la popularidad entre jóvenes y la profesionalización de los torneos [1][2].

Los esport, definidos como la práctica competitiva de videojuegos a nivel profesional y amateur, se han posicionada como un campo interdisciplinario de investigación académica en los últimos años. Revisiones sistemáticas como la de Reitman et al, destaca el crecimiento en los trabajos científicos publicados a la fecha, abarcando temáticas como psicología, negocios, derecho y ciencias del deporte, como un crecimiento importante luego de la pandemia [3]. Aunque el crecimiento de los trabajos es considerable, aun se observa brechas importantes en metodologías estandarizadas y estudios transversales [4]

En esta línea, la evolución histórica de la investigación inicia formalmente cerca de 2002, con trabajos pioneros en economía del juego, psicología y dinámica de grupos [5]. Los ecosistemas de esport, se identifican como redes multifactoriales formada por jugadores, organizadores, patrocinadores y plataformas de streaming. En este sentido, también propone una descripción del fenómeno más estructural vinculándolo con la infraestructura tecnológica y de gobernanza [6]

Respecto a la educación superior, los ecosistemas universitarios integran aspectos relacionadas con la formación, la competencia y la investigación [7]. Por otro lado, autores como Anderson et al proponen modelo que podrían complementar el conjunto de acciones vinculadas a los índices de retención estudiantil y la representatividad institucional [8]

En este contexto, diversas instituciones han comenzado a explorar su potencial como herramienta para promover el pensamiento crítico y la formación tecnología [9]. En América latina, y particularmente en argentina, la integración de los esports en el ámbito académico se encuentra en una etapa incipiente. Si bien existen algunas iniciativas, todavía es limitada la documentación formal que sistematice la experiencia del deporte electrónico en el nivel superior [7]

Asimismo, no solo la experiencia, es uno de los factores más importante de este modelo. En este sentido, es también un medio importante para trabajar en el fortalecimiento de los insumos para mejorar el índice de retención estudiantil. Aunque hasta el momento, es un aspecto no profundizado desde la bibliografía científica, se relevan experiencias de universidades donde destacan el impacto positivo de los deportes y los vinculan con el éxito estudiantil y el rendimiento académico [10][11][12]

En este marco, La Universidad Abierta Interamericana ha desarrollado durante los últimos años una serie de iniciativas orientadas a motivar y promover el ecosistema esport dentro de la comunidad universitaria. Estas actividades, han incluido la organización de eventos, torneos y una importante cantidad de actividades académicas y tecnológicas de capacitación. Con estas acciones, se busca generar espacios de formación y la articulación con diferentes actores de la industria de los videojuegos y el entretenimiento digital. Asimismo, se promueve la creación de espacios interdisciplinarios de aprendizaje, y experimentación sostenida por la investigación formal,

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto se radica en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la UAI y se vincula con la línea de investigación Sociedad del conocimiento y tecnologías aplicadas a la educación adopta un enfoque documental, de tipo aplicado, dado que busca identificar y promover el aporte de este contexto a la mejora constante de competencias y habilidades educativas, al mismo tiempo, desarrollar un modelo con aplicaciones prácticas en contextos académicos universitarios.

Para el desarrollo del proyecto, se adopta un enfoque práctico, sostenido en análisis estadístico y cuantitativo, con el fin de construir y evaluar un modelo de trabajo capaz de favorecer la adaptación universitaria y la integración social. El diseño de la metodología es no experimental, y transversal, basado en el análisis cualitativo, partiendo de la recolección

de experiencias de los estudiantes y su impacto en el fortalecimiento de la identidad institucional.

3. RESULTADOS OBTENIDOS / ESPERADOS

Los objetivos del presente trabajo se orientan a fundamentar la relevancia de desarrollar, implementar y analizar un enfoque integral para la construcción de ecosistemas de deportes electrónicos en la universidad, que articule dimensiones académicas, tecnológicas, y comunitarias. De este modo, se busca no solo estandarizar las experiencias desarrolladas, sino también crear un marco conceptual que permita comprender el impacto de la propuesta en la formación de alumnos, el lugar de la institución y el vínculo con la industria.

Para lograr este propósito, el proyecto propone realización de un relevamiento y luego el análisis de la literatura científica sobre deportes electrónicos en el nivel superior, con un claro énfasis en los modelos de implementación, experiencias universitarias y los enfoques interdisciplinarios. Luego de esta revisión, se evaluarán las buenas prácticas y lineamientos que demuestren y generen resultados significativos en contextos académicos.

Luego, se trabajará en la sistematización de las actividades desarrolladas en la Universidad Abierta Interamericana, integrando información relativa a la formación del estudiante, la participación en competencias, y la vinculación con actores del ecosistema de videojuegos y esports. Como complemento, se diseñarán y aplicarán rubricas de evaluación que permitan medir el impacto de los esports en los procesos de formación y sistematización institucional [14]. Este proceso posibilitará estructurar un modelo de análisis basado en dimensiones concretas para poder comprender el fenómeno en contextos universitarios.

Además, se analizará el impacto de estas iniciativas en términos cuantitativos, con variables que midan la participación, el crecimiento de la comunidad y posicionamiento institucional en el sector competitivo. De manera complementaria, se analizarán herramientas de análisis y visualización que permitan presentar de forma clara la evolución y alcance de las actividades implementadas.

Por último, al recolectar los resultados obtenidos, se diseñará un decálogo de lineamientos orientados al diseño y fortalecimiento de programas de esports en la educación superior, con el objetivo de sugerir un modelo replicable y transferible que pueda ser adoptado por otras universidades que presenten interés en integrar estas prácticas en sus estrategias de formación académica.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente trabajo se desarrolla en el marco de del proyecto interdisciplinario en la facultad de tecnología informática de la UAI y es llevado adelante bajo la dirección del director de laboratorio *de Desarrollo de Software y Servicios a Alumnos*: Federico Peralta (alumno), Francisco Chappani (alumno), Franco Gualdrini (alumno), Nahuel Pardal (alumno) y Belén Herrera (graduada)

BIBLIOGRAFÍA

[1] ASO World. (2025). Global Esports Market Report (2025). Recuperado de https://asoworld.com/blog/global-esports-market-report-2025/.

[2] Pierre-Jean Benghozi, Polytechnique Insights, oct. 2025. https://www.polytechnique-insights.com/en/columns/digital/esports-everything-you-need-to-know-about-this-exploding-digital-market (accedido el 18 de marzo 2026).

[3] Reitman, J. G., Anderson-Coto, M. J., Wu, M., Lee, J. S., & Steinkuehler, C. (2020). Esports research: A literature review. *Games and Culture, 15*(5), 498-

528. https://doi.org/10.1177/1555412019840892

[4] Wagner, M. G. (2006). On the scientific relevance of eSports. In International Conference on Internet Computing & Conference on Computer Games Development (pp. 437–442). CSREA Press

[5] Pedraza-Ramírez, I., Sharpe, B., Behnke, M., Toth, A. J., & Poulus, D. R. (2025). The psychology of esports: Trends, challenges, and future directions. *Psychology of Sport and Exercise, 78*, Article 102967. https://doi.org/10.1016/j.psychsport.2025.102967

[6] Taylor, T. L. (2012). *Raising the stakes: E-sports and the professionalization of computer gaming*. MIT Press. https://doi.org/10.7551/mitpress/8624.001.0001

[7] Jenny, S. E., Bowman, M. M., & Bingham, T. (2021). An inventory and analysis of global higher education esports programs. *International Journal of Esports, 1*, Article 59. https://www.ijesports.org/article/59/html

[8] Anderson, C. G., Tsaasan, A. M., Reitman, J., Lee, J. S., Wu, M., Steel, H., Turner, T., & Steinkuehler, C. (2018). *Understanding esports as a STEM career ready curriculum in the wild*. En *Proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)* (pp. 1-9). IEEE. DOI: 10.1109/VS-Games.2018.8493445

[9] Hernández-Martín, A., López-Fernández, J., Addolorato, S., & Díaz-Benito, V. J. (2024). Realización de un torneo de esports: aprender basándose en proyectos. In *Construyendo el futuro de la educación superior en la era digital* (pp. 61-70). Dykinson.

[10] ounews. (2025). OU Collaborates on Groundbreaking National Study on Collegiate Esports. Recuperado de https://www.ou.edu/news/articles/2025/october/ou-collaborates-on-groundbreaking-national-study-on-collegiate-esports

[11] Gomez, R Consejo Superior de Deportes (CSD). (2018). *Estudio sobre los modelos de deporte universitario de Estados Unidos, Canadá y Australia*. https://www.csd.gob.es/sites/default/files/media/files/2018-09/Texto%20del%20estudio%20en%20PDF....pd

[12] Olarte Aldana, F. A. ., y Medina Cervantes, O. . (2021). PROGRAMA DEPORTIVO ENFOCADO A LOS DEPORTES ELECTRONICOS EN LA UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS. *Boletín Semillas Ambientales*, *15*(1). https://revistas.udistrital.edu.co/index.php/bsa/article/view/18896

[13] B. T. Sharpe and D. R. Poulus, The Psychology of Esports Performance. London: Routledge, 2026. doi: https://doi.org/10.4324/9781003663676.

[14] Neil, C., Battaglia, N., De Vincenzi, M. (2024). Rubric-Driven Competency Development: A Case Study. In: Pesado, P., Panessi, W., Fernández, J.M. (eds)

Computer Science – CACIC 2023. CACIC 2023. Communications in Computer and Information Science, vol 2123. Springer, Cham. https://doi.org/10.1007/978-3-031-62245-8_24

Diplomatura en Deportes Electrónicos UNLP

Cesar Estrebou [(1)], Santiago Medina[(1,2)], Federico Ruera[(2)], Patricia Pesado[(1)], Armando De Giusti[(1)]

[(1)] Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 - La Plata, Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 - La Plata, Buenos Aires
[(2)] Proyecto Esports UNLP
(cesarest, smedina, fruera, ppesado, degiusti) @lidi.info.unlp.edu.ar

RESUMEN

Los “deportes electrónicos” son básicamente competiciones de videojuegos, que se realizan en forma individual o por equipos. Su crecimiento ha sido explosivo en los últimos años y se ha discutido su inclusión en los Juegos Olímpicos.
En este contexto hay mucho interés de los jóvenes por participar en todas las actividades vinculadas con los Esports (desde el desarrollo de los videojuegos mismos hasta la gestión de torneos o la conformación y entrenamiento de equipos que compiten).
El desarrollo de la actividad (en Argentina y en el mundo) indica posibilidades laborales en diferentes áreas disciplinarias asociadas con los Esports.

La Diplomatura en Deportes Electrónicos de la Facultad de Informática de la Universidad Nacional de La Plata es el instrumento de formación académica del Proyecto Esports UNLP que se desarrolla en el III-LIDI de la Facultad de Informática de la UNLP.
El plan de estudios de la diplomatura permite integrar contenidos que abarcan las diferentes áreas de los Esports, de modo de abrir un marco de posibilidades para alumnos universitarios y pre-universitarios interesados en el tema y que motiva el desarrollo de vocaciones en las disciplinas que conforman el ecosistema.
Tiene su foco principal aplicado en la interdisciplinaridad y la interacción de estudiantes que están cursando las distintas carreras de la UNLP.
Los estudiantes de la carrera interactúan en el Centro de Entrenamiento en Deportes Electrónicos, espacio destinado a la práctica de deportes electrónicos, con los planteles oficiales de la Universidad, que compiten en torneos nacionales e internacionales de videojuegos.
Desde la Facultad de Informática se entiende que la carrera es un elemento motivador para desarrollar estudios universitarios en diferentes disciplinas (en particular Informática) asociadas con los Deportes Electrónicos.

Palabras clave: deportes electrónicos, interdisciplina, videojuegos, diplomatura, formación.

CONTEXTO

La Universidad Nacional de La Plata ha iniciado en 2021 el Proyecto Esports UNLP que se desarrolla en el Instituto de Investigación en Informática (III-LIDI) de la Facultad de Informática, fomentando la formación en temas relacionados a los deportes electrónicos por sus alumnos, docentes y nodocentes. En este Proyecto colaboran alumnos/especialistas de diferentes disciplinas (Informática, Ciencias Económicas, Comunicación, Psicología, etc).
En este marco, en 2023 se lanzó la Diplomatura en Deportes Electrónicos. Asimismo, el Proyecto cuenta con un Centro de Entrenamiento instalado en el Centro de Innovación y Transferencia Tecnológica de la Facultad de Informática de la UNLP.

1. INTRODUCCIÓN

Los “deportes electrónicos” son básicamente competiciones de videojuegos, que se realizan en forma individual o por equipos.
En un breve recorrido, los primeros torneos profesionales se pueden rastrear a partir de la

década del 80, aunque recién en los últimos años cobraron gran popularidad, en parte gracias a las plataformas de streaming.
Desde un punto de vista deportivo, en 2018 se incluyeron a los deportes electrónicos como un deporte de demostración en los Asian Games, y en la edición del 2022 pasaron a ser deporte oficial. Incluso Japón pretendió que se los incluya en los Juegos Olímpicos 2021. Además, en Estados Unidos y China se ofrecen becas para universidades que cuentan con sus propias ligas nacionales universitarias.
Por otra parte, también se llevan a cabo campeonatos nacionales, como es el caso del torneo organizado por la Asociación de Fútbol Argentino (AFA), la Bundesliga, la Liga de España y la Premier League. Todas cuentan con sus versiones de torneos de esports, donde participan la mayoría de sus equipos.

El interés de la Facultad de Informática, desde el III-LIDI ha sido no sólo "utilizar" juegos y plataformas para videojuegos y competencias, sino también desarrollar capacidades para generar los mismos, disponer de una infraestructura propia para torneos y gestión de equipos, así como investigar en temas asociados tales como Realidad Extendida, análisis de datos o Inteligencia Artificial aplicada a las estrategias de diferentes juegos de Esports.
Por otra parte, se busca generar una oferta académica que resulte un elemento motivador para vocaciones universitarias (en particular en Informática) y también para presentar un área de posible salida laboral a estudiantes de diferentes disciplinas.
En ese contexto la Facultad de Informática implementó la Diplomatura en Deportes Electrónicos, que está destinada a quienes hayan finalizado la escuela secundaria y de forma particular a estudiantes de la Universidad Nacional de La Plata, tomando como base la Ordenanza 290/16 de la UNLP referida a las Diplomaturas y la Ordenanza 334/22 de la Facultad de Informática de la UNLP "REGLAMENTO PARA LAS DIPLOMATURAS EN LA FACULTAD DE INFORMÁTICA".
Esta formación busca que el diplomado tenga una visión del ecosistema de los deportes electrónicos, es decir, sobre los jugadores, equipos, organizadores, comunicadores, entrenadores, publicidad, entre otros aspectos. Además, que conozca los fundamentos tecnológicos de soporte para los Esports.
Se combinan contenidos tecnológicos de impacto en los Esports que son conocidos y tratados en la Facultad, en particular en el III-LIDI con contenidos relacionados a la gestión de proyectos de Esports (Gestión/ Psicología/ Marketing/ Comunicación/ Dirección de proyectos de Esports) los cuales caracterizan varias de las carreras existentes en el país.
La Diplomatura adoptó la modalidad "a distancia con instancias presenciales", tratándose de un programa que se pueda hacer "a distancia sincrónica" combinado con actividades presenciales, de modo de utilizar el equipamiento y casos prácticos de Esports en el CIyTT de la Facultad de Informática. Para su desarrollo se utiliza la legislación y recursos de educación a distancia de la Facultad de Informática. Las actividades se realizan en el entorno de enseñanza aprendizaje IDEAS de la Facultad, con tutores para atención de los alumnos y posibilidad de consultas presenciales atendidas por el equipo del proyecto Esports. La iniciativa contempla que se realicen actividades concretas en el ámbito del proyecto.

2. PLAN DE FORMACIÓN

El plan de formación de la diplomatura consta de 8 módulos orientados a cubrir cada uno de los aspectos importantes en el armado, organización y desarrollo de un proyecto Esports**.**
Como cierre de la cursada se debe cumplir la aprobación de una Práctica Profesional Supervisada con diferentes modalidades.
Asimismo, los estudiantes cuentan con el Centro de Entrenamiento en Deportes Electrónicos, instalado en el Centro de Innovación y Transferencia Tecnológica de la Facultad de Informática de la UNLP desde 2022. Este espacio, está destinado a la práctica de los planteles oficiales de la Universidad, que compiten en torneos nacionales e internacionales de deportes electrónicos, al desarrollo de torneos y al desarrollo de prácticas de los estudiantes de la Diplomatura.

Plan de Estudios de la Diplomatura en Deportes Electrónicos

Espacios Curriculares	
Introducción. Aspectos generales de los Esports	**20hs.**
Tecnología asociada con Esports	**18hs.**
Proyectos de Esports	**18hs.**
Dirección de Proyectos de Esports	**20hs.**
Aspectos tecnológicos en Proyectos de Esports	**20hs.**
Informática y Esports	**20hs.**
Aspectos de Psicología y Comunicación	**18hs.**
Aspectos Complementarios de Informática y Tecnología	**19hs.**
Práctica Profesional Supervisada	**80hs.**
TOTAL CARGA HORARIA	**232hs.**

1- Introducción. Aspectos generales de los Esports.

Este módulo es de formación general y pretende que el alumno tenga un panorama de todo el espectro que abarca los deportes electrónicos. A partir de la definición de videojuegos a los Esports, considerando el tema de la industria vinculada con los Esports y los modelos de gestión/negocios, así como los desafíos actuales del sector (tanto tecnológicos como legales y de administración).

2- Tecnología asociada con Esports

Presenta la infraestructura básica que se requiere para el desarrollo de Esports. Se desarrollan los conceptos de servidores locales y en la nube, con sus ventajas y desventajas.

Expone los temas de comunicación, streaming y trasmisión en tiempo real, reflexionando sobre los tiempos de respuesta en función de los usuarios simultáneos que estén conectados a un juego.

Detalla el equipamiento especial que se emplea en los videojuegos asociados con Esports.

3- Proyectos de Esports.

Se plantean los conceptos básicos de un proyecto que se enfoca en un producto, tal como pueden ser los torneos de Esports, poniendo énfasis en la planificación de los mismos. Expone las nociones de costos/presupuestos/análisis financiero y valorización de los proyectos y productos. Relaciona los Esports con la gestión emprendedora y analiza los riesgos y beneficios potenciales del negocio.

4- Dirección de Proyectos de Esports

Se presentan el modelo y actividad de un proyecto de negocio. El caso de los proyectos de Esports. Analiza el plan de operaciones, en particular el Plan de Marketing. Discute la conformación del equipo de recursos humanos y los criterios para su elección. Expone los conceptos de Plan de Inversiones y el marco legal de un proyecto de Esports.

5- Aspectos tecnológicos en Proyectos de Esports

Expone aspectos centrales en la elaboración de videojuegos (desde el punto de vista tecnológico): visualización / manejo de imágenes / ambientes inmersivos y combinaciones de realidad virtual y realidad aumentada.

Plantea el manejo de los datos, la seguridad de los mismos y la recolección de datos para análisis de torneos y también de estrategias de juegos.

6- Informática y Esports

Analiza los aspectos informáticos en el diseño y desarrollo de videojuegos, en particular su programación.

Relaciona juegos serios con videojuegos y discutir su aplicación en Educación. Discutir casos de estudio concretos.

7- Aspectos de Psicología y Comunicación.

Analiza aspectos significativos de psicología organizacional en Esports. Relaciona con la psicología deportiva en general y el entrenamiento de los jugadores.

Presenta aspectos de la comunicación formal e informal. Los relaciona con el plan de comunicación de un torneo de Esports.
Discute el empleo de las redes sociales y los canales de distribución de contenidos en general, así como los planes de comunicación desde la gestión de un proyecto de Esports.

8- Aspectos Complementarios de Informática y Tecnología

Presenta temas de importancia para perfeccionar aplicaciones de Esports desde la tecnología informática: el tratamiento de grandes volúmenes de datos (Big Data) así como el perfeccionamiento de estrategias de juego a partir de algoritmos de Inteligencia Artificial.
Expone el concepto de "sistema inteligente / experto" y como incorporar estos criterios a un juego individual o colectivo.
Presenta casos de estudio concretos, para comprender la importancia de la Inteligencia Artificial en la temática de los Esports.

3- PRACTICA PROFESIONAL SUPERVISADA

La Práctica Profesional Supervisada tiene 3 modalidades para su aprobación:

- Trabajo de investigación relacionado con temas de esports (profundizando un tema o abordando alguno no incluido en la currícula).
- Trabajo Práctico Integrador
- Talleres de formación complementaria

La idea principal de la PPS es vincular al alumno de la Diplomatura con el ámbito de desarrollo del proyecto que es un modo de integrarlos a la realidad de los Esports, tanto desde el punto de vista de las competencias como de la tecnología o la gestión de un proyecto de Esport.

1- Trabajo de Investigación

Para el desarrollo del trabajo se define un tutor docente y un tema que puede ser propuesto por el alumno o elegido de la lista de temas propuestos desde la coordinación, como por ejemplo:

- Análisis informático de videojuegos.
- Análisis de Juegos Educativos.
- Aplicaciones móviles orientadas a juegos educativos.
- Estrategias inteligentes orientadas a deportes electrónicos.
- Servidores y comunicaciones en sistemas de deportes electrónicos.
- Limitaciones de la tecnología en Esports.
- Análisis de la estructura deportiva de un equipo competitivo.
- Análisis comparativo de juegos para Esports.
- Planificación de Torneos (Financiera / Marketing).
- Herramientas de gestión organizacional.
- Estrategias de comunicación externa.
- Análisis del desarrollo de Esports en Argentina y el mundo.

La aprobación de la PPS en esta modalidad contará con una breve síntesis escrita y un coloquio técnico.
Estos trabajos pueden ser realizadas en forma individual o en grupo de hasta 3 alumnos, para realizar un estudio complementario / estudio de tecnologías / análisis comparativo de alternativas / investigación y/o desarrollo dentro del proyecto de Esports UNLP, etc.

2- Trabajo Práctico Integrador

Esta modalidad permite al alumno aprobar su PPS progresivamente mientras va cursando cada módulo.
Se plantea un enunciado general relacionado con el ecosistema de Esports y sus competencias/torneos, que el alumno tiene que analizar con el enfoque de cada módulo.
Esta modalidad tiene la intención de desafiar a los estudiantes con la aplicación de los conceptos de cada tema sobre un mismo problema transversal.

3- Talleres de formación Complementaria

A partir del año 2025 con la cuarta cohorte se presentó una nueva modalidad que propone la posibilidad de cursar y aprobar al menos 3 talleres de formación complementaria.
Estos talleres son dictados por egresados y docentes de la diplomatura con la intención de agregar contenido a la formación de los estudiantes. Por ejemplo:

- Taller de Streaming
- Gestión de Redes Sociales
- La interdisciplinaridad en el desarrollo de video juegos
- Herramientas de Inteligencia Artificial
- Diseño de juegos competitivos.

4. RESULTADOS ESPERADOS/OBTENIDOS

La diplomatura tuvo hasta el momento cuatro cohortes, en cada inscripción se recibieron más de 150 interesados. Las dos primeras cohortes se dictaron durante 2023, la tercera se dictó en 2024 y la cuarta en 2025.
Al momento han egresado 131 estudiantes.

Los objetivos para 2026 son:

- Dictado de la 5ta cohorte de la diplomatura.
- Agregado de 2 talleres de formación completaría con temáticas nuevas.
- Generación de un repositorio con los trabajos monográficos entregados durante todas las cohortes.
- Generación de bibliografía propia.

Los alumnos (que cursan carreras en distintas facultades de la UNLP) han logrado:

- Tener una visión clara del ecosistema de deportes electrónicos: jugadores, equipos, organizadores, comunicadores, entrenadores, publicidad, etc.
- Conocer los fundamentos tecnológicos de soporte para la actividad de Esports, en particular la tecnología para el desarrollo de videojuegos y su gestión y perfeccionamiento.
- Incorporar herramientas para desenvolverse en el entorno de los deportes electrónicos de diferentes áreas del conocimiento aplicadas a los Esports.
- Realizar una práctica concreta en el ámbito del proyecto de Esports de la UNLP, en los temas de su interés.

5. FORMACION DE RECURSOS HUMANOS

Han egresado de la Diplomatura 131 estudiantes En torno a la cursada de cada cohorte se fueron dando diferentes experiencias, colaboraciones y participaciones de los alumnos en el proyecto Esports y la diplomatura.
Al menos 5 egresados han participado del dictado de los módulos y los talleres complementarios dándole una entidad mayor al diploma obtenido y una motivación extra a los estudiantes.
El trabajo interdisciplinario de los docentes e integrantes del proyecto esports permite la participación de los alumnos en la organización de torneos, el entrenamiento de equipos, las actividades de comunicación y difusión. Esto complementa sus estudios y les da el marco adecuado para que puedan extender su formación.

6. BIBLIOGRAFÍA

- Sitio web Proyecto Esports UNLP. https://esports.info.unlp.edu.ar/
- Sitio web Diplomatura en Deportes electrónicos. https://www.info.unlp.edu.ar/deportes-electronicos/
- Estrebou, C.(2025) “Deportes electrónicos: un tema para despertar vocaciones a la informática”. Cesar Estrebou, Santiago Medina, Federico Arteca, Federico Ruera. Libro de actas XX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2025). Págs. 441-444. ISBN: 978-987-8994-42-0. Junio 2025. https://sedici.unlp.edu.ar/handle/10915/184997
- Ordenanza 334/22 de la Facultad de Informática de la UNLP “REGLAMENTO PARA LAS DIPLOMATURAS EN LA FACULTAD DE INFORMÁTICA”.
- Antón, M. (2018). “Los deportes electrónicos: el espectáculo en las competiciones de videojuegos” [Tesis de Doctorado, Universidad Complutense de Madrid]. E-Prints Complutense. https://eprints.ucm.es/id/eprint/54986/1/T41009.pdf
- AEVI - Spanish Association of Videogames. (2020). “Los esports en España: situación actual y posición de la industria. Asociación Española de Videojuegos.” http://www.aevi.org.es/web/wp-content/uploads/2020/12/Informe_esports_ESP_20.pdf
- Carrillo, J. A. (2015). “La dimensión social de los videojuegos' online': de las comunidades de jugadores a los 'esports'.” Index. Comunicación: Revista científica en el ámbito de la Comunicación Aplicada, 5(1), 39-51.
- Taylor, T. L. (2012). “Raising the Stakes: E-sports and the Professionalization of Computer Gaming.” Cambridge: MA: MIT Press
- Gutierrez Gonzalez Ángel. “Tecnologías de la Información - Un enfoque interdisciplinario” Alfaomega Grupo Editor Argentino
- Austerb David. “La Tecnología del Streaming de Vídeo y Audio” Editorial Escuela de Cine y Vídeo de Andoaín, S.L.

- Angel Arias. “Computación en la Nube” Editorial: CreateSpace Independent Publishing Platform; 2nd edición
- Guillén Muñoz, Á. (2020). “La planificación estratégica en equipos de esports”. Universidad de Valladolid. https://uvadoc.uva.es/bitstream/handle/10324/43529/TFG-N.%201476.pdf
- Pérez-Sagarzazu, G. (2019).” Introducción de los deportes electrónicos en la Universidad y diseño de su estructura organizativa” UNAV Esports.
- Nielsen Esports (2022). “The ROI of Esports.” The Nielsen Company.
- Scholz & Barlow. (2019). “Esports is Business.” Springer International Publishing.
- Cáliz Ortiz, J. S., & Paredes Solano, A. D. (2020).”Cómo facilitar el proceso de reclutamiento de aspirantes latinoamericanos a equipos de esports.” Biblioteca Digital Icesi.
- Cristòfol, F., Martínez-Ruiz, A., Román, I. & Rodríguez, C. (2020). “Evolución de las estrategias de patrocinio en los esports en España: 2013-2021.” Ámbitos: Revista Internacional de Comunicación 48, pp. 188-204. Es
- Fanjul, C., González, C., & Peña, P. J. (2019). “La influencia de los jugadores de videojuegos online en las estrategias publicitarias de las marcas: comparativa entre paña y Corea.” Comunicar, 58.
- Nallar, D. (2016). “Diseño de juegos en América Latina II: Diseño y narrativa transmedia.”
- Observatorio de Industria de Videojuegos (2021). Resumen Anual 2021. UNRaf. https://www.unraf.edu.ar/images/INVESTIGACION/OBSERVATORIO_VideoJuegos/PDFinformes/Observatorio_Videojuegos_UNRaf_2021.pdf
- “Aprende a crear tus propios videojuegos con Unity 3D Volumen 1: Introducción práctica a Unity 3D” (Spanish Edition) Edición Kindle. Editorial TRINIT Asoc. Informáticos de Zaragoza. Año 2018.
- Luis Ruelas. “Desarrollo de Videojuegos y Realidad Virtual con Unity 3D.” Editorial Ra-Ma. ISBN 978-84-18551-90-1
- Jose M. Martinez, Fernando Navarro, Antonio Martinez · 2018. “Realidad Virtual y Realidad Aumentada.” Editorial Ra-Ma. ISBN13 9788499647395
- Cabero Almenara, J. y Jiménez García, F. “Realidad Aumentada. Tecnología para la Formación.” Editorial Síntesis. 2016. ISBN: 978-84-9077-258-4
- Documentación en Videojuegos: Documento de diseño (GDD) en https://eldocumentalistaudiovisual.com/2015/02/06/documentacion-en-videojuegos-documento-dediseno-gdd/ RA-MA S.A. Editorial y Publicaciones
- David Ballejo Fernández , Cleto Martín Angelina. “Desarrollo de Videojuegos 1: Arquitectura del Motor de Videojuegos” “Desarrollo de Videojuegos 2: Programación Gráfica” “Desarrollo de Videojuegos 3: Técnicas Avanzadas” “Desarrollo de Videojuegos 4: Desarrollo de Componentes”. Editados por la Universidad de Castilla La Mancha
- Ana Belén Gomez Sanz . “Gamificación y Juegos Serios. Curso Práctico” RA-MA S.A. Editorial y Publicaciones
- Ricardo Casañ Pitarch. “Videojuegos Serios en educación” Editor: Universitat Politècnica de València
- Antón, M; García, F (2014) “Deportes electrónicos. Una aproximación a las posibilidades comunicativas de un mercado emergente.” Questiones publicitarias, 1, pp.98-115
- Mendoza, G. (2019). “El trabajo psicológico online con equipos de esports: experiencias en Vodafone Giants”. Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 4(1), 1-7.
- Pedraza-Ramirez, I. (2019). “Generación LoL: entrenamiento psicológico mediante una propuesta holística con un equipo profesional de Esports.” Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 1-10.
- Vaamonde, A. G. N., Toribio, M. J., Molero, B. T., & Suárez, A. (2018). “Beneficios cognitivos, psicológicos y personales del uso de los videojuegos y esports: una revisión” Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 3(2), 1-14.
- Pablo L. Gutierrez Utande – Angel Arroyo Castillo (Tutor) “Aplicación de Inteligencia Artificial en Videojuegos” Universidad Politécnica de Madrid – (2017)
- Bernard Marr - “BIG DATA: La utilización del Big Data, el análisis y los parámetros SMART para tomar mejores decisiones y aumentar el rendimiento” Editorial TEELL - 2016
- Eric Sadin - “La Inteligencia Artificial o el Desafío del siglo” Colección Futuros Próximos – 2020

Experiencia del Taller de e-Sports de la Universidad Nacional de Hurlingham

Germán Beneditto[1][0009-0006-1085-6760], Marisa Panizzi[1] [0000-0002-9699-1544], Fernando Puricelli[1] [0009-0004-5410-4611]

[1] Instituto de Tecnología e Ingeniería, Universidad Nacional de Hurlingham, Av. Vergara 2222 (B1688GEZ), Villa Tesei - Bs. As. Argentina

Esports@unahur.edu.ar, Marisa.panizzi@unahur.edu.ar, Fernando.puricelli@unahur.edu.ar

Resumen

En este artículo se presenta la iniciativa del Taller de e-Sports que se imparte en la Universidad Nacional de Hurlingham. Se detallarán los objetivos generales y específicos delineados en esta propuesta, así como también los dos ejes de contenidos sobre los que se sustenta este taller. El objetivo del taller no es únicamente mejorar las habilidades técnicas, sino también fomentar competencias sociales, el pensamiento estratégico colectivo y la adquisición de herramientas para manejar la frustración y el estrés, consolidando un entorno de juego positivo y constructivo. Además, se presentan algunas conclusiones del dictado de este en años anteriores, desde el 2023.

Palabras clave: e-Sports, psicología del deporte, videojuegos, salud mental, bienestar, UNAHUR.

Contexto

En la actualidad, dentro de la Dirección de Deportes de la Universidad Nacional de Hurlingham, bajo la órbita de la Secretaría de Servicios a la Comunidad se consolidó la iniciativa de crear un Taller de e-Sports que estuviera en línea con el resto de los Talleres de Deportes, donde se busca generar un espacio de interés comunitario, colaborativo y de aprendizaje. Esta propuesta tiene como finalidad evolucionar la experiencia desarrollada durante el año anterior, articulando un espacio formativo y recreativo que no solo responde al creciente interés del alumnado, sino que también se alinea con los objetivos estratégicos de la Universidad de fomentar la integración y el bienestar estudiantil.

Introducción

Debido al auge y creciente interés en la industria de los e-Sports en los últimos años, se propone generar un espacio de e-Sports con formato de taller, dentro del marco de las actividades de deportes dependiente de la Secretaría de Servicios a la Comunidad. Los e-Sports, abreviatura de deportes electrónicos, se refieren a las competiciones de videojuegos organizadas, tanto a nivel profesional como amateur, caracterizadas por su alto nivel competitivo y estructuradas en ligas o torneos similares a los deportes tradicionales, donde los jugadores compiten en diversos géneros. Todos los e-Sports son videojuegos, pero no todos los videojuegos son e-Sports [1].

En la última década, los e-Sports han experimentado un crecimiento explosivo a nivel global, capturando la atención de millones de personas [2].

Algunos de los videojuegos preferidos en los torneos y con más fama son [1]:

- Juegos multijugador de arena de batalla (MOBA en inglés): League of Legends, Dota 2, etc.

- Juegos de disparo en primera persona (shooters): Counter Strike, Call of Duty Warzone, etc.
- Juegos de disparo con héroes (Hero Shooters): Overwatch, Apex Legends, Valorant, etc
- Juego de pelea (Fighting Games): Street Fighter, Tekken, Mortal Kombat, etc.
- Juegos de estrategia en tiempo real (RTS): Starcraft II, Warcraft III, Clash Royale, etc
- Juegos deportivos (Sport Simulator): (EA Sports FC, Rocket League, etc.)

El taller que se menciona en este artículo ha sido organizado en dos ejes temáticos transversales y complementarios. El primer eje se centra en el aprendizaje técnico y estratégico dentro de los videojuegos, mientras que el segundo aborda la psicología aplicada al deporte y el bienestar integral del jugador, garantizando una formación completa.

El primer eje temático denominado "Desarrollo de Habilidades y Estrategia en el Juego" tiene como objetivo brindar todos los conocimientos técnicos y tácticos necesarios para comprender y mejorar en los juegos abordados. En la tabla 1, se presentan los contenidos que se abordan en el eje temático.

El segundo eje temático denominado "Psicología aplicada al Deporte y Bienestar del Jugador" cuyo objetivo es integrar los contenidos de psicología deportiva y bienestar general, proporcionando herramientas para un desarrollo saludable tanto para jugadores competitivos como recreativos. En la tabla 2, se presentan los contenidos que se abordan en este eje temático.

Tabla 1. Eje temático 1 "Desarrollo de Habilidades y Estrategia en el Juego".

Fundamentos del Juego:
● Historia, contexto e interfaz de cada título. ● Mecánicas básicas de control y movimiento. ● Introducción a personajes, agentes y sus habilidades. ● Análisis detallado de mapas y objetivos principales/secundarios.
Tácticas de Equipo:
● Técnicas de comunicación efectiva en tiempo real. ● Estrategias básicas (fase de líneas, control de mapa) y avanzadas (rotaciones, emboscadas). ● Análisis del "meta" actual del juego y su evolución.
Roles Específicos:
Estudio y práctica focalizada en los distintos roles de cada juego, como *Support, Mid, ADC,* y *Jungla* (League of Legends); *Duelistas, Iniciadores, Controladores y Centinelas* (Valorant); y Tank, DPS, y Support (Overwatch).

Tabla 2. Eje temático 2 "Psicología aplicada al Deporte y Bienestar del Jugador".

Temas para el Desarrollo Competitivo [3]	Temas para la Salud y el Bienestar General
Liderazgo y Roles dentro del grupo [7]	Manejo de la Frustración
Gestión emocional y control de presiones	Fluir, Mindfulness y Concentración [4] y [9]
Toma de decisiones y Creatividad	Ergonomía y Descanso
Motivación y Confianza	Uso responsable de videojuegos y prevención de adicciones [8]
Entrenamiento Invisible (Nutrición, Sueño)	Salud mental y convivencia en ambientes digitales [5] y [6]
Proyecto de vida y profesionalización	Hábitos saludables

El taller se estructura en torno a juegos pilares, con flexibilidad para incorporar otros títulos según el interés y la demanda de la comunidad.

- Juegos Pilares: League of Legends, Valorant y Overwatch.
- Otros Juegos contemplados: EA Sports FC, Clash Royale, Teamfight Tactics (TFT), Rocket League y otros según la demanda.

Estos ejes de contenido se articulan a través de una metodología de trabajo que garantiza la aplicación práctica de los conocimientos y el fomento de una comunidad activa y comprometida.

Líneas de Investigación, Desarrollo e Innovación

Actualmente, UNAHUR no cuenta con una línea de I+D+i en relación con los e-Sports. Como se mencionó anteriormente, bajo la órbita de la Secretaría de Servicios a la Comunidad dentro de la Dirección de Deportes de la Universidad Nacional de Hurlingham, se ha gestado el espacio para el dictado del taller de e-Sports. A partir de la iniciativa de la RedUNCI[1] denominada Liga Interuniversitarias de deportes electrónicos (LIDE[2]) cuyo objetivo es indagar e incorporar esta temática en los futuros eventos científicos así como también coorganizar las competencias en distintas Universidades, el Instituto de Tecnología e Ingeniería ha decidido implementar un espacio para desarrollar actividades académicas que permita contribuir a esta iniciativa.

Estructura organizativa del Taller

Para poder llevar adelante las actividades del taller, hay un profesor a cargo y un alumno asistente para poder gestionar y monitorear las necesidades del mismo. Asimismo, para poder llevar a cabo el eje temático 1, se convoca a estudiantes con conocimientos y experiencia avanzados en los juegos que se trabajan, para que oficien con el rol de Referentes de los mismos [4]. Dentro de las tareas de este rol se fomenta el dictado de clases introductorias, temáticas, avanzadas y su aplicación práctica de cada juego. También se trabaja la organización de torneos intra taller, de toda la UNAHUR, y la creación de equipos que representen a la Universidad con miras a poder competir en las ligas interuniversitarias.

Resultados y Objetivos

Los resultados concretos de esta experiencia ha sido el dictado del taller en los años 2023, 2024 y 2025. Se fomentó la participación activa de los estudiantes en un espacio de intercambio y aprendizaje, con componentes de desarrollo técnico, salud y bienestar, y el abordaje de temáticas transversales como streaming, kinesiología, ejercitación física, nutrición, ergonomía y la locución de partidas.

Se propone como objetivo a futuro el dictado de un taller para realizar una transferencia a la comunidad universitaria intra y extramuros sobre los aspectos psicológicos que inciden en los jugadores que participan este tipo de competencias que permitan manejar la frustración y el estrés, consolidando un entorno de juego positivo y constructivo.

Formación de Recursos Humanos

En relación con la formación de RRHH, se puede mencionar la cantidad de estudiantes de las distintas carreras de UNAHUR que han realizado el taller durante los años mencionados anteriormente, que asciende a un

1 Red de universidades con carreras en Informática
2 https://redunci.info.unlp.edu.ar/lide-1/

total de aproximadamente 200 estudiantes, sin contar aquellos que por una cuestión de cupo quedaron en formato virtual, accediendo a las clases virtuales y los materiales dictados en el salón presencial.

Referencias

[1] IEEE Games Technical Committee Argentina- https://r9.ieee.org/argentina-cis/inicio/jb1day/e-sports-nacimiento-y-evolucion/

[2] – Agencia Córdoba Deportes, disponible en: https://deportes.cba.gov.ar/programas/e-sports-cordoba

[3] - Lifschitz, Tomás. Game On. Entrená Tu Mente y mejora tu rendimiento en los eSports, Editorial Planeta. ISBN 9789504969525 (2020).

[4] Csikszentmihalyi, Mihaly. Fluir, Una psicología de la felicidad, Editorial Kairós. ISBN 9788472453722 (1999).

[5] Stamateas, Bernardo. Gente tóxica. Editorial B de Bolsillo. ISBN-13: 9788413147918 (2023).

[6] Garrote, Marta. Tipos de gamers tóxicos, disponible en: https://www.marca.com/videojuegos/esports/2020/12/27/5fe84bfe46163f78788b463f.html

[7] Pichon Riviére, Enrique. El proceso grupal, Ed. Nueva Visión. ISBN-13: 9789506020002 (1997).

[8] Carbonell, Xavier. El diagnóstico de adicción a videojuegos en el dsm-5 y la cie-11: retos y oportunidades para clínicos, disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=7790909

[9] Bavelier, Daphne. El cerebro bajo la influencia de los videojuegos, charla TED disponible en: https://youtu.be/CNVBxALxkKE?si=Wr7P9in6CO68dFiO

Grupo de interés en Deportes Electrónicos

Patricia Pesado, Jorge Eterovic
Coordinación RedUNCI
ppesado@lidi.info.unlp.edu.ar, eterovic@unlam.edu.ar

Resumen

Existe una importante relación entre los deportes electrónicos y la Informática. El ámbito de los deportes electrónicos muestra el impacto interdisciplinario de la Informática, al combinar conocimientos de múltiples disciplinas.
En 2025 la RedUNCI decidió poner foco en los Esports como un tema interdisciplinario que permite la interacción de los alumnos de las distintas universidades, como también la investigación sobre distintos aspectos del ecosistema de los deportes electrónicos.

Palabras clave: deportes electrónicos, interdisciplina, videojuegos, competencias.

Contexto

Luego de varias acciones informales en el tema, a fines de 2025 se creó un Grupo de Interés de la RedUNCI en el que participan: UNLP, UNNOBA, UNSE, UNPA, UNER, UNAHUR, UAI, UNNE, UNSL, UNLaM, UNCOMA, UNS, UTN, UNRN Sede Andina, UNCU, UNLaR, UNICEN y UNLu. Este Grupo de Interés está abierto a otras universidades de la RedUNCI.

I. Introducción

Los “deportes electrónicos” son básicamente competiciones de videojuegos, que se realizan en forma individual o por equipos. Su crecimiento ha sido explosivo en los últimos años y se ha discutido su inclusión en los Juegos Olímpicos.

Es importante comprender la relación entre los deportes electrónicos y la Informática: desde los algoritmos propios de los juegos, el manejo de realidad virtual/aumentada/ extendida, la posibilidad de manejar aplicaciones en red, los mecanismos de interacción humano computador, el empleo de IA en el perfeccionamiento de los juegos y la gestión distribuida de juegos que pueden involucrar a diferentes participantes geográficamente distribuidos y conectados por Internet.

En cierto modo el ámbito de los deportes electrónicos muestra el impacto interdisciplinario de la Informática, al combinar conocimientos de múltiples disciplinas.

En este contexto hay mucho interés de los jóvenes por participar en todas las actividades vinculadas con los Esports (desde el desarrollo de los videojuegos mismos hasta la gestión de torneos o la conformación y entrenamiento de equipos que compiten).

El desarrollo de la actividad (en Argentina y en el mundo) indica posibilidades laborales en diferentes áreas disciplinarias asociadas con los Esports.

II. Acciones

El trabajo comenzó con un relevamiento de las universidades que tenían algún proyecto de deportes electrónicos, carreras relacionadas o simplemente les interesaba incursionar en el tema.
En ese contexto las universidades interesadas han avanzado en un plan de actividades para 2025:

- Se creó una Liga Interuniversitaria de Deportes Electrónicos (LIDE) para fortalecer la interacción entre Universidades del país, alentando la participación de estudiantes de las Universidades de la RedUNCI en los equipos.
  Esta propuesta debiera servir como un incentivo a las vocaciones en Informática y también en otras áreas universitarias resultantes de la interacción disciplinar con Informática.

- Se creó un área en WICC 2026 sobre Esports donde se presenten líneas de investigación/desarrollo relacionadas con el ecosistema de los deportes electrónicos y su impacto social y académico.
- Se está analizando la posibilidad de hacer algún Taller sobre el impacto de los Deportes Electrónicos en aspectos sociales en el ámbito de TEYET 2026.
- Se analiza realizar alguna actividad relacionada a los Esports en el marco de la Escuela de CACIC 2026.
- Se han visibilizado estas acciones en el sitio de la RedUNCI

III. LIDE

El Grupo de Interés discutió que aspectos son necesarios para la organización de la LIDE:

- Disciplinas que se pueden incluir (deportivas, de estrategia, de disparos)
- Calendario anual y relación con las actividades de la RedUNCI
- Implementación de un torneo con participantes distribuidos geográficamente
- Coordinación de la organización e implementación del torneo.
- Final presencial

Esta iniciativa impulsada por la UNLP pretende darle un canal de competencia (no comercial) a los equipos de las Universidades, fortaleciendo las relaciones entre estudiantes de diferentes Universidades y canalizando una actividad que no tiene una regularidad actual en Argentina.

Se han definido las características de la LIDE I:

- Disciplina deportiva: FIFA
- entre abril (WICC) y junio (TEYET) 2026 - 6/8 sábados
- participantes de todas las universidades de la RedUNCI (deberán presentar un certificado de alumno regular de la universidad a la que representan y aval del representante)
- final presencial en el Centro de Entrenamiento de Deportes Electrónicos de la Facultad de Informática de la UNLP.
- Coordinación de la organización e implementación del torneo a cargo de la UNLP con la colaboración de las universidades que conforman este grupo de interés de la RedUNCI.
- Inscripción a través de formulario WEB desde febrero y hasta 15/marzo

Asimismo se prevé la realización de la LIDE II en el segundo semestre de 2026 con una Disciplina de estrategia: League of Legends.

Referencias

- Estrebou, C.(2025) “Deportes electrónicos: un tema para despertar vocaciones a la informática”. Cesar Estrebou, Santiago Medina, Federico Arteca, Federico Ruera. Libro de actas XX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2025). Págs. 441-444. ISBN: 978-987-8994-42-0. Junio 2025. https://sedici.unlp.edu.ar/handle/10915/184997
- Antón, M. (2018). “Los deportes electrónicos: el espectáculo en las competiciones de videojuegos” [Tesis de Doctorado, Universidad Complutense de Madrid]. E-Prints Complutense. https://eprints.ucm.es/id/eprint/54986/1/T41009.pdf
- Carrillo, J. A. (2015). “La dimensión social de los videojuegos' online': de las comunidades de jugadores a los 'esports'.” Index. Comunicación: Revista científica en el ámbito de la Comunicación Aplicada, 5(1), 39-51.
- Guillén Muñoz, Á. (2020). “La planificación estratégica en equipos de esports”. Universidad de Valladolid. https://uvadoc.uva.es/bitstream/handle/10324/43529/TFG-N.%201476.pdf
- Pérez-Sagarzazu, G. (2019).” Introducción de los deportes electrónicos en la Universidad y diseño de su estructura organizativa” UNAV Esports.
- Nielsen Esports (2022). “The ROI of Esports.” The Nielsen Company.
- Sitio web Proyecto Esports UNLP. https://esports.info.unlp.edu.ar/

Proyecto Esport UNLP

Santiago Medina[1,2], Federico Ruera[2], Cesar Estrebou [1], Patricia Pesado[1], Armando De Giusti[1]

[1] Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 - La Plata, Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 - La Plata, Buenos Aires
[2] Proyecto Esports UNLP
(cesarest, smedina, fruera, ppesado, degiusti) @lidi.info.unlp.edu.ar

RESUMEN

El Proyecto Esports UNLP, desarrollado en el Instituto de Investigación en Informática LIDI (III-LIDI) de la Facultad de Informática de la UNLP desde 2021, articula formación académica e infraestructura en un ecosistema competitivo como entorno de innovación educativa y tecnológica.

El proyecto integra una Diplomatura en Deportes Electrónicos de carácter interdisciplinario, un Centro de Entrenamiento con infraestructura específica y un circuito competitivo estructurado en torno a la Liga Esports Interfacultades (LEIF), configurando un dispositivo institucional que opera simultáneamente como espacio académico, experimental y de práctica profesional supervisada. Acompañado por un plan de estudios que abarca las diferentes áreas laborales de los deportes electrónicos, pero con una fuerte impronta informática de base, busca abrir un marco de posibilidades para alumnos universitarios y pre-universitarios, motivando el desarrollo de vocaciones en las disciplinas que conforman el ecosistema.

Palabras clave: interdisciplinaridad, deportes electrónicos, videojuegos, competencias, diplomatura.

CONTEXTO

El desarrollo de los deportes electrónicos basados en competiciones individuales o por equipos ha tenido un crecimiento exponencial en todas sus versiones, desde la formación de equipos amateurs hasta la profesionalización en competiciones en diferentes circuitos, tanto nacionales e internacionales.

Los deportes electrónicos ofrecen una gran variedad de oportunidades laborales en Argentina y en el mundo: involucran infraestructura de redes, procesamiento en tiempo real, análisis de datos, producción audiovisual, gestión de comunidades, psicología deportiva y coordinación de equipos competitivos.

En este marco, la Universidad Nacional de La Plata impulsa desde 2021 el Proyecto Esports UNLP, desarrollado en el III-LIDI de la Facultad de Informática. La propuesta surge como iniciativa interdisciplinaria destinada a fomentar la práctica institucional de deportes electrónicos por sus alumnos, docentes y nodocentes, y simultáneamente transformar ese espacio en una plataforma formativa y tecnológica.

Desde su origen, el proyecto se planteó con doble orientación: consolidar equipos representativos de la UNLP que participen en competencias nacionales, y utilizar el ecosistema competitivo como entorno pedagógico capaz de despertar vocaciones a la informática e integrar saberes técnicos con habilidades organizacionales.

1. INTRODUCCIÓN

Desde hace unos años la UNLP cuenta con un espacio destinado a la práctica de deportes electrónicos. La iniciativa permite contar con equipos oficiales de la Universidad, que compiten habitualmente en torneos nacionales e internacionales de videojuegos.

Si bien las primeras experiencias competitivas organizadas dentro del ecosistema de los

deportes electrónicos pueden rastrearse en las últimas décadas del siglo XX, su consolidación contemporánea responde a la convergencia entre infraestructura de redes de alta velocidad, plataformas de transmisión en vivo y modelos organizacionales profesionalizados. Este recorrido dio lugar a los actuales circuitos competitivos, ligas estructuradas y ecosistemas empresariales que articulan tecnología, producción y gestión deportiva.

En los últimos años, esta institucionalización ha dado lugar a la incorporación de los deportes electrónicos en eventos multideportivos, al reconocimiento formal por parte de federaciones nacionales e internacionales e incluso de asociaciones deportivas tradicionales. En el ámbito universitario, toma formas por demás diversas: desde una oferta lúdica extracurricular hasta la formalización de programas de becas y ligas interuniversitarias asentadas en la articulación de cultura digital juvenil y profesionalización.

En ese sentido, la Facultad de Informática, desde el III-LIDI busca generar un ambiente donde los integrantes de la comunidad educativa puedan participar y formarse en temas relacionados con los Esports. Se pone a disposición del proyecto infraestructura propia para torneos y gestión de equipos, buscando también potenciar el interés en la investigación de temas asociados.

El deporte electrónico universitario constituye así un espacio de competencia digital multifacético, con objetivos formativos, prácticas colaborativas y dispositivos institucionales de gran interés para el desarrollo de competencias laborales. La existencia de ligas universitarias nacionales e internacionales, así como de programas académicos formales, evidencia una transición hacia estructuras reguladas en las que trabajar desde una perspectiva I+D+i.

La consolidación institucional del Proyecto Esports UNLP, en el III-LIDI y la Facultad de Informática, abre líneas de investigación en infraestructura tecnológica, investigación en videojuegos, inteligencia artificial e interdisciplina. El Centro de Entrenamiento constituye un espacio de integración de redes, estaciones de alto rendimiento, sistemas de transmisión en vivo y herramientas de producción multimedia. Las instancias de entrenamiento táctico, el análisis técnico y las transmisiones en directo generan un contexto de uso intensivo de infraestructura informática en tiempo real, operando como laboratorio experimental donde aplicar conocimiento vinculados a arquitectura en sistemas, soporte y gestión tecnológica.

Asimismo, abre líneas explícitas de investigación, aprovechando el circuito competitivo como banco de pruebas de análisis de desempeño, modelización de decisiones en entornos dinámicos y evaluación de comportamientos en sistemas complejos. Se vuelve un entorno de simulación profesional que convoca estudiantes de las 17 facultades de la Universidad Nacional de La Plata, funcionando como laboratorio sociotécnico interdisciplinario en tres dimensiones complementarias: (i) diseño y evaluación de entornos competitivos universitarios mediados por tecnologías digitales, (ii) desarrollo y adaptación de herramientas informáticas de gestión y transmisión de torneos, y (iii) formación de recursos humanos con perfil interdisciplinario.

2. ESTRUCTURA Y ACTIVIDADES DEL PROYECTO

La estructura del proyecto se apoya en tres dispositivos complementarios. Un eje académico, a través de la Diplomatura en Deportes Electrónicos creada en 2022 como propuesta de formación con fuerte impronta informática. Un Centro de Entrenamiento en Esports, instalado en el Centro de Innovación y Transferencia Tecnológica de la Facultad de Informática (CIYTT-UNLP), y concebido como espacio híbrido -aula, laboratorio y estudio de producción audiovisual-. Y una Liga Esports Interfacultades (LEIF), como circuito competitivo institucional, desarrollada desde 2023 de manera presencial con regularidad semestral.

Esta articulación institucional busca resignificar la competencia como un entorno de aprendizaje: lejos de constituir una actividad extracurricular aislada, configura un dispositivo estructurado donde la práctica

competitiva funciona como escenario controlado para la aplicación de conocimientos interdisciplinarios a través de:

- Reclutamiento y evaluación por competencias de jugadores.
- Conformación de grupos de gestión de equipos, entrenadores, psicólogos deportivos, comunicadores, etc.
- Organización de entrenamientos y competiciones internas.
- Participación en torneos nacionales e internacionales.
- Gestión de canales de comunicación, transmisión de contenidos en vivo y diseño de identidades visuales.

3. RESULTADOS ESPERADOS/OBTENIDOS

Desde su implementación, el Proyecto Esports UNLP ha logrado consolidar una estructura institucional sostenida en el tiempo. Mientras que la Diplomatura en Deportes Electrónicos ha alcanzado ya cuatro ediciones, el Centro de Entrenamiento -operativo desde octubre de 2022- se ha constituido como soporte físico y tecnológico para las actividades de los equipos representativos y eventos de la comunidad.

Desde un punto de vista cualitativo, el principal resultado radica en esa consolidación de un entorno articulado: la competencia ha dejado de ser un fin en sí mismo, convirtiéndose en espacio de aplicación, experimentación y validación de conocimientos.

➢ Equipos UNLP. Centro de Entrenamiento.

2025 representa un año de inflexión para el Proyecto esports UNLP a nivel equipos, principalmente por el impacto de la incorporación de nuevo equipamiento: la ampliación del alcance al incorporar videojuegos de consola (PS5) permitieron ampliar también el número de disciplinas. A las cinco disciplinas fundacionales, League of Legends, Age of Empires, Valorant, Counter Strike y Teamfight Tactics (elegidas oportunamente por representar circuitos con oferta competitiva en PC a nivel nacional), se le suman Rocket League, EA FC (ex FIFA) y Fornite, así como experiencias lúdicas comunitarias en fighting games como 2XKO.

El interés en estas disciplinas abre el camino para la conformación de nuevos equipos de trabajo, siempre bajo la estructura formal de un cuerpo técnico (entrenador y analista) y un psicólogo deportivo acompañando a los jugadores.

Por otro lado, la mejora en términos de infraestructura para el centro de streaming permitió ampliar también el grupo de trabajo por fuera de lo estrictamente deportivo, dando lugar a nuevos roles de diseño, comunicación, operación de herramientas de transmisión en directo y producción de materiales audiovisuales.

➢ Participación de la UNLP en torneos.

En términos competitivos, la participación de la Universidad Nacional de La Plata a través del proyecto Esports UNLP se ha consolidado en competencias nacionales y continentales tanto universitarias como híbridas (de índole privada y semiprofesional). Esta participación no solo fortalece la práctica local de deportes electrónicos, sino que articula a estudiantes con experiencias competitivas estructuradas y transversales. En el plano internacional, la UNLP participó en la edición 2025 de los **FISU America Esports**, campeonato continental universitario de EA FC25 (fútbol electrónico), compitiendo frente a representantes de otras instituciones de América. De igual manera, los representantes de la disciplina Rocket League participaron del **RLCS Open Sudamérica**: el torneo oficial de mayor prestigio a nivel continental y clasificatorio para el campeonato mundial. A nivel local, por su parte, los ganadores de la liga interna de la UNLP participaron de la **Liga del Paraná**, circuito competitivo nacional semiprofesional de League of Legends, y de las clasificatorias de Counter Strike 2 para la **Argentina Game Show**.

➢ Organización de torneos

La Liga Esports Interfacultades (LEIF) suma seis ediciones presenciales entre 2023 y 2025, registrando 454 participaciones totales.

Liga Esports Interfacultades	Jugadores
2023 League of Legends	**111**
2023 Age of empires	**32**
2024 League of Legends	**97**
2024 Valorant	**40**
2024 Counter Strike	**84**
2025 League of Legends	**48**
2025 Aram	**42**
TOTAL JUGADORES ACTIVOS	**454**

La participación de estudiantes, docentes y nodocentes provenientes de múltiples unidades académicas (16 facultades sobre un total de 17 con al menos un representante en 3 años) confirma la transversalidad institucional del proyecto. Mientras que la colaboración de estudiantes y graduados de muchas de estas unidades confirma, además, su capacidad de integración interdisciplinaria. Visto como laboratorio de experiencias y prácticas profesionales supervisadas, destacan en la coordinación de estos torneos los perfiles de gestión, diseño, programación y comunicación social.

Para este 2026, junto con la organización de la séptima edición de la LEIF, se formalizó la propuesta de una Liga Interuniversitaria de Deportes Electrónicos (LIDE) a través de la RedUNCI. Propuesta que busca fortalecer la interacción entre Universidades del país, alentando la participación de estudiantes de las Universidades de la RedUNCI en este ecosistema híbrido, buscando servir como un incentivo a las vocaciones en Informática y también en otras áreas universitarias resultantes de la interacción disciplinar con Informática.

- **Investigación y Desarrollos asociados con los deportes electrónicos.**

Independientemente de otras actividades de difusión/información/competencias, interesa potenciar el crecimiento del tema de los deportes electrónicos desde la Investigación y Desarrollo. En el mediano plazo, se proyecta profundizar las líneas de investigación asociadas al análisis de datos, fortalecer el desarrollo de plataformas tecnológicas propias y avanzar en propuestas interuniversitarias en el ámbito de la RedUNCI. Por otro lado, se prevé trabajar en la sistematización de indicadores que permitan evaluar con mayor precisión el impacto del proyecto en la generación de vocaciones tecnológicas.

- **Difusión en Paneles/ Charlas/ Conferencias/ Jornadas.**

Durante 2025 los integrantes del proyecto participaron en diferentes eventos que permitieron la difusión de las actividades realizadas:

- Expo UNLP
- I Jornada de Deportes Electrónicos de la Facultad de Informática, Universidad Nacional de La Plata, Conferencia a cargo del Esp. Cesar Estrebou “Inteligencia Artificial en los Deportes Electrónicos”
- TEyET, Universidad Nacional de Avellaneda, Demo “Deportes Electrónicos: un tema para despertar vocaciones a la informática”
- CACIC, Universidad Nacional de Rio Negro, presentación de proyecto y propuesta de LIDE

4. FORMACION DE RECURSOS HUMANOS

La contribución del Proyecto Esports UNLP en términos de formación de recursos humanos ofrece múltiples aristas. Por un lado, la Diplomatura en Deportes Electrónicos funciona como formación estructurada, integrando contenidos tecnológicos, organizacionales y estratégicos con instancias de práctica profesional supervisada. Por otro, la LEIF y el Centro de Entrenamiento constituyen entornos de aprendizaje situado donde estudiantes asumen responsabilidades reales en organización, soporte técnico y producción multimedia. Este esquema responde a un modelo de cognición situada, favoreciendo el desarrollo de competencias técnicas y transversales en un marco institucional guiado.

Asimismo, el desarrollo de software específico y otras iniciativas tecnológicas vinculadas al proyecto ponen en evidencia la articulación entre formación de grado e innovación, consolidando un espacio de convergencia entre investigación, docencia y desarrollo. Aparecen allí casos puntuales, desde egresados de la Diplomatura que actualmente imparten talleres de formación profesional específica en materia de producción audiovisual, diseño de identidades digitales y manejo de redes sociales. O proyectos de índole más deportiva, como la academia de nuevos jugadores de League of Legends, donde se ponen en juego habilidades de transmisión de conocimiento, comunicación y liderazgo.

5. BIBLIOGRAFÍA

- Sitio web Proyecto Esports UNLP. https://esports.info.unlp.edu.ar/
- Sitio web Diplomatura en Deportes electrónicos. https://www.info.unlp.edu.ar/deportes-electronicos/
- Estrebou, C.(2025) “Deportes electrónicos: un tema para despertar vocaciones a la informática”. Cesar Estrebou, Santiago Medina, Federico Arteca, Federico Ruera. Libro de actas XX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2025). Págs. 441-444. ISBN: 978-987-8994-42-0. Junio 2025. https://sedici.unlp.edu.ar/handle/10915/184997
- Ordenanza 334/22 de la Facultad de Informática de la UNLP “REGLAMENTO PARA LAS DIPLOMATURAS EN LA FACULTAD DE INFORMÁTICA”.
- Antón, M. (2018). “Los deportes electrónicos: el espectáculo en las competiciones de videojuegos” [Tesis de Doctorado, Universidad Complutense de Madrid]. E-Prints Complutense. https://eprints.ucm.es/id/eprint/54986/1/T41009.pdf
- AEVI - Spanish Association of Videogames. (2020). “Los esports en España: situación actual y posición de la industria. Asociación Española de Videojuegos.” http://www.aevi.org.es/web/wp-content/uploads/2020/12/Informe_esports_ESP_20.pdf
- Antueno, H., & Ruera, F. (2023). Esports y Universidad Pública: la experiencia del trabajo psicológico con equipos universitarios [Contribución en mesa redonda]. Congreso Nacional de Psicología Aplicada al Deporte y a la Actividad Física, Asociación de Psicología del Deporte Argentina, Buenos Aires — Universidad de Belgrano.
- Carrillo, J. A. (2015). “La dimensión social de los videojuegos' online': de las comunidades de jugadores a los 'esports'.” Index. Comunicación: Revista científica en el ámbito de la Comunicación Aplicada, 5(1), 39-51.
- Taylor, T. L. (2012). “Raising the Stakes: E-sports and the Professionalization of Computer Gaming.” Cambridge: MA: MIT Press
- Gutierrez Gonzalez Ángel. “Tecnologías de la Información - Un enfoque interdisciplinario” Alfaomega Grupo Editor Argentino
- Austerb David. “La Tecnología del Streaming de Vídeo y Audio” Editorial Escuela de Cine y Vídeo de Andoaín, S.L.
- Angel Arias. “Computación en la Nube” Editorial: CreateSpace Independent Publishing Platform; 2nd edición
- Guillén Muñoz, Á. (2020). “La planificación estratégica en equipos de esports”. Universidad de Valladolid. https://uvadoc.uva.es/bitstream/handle/10324/43529/TFG-N.%201476.pdf
- Pérez-Sagarzazu, G. (2019).” Introducción de los deportes electrónicos en la Universidad y diseño de su estructura organizativa” UNAV Esports.
- Nielsen Esports (2022). “The ROI of Esports.” The Nielsen Company.
- Scholz & Barlow. (2019). “Esports is Business.” Springer International Publishing.
- Cáliz Ortiz, J. S., & Paredes Solano, A. D. (2020).”Cómo facilitar el proceso de reclutamiento de aspirantes latinoamericanos a equipos de esports.” Biblioteca Digital Icesi.
- Fanjul, C., González, C., & Peña, P. J. (2019). “La influencia de los jugadores de videojuegos online en las estrategias publicitarias de las marcas: comparativa entre paña y Corea.” Comunicar, 58.
- Luis Ruelas. “Desarrollo de Videojuegos y Realidad Virtual con Unity 3D.” Editorial Ra-Ma. ISBN 978-84-18551-90-1
- Jose M. Martinez, Fernando Navarro, Antonio Martinez · 2018. “Realidad Virtual y Realidad Aumentada.” Editorial Ra-Ma. ISBN13 9788499647395
- Cabero Almenara, J. y Jiménez García, F. “Realidad Aumentada. Tecnología para la Formación.” Editorial Síntesis. 2016. ISBN: 978-84-9077-258-4
- David Ballejo Fernández , Cleto Martín Angelina. “Desarrollo de Videojuegos 1: Arquitectura del Motor de Videojuegos” “Desarrollo de Videojuegos 2: Programación Gráfica” “Desarrollo de Videojuegos 3: Técnicas Avanzadas” “Desarrollo de Videojuegos 4: Desarrollo de Componentes”.

Editados por la Universidad de Castilla La Mancha
- Ana Belén Gomez Sanz . “Gamificación y Juegos Serios. Curso Práctico” RA-MA S.A. Editorial y Publicaciones
- Ricardo Casañ Pitarch. “Videojuegos Serios en educación” Editor: Universitat Politècnica de València
- Antón, M; García, F (2014) “Deportes electrónicos. Una aproximación a las posibilidades comunicativas de un mercado emergente.” Questiones publicitarias, 1, pp.98-115
- Mendoza, G. (2019). “El trabajo psicológico online con equipos de esports: experiencias en Vodafone Giants”. Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 4(1), 1-7.
- Pedraza-Ramirez, I. (2019). “Generación LoL: entrenamiento psicológico mediante una propuesta holística con un equipo profesional de Esports.” Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 1-10.
- Vaamonde, A. G. N., Toribio, M. J., Molero, B. T., & Suárez, A. (2018). “Beneficios cognitivos, psicológicos y personales del uso de los videojuegos y esports: una revisión” Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 3(2), 1-14.
- Pablo L. Gutierrez Utande – Angel Arroyo Castillo (Tutor) “Aplicación de Inteligencia Artificial en Videojuegos” Universidad Politécnica de Madrid – (2017)
- Bernard Marr - “BIG DATA: La utilización del Big Data, el análisis y los parámetros SMART para tomar mejores decisiones y aumentar el rendimiento” Editorial TEELL - 2016
- Eric Sadin - “La Inteligencia Artificial o el Desafío del siglo” Colección Futuros Próximos – 2020

Redes de Cooperación Científica Internacionales

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

Acciones de colaboración en el Consorcio de Cloud Computing, Big Data y Emerging Topics durante 2025 (CCC-BD&ET)

III-LIDI – Instituto de Investigación en Informática LIDI (UNLP – Argentina)
LISSI–Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información (UNS – Argentina)
VyGLab – Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (UNS– Argentina)
LIDIC – Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (UNSL – Argentina)
HPC4EAS – High Performance Computing for Efficient Applications and Simulation (UAB – España)
SMILe – Soft Management of Internet and Learning (Universidad de Castilla-La Mancha – España)
ArTeCS – Group of Architecture and Technology of Computing Systems (UCM – España)
LITRP – Laboratorio de Investigaciones Tecnológicas en Reconocimiento de Patrones (UCM – Chile)
LCG – Laboratorio de Computación Gráfica (UNSL – Argentina)
ITIC – Instituto para las Tecnologías de la Información y las Comunicaciones (UNCu – Argentina)
DisCo – Grupo de I+D+I en Computación Distribuida (Universidad de Zaragoza – España)
SAL – Grupo de Supercomputación- Algoritmos (Universidad de Almería – España)
ITT – Instituto de Transferencia de Tecnología (UNNOBA – Argentina)
& Investigadores Asociados al CCC- BD&ET

Resumen

El Consorcio de I+D+i en *Cloud Computing, Big Data & Emerging Topics* (CCC-BD&ET) [1] es una iniciativa para fomentar y formalizar la colaboración existente entre grupos de investigación de varias universidades en temáticas vinculadas a Cloud Computing, al Análisis Masivo de Datos y a Tópicos Emergentes, como las tecnologías 4.0 o los gemelos digitales aplicados a distintos ámbitos, entre otros. Estas temáticas, y su integración, han adquirido creciente importancia por su aplicación en dominios de alto impacto como las ciudades inteligentes, la transformación digital, los sistemas de E-health y los basados en tecnologías tales como IoT, IIoT y Blockchain.

Los integrantes del consorcio, provenientes mayoritariamente de Argentina, Chile y España han tenido, a lo largo de los años, diversas experiencias de trabajo conjunto que fueron consolidadas a partir de la organización y realización de las *Jornadas de Cloud Computing-Big Data & Emerging Topics* (JCC-BD&ET) llevadas a cabo en la Universidad Nacional de La Plata (Argentina). La constitución de este Consorcio, reafirma y formaliza estas líneas de colaboración proponiendo acciones de cooperación académica vinculadas con la formación de recursos humanos, la formulación y ejecución de proyectos conjuntos, y la vinculación con empresas y organismos relacionados con la industria informática, entre otras.

Palabras clave: Cloud Computing, Big Data, Colaboración interinstitucional, Cooperación internacional.

Contexto

El Consorcio de I+D+i en *Cloud Computing, Big Data & Emerging Topics* (CCC-BD&ET) resulta de la cooperación, llevada a cabo durante varios años, entre grupos de investigación, desarrollo e innovación de Universidades vinculadas con las realización anual de las *Jornadas de Cloud Computing-Big Data & Emerging Topics* (JCC-BD&ET), organizadas por la Universidad Nacional de La Plata (UNLP), en Argentina.

I. Introducción

Las Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET) son un encuentro anual de intercambio de ideas, proyectos, resultados científicos y aplicaciones concretas en diferentes áreas relacionadas con Cloud Computing, Inteligencia de Datos, Big Data y Tecnologías Emergentes.

Las Jornadas integran ponencias científicas con experiencias de desarrollos y aplicaciones, fomentando la interacción entre la academia y los sectores productivos/industriales, en el área

temática de Cloud Computing, Big Data y Tecnologías Emergentes.

En el marco de las Jornadas también se desarrollan Conferencias, Paneles y Cursos de Posgrado específicos.

Las JCC-BD&ET son organizadas por el Instituto de Investigación en Informática LIDI (III-LIDI) y la Secretaría de Posgrado de la Facultad de Informática de la UNLP en colaboración con Universidades de Argentina y del exterior. Cuentan con diferentes auspicios de organismos de Ciencia y Tecnología de Argentina y de sectores representativos de la industria del Software de Argentina. Se iniciaron en 2013 con foco en los temas de Cloud Computing. A partir del 2015, pasaron a ser las Jornadas de Cloud Computing & Big Data y, en el año 2020, se constituyeron en JCC-BD&ET.

A partir de estos encuentros fueron surgiendo distintas actividades en colaboración, que no sólo se formalizaron mediante acuerdos entre los distintos actores. Los resultados de las colaboraciones se encuentran reflejados en las publicaciones detalladas en las páginas mencionadas en las referencias.

Objetivos

Como corolario de las distintas actividades de colaboración realizadas por los participantes a las JCC-BD&ET es que surge la propuesta de conformar un Consorcio de I+D+i en Cloud Computing, Big Data & Emerging Topics que permita afianzar y proyectar a futuro las relaciones existentes [1]. Bajo el amparo de las JCC-BD&ET nuestro objetivo es seguir constituyendo un foco anual de encuentro, independientemente del trabajo académico y científico que se realice durante el año para la concreción de los objetivos del Consorcio, el fortalecimiento de la cooperación, la formación de recursos humanos y la contribución a la comunicación científica en el área.

Cooperaciones actuales entre Universidades del Consorcio

En el marco del consorcio, sus integrantes han establecido colaboraciones plasmadas en los siguientes proyectos que han sido formalizados en diferentes convocatorias:

a) Proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de Rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real.” UNP 2023-2026, con la colaboración de UAB, UCM, UNRN, UNGS, UNCOMA y UNDEC.
b) Proyecto “Inteligencia de Datos. Técnicas y Modelos de Machine Learning” UNLP, 2023-2026, con la colaboración de UCLM y URV.
c) Proyecto “Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital” UNLP 2023-2026, con la colaboración de UNS, UCLM y UNIZAR.
d) Proyecto “Computación de Altas Prestaciones y algoritmos heurísticos para la gestión de sistemas sanitarios inteligentes” UAB, con la colaboración de UNLP y UNSL.
e) Proyecto UNI-UEAR “Alianza Universitaria Argentina Europea para la Transformación Digital”. Consorcio internacional con participación de UNLP y UNS. (desde 2022)
f) Programa GEMINAE - UNITA. “Cooperación entre universidades asociadas en Geminae y los integrantes de UNITA para las mejores prácticas educativas.” Consorcio internacional con participación de UNIZAR y UNLP (desde 2023)
g) Proyecto TED: “Ecosistema tecnológico para el reconocimiento del estado de ánimo de los pacientes durante el proceso de rehabilitación cardíaca”. Acrónimo del proyecto: TEMOR” UNIZAR, con la participación de UNLP.

h) Proyecto "Experiencias Lúdicas con Agentes Sociales Interactivos y Robots: Aprendizaje Social y Comunicación Intergeneracional (PLEISAR)" UNIZAR, con la participación de UNLP.
i) Proyecto "Soluciones computacionales para la gestión de datos", UNSL con la colaboración de UAB.

Resultados Obtenidos

Los resultados esperados están relacionados con el fortalecimiento de las capacidades de los grupos de investigación del Consorcio, en términos de sus RRHH, redes de trabajo, acceso a recursos organizacionales y mejoramiento de la calidad de los resultados producidos, producto de la colaboración interdisciplinaria e inter-universitaria promovida por el Consorcio. Los resultados mas relevantes se mencionan en las siguientes subsecciones.

Formación de RRHH

Los integrantes del CCC-BD&ET están abocados a la formación de recursos humanos tanto a nivel de grado como de posgrado, en temáticas afines al Consorcio, en particular en muchos casos se realiza colaboración entre los distintos grupos, a través de participación en los comités académicos de carreras, dictado de cursos, dirección y evaluación de tesis.
Por otro lado, fruto de las interacciones entre investigadores en diferentes temas, han surgido cursos e incluso nuevas carreras de Postgrado que tienen intervención de grupos de I+D+I del Consorcio.
Entre estas carreras y cursos se pueden mencionar:

a) Maestría "Cómputo de Altas Prestaciones" (UNLP); Especialización "Cómputo de Altas Prestaciones y Tecnología GRID" (UNLP) y la Especialización "Cómputo de Altas Prestaciones" (UNLP) se dictan con participación de UNS, UNSL, UAB y UCM.
b) Maestría "Gestión y tecnología de ciudades inteligentes" (UNS-UNLP con colaboración de UCLM).
c) Maestría "Tecnología Informática Aplicada en Educación" (UNLP) y la Especialización "Tecnología Informática Aplicada en Educación" (UNLP) se dictan con participación de UAB y UNIZAR.
d) Maestría "Inteligencia de datos orientada a Big Data" y la Especialización "Inteligencia de datos orientada a Big Data" (UNLP) se dictan con participación de UCLM y URV.
e) Maestría "Transformación Digital" (UNLP-UNNOBA)
f) Especialización "Computación Gráfica, Imágenes y Visión por Computadora" (UNLP) se dicta con participación de UNS y UNSL.
g) Diplomatura de Postgrado "Inteligencia Artificial Generativa" se dicta con participación de UCLM y URV.
h) Diplomatura de Postgrado "Tecnologías Digitales en Escenarios Educativos" (UNLP) se dicta con participación de UAB y UNIZAR.
i) Diplomatura de Postgrado "Transformación Digital" (UNLP-UNNOBA)

En el mismo sentido, se resaltan las colaboraciones establecidas para la (co-) dirección de estudiantes de posgrado:

a) Dirección Tesis Doctoral "Modelización y gestión del consumo energético en un sistema de altas prestaciones con tolerancia de fallos" M. Morán.(UNLP-UAB-UNCOMA). Egresada 2025.
b) Dirección Tesis Doctoral "Diseño de un modelo computacional basado en agentes para simular servicios de telemedicina" E. Solís Acosta.(UNLP-UAB). En curso.
c) Dirección Tesis Doctoral "Modelizacion y Simulacion basada en Agentes Aplicada a la Arquitecturas de Entrada/Salida de los computadores Paralelos" D. Encinas.(UNLP-UAB). En evaluación.
d) Dirección Tesis Doctoral "Desarrollo de un modelo de relaciones entre videos educativos, perfiles de estudiantes y emociones" G. Astudillo. (UNLP-UNIZAR) En curso
e) Dirección de Tesis Doctoral "Modelo de

predicción orientado a la generación de residuos electrónicos con un enfoque de Ciudades Inteligentes Sostenibles desde la perspectiva de la Inteligencia Ambiental" Facuy Delgado Jussen Paul. (UNLP-UNS) En curso

f) Dirección de Tesis Doctoral "Modelo digital para la reducción del impacto ambiental de las actividades agrícolas utilizando variables climatológicas en tiempo real aplicado al cordón frutihortícola del gran La Plata" Leonardo Navarria (UNLP-UNS) En curso
g) Dirección de Trabajo Final Integrador de Especialización en Tecnología Informática Aplicada en Educación "Revisión y análisis del impacto de la utilización de entornos virtuales de enseñanza y aprendizaje en la resiliencia de estudiantes universitarios" María Inés Ragone (UNLP-UAB) Egresada 2025

Asimismo, los grupos tienen reuniones de intercambio de posibilidades de apoyo en actividades de postgrado para alumnos de Maestría y Doctorado.
Esto incluye financiamiento para estadías de investigación y dictado de cursos abiertos a las Universidades del Consorcio, así como la definición de Ciclos y Estadías Postdoctorales.

Producción científica conjunta (2025)

a) "Adquisición y almacenamiento de variables meteorológicas provenientes de diferentes dispositivos electrónicos para la construcción de gemelos digitales" Leonardo Navarria (UNLP), Ariel Pasini (UNLP), Elsa Estévez (UNS). Libro de Actas del XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025). ISBN: 978-987-8258-99-7. Octubre 2025. http://sedici.unlp.edu.ar/handle/10915/189846
b) "Una Arquitectura para Gemelos Digitales de Tráfico basada en Aprendizaje por Refuerzo Federado Multiagente" Franco Vinciguerra (CAECE), Armando De Giusti (UNLP), Marcelo Errecalde (UNSL), Miguel Méndez Garabetti (CAECE). Libro de Actas del XXXI Congreso Argentino de Ciencias de la Computación (CACIC 2025). ISBN: 978-987-8258-99-7. Octubre 2025. http://sedici.unlp.edu.ar/handle/10915/189846.
c) "An Extended Technological Ecosystem for the Detection of Risk of Dehydration in Cattle Based on Precision Livestock Farming". Molina, S.M. (UNSL), Luque, E. (UAB), Rexachs, D. (UAB). In: Pesado, P., Thomas, P. (eds) Computer Science – CACIC 2024. CACIC 2024. Communications in Computer and Information Science, vol 2520. Springer, Cham. https://doi.org/10.1007/978-3-032-00718-6_5
d) "Enhancing Patient-Centric Synthetic Data Generation: A Comparative Analysis for Chronic Kidney Disease". Alvarez, C. (UNCOMA), Ibeas, J. (UNCOMA), Balladini, J. (UNCOMA), Suppi, R. (UAB). In: Pesado, P., Thomas, P. (eds) Computer Science – CACIC 2024. CACIC 2024. Communications in Computer and Information Science, vol 2520. Springer, Cham. https://doi.org/10.1007/978-3-032-00718-6_4
e) "WEEE Prediction Model Based on Neural Networks" J. Facuy (UAE – Ecuador), A. Pasini (UNLP – Argentina), E. Estévez (UNS / UNLP – Argentina), C. Moran (UAE – Ecuador). 13th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2025). Págs. 63-73. ISBN: 978-950-34-2583-1. Junio 2025. https://doi.org/10.35537/10915/182427
f) "Análisis de la utilización de entornos virtuales de enseñanza y aprendizaje y su relación con la resiliencia de estudiantes universitarios" María Inés Ragone (UNLP), Dolores Rexachs (UAB), Cecilia Sanz (UNLP). Libro de actas XX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2025). Págs. 189-202. ISBN: 978-987-8994-42-0. Junio 2025. http://sedici.unlp.edu.ar/handle/10915/184

997
g) “State of Practice of Digital Transformation in Argentina – Experiences of the UNI-UEAR Project” Elsa Estevez (UNS), Martín Larrea (UNS), Armando De Giusti (UNLP), Mario Gimelli (CIN). 2025 Eleventh International Conference on eDemocracy & eGovernment (ICEDEG). Págs. 62-67. Junio 2025. https://doi.org/10.1109/ICEDEG65568.2025.11081692
h) “Acciones de colaboración en el Consorcio de Cloud Computing, Big Data y Emerging Topics durante 2024 (CCC-BD&ET)” Instituto de Investigación en Informática LIDI III-LIDI (UNLP), Lab de I-D en Ing de Soft y Sist de Información LISSI (UNS), Lab de I-D en Visualización y Computación Gráfica VyGLab (UNS), Lab de I-D en Inteligencia Computacional LIDIC (UNSL), High Performance Comp for Efficient App and Sim HPC4EAS (UAB), Soft Management of Internet and Learning SMILe (UCLM), Group of Architecture and Tech of Comp Systems ArTeCS, Lab de Inv Tecn en Reconocimiento de Patrones LITRP (UCM), Lab de Computación Gráfica LCG (UNSL), Instituto para las Tecn. de la Inf. y las Comunic. ITIC (UNCUYO), Grupo de I+D+I en Comp. Ditrib. DisCo (UZ).XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025, Mendoza). Abril 2025.
i) “UNI-UEAR – Alianza Universitaria Argentina Europea para la Transformación Digital. Avances de ejecución” Mario Gimelli (CIN), Elsa Estevez (UNS), UBA Universidad de Buenos Aires, UNLP Universidad Nacional de La Plata, UNQ Universidad Nacional de Quilmes, UNS Universidad Nacional del Sur, UNSAM Universidad Nacional de San Martín.XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025, Mendoza). Abril 2025.

Integrantes del Consorcio

El Consorcio está integrado tanto por grupos de investigación como por investigadores asociados. A continuación se detallan los mismos.

Grupos que conforman el CCC- BD&ET

III-LIDI – Instituto de Investigación e Informática LIDI (Universidad Nacional de La Plata – Argentina) [16]
Directora: *Lic. Patricia Pesado*

LISSI – Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información (Universidad Nacional del Sur – Argentina) [17]
Director: *Dr. Pablo Fillottrani*

VyGLab – Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (Universidad Nacional del Sur – Argentina) [18]
Directora: *Dra. María Luján Ganuza*

LIDIC – Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (Universidad Nacional de San Luis – Argentina) [19]
Director: *Dr. Marcelo Errecalde*

HPC4EAS – High Performance Computing for Efficient Applications and Simulation Research Group (Universidad Autónoma de Barcelona – España) [20]
Director: *Dr. Emilio Luque*

SMILe – Soft Management of Internet and Learning (Universidad de Castilla-La Mancha – España) [21]
Director: *Dr. José A. Olivas Varela*

ArTeCS – Group of Architecture and Technology of Computing Systems (Universidad Complutense de Madrid – España) [22]
Director: *Dr. Francisco Tirado*

LITRP – Laboratorio de Investigaciones Tecnológicas en Reconocimiento de Patrones (Universidad Católica de Maule – Chile) [23]
Representante para el Consorcio: *Dr. Ricardo Barrientos*

LCG – Laboratorio de Computación Gráfica (Universidad Nacional de San Luis–Argentina) [24]

Director: *Mg. Roberto Guerrero*

ITIC – Instituto para las Tecnologías de la Información y las Comunicaciones (Universidad Nacional de Cuyo – Argentina)
[25].
Director: *Dr. Carlos García Garino*

DisCo – Grupo de I+D+I en Computación Distribuida (Universidad de Zaragoza – España) [26].
Coordinador: *Dr. Pedro Javier Álvarez Pérez-Aradros*

SAL – Grupo de Supercomputación- Algoritmos (Universidad de Almería – España) [27]
Director: *Gracia Ester Martín Garzón*

ITT – Instituto de Transferencia de Tecnología (UNNOBA – Argentina) [28]
Director: *Hugo Ramón*

Investigadores Asociados al CCC- BD&ET

Dr. Aurelio Fernández (Universidad Rovira i Virgili – España).
Dr. Emmanuel Frati (Universidad Nacional de Chilecito – Argentina).
Dr. Javier Balladini (Universidad Nacional de Comahue – Argentina).
Dra. Adriana Gaudiani (Universidad Nacional de General Sarmiento – Argentina).
Dra. Mónica Denham (Universidad Nacional de Río Negro – Argentina).
Lic. Nelson Rodríguez (Universidad Nacional de San Juan – Argentina).

Referencias

[1] Consorcio de I+D+I en Cloud Computing, Big Data & Emerging Topics (CCC-BD&ET), https://jcc.info.unlp.edu.ar/consorcio-de-idi-en-cloud-computing-big-data-emerging-topics/.
[2] Consorcio para la colaboración en I+D+I en temas de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET)- WICC 2021. 978-987-24611-3-3; 978-987-24611-4-0. Páginas: 778-783.
[3] "Consorcio para la colaboración en I+D+I en Temas de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET) Proyecto Integrador: "Transformación Digital en la incorporación de la Resiliencia como un Key Performance Indicator de Prestaciones Sociales (KPIS)". III-LIDI (UNLP) LISSI (UNS), VyGLab (UNS), LIDIC (UNSL), HPC4EAS (UAB), SMILe (UCLM), ArTeCS (UCM, España), LITRP (UCM, Chile), LCG (UNSL), ITIC (UNCu), DisCo (UZ) & Investigadores Asociados al CCC- BD&ET. XXIV WICC 2022. ISBN 978-987-48222-3-9. Págs. 574-578.
[4] "Avances del Consorcio para la colaboración en I+D+I en Temas de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET)" III-LIDI (UNLP), LISSI (UNS), VyGLab (UNS), LIDIC (UNSL), HPC4EAS (UAB), SMILe (UCLM), ArTeCS (UCM, España), LITRP (UCM, Chile), LCG (UNSL), ITIC (UNCu), DisCo (UZ) & Investigadores Asociados al CCC-BD&ET. WICC 2023. ISBN 978-987-3724-66-4.
[5] "Acciones de colaboración en el Consorcio de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET)" III-LIDI (UNLP), LISSI (UNS), VyGLab (UNS), LIDIC (UNSL), HPC4EAS (UAB), SMILe (UCLM), ArTeCS (UCM, España), LITRP (UCM, Chile), LCG (UNSL), ITIC (UNCu), DisCo (UZ) & Investigadores Asociados al CCC-BD&ET. WICC 2024. ISBN: 978-987-8352-57-2. Págs. 699-705.
[6]"Acciones de colaboración en el Consorcio de Cloud Computing, Big Data y Emerging Topics durante 2024 (CCC-BD&ET)" Instituto de Investigación en Informática LIDI III-LIDI (UNLP), Lab de I-D en Ing de Soft y Sist de Información LISSI (UNS), Lab de I-D en Visualización y Computación Gráfica VyGLab (UNS), Lab de I-D en Inteligencia Computacional LIDIC (UNSL), High Performance Comp for Efficient App and Sim HPC4EAS (UAB), Soft Management of Internet and Learning SMILe (UCLM), Group of Architecture and Tech of Comp Systems ArTeCS, Lab de Inv Tecn en Reconocimiento de Patrones LITRP (UCM), Lab de Computación Gráfica LCG (UNSL), Instituto para las Tecn. de la Inf. y las Comunic. ITIC (UNCUYO), Grupo de I+D+I en Comp. Ditrib. DisCo (UZ).XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025, Mendoza). Abril 2025.
[16]http://weblidi.info.unlp.edu.ar
[17]https://lissi.cs.uns.edu.ar
[18]http://vyglab.cs.uns.edu.ar
[19]http://lidic.unsl.edu.ar
[20] https://grupsderecerca.uab.cat/hpc4eas/
[21]http://smile.esi.uclm.es
[22]https://artecs.dacya.ucm.es
[23]http://www.litrp.cl
[24]http://www.lcg.unsl.edu.ar/
[25] http://itic.uncu.edu.ar/
[26] http://webdiis.unizar.es/DISCO/
[27] http://www.hpca.ual.es/
[28] https://itt.unnoba.edu.ar

Consorcio UNI-UEAR - Alianza Universitaria Argentina Europea para la Transformación Digital Cursos de Capacitación Docente RedUNCI

Adrian Pousa, Ariel Pasini, Cecilia Sanz, Facundo Quiroga, Federico Walas, Franco Ronchetti, Ismael Rodriguez, Juan Fernandez Sosa, Leandro Libutti, Pablo Thomas, Santiago Medina, Patricia Pesado, Armando De Giusti

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{apousa, apasini, csanz, fquiroga, fwalas, fronchetti, ismael, jfernandez, llibutti, pthomas, smedina, ppesado, degiusti}@lidi.info.unlp.edu.ar

RESUMEN

Este proyecto internacional tiene como objetivo contribuir al desarrollo y promoción de los derechos digitales en Argentina, incrementar la cooperación entre Europa y Argentina en temas de Transformación Digital mediante la generación y difusión de conocimiento en el tema, y promover alianzas del sector académico con instituciones del sector público, privado, y de la sociedad civil en temas de transformación digital. Cofinanciado por la Unión Europea, el proyecto se encuentra bajo la coordinación del CIN y participan 5 Universidades Nacionales y 4 Instituciones Europeas.
En este contexto se describen los 3 cursos que la UNLP dictará como Capacitaciones para docentes de la RedUNCI durante 2026.

Palabras Claves
Transformación Digital – Tecnologías para la TD – Innovación – Gobernanza de Datos

CONTEXTO

El proyecto "*UNI-UEAR - Alianza Universitaria Argentina Europea para la Transformación Digital*" cofinanciado por la Unión Europea, es ejecutado por el Consejo Interuniversitario Nacional (CIN) en colaboración con un consorcio de cinco universidades nacionales: Universidad de Buenos Aires (UBA), Universidad Nacional de La Plata (UNLP), Universidad Nacional de Quilmes (UNQ), Universidad Nacional del Sur (UNS) y Universidad Nacional de San Martín (UNSAM) y cuatro instituciones europeas: European University Institute (Italia), Gdansk University of Technology (Polonia), Technical University of Delft (Países Bajos), y University for Continuous Education Krems (Austria). Tiene como objetivo principal contribuir a la promoción de los derechos digitales en Argentina; y en particular, busca fortalecer la creación y difusión de conocimiento aplicado para la transformación digital; y desarrollar alianzas entre la academia, los sectores público y privado y la sociedad civil para la transformación digital.

El Proyecto se ejecuta bajo el acuerdo NDICI HR INTPA/2022/437-643 -- OPSYS: PC-20191.

1. OBJETIVOS DEL PROYECTO

El objetivo general del proyecto es contribuir al desarrollo y promoción de los derechos digitales en Argentina. A partir de

este, se derivaron tres objetivos específicos:

- Incrementar la cooperación entre Europa y Argentina en temas de transformación digital.
- Generar conocimiento en temas de transformación digital y difundirlo.
- Desarrollar alianzas del sistema académico con diferentes actores públicos y privados, en proyectos y desarrollos vinculados con la transformación digital.

2. RESULTADOS OBTENIDOS/ESPERADOS

La UNLP participó en la confección de 3 cursos desarrollados en el marco del Proyecto UNIUEAR:

➢ Tecnologías para la TD

El curso se enfoca en discutir las definiciones de Transformación Digital y su relación con los vectores de cambio tecnológico en la sociedad actual, analizando las nuevas tecnologías como vectores del cambio tecnológico que impulsan la Transformación Digital (Inteligencia Artificial, Computación en la Nube, Big Data, Internet de las Cosas, Blockchain) y el cambio de procesadores y comunicaciones que sustentan estas tecnologías y sus aplicaciones en diferentes áreas de la sociedad (Industria / Empresas / Organizaciones Públicas / Educación), discutiendo casos concretos de Transformación Digital.

Objetivos de aprendizaje:

- Discutir las definiciones de Transformación Digital y su relación con los vectores de cambio tecnológico en la sociedad actual.
- Analizar las nuevas tecnologías como vectores del cambio tecnológico que impulsan la Transformación Digital (Inteligencia Artificial, Computación en la Nube, Big Data, Internet de las Cosas, Blockchain) y el cambio de procesadores y comunicaciones que sustentan estas tecnologías y sus aplicaciones.
- Discutir el impacto del empleo de estas nuevas tecnologías en diferentes áreas de la sociedad (Industria / Empresas / Organizaciones Públicas / Educación)
- Analizar casos concretos como ejemplos del proceso de Transformación Digital, con el empleo de nuevas tecnologías.

➢ Gestión de la innovación

Este curso trata de distinguir entre las diferentes formas de innovación, ilustrando el rol de la innovación en los procesos de transformación digital. Se analizan casos de innovación en productos y servicios, examinando el papel y el impacto de la investigación en el desarrollo de nuevos productos. Asimismo se discuten diferentes estrategias de transferencia de tecnología para productos/servicios innovadores. Se pone énfasis en detallar los pasos del proceso de desarrollo de un nuevo producto/servicio y se analizan casos concretos en la industria, la educación y los servicios sociales. Por último se expone la importancia de las habilidades requeridas en los recursos humanos involucrados.

Objetivos de aprendizaje:

- Distinguir entre las diferentes formas de innovación
- Ilustrar el papel de la innovación en la transformación digital.
- Demostrar conocimiento en los últimos desarrollos contemporáneos en innovación de productos/servicios. Instrumentos de las organizaciones.
- Examinar el papel y el impacto de la investigación en el desarrollo de nuevos productos.
- Contrastar diferentes estrategias de transferencia de tecnología para productos/servicios innovadores.
- Analizar el papel de la gestión de nuevos productos/servicios en el proceso de innovación.
- Construir los pasos en el proceso de desarrollo de nuevos productos.
- Inspeccionar las características específicas del desarrollo de nuevos servicios.
- Analizar la aplicación de innovación a los servicios sociales, la industria y la educación. Análisis de casos.
- Emplear habilidades de resolución de problemas, conocimientos esenciales y

habilidades interpersonales para trabajar y tomar decisiones de manera efectiva como equipo.

- Utilizar "design thinking" en el proceso de innovación.

➢ **Gobernanza de datos**

El curso se enfoca en comprender la importancia de la Gobernanza de Datos en el marco de la Transformación Digital. Se analiza el "ciclo de vida" de los datos, así como herramientas utilizables en la gestión de datos. Se pone énfasis en los aspectos de seguridad y privacidad de los datos, así como en aspectos del análisis inteligente de datos, en particular de big data. Por último se discuten casos en los que la aplicación de la gobernanza de datos conduce a una mejor toma de decisiones en diferentes organizaciones.

Objetivos de aprendizaje:

- Comprender la importancia del gobierno de datos en el marco de la Transformación Digital. Interpretar el concepto de "ciclo de vida de los datos"
- Demostrar conocimiento de las tecnologías de manejo de datos (modelos/herramientas) en diferentes ámbitos.
- Analizar los temas de seguridad y privacidad de los datos. Gobierno, gestión y operación de la seguridad.
- Analizar riesgos en la seguridad y privacidad de los datos. Estudio de casos de incidentes.
- Comprender la Inteligencia de datos: la importancia para la Transformación Digital de ir de "datos" a "conocimiento" para la toma de decisiones.
- Analizar diferentes tecnologías de Minería de Datos. Impacto en casos de Transformación Digital.
- Estudiar conceptos y aplicaciones de Big Data en diferentes organizaciones.
- Estudiar casos de Análisis Inteligente de Datos (con diferentes herramientas tecnológicas) y su impacto en el proceso de Transformación Digital en distintas organizaciones.

3. FORMACION DE RECURSOS HUMANOS

Los 3 cursos serán dictados en forma virtual durante 2026.

- ➢ Tecnologías para la TD se dictarán en Marzo/26
- ➢ Gobernanza de datos se dictará en Mayo/26
- ➢ Gestión de la innovación se dictará en Septiembre/26

4. AGRADECIMIENTOS

Proyecto co-financiado por la Unión Europea. Acuerdo: NDICI HR INTPA/2022/437-643 -- OPSYS: PC-20191

5. BIBLIOGRAFIA

- Fawad A. Khan (2021). Digital Transformation using Emerging Technologies: A CxO's Guide to Transform your Organization. Jason M. Anderson (Author), 2021.
- PMI (2023). The Digital Transformation Playbook: What You Need to Know and D. Projec Management Institute PMI (Author) – PMI Editor – 2023.
- Siu Loon Hoe (2022). Digital Transformation: Strategy, Execution and Technology. Auerbach Publications, 2022,
- Borek, A., Prill, N. (2023). Driving Digital Transformation through Data and AI: A Practical Guide to Delivering Data Science and Machine Learning Products. Kogan Press, 2023.
- Cubo, J. L. Hernández Carrión, M. Porrúa, B. Roseth, *Guía de transformación digital del gobierno*.
- Banco Interamericano de Desarrollo (BID), Agosto 2022. DOI: http://dx.doi.org/10.18235/0004402
- Ndulu, E. Stuart, S. Dercon, P. Knaack, *Driving Digital Transformation: Lessons from Seven Developing Countries*. Oxford University Press, 2023.
- S. Loon Hoe, *Digital Transformation: Strategy, Execution and Technology*. New York: Auerbach Publications, 2022.
- N. Chopra, *Accelerated Digital Transformation: How Established Organizations Can Gain Competitive Advantage in the Digital Age*. Kogan Press, 2023.
- Kaplan, *Digital Transformation and Disruption of Higher Education.* Cambridge Press, 2022.
- Evergreen, *Autonomous Transformation: Creating a More Human Future in the Era of Artificial Intelligence*. Wiley, 2023.
- N. J. Medoff, B. K. Kaye, *Now Media: The*

Evolution of Electronic Communication. Routledge, 2021.

§ Bertot, J. C., Estevez, E., & Janowski, T. (2016). Digital Public Service Innovation: Framework proposal. ACM International Conference Proceeding Series, 01-03-March-2016(March), 113–122. https://doi.org/10.1145/2910019.2910108

§ Bertot, J., Estevez, E., & Janowski, T. (2016). Universal and contextualized public services: Digital public service innovation framework. Government Information Quarterly, 33(2), 211–222. https://doi.org/10.1016/j.giq.2016.05.004

§ Sánchez, C., Lasagna, M., & Marcet, X. (2013). Innovación pública: RIL® editores.

§ Herbert Endres · Stefan Huesig · Robin Pesch. "Digital innovation management for entrepreneurial ecosystems: services and functionalities as drivers of innovation management software adoption" Review of Managerial Science (2022) 16:135–156 https://doi.org/10.1007/s11846-021-00441-4 Springer.

§ Rowan Gibson "The Four Lenses of Innovation: A Power Tool for Creative Thinking" John Wiley & Sons, Inc. 2015

§ Seppo Poutanen, Anne Kovalainen "Skills, Creativity and Innovation in the Digital Platform Era: Analyzing the New Reality of Professions and Entrepreneurship" Routledge Studies in Entrepreneurship) 1st Edición. 2023

§ Basco, A., Beliz, G., Coatz, D. y Garnero, P. (2018). industria 4.0: fabricando el futuro. bid-intal-uia.

§ Wollschlaeger, M., Sauter, T., Jasperneite J,. (2017) The Future of Industrial Communication. Automation Networks in the Era of the Internet of Things and Industry 4.0. march 2017 ■ IEEE industrial electronics magazine

Agencia de Acceso a la Información Pública (AAIP). (2024). Guía para entidades públicas y privadas en materia de transparencia y protección de datos personales para una IA responsable. https://www.argentina.gob.ar/sites/default/files/aaip-argentina-guia_para_usar_la_ia_de_manera_responsable.pdf

APRA. (2013). Prudential Practice Guide CPG 235 — Managing Data Risk. https://www.apra.gov.au/managing-data-risk (PDF: https://www.apra.gov.au/sites/default/files/Prudential-Practice-Guide-CPG-235-Managing-Data-Risk_1.pdf)

Argentina. (2000). Ley 25.326 de Protección de los Datos Personales (texto actualizado). InfoLEG. https://servicios.infoleg.gob.ar/infolegInternet/verNorma.do?id=64790

Data Governance Institute. (s. f.). Principles & Frameworks. https://datagovernance.com/

Ramez Elmasri, Shamkant Navathe. Fundamentals of Database Systems 7th Edition, Pearson, 2015. ISBN: 978-0133970777.

Bertone, R., Thomas, P. (2011). Introducción a las bases de datos: fundamentos y diseño. Prentice Hall / Pearson Educación, 2011.

Netscher, S., Anders, I., & Henzen, C. (2022). Activities and Challenges in developing Discipline-Specific Data Management Plan Templates: From Vertical to Horizontal Integration of RDM Practices. Bausteine Forschungsdatenmanagement, (1), 13-25.

European Commission. (2016). H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020 [Guide]. European Commission.

European Research Council Scientific Council. (2022, April 20). Open Research Data and Data Management Plans: Information for ERC grantees (Version 4.1). ERC.

Data Mining: Practical Machine Learning Tools and Techniques. Ian Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal. Morgan Kaufmann Publishers. 2017.

Introducción a la Minería de Datos. Hernández Orallo, Ramirez Quintana, Ferri Ramirez. Pearson – Prentice Hall. 2004

Learning Data Mining with Python. Robert Layton. Packt Publishing. 2017 (segunda edición)

Big Data for dummies. Judith Hurwitz, Alan Nugent, Dr. Fern Halper and Marcia Kaufman. John Wiley & Sons, Inc. ISBN 978-1-118-50422-2. 2013.

Participación en el Programa UNITA GEMINAE

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC - 526 e/ 10 y 11 La Plata Buenos Aires

Directora III-LIDI: Patricia Pesado [0000-0003-0000-3482]
Coordinadora en GEMINAE por el III-LIDI: Cecilia Sanz (1) [0000-0002-9471-0008]
(ppesado, csanz) @lidi.info.unlp.edu.ar

Resumen

Geminae es un programa de la Alianza Universitaria UNITA que busca crear una red de universidades para compartir buenas prácticas. El objetivo es mejorar la movilidad de los estudiantes entre las instituciones y fomentar que se establezcan en las regiones de UNITA. La red de universidades se basa en la calidad de las instituciones y en la cooperación activa entre universidades fuera de la Unión Europea y los socios de UNITA.

Palabras clave: Buenas prácticas – innovación en educación

Contexto

Desde el año 2023 el Instituto de Investigación en Informática LIDI (III-LIDI) participa en el programa UNITA Geminae.

Introducción

El Programa GEMINAE UNITA se orienta a generar una red de universidades para lograr el intercambio de buenas prácticas desarrolladas en los ámbitos de formación, investigación, innovación y vinculación con la sociedad.
El programa se basa en la cooperación activa entre universidades fuera de la Unión Europea: UNLP, ITBA, UNCuyo (Argentina), FBS, UB, UEC, UEOP, UFB, UFMG, UFV, UFRJ, URB (Brasil), UQTR, US (Canadá), ULH (Cuba), UCC, UA, UR, UNC (Colombia), UASLP, UM (México), UASD (Rep. Dominicana); UB (Argelia), UAC (Benín), 2IE (Burkina Faso), UCV (Cabo Verde), UA (Madagascar), UH2C (Marruecos), UPM (Mozambique), UNTL (Timor-Leste)) y los socios de UNITA (Agencia Universitaria de la Francofonía – AUF, Organización de Estados Iberoamericanos – OEI, Asociación de Universidades de Lengua Portuguesa – AULP.

Objetivos

Su objetivo es proporcionar competencias y habilidades que permitan a los estudiantes convertirse en futuros actores de economías sostenibles, basadas en la inclusión y la igualdad, teniendo un impacto local, regional y global.

En particular:

- Intercambiar buenas prácticas en formación, investigación, innovación y vinculación con la sociedad
- Facilitar la movilidad de estudiantes entre las instituciones
- Fomentar que los estudiantes se establezcan en las regiones de UNITA
- Crear currículums flexibles y personalizados
- Fomentar el compromiso cívico
- Introducir pedagogías innovadoras
- Lograr un impacto sostenible y un crecimiento azul

[1] Investigador Asociado de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

Resultados Obtenidos

- En el año 2025 se participó en el Congreso Geminae 2025: innovación, internacionalización y cooperación académica. Dos investigadores del III-LIDI realizaron una estancia en la Universidad de Zaragoza y realizaron dos presentaciones en el marco del Congreso Geminae. Al mismo tiempo, se recibió uno de los *Best practices Awards*[2] por el proyecto: "Abordaje de temas de salud mediados por tecnologías digitales. Trabajo interdisciplinario y de articulación entre universidad y escuelas". En este proyecto se co-diseñan de manera interdisciplinaria diferentes tecnologías digitales que ponen el foco en el aprendizaje de temas para la salud, y se ponen a disposición a través de talleres, charlas, muestras que permiten acercarlos a distintos jóvenes y escuelas.
- Se llevaron adelante reuniones periódicas con las universidades participantes y en la estancia en Zaragoza se fortalecieron vínculos con otras universidades como la Universidad de Pau (Francia) y la de Rumania.
- Se obtuvo un Erasmus KA171 para movilidad y fortalecimiento de los temas de interés de la red. Una de las movilidades al congreso fue fruto de este Erasmus. Se está realizando una nueva presentación de KA171 para este año.

Formación de Recursos Humanos

En vinculación con la Universidad de Zaragoza, que ha sido el nexo con el Programa GEMINAE UNITA se cuenta con una tesis doctoral de co-tutela en temas de interés conjunto. Se finalizó una tesis de maestría co-dirigida por una investigadora de Zaragoza. Se está por iniciar una nueva tesis doctoral en co-tutela. Se focalizan acciones de vinculación con el medio y transferencia en función de temas de interés de las universidades que participan en este programa. En particular, en relación a la Universidad de La Plata, la de Zaragoza y el resto de las universidades participantes de Latinoamérica.

[2] https://weblidi.info.unlp.edu.ar/conferencia-geminae-2025-premio-a-las-mejores-practicas-del-programa-geminae-unita/

Referencias

- Sitio web Programa UNITA GEMINAE. https://univ-unita.eu/Sites/UNITA/en/
- Astudillo, G., Asiain Ansorena, D., Sanz, C., Ponce De León, J., Beltrán Blázquez, J.R. and Baldassarri. S. 2024. Capturing physiological data for detecting boredom in educational contexts. In Proceedings of the XXIII International Conference on Human Computer Interaction (Interacción '23). Association for Computing Machinery, New York, NY, USA, Article 27, 1–5. https://doi.org/10.1145/3612783.3612811 **York, NY, USA, Article 27, 1–5. https://doi.org/10.1145/3612783.3612811**
- Astudillo, G. J., Sanz, C. V., & Baldassarri, S. (2025). Emotionally-tagged video database: a systematic review. Data Technologies and Applications, 1-26.
- Ferreyra, P., Coma-Roselló, T., & Sanz, C. V. (2024). Revisión sistemática sobre videojuegos como método de evaluación sigilosa para analizar rasgos de personalidad. In XIX Congreso Nacional de Tecnología en Educación y Educación en Tecnología-TE&ET 2024 (Morón, 13 y 14 de junio de 2024).
- Ferreyra, P., Dieser, P., Sanz, C., Coma-Roselló, T. (2026). Evaluación del rasgo responsabilidad a través de un videojuego serio de química y matemática. Revista Iberoamericana De Tecnología En Educación Y Educación En Tecnología, 42, e3. https://doi.org/10.24215/18509959.42.e3

Tesis Doctorales

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación

ANÁLISIS VISUAL DE DATOS MULTIDIMENSIONALES

Autora: Dra. Antonella S. Antonini
Directoras: Dra. Silvia Mabel Castro / Dra. María Luján Ganuza
Unidad académica: Departamento de Ciencias e Ingeniería de la Computación (Universidad Nacional del Sur)
Fecha de exposición: 26 de junio de 2025

MOTIVACIÓN

En la actualidad, nos encontramos inmersos en la era de los datos, donde la cantidad generada y almacenada crece a un ritmo sin precedentes. Los avances tecnológicos han impulsado un incremento exponencial en la producción de datos, llevando a volúmenes que superan cualquier estimación previa. En 2003, se calculaba que cada persona almacenaba aproximadamente 800 megabytes de datos digitales, duplicándose esta cifra cada tres años [LV03]. Hoy en día, esta tendencia no solo persiste, sino que se acelera: en 2023, la cantidad total de datos generados alcanzó los 120 zettabytes, y se proyecta que para 2025 esta cifra ascienda a 181 zettabytes.

Este crecimiento en la generación de datos no solo ha transformado la manera en que se almacenan y gestionan estos datos, sino que también ha introducido desafíos significativos en su procesamiento, análisis y aprovechamiento. Para que estos conjuntos de datos sean comprensibles y útiles para los seres humanos, es fundamental convertirlos en representaciones visuales adecuadas que permitan extraer información de forma clara y efectiva. Sin este paso, los datos crudos pueden resultar ininteligibles y difíciles de interpretar, incluso cuando su volumen es relativamente pequeño. En este sentido, la visualización de datos juega un papel crucial, ya que facilita la identificación de patrones, la formulación de hipótesis y la toma de decisiones fundamentadas, aprovechando así esta invaluable fuente de conocimiento.

En particular, en diversas áreas es habitual visualizar conjuntos de datos caracterizados por múltiples dimensiones, lo que introduce desafíos significativos en su representación gráfica. A medida que la dimensionalidad de los datos aumenta, lograr una visualización efectiva se vuelve más complejo. Las técnicas tradicionales, diseñadas para datos de pocas dimensiones, muchas veces resultan insuficientes para capturar la estructura de estos conjuntos de datos. Uno de los principales obstáculos en la visualización de datos multidimensionales es la sobrecarga visual, donde la visualización se vuelve tan densa y compleja que los usuarios no pueden procesar toda la información de manera efectiva.

Otro desafío considerable es la pérdida de información que puede ocurrir al reducir las dimensiones, lo que afecta la calidad de la interpretación de los datos. Para facilitar el análisis de grandes volúmenes de datos, en muchos casos es necesario representarlos en un espacio de baja dimensionalidad, generalmente de dos o tres dimensiones [DKZ13]. Las representaciones bidimensionales son más fáciles de explorar, y la percepción humana está naturalmente adaptada al mundo tridimensional que nos rodea. En algunos casos, puede ser aceptable reducir la dimensionalidad eliminando variables, lo que Inselberg denominó “mutilar dimensiones” [GLI98]. Sin embargo, limitarse a un subconjunto de dimensiones podría resultar en una representación

simplificada que no capture adecuadamente la complejidad inherente de los datos. Además, la selección de las dimensiones más relevantes es un proceso subjetivo, que depende del conocimiento profundo del dominio en cuestión y de los objetivos del análisis. En otros escenarios, es crucial analizar todas las dimensiones de manera conjunta, ya que solo al considerarlas en su totalidad se pueden revelar relaciones, patrones complejos y estructuras subyacentes que no serían evidentes si se redujeran o eliminaran dimensiones. En este contexto, la reducción de la dimensionalidad se presenta como una herramienta útil, al proyectar los elementos de datos desde un espacio de alta dimensión a uno de menor dimensión preservando la mayor cantidad posible de la información original.

La representación visual de las relaciones entre las variables también puede ser un desafío en el análisis de conjuntos de datos multidimensionales. Como señaló Jacques Bertin [Ber83], la información no solo consiste en elementos de datos individuales, sino también en las relaciones entre ellos. Sin embargo, a medida que el número de variables aumenta, muchas de las técnicas de visualización existentes no logran representar adecuadamente estas relaciones.

A pesar de estos desafíos, los avances en el poder computacional y la creciente disponibilidad de herramientas y algoritmos especializados han facilitado considerablemente el análisis de datos multidimensionales. Comprender en profundidad los métodos disponibles, identificar sus capacidades y limitaciones, seleccionar cuidadosamente las técnicas más adecuadas, y desarrollar soluciones alternativas es crucial para extraer información relevante de conjuntos de datos complejos y aprovechar todo su potencial.

APORTES DE LA INVESTIGACIÓN

La era digital ha traído consigo una explosión sin precedentes en la generación y acumulación de datos, presentando retos significativos para su procesamiento y análisis. La visualización de datos desempeña un papel crucial, ya que permite convertir grandes y complejos conjuntos de datos en representaciones visuales comprensibles. Sin embargo, a medida que la dimensionalidad de los datos aumenta, el diseño de visualizaciones efectivas se convierte en un desafío cada vez mayor.

En este contexto, esta tesis propone soluciones innovadoras que optimizan el proceso de visualización mediante el desarrollo de un marco metodológico robusto, la realización de un análisis exhaustivo de las técnicas de visualización más utilizadas y sistemas de recomendación de técnicas existentes, y la implementación de herramientas innovadoras. A lo largo de este trabajo se abordan los aspectos fundamentales de la visualización, estructurados bajo un marco metodológico basado en las tres preguntas clave *"¿Qué? - ¿Por qué? - ¿Cómo?"* definidas por Munzner [Mun14], que estructuran el proceso de visualización en función de los datos, su representación visual y los objetivos analíticos. Este enfoque proporciona una base sólida para diseñar visualizaciones efectivas, adaptadas a las necesidades y requerimientos específicos de los usuarios.

Además, se lleva a cabo un análisis exhaustivo de las técnicas de visualización más utilizadas, las cuales se agrupan en una nueva taxonomía que incluye enfoques basados en geometría, íconos, pixeles, jerarquías y grafos, y se propone una taxonomía unificada de tareas que facilita la interpretación y aplicación de las tareas analíticas en el diseño de visualizaciones. Asimismo, se diseña y desarrolla un sistema de recomendación integral que optimiza la selección de técnicas de visualización, considerando las propiedades de los datos, las preferencias de representación visual y los objetivos analíticos del usuario. Por otra parte, se desarrollan e implementan enfoques y herramientas especializadas para la visualización de datos multidimensionales, con un enfoque

particular en las Ciencias Geológicas. Estas soluciones combinan técnicas tradicionales de visualización con métodos avanzados de aprendizaje automático para mejorar el análisis de datos complejos.

A continuación, se detallan los principales aportes de este trabajo de investigación:

Relevamiento y análisis de los modelos de visualización y definición del marco metodológico.

Se realiza un relevamiento y análisis de los modelos de visualización con el objetivo de definir un marco metodológico que estructure el proceso de visualización. Para ello se adopta el enfoque basado en las tres preguntas clave propuestas por Munzner [Mun14] *"¿Qué? - ¿Por qué? - ¿Cómo?"* que permiten relacionar los datos, las tareas analíticas y la representación visual.

En este contexto se estudian distintos modelos conceptuales de la literatura que explican cómo los datos se transforman en visualizaciones comprensibles. El pipeline de Card [CMS99] describe un proceso secuencial desde los datos crudos hasta su interpretación mediante transformaciones de datos, mapeos visuales e interacción. El modelo de estados de Chi [Chi00] amplía esta visión permitiendo intervenir en etapas intermedias del proceso y no solo en la vista final. Por su parte, el Modelo Unificado de Visualización (UVM) formaliza el flujo en estados bien definidos desde los datos originales hasta los datos visualizados [MCFE03]. En conjunto, estos modelos consideran la visualización como un proceso de transformación guiado por los datos, los objetivos del usuario y la representación gráfica.

A partir de este análisis se describe en detalle el Modelo Unificado de Análisis Visual (UVAM), que integra las características del Modelo Unificado de Visualización con aspectos específicos del análisis visual [GULC23, MCFE03]. El UVAM constituye el marco conceptual central de la tesis y la base para los desarrollos posteriores (ver Figura 1).

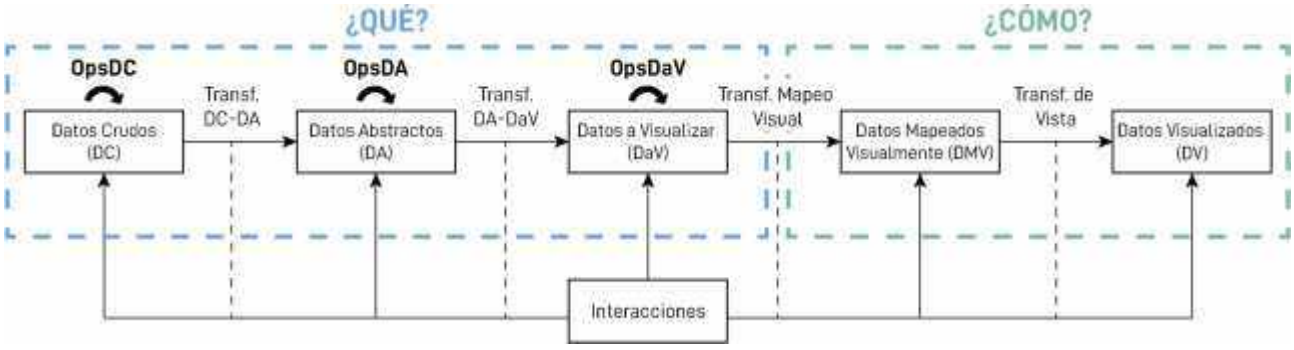


Figura 1: Modelo Unificado de Análisis Visual (UVAM) presentado por Ganuza *et al.* en el contexto de los tres pilares de la visualización. Figura adaptada de [GULC23].

Análisis de los tres pilares del marco *"¿Qué? - ¿Por qué? - ¿Cómo?"*.

Se lleva a cabo un análisis exhaustivo de los tres pilares fundamentales del marco propuesto: los datos, su representación visual y los objetivos dentro del proceso de visualización (ver Figura 2). Este análisis incluye una revisión detallada de la taxonomía de los datos, el estudio de los componentes que conforman la estructura visual, y un relevamiento exhaustivo de las diferentes taxonomías de tareas presentes en la literatura.

En primer lugar, se caracterizan los datos, analizando su estructura, tipo y dimensionalidad. Se describen distintas clasificaciones relevantes para la visualización, tales como datos categóricos, ordinales y cuantitativos, así como estructuras tabulares, espaciales, temporales y multidimensionales. Esta caracterización resulta esencial porque condiciona qué transformaciones son posibles y qué representaciones pueden utilizarse sin introducir ambigüedad o pérdida significativa de información.

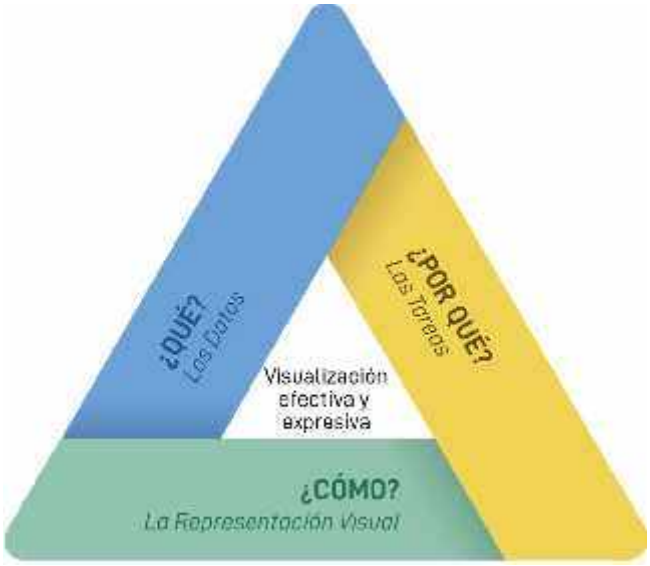


Figura 2: Los tres pilares de la visualización. Los modelos de visualización pueden estructurarse utilizando el marco *"¿Qué? - ¿Por qué? - ¿Cómo?"*: los datos, las tareas y la representación visual, para diseñar visualizaciones tanto efectivas como expresivas.

Luego se estudian las tareas analíticas, es decir, los objetivos que el usuario busca resolver mediante la visualización. Se identifican tareas generales como búsqueda, comparación, detección de correlaciones, identificación de anomalías, análisis de distribución y agrupamiento. La elección de la visualización no depende únicamente de los datos sino principalmente de la tarea a realizar, ya que una misma técnica puede ser adecuada para ciertos objetivos e ineficiente para otros.

Finalmente se analizan las representaciones visuales, entendidas como el conjunto de codificaciones gráficas utilizadas para comunicar la información. Se estudian los canales visuales (posición, color, tamaño, forma, orientación, textura) y su capacidad perceptual para transmitir distintos tipos de variables. Asimismo, se discute la relación entre codificación visual y precisión perceptual, estableciendo criterios para seleccionar representaciones efectivas.

Este relevamiento sirvió como base para desarrollar una taxonomía de tareas, que unifica las terminologías empleadas en la literatura para describir las diversas tareas analíticas. Además, se exploran las interacciones aplicables a los distintos estados y transformaciones del proceso de visualización.

Relevamiento exhaustivo de técnicas de visualización para datos multidimensionales.

Se describen las técnicas de visualización más empleadas en el análisis de datos multidimensionales y se agrupan en una nueva taxonomía, que abarca enfoques basados en geometría, íconos, pixeles, jerarquías y grafos. El objetivo en este punto consiste en ofrecer una comprensión profunda de cada técnica, sus características y aplicaciones específicas.

Se profundiza en la representación visual como componente fundamental del análisis visual, abordando las técnicas utilizadas para mostrar datos multidimensionales y sus capacidades para soportar distintas tareas analíticas.

Entre ellas se incluyen las técnicas geométricas, que proyectan los datos en espacios de menor dimensión (como proyecciones y coordenadas paralelas); las técnicas basadas en íconos, que representan cada elemento mediante glifos con múltiples atributos visuales; las técnicas basadas en píxeles, orientadas a grandes volúmenes de datos mediante representaciones densas; las técnicas jerárquicas, que organizan la información según estructuras de contención; y las técnicas basadas en grafos, utilizadas para representar relaciones entre entidades.

Para cada una de estas técnicas se analizan sus ventajas, limitaciones y tipos de tareas que soportan, destacando que no existe una visualización universalmente óptima, sino que su efectividad depende del objetivo analítico y de las características de los datos. También se discute el rol de la interacción —filtrado, selección, zoom, reordenamiento— como elemento indispensable para explorar la información y reducir la sobrecarga visual.

Finalmente, se establecen criterios de selección de técnicas basados en la relación entre los datos, las tareas y la representación visual, sentando las bases para el posterior desarrollo del sistema de recomendación propuesto en la tesis.

Relevamiento de sistemas de recomendación de técnicas de visualización.

Se realiza un relevamiento de los sistemas existentes de recomendación de técnicas de visualización, evaluando sus metodologías y criterios de selección, tales como la naturaleza de los datos, los objetivos del análisis, el conocimiento del dominio y las preferencias de visualización del usuario. Este análisis establece el marco conceptual fundamental para la creación del sistema de recomendación propuesto.

Al analizar un conjunto de datos, la elección de una representación visual adecuada no es trivial, especialmente en datos multidimensionales, y que suele depender de la experiencia del analista. Por ello, distintos trabajos han propuesto sistemas automáticos de recomendación de visualizaciones.

En primer lugar, se revisan enfoques existentes en la literatura, los cuales pueden clasificarse según la información que utilizan para recomendar. Algunos sistemas se basan únicamente en las características de los datos (tipo de variables, cardinalidad o dimensionalidad) [Gna81, Mac86, MHS07], otros consideran las tareas analíticas que el usuario desea realizar [RM90, WL90, Cas91], y otros incorporan preferencias del usuario [GW09, MVT17]. Sin embargo, la mayoría de estos enfoques analiza solo una parte del problema y no integra simultáneamente datos, tareas y representación visual, lo que limita la pertinencia de las recomendaciones.

A partir de estas limitaciones, se presenta el sistema de recomendación propuesto en la tesis, denominado CVRS. El sistema se fundamenta en el marco metodológico definido previamente y relaciona explícitamente los tres pilares del análisis visual: los datos, las tareas y la representación

visual. Para ello utiliza reglas de conocimiento derivadas del modelo conceptual, evaluando tipo de datos, número de dimensiones, estructura, tareas a resolver y preferencias del usuario.

CVRS supera las limitaciones de enfoques previos al integrar múltiples factores del proceso de visualización, contribuyendo a mejorar la interpretabilidad, la reproducibilidad del análisis y la accesibilidad para usuarios no expertos.

Diseño e implementación de un sistema de recomendación integral de visualización.

Se diseña y desarrolla un sistema de recomendación de técnicas de visualización que integra factores clave para la selección de las representaciones visuales más apropiadas. Este sistema se destaca por considerar los tres pilares fundamentales de la visualización en su recomendación: las características del conjunto de datos, los objetivos analíticos del usuario y sus preferencias de representación visual.

CVRS se basa en una base de conocimiento construida a partir del modelo de análisis visual, donde se establecen relaciones explícitas entre características de los datos, las tareas analíticas y las técnicas de visualización. Para ello se definen parámetros de entrada que describen el problema de análisis, tales como tipo y estructura de los datos, dimensionalidad, objetivo del usuario y preferencias de exploración. A partir de estos parámetros, el sistema evalúa un conjunto de reglas que determinan qué representaciones son adecuadas, cuáles son posibles, pero no óptimas y cuáles deben descartarse. Ante una posible configuración del usuario, el sistema pondera alternativas y genera un conjunto ordenado de recomendaciones (ver Figura 3).

El sistema cuenta además con una interfaz simple e intuitiva de interacción, que permite al usuario especificar su problema de análisis y recibir sugerencias junto con una justificación basada en las características detectadas.

El trabajo realizado incluye ejemplos de funcionamiento donde el CVRS recomienda distintas técnicas para un mismo conjunto de datos dependiendo de la tarea planteada, demostrando su capacidad de adaptación al contexto analítico. De esta manera, el sistema no solo automatiza la selección de visualizaciones, sino que actúa como guía conceptual durante el proceso de análisis.

Diseño e implementación de sistemas especializados para la visualización de datos multidimensionales.

Además, se diseñan e implementan diversos enfoques y herramientas innovadoras para la visualización de datos multidimensionales, con un enfoque particular en las Ciencias Geológicas. Estos sistemas combinan técnicas tradicionales de visualización con métodos avanzados de aprendizaje automático, ofreciendo soluciones integradas para la representación y análisis de conjuntos de datos complejos. Herramientas como Spinel Web [AGF+21], VISUEL [AGC22], GLC-Frame [LAGC24] y SpinelVA [ALF+24] se destacan como contribuciones clave, al facilitar la exploración y el análisis visual interactivo de grandes volúmenes de datos.

Además, se ha formalizado una taxonomía que abarca todas las características capaces de describir cualquier técnica de visualización NP-GLC en 2D, así como las interacciones asociadas [ALGC23]. Esta taxonomía ha sido fundamental para estructurar el análisis de estas técnicas, y guiar el diseño y desarrollo de nuevas soluciones NP-GLC. Además, a través de casos de estudio específicos, se ha resaltado el potencial y el impacto de estos sistemas en el análisis de datos multidimensionales.

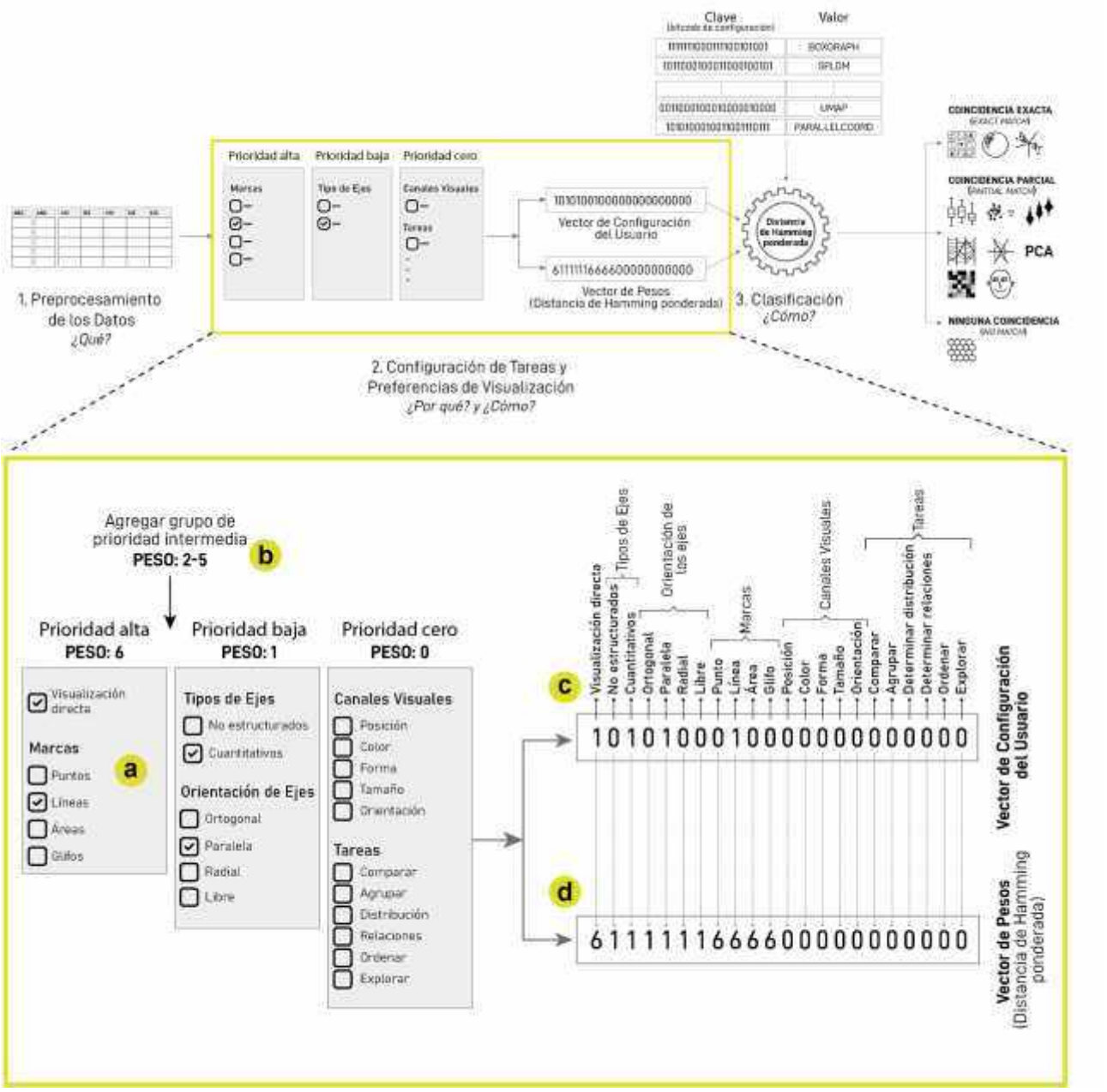


Figura 3: Arquitectura del sistema CVRS. (1) Pre procesamiento de los datos: se clasifican los datos en categóricos y numéricos, permitiendo al usuario seleccionar las columnas clave para su visualización. (2) Configuración de tareas y preferencias de visualización: el usuario define los parámetros que orientarán la visualización de los datos de acuerdo con sus preferencias de representación visual y objetivos de análisis. (3) Clasificación: el sistema selecciona las técnicas de visualización más adecuadas para comunicar de manera efectiva la información.

Se describen las distintas aplicaciones implementadas para validar en la práctica los modelos y métodos propuestos en la tesis. Estas herramientas constituyen casos reales de uso donde se integran técnicas de visualización multidimensional, interacción y análisis visual asistido, principalmente en el dominio geoquímico.

Spinel Web permite la exploración de datos geoquímicos desde un entorno web interactivo, facilitando el acceso y análisis remoto. VISUEL funciona como un entorno general de análisis visual que integra múltiples técnicas de visualización y mecanismos de interacción. GLC-Frame está orientado al análisis composicional y al estudio de relaciones entre elementos químicos, mientras que SpinelVA incorpora técnicas de aprendizaje automático para apoyar la interpretación de patrones.

Cada aplicación explota diferentes representaciones visuales y estrategias de interacción, demostrando cómo los modelos conceptuales definidos en la tesis pueden materializarse en herramientas concretas. Los sistemas permiten detectar agrupamientos, correlaciones y comportamientos anómalos en datos reales, mostrando mejoras respecto de análisis estadísticos tradicionales.

LÍNEAS DE INVESTIGACIÓN FUTURAS

En base al trabajo realizado, se identificaron diversas direcciones futuras de investigación que complementan y amplían el enfoque desarrollado.

A continuación, se detallan algunas de estas posibles líneas de investigación:

Avance en el Desarrollo e Implementación de Nuevas Técnicas de Visualización para el Análisis de Datos Multidimensionales. El propósito es ampliar y perfeccionar las técnicas de visualización actuales, creando nuevos modelos, herramientas y enfoques que faciliten la exploración de datos por geólogos y otros expertos. Aunque el enfoque principal está dirigido a las Ciencias Geológicas, estas metodologías tienen una aplicación más amplia y pueden ser empleadas en diversas disciplinas y áreas de aplicación que también demandan el análisis de datos multidimensionales.

Extensión del CVRS para el Análisis y Representación de Redes. En el prototipo inicial de CVRS, el sistema de recomendación de técnicas de visualización se ha diseñado principalmente para trabajar con conjuntos de datos tabulares, un formato accesible, fácil de comprender y ampliamente utilizado en múltiples áreas de aplicación. Sin embargo, el análisis de datos no siempre se ajusta a una estructura tabular. En algunos contextos, se requiere el análisis de datos organizados como redes, en los cuales los elementos (nodos) están interconectados mediante enlaces, formando estructuras complejas. Como parte del trabajo futuro, se plantea ampliar las capacidades del sistema CVRS para incluir el análisis y la recomendación de técnicas de visualización dirigidas a conjuntos de datos estructurados como redes. Esta extensión permitirá que el sistema se pueda aplicar en una mayor variedad de contextos y disciplinas.

Diseño e Implementación de un Sistema de Recomendación Integral de Visualización para el Análisis de Datos Geoquímicos. Se propone el diseño y desarrollo de un sistema integral para la recomendación de técnicas de visualización dirigidas al análisis de datos geoquímicos. Este sistema incorporará gráficos especializados utilizados por geólogos, así como métodos de visualización apropiados para representar datos multidimensionales. Además, para ofrecer recomendaciones más precisas, el sistema incorporará un cuarto factor clave en el proceso de recomendación, sumado a los ya considerados en el sistema propuesto anteriormente: el conocimiento del dominio, tal como lo destacan Kaur y Owonibi [KO17]. De este modo, el sistema no solo sugerirá las técnicas de visualización más apropiadas en función de los datos, los objetivos del análisis y las preferencias de representación visual, sino que también asegurará una alineación precisa con las prácticas y requisitos específicos del análisis geoquímico.

CONCLUSIONES

La visualización, como un puente entre los datos y la interpretación humana, emerge como una herramienta clave para enfrentar estos desafíos. No obstante, el diseño y la selección de representaciones visuales efectivas y expresivas, requieren no solo una comprensión profunda de las características intrínsecas de los datos, sino también una comprensión integral de los objetivos analíticos y las preferencias del usuario.

En este contexto, la presente tesis integra avances conceptuales y prácticos destinados a abordar los desafíos de la visualización de datos multidimensionales. Este trabajo propone soluciones que optimizan el proceso de visualización mediante el desarrollo de un marco metodológico robusto, la realización de un análisis exhaustivo de técnicas de visualización y sistemas de recomendación existentes, y la implementación de herramientas innovadoras. Este enfoque integral no solo mejora la representación y comprensión de los datos multidimensionales, sino que también sienta las bases para futuros avances en el análisis de datos en un entorno cada vez más dinámico y desafiante.

BIBLIOGRAFÍA

[AGC22] Antonini, A. S., Ganuza, M. L., and Castro, S. M. Visuel - a web dynamic dashboard for data visualization. *Journal of Computer Science and Technology 22*, 1 (2022).

[AGF+21] Antonini, A. S., Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Matkovic, K., Gröller, E., Bjerg, E. A., and Castro, S. M. Spinel web: An interactive web application for visualizing the chemical composition of spinel group minerals. *Earth Science Informatics 14*, 1 (2021), 521–528.

[ALF+24] Antonini, A. S., Luque, L., Ferracutti, G. R., Bjerg, E. A., Castro, S. M., and Ganuza, M. L. Spinelva. a new perspective for the visual analysis and classification of spinel group minerals. *Earth Science Informatics 17*, 4 (2024), 3851–3861.

[ALGC23] Antonini, A. S., Luque, L., Ganuza, M. L., and Castro, S. M. Toward a taxonomy for 2d non-paired general line coordinates: A comprehensive survey. *International Journal of Data Science and Analytics 15*, 2 (2023), 133–158.

[Ber83] Bertin, J. Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press, Madison, WI, 1983.

[Cas91] Casner, S. M. Task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics (TOG) 10*, 2 (1991), 111–151.

[Chi00] Chi, E. A taxonomy of visualization techniques using the data state reference model. In IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings (2000), pp. 69–75.

[CMS99] Card, S., Mackinlay, J., and Shneiderman, B. Readings in Information Visualization: Using Vision To Think. Morgan Kaufmann, 1999.

[DKZ13] Dzemyda, G., Kurasova, O., and Zilinskas, J. Multidimensional Data Visualization: Methods and Applications, vol. 75 of Springer Optimization and Its Applications. Springer, 2013.

[GLI98] Grinstein, G., Laskowski, S., and Inselberg, A. Key problems and thorny issues in multidimensional visualization. In Proceedings of the conference on Visualization '98 (1998), IEEE Computer Society Press, pp. 505–506.

[Gna81] Gnanamgari, S. *Information Presentation Through Default Displays*. PhD thesis, University of Pennsylvania, USA, 1981.

[GULC23] Ganuza, M., Urribarri, D., Larrea, M., and Castro, S. Uvam: The unified visual analytics model. the unified visualization model revisited. In Cloud Computing, Big Data & Emerging Topics (2023), Springer Nature Switzerland, pp. 189–203.

[GW09] Gotz, D., and Wen, Z. Behavior-driven visualization recommendation. In *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI '09)* (New York, NY, USA, 2009), Association for Computing Machinery, pp. 315–324.

[KO17] Kaur, P., and Owonibi, M. A review on visualization recommendation strategies. In Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017) - IVAPP (2017), vol. 3, SciTePress, pp. 266–273

[LAGC24] Luque, L., Antonini, A. S., Ganuza, M. L., and Castro, S. M. GLC-Frame: A framework and library for exploration of multidimensional data with general line coordinates. *Journal of Computer Science and Technology*
24, 1 (2024).

[LV03] Lyman, P., and Varian, H. R. How much information? The Journal of Electronic Publishing 9, 1 (2003).

[Mac86] Mackinlay, J. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics (TOG) 5*, 2 (1986), 110–141.

[MCFE03] Martig, S. R., Castro, S. M., Fillottrani, P. R., and Estevez, E. C. Un modelo unificado de visualización. In IX Congreso Argentino de Ciencias de la Computación (2003), Universidad Nacional de La Plata, pp. 881–892.

[MHS07] Mackinlay, J., Hanrahan, P., and Stolte, C. Show me: Automatic presentation for visual analysis. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1137–1144.

[Mun14] Munzner, T. Visualization Analysis and Design. AK Peters Visualization Series. CRC press, 2014.

[MVT17] Mutlu, B., Veas, E., and Trattner, C. Tags, titles or qas? choosing content descriptors for visual recommender systems. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media (HT '1)* (New York, NY, USA, 2017), Association for Computing Machinery, pp. 265–274.

[RM90] Roth, S. F., and Mattis, J. Data characterization for intelligent graphics presentation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90)* (New York, NY, USA, 1990), Association for Computing Machinery, pp. 193–200.

[WL90] Wehrend, S., and Lewis, C. A problem-oriented classification of visualization techniques. In *Proceedings of the 1st Conference on Visualization (VIS '90)* (Washington, DC, USA, 1990), IEEE Computer Society Press, pp. 139–143.

Análisis Visual de Datos provenientes de Registradores Oculares

Autor: Leandro Emanuel Luque

Directoras: Dra. Silvia Castro - Dra. Maria Luján Ganuza

Institución: Universidad Nacional del Sur

Fecha de defensa: 10 de Julio de 2025

Motivación

En los últimos años el uso de registradores oculares se ha extendido en distintos campos de aplicación como la psicología, la medicina, el marketing y los videojuegos, entre los más relevantes. La capacidad de poder obtener datos tan relevantes como el tamaño de la pupila y la posición de los ojos de forma no invasiva, permite extender nuestra comprensión del comportamiento de las personas en determinadas circunstancias. Con estos resultados se pueden diseñar nuevas experiencias que exploten los comportamientos específicos asociados a determinada tarea. Un ejemplo concreto de esto es la optimización de renderizado de videojuegos, que busca aprovechar el campo de atención visual del jugador para mejorar la calidad de esa región puntual de la imagen, en tanto que el resto se procesa con una calidad más baja.

Paralelamente, el aumento exponencial de estímulos y de datos producidos actualmente ha derivado en la necesidad de emplear técnicas de análisis visual para mejorar la exploración y comprensión de los patrones de comportamiento. Para el caso de *eye tracking*, la rama de análisis visual asociada a este campo ha tomado relevancia dentro de la comunidad científica por su utilidad para realizar análisis comparativos de grupos. La facilidad de poder condensar grandes cantidades de información en una representación visual útil y efectiva, brindando interacciones adecuadas, mejora de manera sustancial la forma de tomar decisiones, comunicar resultados y detectar tendencias y/o patrones ocultos en los datos.

El Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC) se ha enfocado en los últimos años en la caracterización de la lectura como tarea cognitiva de orden superior, dado que involucra un gran número de mecanismos como la memoria, la decodificación de símbolos y el razonamiento, entre otros, para procesar y comprender un texto. Mediante el uso de dispositivos de registro ocular (*eye trackers*) es posible medir de manera no invasiva y precisa cómo una persona procesa el texto a medida que lee. Usando estas mediciones es posible crear modelos y marcadores cognitivos que permitan modelar el comportamiento de un grupo de personas, y detectar aquellas que poseen un cierto deterioro cognitivo actual o potencial.

Mientras más complejos y de mayor alcance sean los experimentos, la necesidad de encontrar nuevas maneras de analizar la información juega un rol importante. En este sentido, el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) desempeña un rol de colaborador, al asistir en el diseño e implementación de soluciones basadas en el análisis visual de datos, orientadas a optimizar la gestión de grandes volúmenes de datos y a agilizar el proceso exploratorio. Estas soluciones no solo se encuentran restringidas a la fase exploratoria del *pipeline* de análisis visual, sino que pueden ser usadas para la formulación y testeo de hipótesis de investigación. Para el caso particular del campo del *eye tracking*, las técnicas de visualizaciones no son empleadas como reemplazo de los modelos de estadística tradicional, sino que sirven como soporte y complemento de los resultados obtenidos.

El objetivo general de esta tesis es el diseño y desarrollo de técnicas de análisis visual para datos espacio-temporales obtenidos a partir de dispositivos de seguimiento ocular. Estas técnicas deberán poder aplicarse a distintos contextos que resulten relevantes desde el punto de vista cognitivo (como la lectura y la exploración de estímulos dinámicos), para analizar y evaluar el comportamiento de las personas.

Aportes

En este trabajo de tesis se detallan y analizan diversas técnicas de análisis visual de datos que fueron diseñadas y desarrolladas para realizar diversas tareas en dominios en los que se emplea eye tracking. Para el caso particular de esta tesis, la colaboración con el Laboratorio de Desarrollo en Neurociencias Cognitivas, ha permitido comprender y desarrollar soluciones para necesidades y requerimientos basados en tareas que requieren un esfuerzo cognitivo elevado. Estas tareas resultan de interés para los expertos, porque permiten usar los resultados obtenidos de un experimento de *eye tracking* como marcadores cognitivos, que posteriormente pueden ser empleados para caracterizar ciertos conjuntos de individuos o bien servir como complemento para el diagnóstico de patologías neurodegenerativas.

En este contexto, la tesis integra avances para abordar los desafíos inherentes a la visualización de datos procedentes del seguimiento ocular. En este sentido, se proponen soluciones que enriquecen el proceso de visualización al considerar la diversidad de estímulos implicados. En lo que respecta a la lectura, se exponen aportes relevantes que abarcan tanto el empleo del e*ye tracking* en la interpretación de oraciones como en el análisis de micro-historias. Asimismo, se ofrecen contribuciones significativas en lo relativo al análisis visual de estímulos dinámicos, tales como el vídeo. De igual forma, se detallan de manera minuciosa los aportes de este trabajo en lo referido a la visualización de datos multidimensionales, las cuales servirán como base para futuros desarrollos en este campo.

Los principales aportes que se obtuvieron durante el desarrollo de la tesis se pueden enumerar de la siguiente forma:

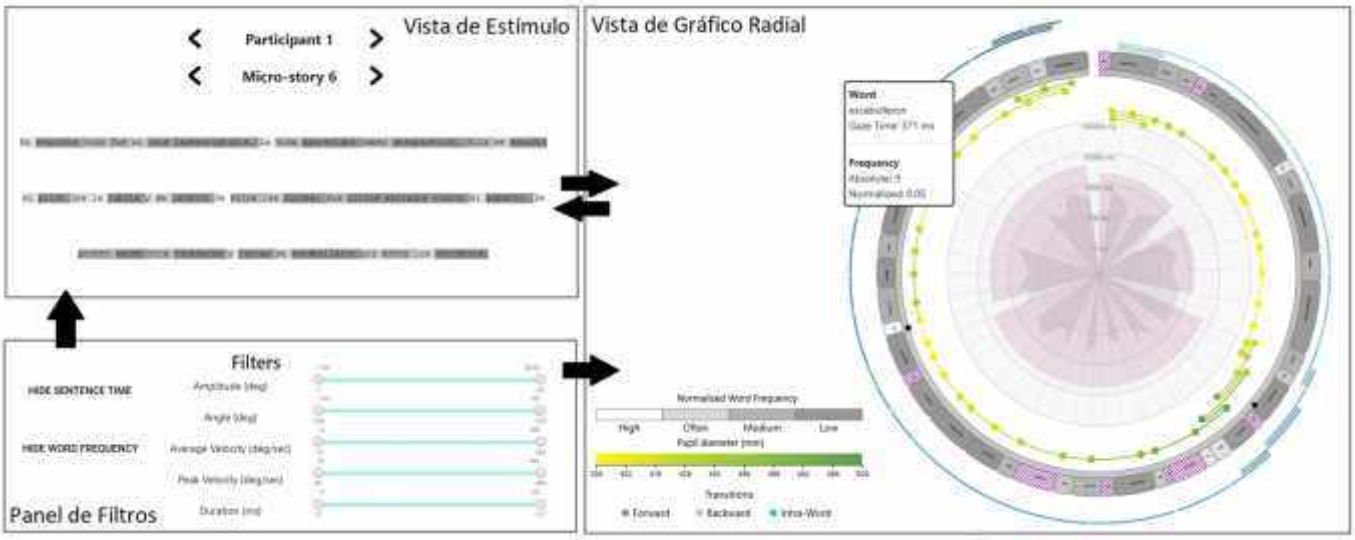
Especificación, diseño y desarrollo de una técnica de análisis visual para analizar movimientos oculares durante la lectura de micro-historias. La lectura ha sido un desafío constante en el ámbito del seguimiento ocular, siendo desarrollada y explorada principalmente por Reichle et al. (2003). Como se explica en el libro de Burch (2022), esta tarea no solo implica considerar cómo se lee un texto, sino también identificar qué palabra u oración es relevante para un participante y por qué. Desde el principio, los experimentos de seguimiento ocular se llevaron a cabo para comprender cómo las personas procesan la información durante una tarea específica. En este contexto, se han utilizado diversos tipos de estímulos—como imágenes, textos, videos—para explorar diferentes aspectos de la cognición. En particular, el texto es un estímulo estático poderoso cuya estructura interna permite explorar la relación entre las palabras y los patrones de atención visual a lo largo del tiempo con un esfuerzo mínimo.

En este trabajo se ha explorado por primera vez el uso de micro-historias como estímulo con alto nivel cognitivo durante la tarea de lectura. Las micro-historias son un tipo particular de texto caracterizado por su brevedad y su alto contenido semántico. Estas propiedades las convierten en un estímulo textual natural interesante para evaluar diferentes procesos cognitivos mediante el análisis del movimiento ocular. Dado que las micro-historias no se han utilizado previamente como estímulo, no es fácil determinar qué conjunto de características debe considerarse en su análisis. Por otro lado, los experimentos de seguimiento ocular en la lectura de microhistorias generan una gran cantidad de datos de diversos tipos, muchos de los cuales tienen una naturaleza espacio-temporal. Por lo tanto, es esencial analizar conjuntamente las diversas características que emergen de estos experimentos para obtener conocimiento sobre ellas y encontrar un subconjunto adecuado que ayude a comprender el proceso cognitivo involucrado en su lectura.. A fin de describir de manera completa este tipo de estímulo por primera vez, se relevó un conjunto de características de interés que resultan útiles para describir el fenómeno de lectura. Gracias a estas características fue posible capturar y describir distintos comportamientos de los participantes durante la lectura de cada una de las micro-historias.

A fin de analizar los datos, se presenta una nueva técnica de análisis visual radial para datos correspondientes a la lectura de micro-historias. La misma busca mostrar cómo un participante lee este tipo de historia haciendo especial énfasis en el cambio pupilar y en el orden en que las palabras son visitadas. Estos dos aspectos resultan de importancia para los expertos ya que les permite determinar la carga cognitiva empleada por el participante (mediante la visualización de la dinámica pupilar) y determinar si se encuentra relacionada con la estrategia de lectura empleada o con características propias del texto.

Por otro lado, el uso de vistas interconectadas entre el estímulo original utilizado durante el experimento y la técnica de visualización, permite mantener el contexto durante los procesos de análisis y exploración de los datos. De esta manera es posible determinar a partir de una palabra del texto cuál es su comportamiento a nivel de movimientos oculares (tiempo total que fue visitada, cantidad de fijaciones, transiciones y dinámica popular). Asimismo, mediante el uso de mecanismos de filtrado y selección es posible añadir más información contextual tanto a la vista de Estímulo como

a la vista Radial principal, para contar con todas las características definidas para el proceso de lectura.

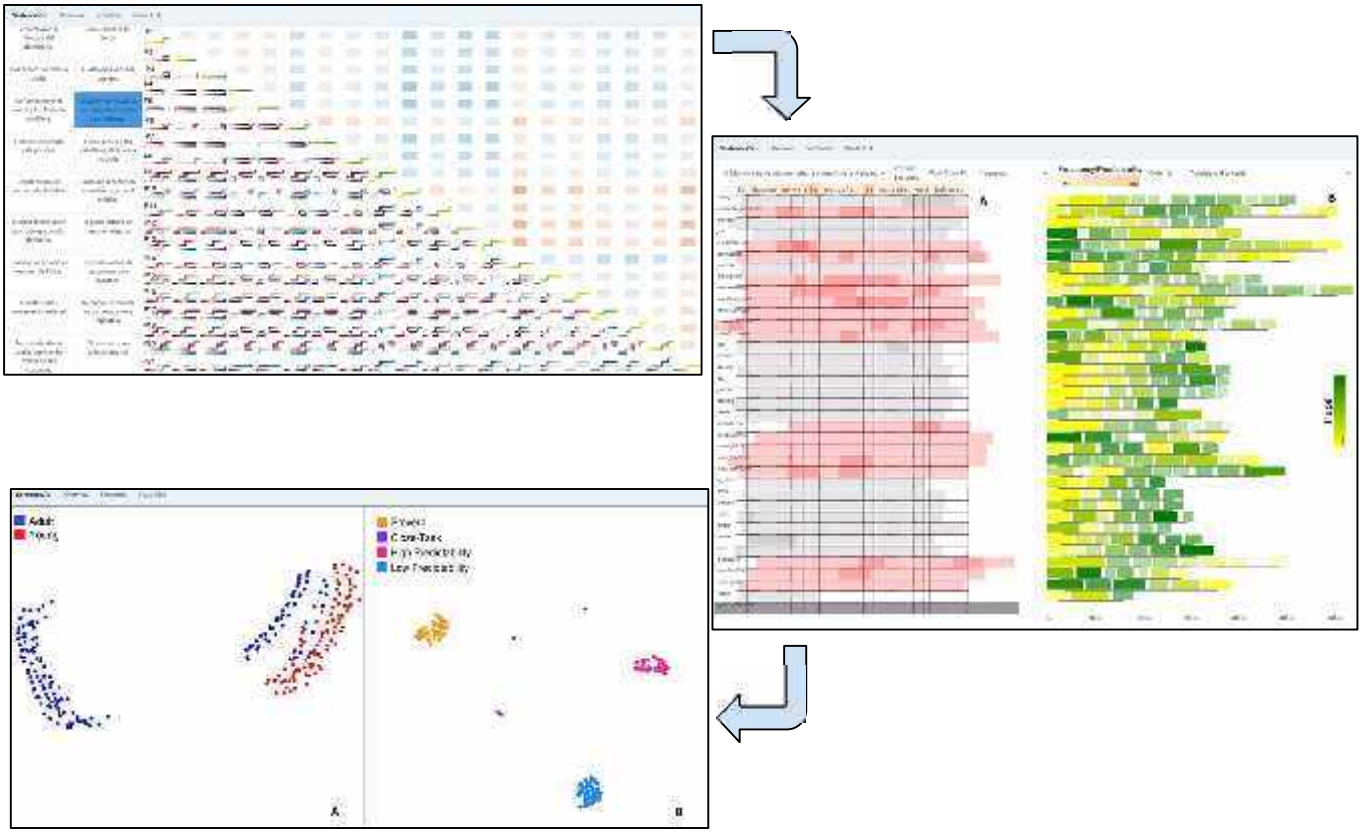


Definición de técnicas de análisis visual para explorar la lectura de oraciones. Mediante el diseño, desarrollo y uso de una herramienta de análisis visual de construcción propia denominada *SentenceVis* fue posible comprender y caracterizar el comportamiento de participantes jóvenes y adultos durante la lectura de oraciones. Dichas oraciones pertenecen al conjunto de datos construido y utilizado por Fernandez et al., 2016 para evaluar carga cognitiva en personas con patologías neurodegenerativas, aunque puede ser empleado para otras tareas vinculadas a desempeño cognitivo.

Este tipo particular de estímulo se encuentra diseñado teniendo en cuenta como principal eje la predictibilidad, que permite evaluar la capacidad de un participante de poder adelantarse al texto que lee. Además, es posible aprovechar las características que ofrece el texto, a fin de poder detectar y extraer marcadores cognitivos de interés para los expertos. Mediante el uso de *SentenceVis* fue posible mejorar el entendimiento de cómo dichos marcadores influyen durante la lectura. El uso de técnicas de visualización nuevas y otras previamente definidas en la literatura, en conjunto con las interacciones propuestas, permitió mejorar el proceso de análisis de datos.

Por otro lado, la herramienta integra un enfoque de aprendizaje automático que permite clasificar las series temporales correspondientes a cada lectura de una oración. Se tienen en cuenta las distintas dimensiones que componen dicha serie, que son la posición en vertical, la posición en horizontal y el valor de pupila registrado. Gracias a esta diferenciación es posible comprender qué factor tiene mayor relevancia en la clasificación en cada uno de los grupos etarios posibles. Por otro lado, estas dimensiones se encuentran plasmadas en distintas técnicas de análisis visual que facilitan la comprensión y manipulación de datos.

Definición de técnicas de análisis visual para explorar estímulos dinámicos. Debido a que diariamente interactuamos con estímulos dinámicos (como navegar en internet, jugar en dispositivos



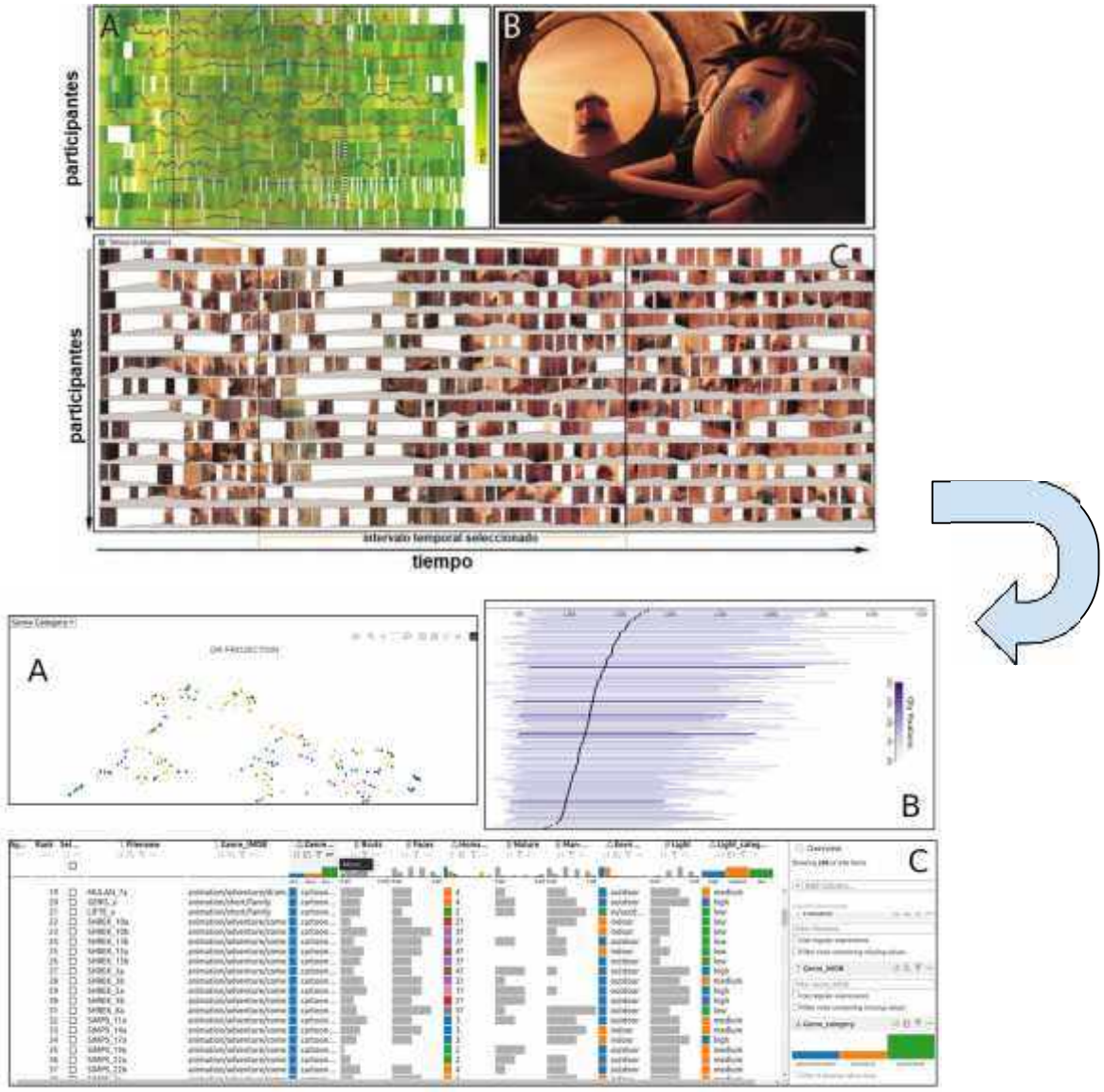
móviles, etc.), resulta interesante analizar, desde el punto de vista del *eye tracking*, cómo se desempeña cognitivamente un individuo mientras mira un estímulo como un video.

Los datos de la pupila en estímulos dinámicos presentan varios desafíos debido a la naturaleza compleja y variable de la dinámica pupilar. Uno de los principales desafíos es la dificultad de visualizar e interpretar estos conjuntos de datos, particularmente en el contexto de estímulos dinámicos. Para abordar este problema, se presenta el diseño y desarrollo de una herramienta de análisis visual para explorar y analizar medidas relacionadas con el movimiento ocular, incluida la dinámica pupilar, en el contexto de estímulos dinámicos. Estos estímulos están centrados específicamente en un conjunto de datos compuesto de fragmentos de videos correspondientes a películas de Hollywood (Costela y Woods, 2019) con diversas características (iluminación, cantidad de rostros humanos, etc.), para determinar qué factores influyen en la atención de la persona.

La herramienta diseñada facilitó el examen y la representación de los cambios en la atención y el esfuerzo cognitivo en entornos dinámicos. Se estudiaron las diferentes alternativas existentes en la literatura para analizar estímulos dinámicos y cómo éstas pueden ser extendidas para incorporar la información relacionada a la pupila. Adicionalmente, se definen escenarios de uso para mostrar cómo puede utilizarse la herramienta para realizar análisis preliminares basados en la dinámica pupilar y en la información relacionada con el número y duración de las fijaciones, permitiendo la comparación de participantes o estímulos. La incorporación de la visualización como un aspecto

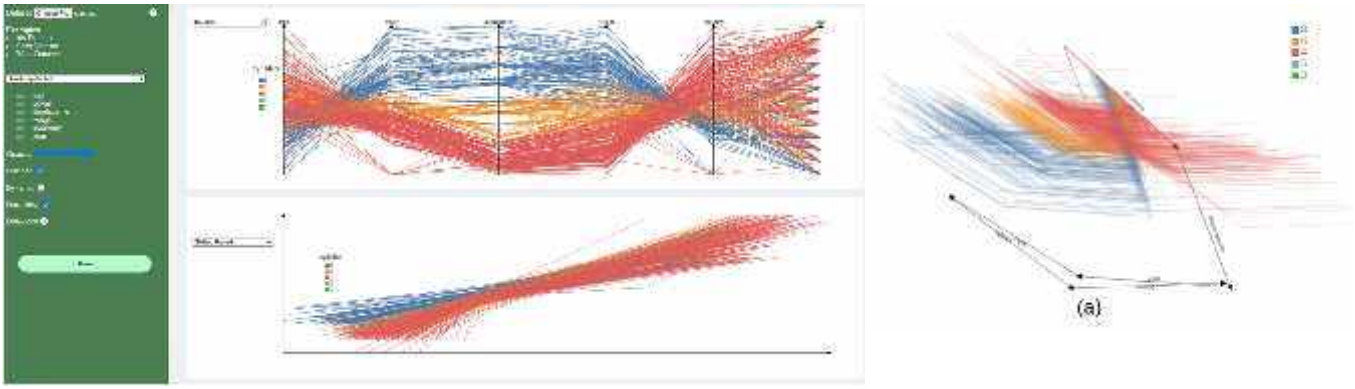
relevante en el análisis de los datos de la pupila en estímulos dinámicos tiene como objetivo explorar el potencial de este enfoque para detectar cambios en el esfuerzo cognitivo.

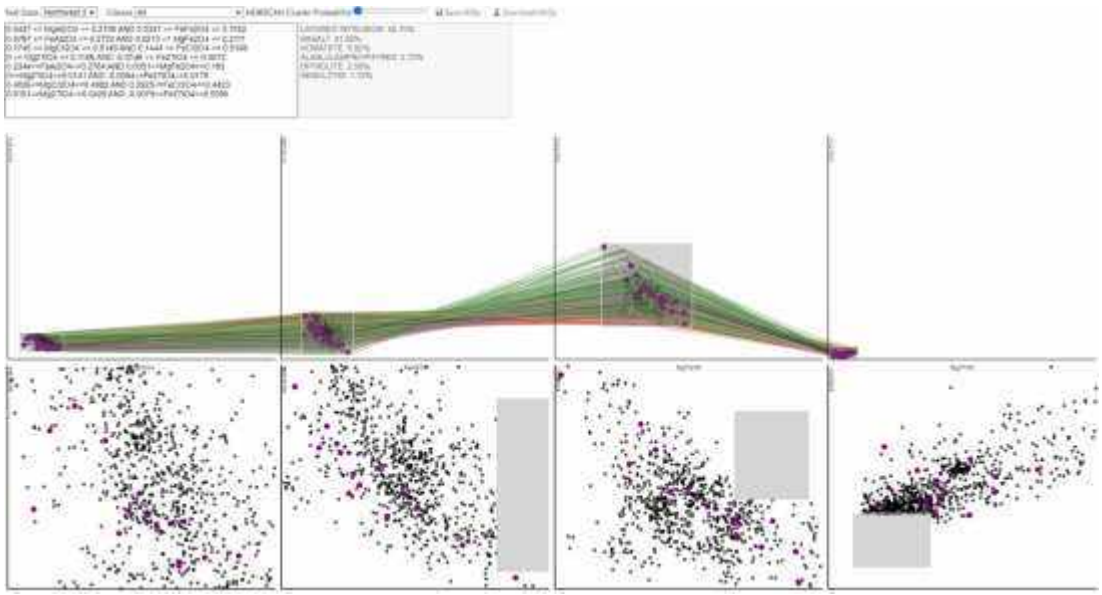
Mediante el uso de la dinámica pupilar como marcador cognitivo, fue posible añadir una nueva dimensión al análisis, que tradicionalmente es realizado teniendo en cuenta las métricas de fijaciones y sacadas. Las métricas se utilizaron de forma conjunta, dado que describen distintos eventos que componen el scanpath de un participante, reflejando así su estrategia de exploración visual durante el procesamiento del estímulo. Al agregar la dinámica pupilar como complemento, es posible determinar cómo se ve influenciada la carga cognitiva a medida que la exploración visual avanza en el tiempo. Esto posibilita la detección de aquellas regiones del estímulo que tienen mayor relevancia para el participante y el poder analizar si la misma se encuentra influenciada por características propias del estímulo o por la estrategia de exploración visual empleada.



Diseño e implementación de herramientas de análisis visual basadas en GLC. Durante el análisis de los datos de *eye tracking* surgió la necesidad de contar con nuevas técnicas visuales que permitieran representar la naturaleza multidimensional de los mismos. A fin de lograr dicho propósito, se exploró un conjunto de técnicas de representación de datos multidimensionales sin pérdida de información denominadas Coordenadas Generales de Línea (GLCs). Estas técnicas han nacido en los últimos años como alternativas a la representaciones multidimensionales clasificas, y han demostrado su utilidad bajo ciertas restricciones y dominios de aplicación.

Se empleó una metodología de trabajo en donde primeramente se definió una taxonomía para dichas técnicas con el fin de conceptualizar y clasificar las propiedades de las mismas. Posteriormente, se diseñaron y desarrollaron herramientas que permitieran emplear estas técnicas con sus diferentes configuraciones, y evaluar y comparar cómo las distintas propiedades afectan la representación visual de los datos. Para concluir, se estableció un caso de estudio del mundo vinculado al dominio de la ciencias geológicas donde se evaluó la utilidad y las limitaciones de estas técnicas para resolver tareas vinculadas a clasificación, detección de valores atípicos y caracterización de diferentes grupos de datos.





Gracias a los resultados obtenidos y presentados en esta tesis, fue posible avanzar en el área de visualización de datos provenientes de registradores oculares. De esta manera se logró el diseño y desarrollo de herramientas y técnicas de análisis visual que resultan útiles para aplicar en tareas como la lectura de texto y la exploración de estímulos dinámicos.

Por otro lado, el relevamiento y análisis de técnicas multidimensionales como complemento a las técnicas propias de *eye tracking,* condujo al diseño y desarrollo de una herramienta de análisis visual denominada *SpinelVA* que combina varias de las mismas para dar soporte a tareas vinculadas al campo de las ciencias geológicas. Además, dicho trabajo hace uso de un enfoque de integración de aprendizaje automático con visualización, para asistir en la resolución de tareas vinculadas al dominio de aplicación, que posteriormente sirvió como punto de partida para el desarrollo de la investigación referida al análisis visual de la lectura de oraciones.

Líneas de investigación futuras

Esta tesis ofrece una visión general sobre el diseño y desarrollo de técnicas de análisis visual de datos provenientes de *eye trackers*, permitiendo tener en cuenta nuevos aspectos en tareas vinculadas a lectura y exploración de videos. En este sentido, se proponen las siguientes líneas de trabajo futuro que se podrían abordar sobre dicho tema:

Extensión de las técnicas de visualización de datos multidimensionales.

Como se señaló anteriormente, las técnicas de visualización basadas en GLCs han demostrado un notable potencial para la representación de datos multidimensionales; sin embargo, presentan una limitación en cuanto al manejo de la dimensión temporal en su definición.

Si bien el marco matemático que da soporte a las GLCs permite el manejo del tiempo como una dimensión extra que es factible de modelar, su concepción está íntegramente desarrollada para modelos financieros. Dichos modelos tienen solamente en cuenta el aspecto relacionado con pronosticar ciertos eventos futuros que pueden alterar un indicador financiero. Dado que estos modelos no se ajustan a los requerimientos que son necesarios para datos de *eye tracking*, es necesario una redefinición y/o extensión de estas técnicas para que puedan ser empleadas en las tareas típicas de clasificación y detección de anomalías, que son de interés para los expertos.

Se deberá tener en cuenta que la naturaleza de los datos es espacio-temporal y no sólo temporal, como es el caso de las series temporales financieras, por lo que que la posición espacial p_{xy} asociada a un instante de tiempo t deberá conservar su relación y significado en la representación final.

Incorporación de métodos de inteligencia artificial al análisis visual exploratorio.

Actualmente el desarrollo sostenido y masivo de soluciones basadas en inteligencia artificial posibilita la formulación de herramientas de análisis visual integrales y completas para la exploración

de datos derivados del *eye tracking*. Existe un nuevo enfoque vinculado al uso de modelos generativos como mecanismos de asistencia en la creación, personalización y recomendación de diversas técnicas de visualización. Esto posibilita que estas nuevas herramientas puedan añadir al *pipeline* de analítica visual las capacidades que ofrecen dichos modelos y ofrecer soluciones más inteligentes.

De este modo, los expertos de dominio que emplean estas herramientas podrán diseñar análisis personalizados que se ajusten a sus requerimientos específicos e inclusos expandir las capacidades de representación e interacción de dichas técnicas.

Evaluación integral de una técnica en múltiples escenarios.

Si bien una técnica de visualización para *eye tracking* está condicionada por el tipo de tarea a realizar, aún existe un extenso campo de investigación respecto a su impacto en diferentes escenarios. Es decir, es fundamental determinar si una técnica originalmente diseñada para una tarea específica puede extenderse o adaptarse a otra completamente distinta.

Evaluar y cuantificar el esfuerzo requerido para llevar a cabo dicha adaptación permitirá obtener soluciones de carácter más general, aplicables en múltiples dominios, sin necesidad de realizar modificaciones significativas. La evaluación integral de las técnicas en múltiples escenarios permitirá identificar debilidades y fortalezas de cada una, a fin de poder establecer qué aspectos deberán valorarse al momento de realizar la adaptación entre tareas.

Título de la Tesis: Análisis y desarrollo de representaciones generales de audio.
Autor: Dr. Leonardo Daniel Pepino.
Directora: Dra. Luciana Ferrer.
Director Adjunto: Dr. Pablo Riera.
Institución: Laboratorio de Inteligencia Artificial Aplicada (LIAA), Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (UBA).
Fecha de Defensa: 16 de diciembre de 2025.

Resumen: El aprendizaje de representaciones ha sido fundamental para el avance del aprendizaje profundo, permitiendo reutilizar modelos para resolver tareas diversas de manera eficiente en términos de cómputo y datos. En particular, el aprendizaje no supervisado permitió aprovechar la abundancia de datos sin etiquetar para aprender representaciones ricas en escenarios de escasez de datos anotados, simplificando el desarrollo de nuevos modelos, y favoreciendo la democratización de la inteligencia artificial.

El objetivo principal de esta tesis doctoral fue el desarrollo de un modelo basado en redes Transformer [1], cuyas representaciones internas fueran útiles para una gama diversa de tareas de audio, abarcando no solo problemas de habla, sino también relacionados a la música y sonidos ambientales. Este desarrollo culminó con **EnCodecMAE** [2], un modelo basado en Masked Autoencoders [3] y pre-entrenado con señales de audio diversas provenientes de Audioset [4], LibriLight [5] y Free Music Archive [6]. Este modelo tiene la particularidad de utilizar EnCodec [7], un codec de audio neuronal, como señal objetivo. EnCodecMAE alcanzó resultados comparables o superiores al estado del arte en un conjunto de 18 tareas diversas abarcando habla, música y audio general, validando la generalidad de las representaciones aprendidas.

Además de presentar EnCodecMAE y mostrar los resultados de su evaluación, realizamos un estudio de ablación, explorando el efecto que tienen distintos aspectos de su diseño.

Por último, realizamos un análisis del alineamiento entre distintas representaciones de audio, incluyendo EnCodecMAE, y representaciones cerebrales obtenidas a partir de resonancias magnéticas funcionales de la corteza auditiva. Mostramos que estos modelos, a pesar de no ser entrenados explícitamente para aproximar la actividad cerebral, exhiben un alineamiento creciente con la misma durante el pre-entrenamiento, y que el desempeño de las representaciones en distintas tareas se correlaciona con la similaridad con las representaciones cerebrales.

Introducción

Tradicionalmente, el procesamiento de señales de audio y la extracción de información a partir de las mismas, se realizaba mediante un uso extensivo de conocimiento experto, desarrollando técnicas de extracción de atributos y utilizando modelos estadísticos como modelos ocultos de Markov (HMM) o heurísticas, y variando los métodos dependiendo del problema que se deseaba resolver.

Hacia la década del 2010, con el auge del aprendizaje profundo, se comenzaron a utilizar redes profundas que operaban directamente sobre la señal de audio o espectrogramas de la misma. De esta manera, la red neuronal aprendía internamente a extraer atributos que fueran relevantes para el problema a resolver. Sin embargo, el entrenamiento de estos modelos requiere de una gran cantidad de datos anotados, los cuales son costosos de obtener y dificultan escalar los modelos.

Una manera de resolver este problema es mediante transferencia de aprendizaje. En esta técnica, se utiliza un modelo (upstream) que haya sido entrenado con una gran cantidad de datos, y luego se

reutilizan las representaciones internas que aprendió, para poder entrenar modelos más pequeños (downstream), en la tarea de interés, requiriendo una menor cantidad de datos anotados. Un ejemplo en audio es VGGish [8], el cual fue entrenado con unas 5 millones de horas de audio para predecir distintas etiquetas, como "canción", "persona hablando" o "bocina". Una vez entrenado el modelo en esta tarea, se pueden extraer activaciones intermedias y utilizarlas como atributos para entrenar nuevos modelos más pequeños que resuelvan tareas similares, como detección de eventos acústicos. Si bien este enfoque es efectivo, requiere que la tarea que se quiere resolver sea similar a la tarea con la que fue entrenado el modelo upstream. Al mismo tiempo, conseguir miles de horas de audio anotadas es costoso y poco práctico.

Con el auge de los modelos de representaciones contextualizadas de texto basados en Transformers [1], como BERT [9], se comenzaron a trasladar estas técnicas al dominio del habla y audio. Modelos como Wav2Vec 2.0 [10] enmascaran segmentos de habla y entrenan redes transformer para predecir estos segmentos enmascarados. Esta tarea se conoce como **modelado de lenguaje enmascarado**, y permite aprender representaciones contextualizadas de forma auto-supervisada, y por ende sin necesitar anotaciones, y resulta en representaciones útiles para resolver distintas tareas. La mayoría de los modelos que aplican estas técnicas se enfocan en el habla, con algunos otros modelos recientes como AudioMAE [11] y BEATs [12] aplicando estas ideas a audio general. Sin embargo, la mayoría de los modelos de representaciones generales de audio trabajan representando espectrogramas como "patches" o recortes rectangulares, lo cual lleva a una pérdida de resolución temporal, y se entrenan solo con audio de YouTube (Audioset) [4] y utilizando objetivos de regresión, en lugar de clasificación como BERT.

Con nuestro modelo, EnCodecMAE [2], apuntamos a utilizar un objetivo de clasificación como BERT, explorando el uso de representaciones discretas obtenidas mediante un codec de audio neuronal: EnCodec [7]. Al mismo tiempo, enmascaramos en el tiempo, en lugar de utilizar "patches", y exploramos la incorporación de datos más diversos y de etapas de "auto-entrenamiento".

Además, nos interesa analizar qué aprenden las distintas representaciones, explorando conexiones entre el aprendizaje de representaciones de audio y la neurociencia. Encontramos que EnCodecMAE y otros modelos de audio son buenos predictores de la actividad de la corteza auditiva medida mediante fMRI, y que esta capacidad de predicción se adquiere de forma indirecta durante el pre-entrenamiento, y es mayor cuanto más útil son las representaciones internas del modelo para resolver distintas tareas de audición [13].

EnCodecMAE

En la Figura 1 se puede observar la arquitectura de nuestro modelo. La forma de onda, con una duración de 4 segundos y una frecuencia de muestreo de 24000 Hz, es transformada en atributos con menor resolución temporal. En nuestra tesis experimentamos con espectrogramas, melspectrogramas y la utilización de redes convolucionales. Los mejores resultados los obtuvimos utilizando melspectrogramas, con un tamaño de ventana de 640 muestras, un paso de 320 muestras y 256 mels. Luego, esta señal X_I, es proyectada linealmente a D dimensiones y enmascarada. Al igual que en otros trabajos del dominio del habla, enmascaramos una proporción M_p de la señal utilizando bloques contiguos de M_g muestras en el tiempo. En lugar de reemplazar las muestras enmascaradas por un token de máscara como en Wav2Vec 2.0 [10] y otros trabajos del dominio del habla, los cuadros enmascarados son descartados de la secuencia, reduciendo su longitud, y por ende haciendo

que el modelo sea más eficiente computacionalmente. Los cuadros que no fueron enmascarados (X_v) son procesados por una red Transformer con atención bidireccional, produciendo representaciones contextualizadas de los cuadros no enmascarados (X_e). Luego, se introducen tokens de máscara en las posiciones que habían sido descartadas, y se vuelven a sumar embeddings posicionales absolutos. De esta manera, se obtiene una secuencia X_d con la misma longitud que la secuencia original X_f. Finalmente, el decoder debe predecir la información faltante en las posiciones enmascaradas, a partir de las representaciones contextualizadas de las posiciones no enmascaradas. En lugar de realizar regresión sobre melspectrogramas, la función de pérdida se calcula entre la secuencia de salida del decoder P, y una representación discreta Y de la señal de audio original sin enmascarar. Esta representación discreta es obtenida a partir de EnCodec [7], un codec de audio neuronal, que dada una señal de audio, la representa como Q secuencias discretas con 1024 valores posibles. Se utilizó entropía cruzada como función de pérdida, utilizando distintos pesos para cada secuencia de EnCodec, y distinto peso para la reconstrucción de los tokens enmascarados o sin enmascarar.

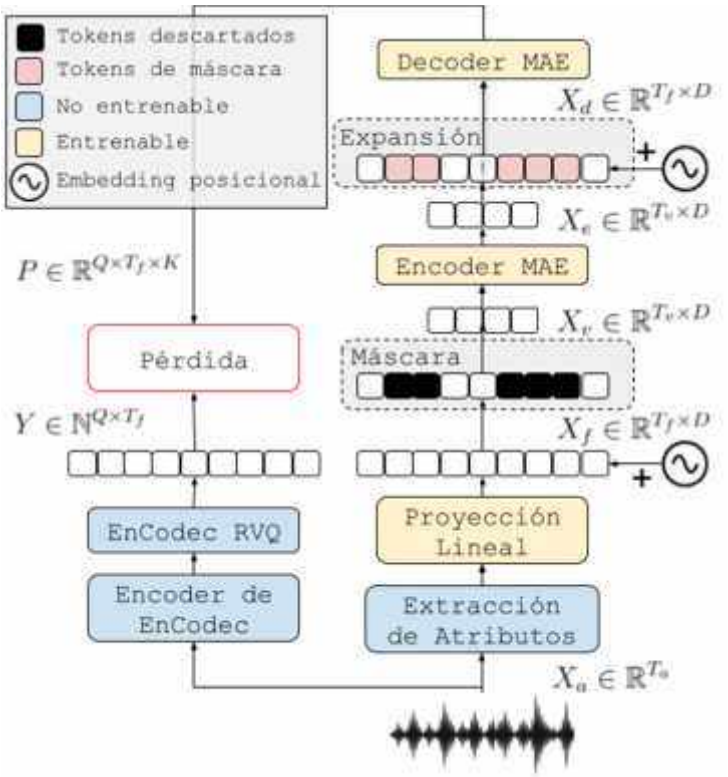


Figura 1. Arquitectura de EnCodecMAE

Los motivos por los que tomamos estas decisiones de diseño son:

- Eficiencia computacional: para el pre-entrenamiento de EnCodecMAE contamos con 2 placas de video RTX 3090, por lo que debimos utilizar segmentos más cortos de audio (4 segundos). A su vez, descartar las partes enmascaradas de la secuencia nos permite reducir aún más la longitud de la secuencia antes de ser procesada por el Encoder, el cual contiene la mayor parte de las capas de Transformer.
- Generalidad: deseamos representaciones generales de audio, que no sean específicas de habla, música o sonidos ambientales, sino que puedan ser utilizadas para tareas en cualquier dominio acústico. Por esto, utilizamos como señal objetivo las representaciones de EnCodec, el cual

permite reconstruir señales diversas con alta fidelidad. De esta forma, nos aseguramos de que la señal objetivo no descarta detalles relevantes de la señal de audio. A su vez, al pasar del dominio continuo al discreto, nuestro trabajo se vuelve más similar a trabajos del campo del procesamiento del lenguaje natural que han sido históricamente exitosos, como BERT.

EnCodecMAE fue entrenado durante 5 días, en un conjunto de datos conteniendo aproximadamente 12000 horas de audio abarcando habla, música y sonidos ambientales. Para esto, combinamos tres conjuntos de datos masivos: Audioset (2 millones de segmentos de 10s de YouTube), LibriLight (6000 horas de habla de audiolibros) y Free Music Archive (unas 800 horas de canciones). Entrenamos varias variantes de EnCodecMAE, variando su tamaño (Small, Base y Large), y realizando una etapa de "auto-entrenamiento" en la que una vez realizados 500000 pasos de entrenamiento, se discretizan las activaciones intermedias de EnCodecMAE utilizando K-Means, y las mismas se utilizan como señal objetivo, reemplazando las secuencias discretas de EnCodecMAE. Esta técnica nos permitió mejorar aún más la calidad de las representaciones.

Una vez que EnCodecMAE se encuentra pre-entrenado, debemos medir cuán buenas y generales son sus representaciones de audio, comparándolas contra el estado del arte. Para esto, utilizamos el benchmark HEAREval [14], el cual evalúa la calidad de las representaciones para resolver un conjunto de 18 tareas diversas de audio. Algunas de estas tareas son el reconocimiento de comandos de voz (Speech Commands), reconocimiento de emociones a partir del habla (CREMA-D), detección de eventos acústicos (FSD50K) y clasificación de sonidos ambientales (ESC-50), clasificación de género musical (GTZAN Genre), transcripción automática de música (MAESTRO) y clasificación de notas musicales (NSynth Pitch). Para la evaluación se extraen activaciones de la salida del encoder de EnCodecMAE y se utilizan como entrada a un perceptrón multicapa, el cual se entrena para resolver las distintas tareas del benchmark. Nuestro modelo alcanza un desempeño comparable o mayor al mejor desempeño en el leaderboard de HEAREval en todas las tareas. Esto se puede observar de forma resumida en la Figura 2, en la que se muestra el desempeño en cada tarea de dos de nuestros modelos (Mel256->EC BaseST y Mel256->EC Large ST) comparado con el modelo estado del arte cuando llevamos a cabo este trabajo (BEATs), y comparado con el mejor desempeño en el leaderboard de HEAREval, y el desempeño promedio en el mismo. Los puntajes se obtienen sustrayendo la media del leaderboard (un puntaje de cero equivale al promedio en HEAREval para la tarea), y dividiendo por el máximo puntaje (un puntaje de uno equivale al mejor desempeño en HEAREval para esa tarea). Podemos observar que nuestros modelos (curvas naranja y verdes) se encuentran por encima de BEATs en la mayoría de las tareas, y en varias ocasiones se acercan o superan al estado del arte (curva violeta). La Figura 2, junto con los resultados completos reportados en la tesis y en nuestra publicación [2], validan la generalidad de EnCodecMAE.

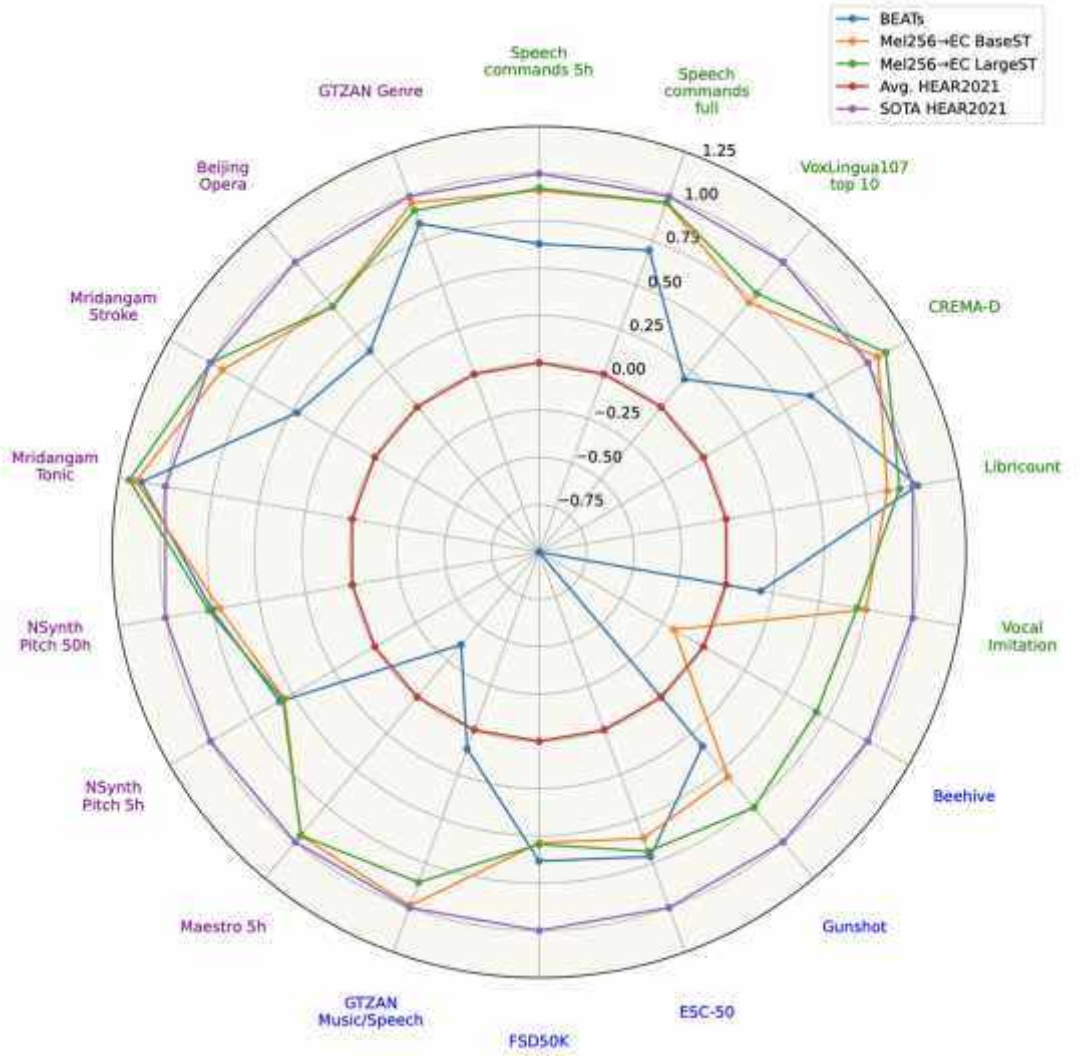


Figura 2. Desempeño normalizado de los modelos propuestos contra el estado del arte.

Representaciones de audio y cerebrales

Una vez que construímos EnCodecMAE, decidimos enfocarnos en entender mejor qué había aprendido y qué estaba representando. Uno de los análisis que trajo resultados más interesantes fue comparar cómo representan sonidos distintos modelos de audio contra cómo lo hace la corteza auditiva de sujetos utilizando mediciones de resonancia magnética funcional (fMRI). Para esto utilizamos dos conjuntos de datos de fMRI existentes [15, 16] en los que los sujetos escucharon los mismos 165 estímulos sonoros abarcando distintos sonidos de la vida cotidiana mientras su actividad cerebral era medida. Al mismo tiempo, procesamos los 165 estímulos sonoros con 36 modelos de audio distintos. De esta forma obtuvimos dos tipos de mediciones:

- Activaciones $X_{m,l} \in \Re^N$ del modelo de audio m y su capa l, para cada uno de los N estímulos.
- Mediciones de fMRI $Y_{s,v}$ pertenecientes al voxel v y sujeto s.

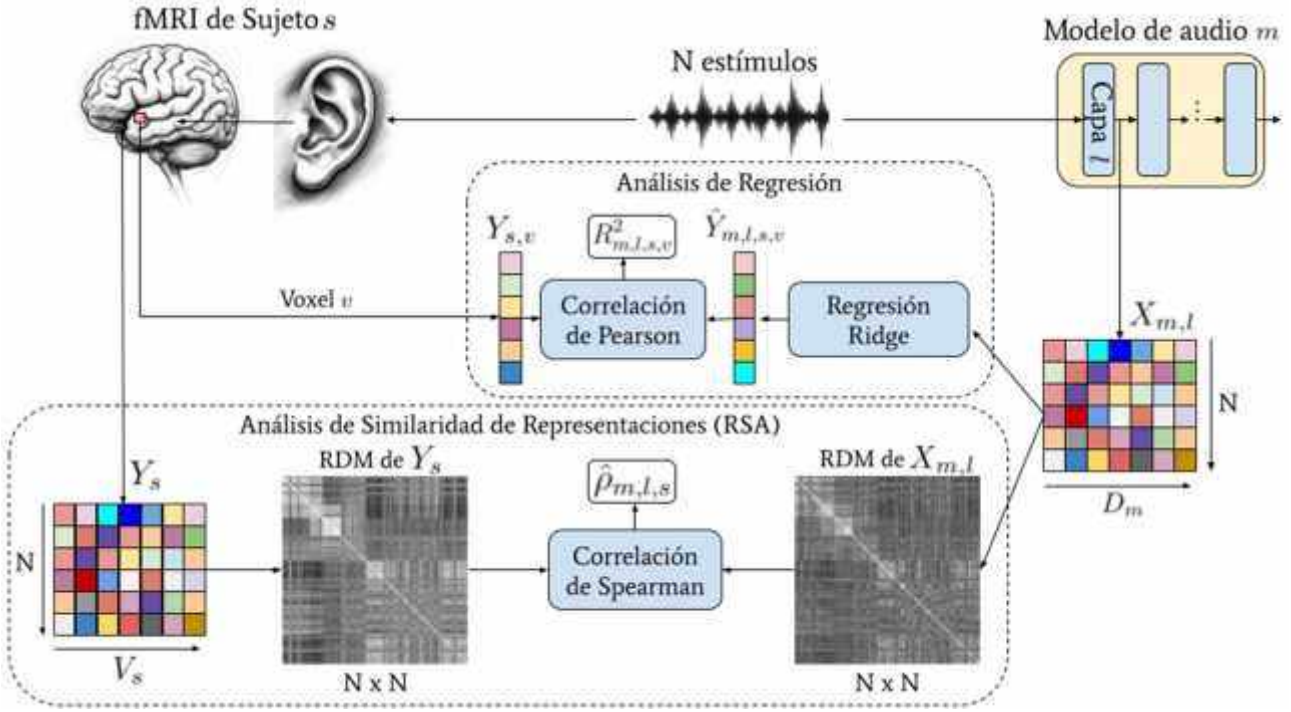


Figura 3. Diagrama resúmen de la metodología utilizada para comparar las representaciones de audio con las mediciones de fMRI.

Utilizamos dos técnicas para comparar ambas mediciones, las cuales se encuentran resumidas en la Figura 3:

- **Regresión ridge**: para cada voxel de cada sujeto se entrena un regresor ridge que intenta predecir su actividad para cada uno de los 165 estímulos, utilizando como entrada las distintas activaciones $X_{m,l}$. Luego, se calcula el coeficiente de determinación de Pearson R^2 entre las predicciones del regresor ridge y las mediciones de fMRI usando datos no observados durante el entrenamiento. De esta manera, se obtiene un tensor $R^2_{m,l,s,v}$ que refleja una medida de la capacidad de las activaciones de la capa l del modelo de audio m para predecir la actividad cerebral del sujeto s en el voxel v.
- **Análisis de similaridad de representaciones (RSA):** para cada sujeto s, se comparan sus mediciones de fMRI Y_s (una matriz con forma $N \times V_s$, en donde V_s es la cantidad de voxels) con las activaciones $X_{m,l}$. Para realizar la comparación se calcularon matrices de disimilaridad a partir de Y_s y $X_{m,l}$. Cada entrada i,j de la matriz de disimilaridad indica cuán disímiles son los estímulos i y j representados mediante Y_s o $X_{m,l}$. Luego, se comparan ambas matrices de disimilaridad utilizando correlación de Spearman, y se obtiene un tensor $\rho_{m,l,s}$ indicando la similaridad entre la representación de audio proveniente de la capa l del modelo m, y la medición de fMRI del sujeto s.

Algunos de los resultados que obtuvimos mediante este análisis son:

- Los modelos más recientes de audio, que utilizan redes transformer y se entrenan de manera auto-supervisada, son los que mejor predicen la actividad cerebral y obtienen un valor de similaridad mediante RSA elevada. De entre estos modelos, EnCodecMAE obtiene los

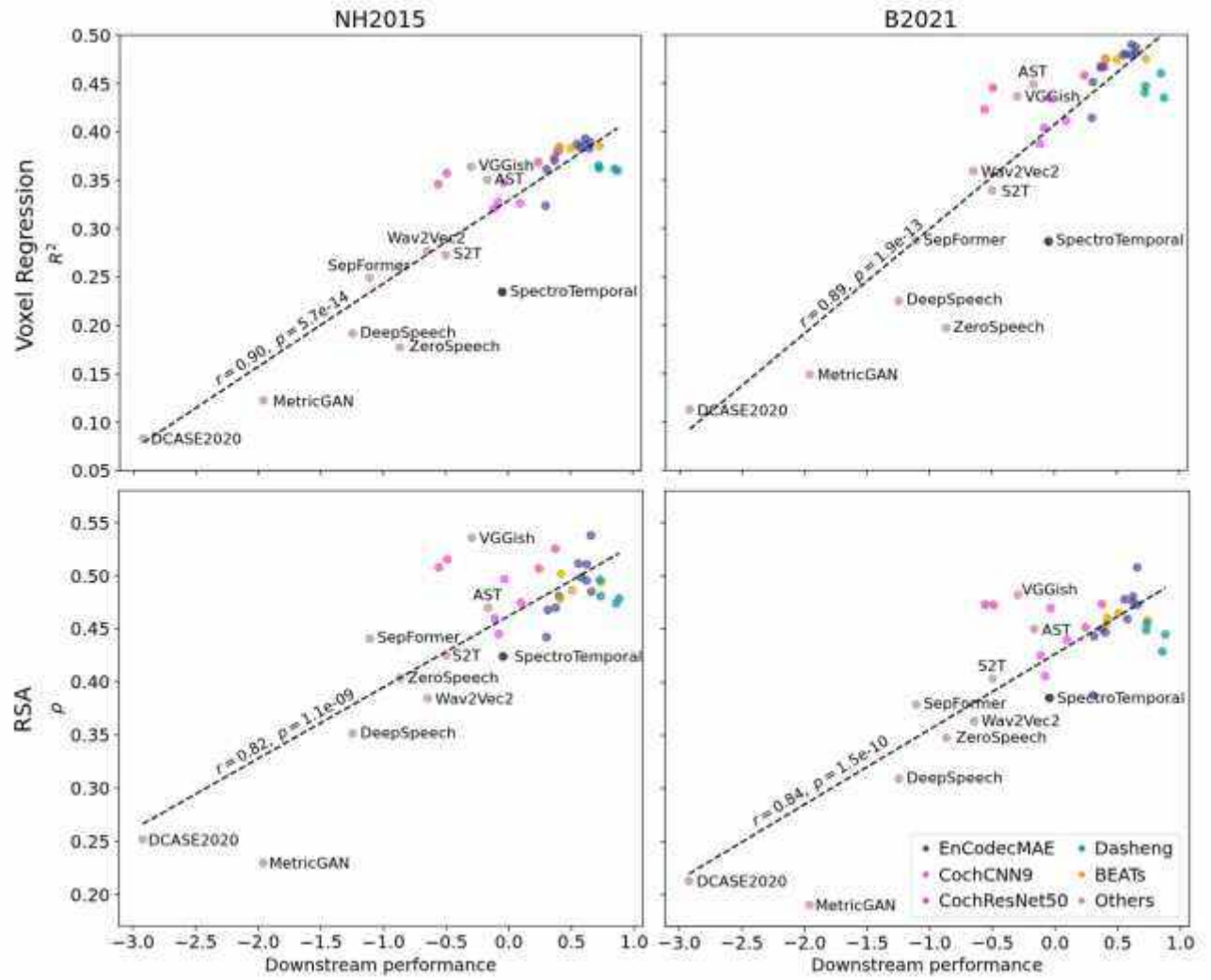


Figura 4. Gráficos de dispersión comparando el desempeño promedio normalizado de cada modelo de audio con su similaridad con las mediciones de fMRI, para ambos conjuntos de datos, y obtenidas tanto mediante RSA como regresión ridge.

mejores valores de R^2 y ρ promedios, reflejando que además de ser útil para resolver tareas de audio, también es un buen predictor de actividad en la corteza auditiva.

- Analizamos cómo evoluciona la similaridad entre las representaciones de EnCodecMAE y las cerebrales a lo largo del pre-entrenamiento. Observamos que tanto el R^2 como ρ crecen a lo largo del pre-entrenamiento, a pesar de que no se realizó ninguna optimización explícita de la similaridad con mediciones cerebrales, y no se utilizaron las mismas durante el pre-entrenamiento. Esto indicaría que el alineamiento entre ambas representaciones es una propiedad que emerge durante el pre-entrenamiento al utilizar un objetivo auto-supervisado.
- Intentamos descubrir qué factores son los que llevan a que un modelo sea mejor predictor de la actividad cerebral. Una de nuestras hipótesis fue que los modelos que eran mejores para resolver distintas tareas de audio, también eran mejores para predecir la actividad cerebral. Para probar esta hipótesis, comparamos los valores de R^2 y ρ de cada uno de los 36 modelos analizados con el desempeño promedio normalizado en HEAREval de estos modelos. De esta forma, obtuvimos los gráficos de dispersión de la Figura 4. Se puede observar que cuanto

mejor es el desempeño de los modelos, también es mayor el valor de R^2 y ρ ($r \geq 0.82$) para ambos conjuntos de datos.

Estos resultados contribuyen al campo surgente del neuroconexionismo que busca vincular modelos computacionales con neurociencia, y aporta nuevo conocimiento en el dominio de la audición computacional. Si bien existían trabajos previos vinculando el desempeño de representaciones en distintas tareas de, por ejemplo, texto y visión, aún no habían resultados en la literatura mostrando este vínculo en el dominio de la audición con modelos de audio. Tampoco encontramos trabajos que muestren explícitamente cómo evoluciona la similaridad a lo largo del pre-entrenamiento de estos modelos.

Conclusiones

Esta investigación doctoral ha permitido avanzar en la frontera del aprendizaje de representaciones de audio, abordando con éxito el desafío de desarrollar modelos generalistas. A continuación se resumen las conclusiones principales:

- **Eficacia de EnCodecMAE:** Se ha demostrado que el uso de un codec de audio neuronal (*EnCodec*) como señal objetivo en un esquema de *Masked Autoencoder* es una estrategia robusta. Al preservar información perceptualmente relevante que otros métodos (como clusters de MFCC) descartan, EnCodecMAE logra representar simultáneamente habla, música y sonidos ambientales.
- **Desempeño y Generalidad:** La evaluación en el benchmark HEAREval confirmó que las representaciones aprendidas son competitivas con el estado del arte, superando incluso a modelos especializados en diversas tareas. El éxito en tareas tan disímiles valida la hipótesis de que es posible aprender una representación general de audio mediante aprendizaje auto-supervisado.
- **Convergencia Neuroconexionista:** Uno de los aportes más interesantes es que, sin supervisión explícita, EnCodecMAE al igual que otros modelos de audio, desarrolla representaciones internas similares a las obtenidas mediante mediciones de la actividad en la corteza auditiva utilizando fMRI.
- **Correlación entre Desempeño y Alineamiento:** El descubrimiento de que una mayor capacidad funcional en tareas de audio se correlaciona con un mayor alineamiento con representaciones de fMRI sugiere que las soluciones óptimas a las que llegan los modelos de redes profundas de audio tienden a converger hacia soluciones biológicamente plausibles.

Líneas futuras de investigación

A partir de los resultados y las limitaciones identificadas en este trabajo, se proponen las siguientes direcciones:

- **Refinamiento de la Señal Objetivo:** Explorar nuevas señales objetivo para el entrenamiento de EnCodecMAE, que incorporen semántica de largo plazo o ensambles de representaciones, por ejemplo combinando MFCC, WavLM y EnCodec, para capturar simultáneamente fonética, contexto y fidelidad acústica.

- **Desacoplamiento de Factores (Disentanglement):** Desarrollar técnicas para separar explícitamente características como el *pitch*, el timbre y el contenido fonético dentro de la representación, permitiendo que los modelos *downstream* seleccionen la información relevante de forma más eficiente.
- **Eficiencia computacional:** Investigar el uso de resoluciones temporales variables (ventanas dinámicas) para representar silencios o sonidos estacionarios de forma más eficiente, y explorar la destilación de conocimiento desde EnCodecMAE Large hacia modelos más livianos basados, por ejemplo, en convoluciones dinámicas.
- **Regularización mediante fMRI:** Utilizar las mediciones cerebrales como parte de la función de costo durante el pre-entrenamiento, forzando al modelo a aprender representaciones que se alineen con la humana.
- **Bioacústica y Generalización entre Especies:** Extender el marco de EnCodecMAE a dominios no humanos (como el canto de aves), analizando si los modelos entrenados en entornos acústicos animales permiten decodificar mejor sus propios sistemas auditivos biológicos, y así entender mejor el procesamiento auditivo en otras especies.

Bibliografía

Principales publicaciones del Doctorado

[2] **Pepino, L.**, Riera, P., & Ferrer, L. EnCodecMAE: leveraging neural codecs for universal audio representation learning. Interspeech 2025.

[13] **Pepino, L.**, Riera, P., Kamienkowski, J., & Ferrer, L. (2025). Better audio representations are more brain-like: linking model-brain alignment with performance in downstream auditory tasks. *arXiv preprint*.

Pepino, L., Riera, P., & Ferrer, L. (2021). Emotion Recognition from Speech Using Wav2vec 2.0 Embeddings. Interspeech 2021.

Pepino, L., Riera, P., Ferrer, L., & Gravano, A. (2020, May). Fusion approaches for emotion recognition from speech using acoustic and text-based features. IEEE *ICASSP 2020*.

Alonso-Jiménez, P., **Pepino, L.**, Batlle-Roca, R., Zinemanas, P., Bogdanov, D., Serra, X., & Rocamora, M. (2024, April). Leveraging pre-trained autoencoders for interpretable prototype learning of music audio. IEEE ICASSP *Workshops* (pp. 833-837).

Pepino, L., Riera, P., & Ferrer, L. (2022, May). Study of positional encoding approaches for audio spectrogram transformers. IEEE *ICASSP 2022*.

Código, modelos y datos asociados: https://github.com/habla-liaa/encodecmae

Otros trabajos

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

[3] He, K., Chen, X., Xie, S., Li, Y., Dollár, P., & Girshick, R. (2022). Masked autoencoders are scalable vision learners. *IEEE/CVF ICCV* (pp. 16000-16009).

[4] Gemmeke, J. F., Ellis, D. P., Freedman, D., Jansen, A., Lawrence, W., Moore, R. C., ... & Ritter, M. (2017, March). Audio set: An ontology and human-labeled dataset for audio events. IEEE ICASSP 2017 (pp. 776-780).

[5] Kahn, J., Riviere, M., Zheng, W., Kharitonov, E., Xu, Q., Mazaré, P. E., ... & Dupoux, E. (2020, May). Libri-light: A benchmark for asr with limited or no supervision. IEEE *ICASSP 2020* (pp. 7669-7673).

[6] Defferrard, M., Benzi, K., Vandergheynst, P., & Bresson, X. (2017). FMA: A Dataset For Music Analysis. En *18th International Society for Music Information Retrieval Conference*.

[7] Défossez, A., Copet, J., Synnaeve, G., & Adi, Y. (2022) High Fidelity Neural Audio Compression. *Transactions on Machine Learning Research*.

[8] Hershey, S., Chaudhuri, S., Ellis, D. P., Gemmeke, J. F., Jansen, A., Moore, R. C., ... & Wilson, K. (2017, March). CNN architectures for large-scale audio classification. IEEE ICASSP 2017 (pp. 131-135).

[9] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. En NAACL 2019*, volumen 1* (pp. 4171-4186).

[10] Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. *NeurIPS*, *33*, 12449-12460.

[11] Huang, P. Y., Xu, H., Li, J., Baevski, A., Auli, M., Galuba, W., ... & Feichtenhofer, C. (2022). Masked autoencoders that listen. *NeurIPS*, *35*, 28708-28720.

[12] Chen, S., Wu, Y., Wang, C., Liu, S., Tompkins, D., Chen, Z., ... & Wei, F. (2023). BEATs: audio pre-training with acoustic tokenizers. En ICML 2023 (pp. 5178-5193).

[14] Turian, J., Shier, J., Khan, H. R., Raj, B., Schuller, B. W., Steinmetz, C. J., ... & Bisk, Y. (2022, July). Hear: Holistic evaluation of audio representations. En *NeurIPS 2021 Competitions and Demonstrations Track* (pp. 125-145). PMLR.

[15] Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*(6), 1281-1296.

[16] Boebinger, D., Norman-Haignere, S. V., McDermott, J. H., & Kanwisher, N. (2021). Music-selective neural populations arise without musical training. *Journal of Neurophysiology*.

Universidad: Universidad Abierta Interamericana (UAI)
Facultad: Facultad de Tecnología Informática
Título a obtener: Doctor en Informática
Tesis: Analíticos Visuales con Inteligencia Artificial Aplicados en Ciencia Marina
Doctorando: Lic. Luis Dario Ceballos
Director: Dr. Marcos Daniel Zárate
Fecha de exposición: 24/02/2026

---

1. INTRODUCCIÓN

1.1. Contexto y Motivación

La investigación en ciencias marinas genera volúmenes masivos de datos provenientes de sensores, expediciones y satélites. Sin embargo, existe una **"brecha de visualización"**: los científicos a menudo poseen datos valiosos pero carecen de las herramientas automatizadas o del tiempo necesario para generar representaciones visuales que respeten los principios de percepción humana. La motivación principal radica en democratizar el acceso a analíticos avanzados, permitiendo que investigadores con distintos niveles de habilidad técnica puedan explorar datos complejos, especialmente aquellos estructurados como **Linked Data (RDF)**, que las herramientas de oficina convencionales no logran interpretar.

1.2. Planteamiento del Problema

El problema se centra en la "brecha de visualización": los científicos a menudo poseen los datos pero carecen de las herramientas o el tiempo para crear visualizaciones que respeten los principios de percepción humana. Además, los datos enlazados (*Linked Data*) presentan estructuras de grafos que no son interpretables por las herramientas de oficina convencionales.

1.3. Objetivos

- **General:** Desarrollar un sistema de recomendación inteligente de visualizaciones para datos marinos.
- **Específicos:** Diseñar una arquitectura que procese datos RDF y tabulares; optimizar la generación de código visual mediante *prompt engineering*; y validar la herramienta con usuarios reales del dominio científico.

2. MARCO TEÓRICO Y ESTADO DEL ARTE

2.1. Fundamentos de Visualización Científica

Se analizan los modelos de procesamiento visual, destacando la importancia de las variables visuales de Bertin (posición, color, tamaño) y su efectividad según el tipo de dato (nominal, ordinal o cuantitativo). Se enfatiza que una visualización incorrecta puede llevar a conclusiones científicas erróneas.

2.2. Tecnologías Emergentes: LLMs y Web Semántica

Se describe el funcionamiento de los Transformadores (GPT) y su capacidad para actuar como "razonadores" sobre esquemas de datos. Se fundamenta el uso de la Web Semántica para proporcionar contexto y significado a los datos marinos, permitiendo que la máquina "entienda" qué representa cada variable.

2.3. Estado del Arte

Se realiza una revisión crítica de herramientas como LIDA, Chat2VIS y GPT4Vis. Se identifica que la mayoría están diseñadas para datos comerciales y carecen de la especialización necesaria para manejar la dimensionalidad espacial y temporal característica de la oceanografía.

3. DESARROLLO DEL SISTEMA VISBOT

3.1. Arquitectura Modular

VisBot se compone de tres módulos centrales:

1. **Capa de Datos:** Maneja la ingesta de archivos CSV y consultas SPARQL sobre *endpoints* RDF.
2. **Capa Lógica (LLM Engine):** Utiliza una técnica de *few-shot prompting* para instruir al modelo sobre cómo transformar metadatos en especificaciones de visualización.
3. **Capa de Presentación:** Una interfaz desarrollada en Streamlit que renderiza gráficos interactivos utilizando la librería Plotly.

3.2. Implementación Técnica y Flujo de Trabajo

El flujo comienza con el perfilado del dataset (estadísticas básicas) que se envía al LLM. El modelo devuelve una justificación del gráfico elegido y el código Python necesario para generarlo. El sistema implementa un mecanismo de "auto-corrección" donde, si el código falla, el LLM recibe el error y lo subsana automáticamente.

4. CASO DE ESTUDIO: CIENCIA MARINA

Este capítulo presenta el caso de estudio desarrollado como validación empírica de las contribuciones teóricas y metodológicas articuladas en capítulos precedentes, documentando exhaustivamente el sistema VisBot (Ceballos, 2025) como materialización concreta de un sistema de recomendación de visualizaciones basado en modelos de lenguaje extensos. El desarrollo completo está explicado en Ceballos et al. (2025), este artículo contiene material complementario proporcionado por los autores disponible en https://doi.org/10.5281/zenodo.15176569. Se puede acceder a la plataforma en https://visbot.streamlit.app/ y el repositorio está disponible públicamente en https://github.com/disenodc/VisBot. El acuerdo de licencia completo está disponible en https://github.com/disenodc/VisBot/blob/main/LICENSE.

4.1. Motivación del Dominio y Datos Utilizados

Se seleccionó como caso de estudio la biodiversidad del Golfo San Jorge. Los datos incluyen registros de expediciones científicas con variables como temperatura, salinidad y abundancia de especies (ej. *Pleoticus muelleri*). Estos datos son inherentemente complejos debido a su componente geoespacial.

4.2. Integración y Resultados en Ciencia Marina

VisBot demostró ser capaz de generar mapas de calor y gráficos de dispersión 3D que relacionan variables físico-químicas con la presencia de especies, permitiendo a los biólogos identificar zonas de alta productividad de manera inmediata.

4.3. Arquitectura y Componentes Funcionales

VisBot (Ceballos, 2025) es una plataforma web modular desarrollada en Python y desplegada en **Streamlit Cloud**. Su arquitectura se fundamenta en tres pilares: modularidad (desacoplamiento de componentes), extensibilidad (capacidad de escalar funciones) y eficiencia computacional (uso de caché y optimización algorítmica).

El sistema sigue un flujo de desarrollo secuencial de cinco etapas que transforman datos crudos en visualizaciones inteligentes:

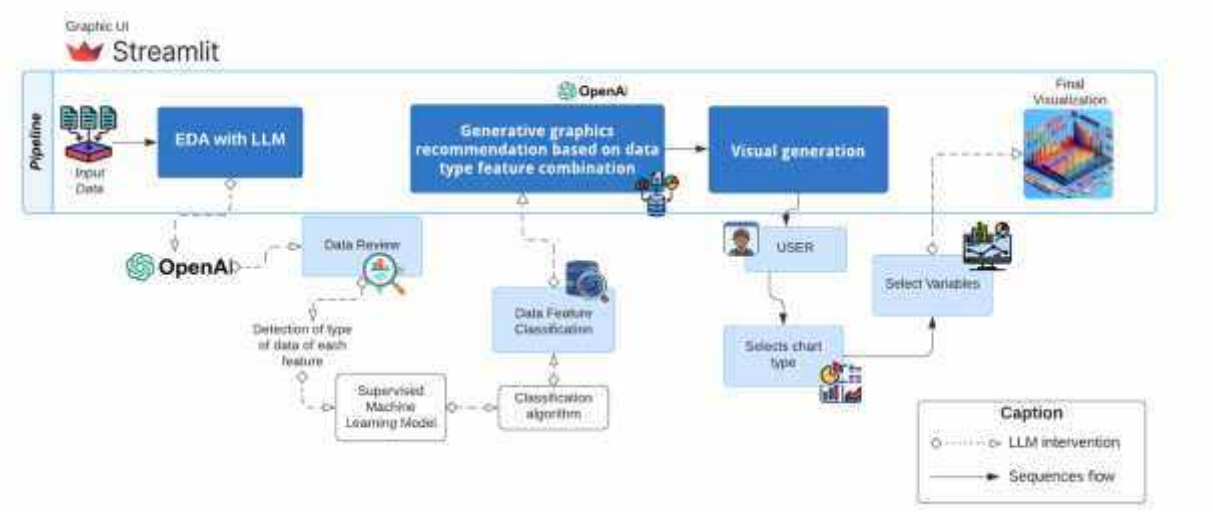


FIGURA 1. Arquitectura de VisBot: 1) EDA con LLM; 2) recomendación de gráficos generativos basada en la combinación de características del tipo de datos; y 3) generación visual y mejoras adicionales; 4) visualización final generada.

Módulo 1: Carga y Parsing de Datos

La ingesta se realiza mediante una interfaz gráfica que abstrae la complejidad técnica. Utiliza la clase `DataLoader` y parsers polimórficos (`CSV`, `Excel`, `JSON`, `RDF`) para estandarizar cualquier entrada en un **Pandas DataFrame**. Soporta además la carga desde URLs remotas y el procesamiento de datos enlazados (RDF) mediante consultas SPARQL.

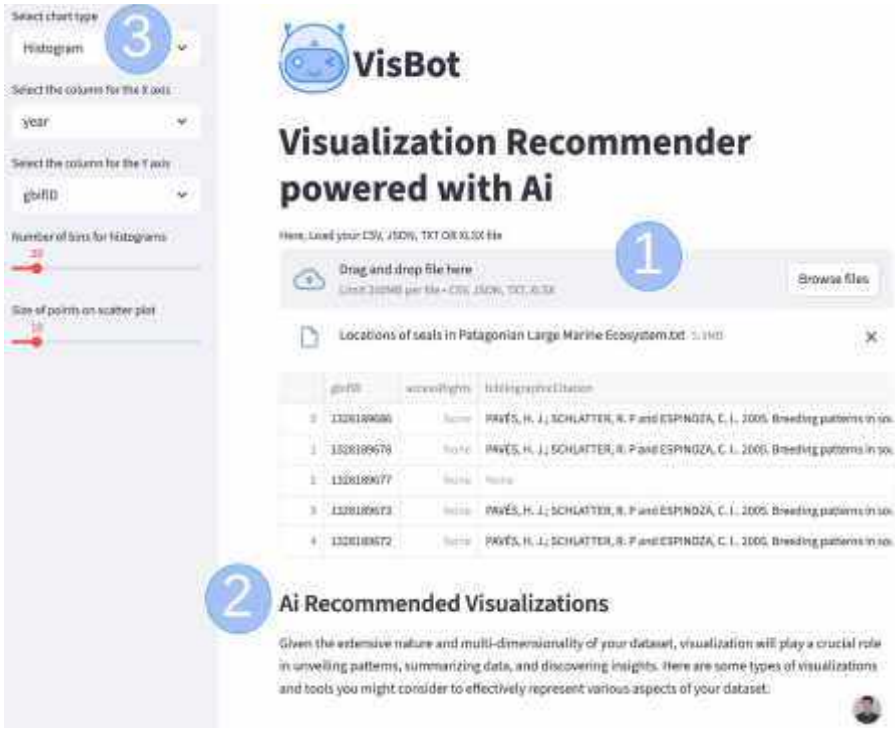


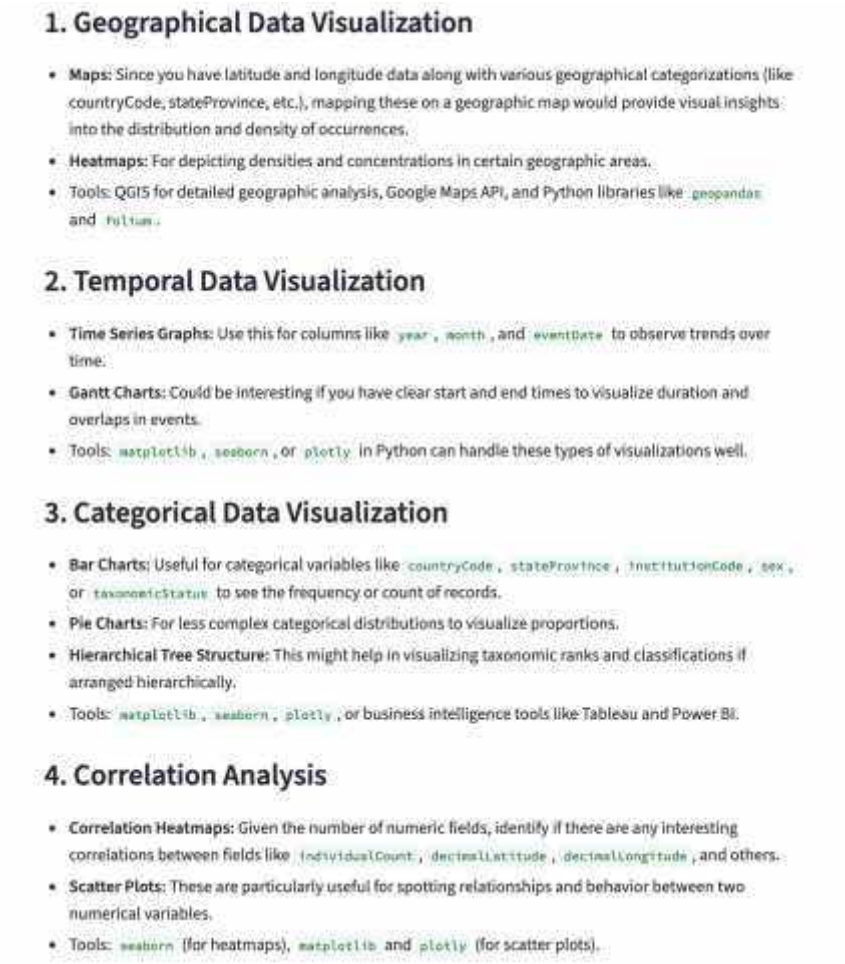
FIGURA 2. Interfaz de usuario de la plataforma VisBot y sus módulos principales: 1) Cargador de datos, 2) Visualizaciones recomendadas por IA, y 3) Módulo de configuración de gráficos.

Módulo 2: Análisis Exploratorio Automatizado (EDA)

A través de la clase `DataProfiler`, el sistema aplica estadística descriptiva y ML no supervisado. Este módulo detecta automáticamente tipos semánticos (fechas, coordenadas geoespaciales, URIs y taxonomías biológicas) mediante expresiones regulares y validación de rangos, permitiendo caracterizar el dataset antes de la recomendación.

Módulo 3: Recomendación mediante LLM (GPT-4)

Es el núcleo de automatización que encapsula la comunicación con la API de OpenAI. Implementa *prompt engineering* jerárquico (System, Context y Query Messages) para razonar sobre el perfil del dataset basándose en principios de percepción visual (Cleveland, McGill, Tufte). Las recomendaciones se generan en formato JSON para garantizar un procesamiento robusto.



1. Geographical Data Visualization

- **Maps:** Since you have latitude and longitude data along with various geographical categorizations (like countryCode, stateProvince, etc.), mapping these on a geographic map would provide visual insights into the distribution and density of occurrences.
- **Heatmaps:** For depicting densities and concentrations in certain geographic areas.
- Tools: QGIS for detailed geographic analysis, Google Maps API, and Python libraries like `geopandas` and `folium`.

2. Temporal Data Visualization

- **Time Series Graphs:** Use this for columns like `year`, `month`, and `eventDate` to observe trends over time.
- **Gantt Charts:** Could be interesting if you have clear start and end times to visualize duration and overlaps in events.
- Tools: `matplotlib`, `seaborn`, or `plotly` in Python can handle these types of visualizations well.

3. Categorical Data Visualization

- **Bar Charts:** Useful for categorical variables like `countryCode`, `stateProvince`, `institutionCode`, `sex`, or `taxonomicStatus` to see the frequency or count of records.
- **Pie Charts:** For less complex categorical distributions to visualize proportions.
- **Hierarchical Tree Structure:** This might help in visualizing taxonomic ranks and classifications if arranged hierarchically.
- Tools: `matplotlib`, `seaborn`, `plotly`, or business intelligence tools like Tableau and Power BI.

4. Correlation Analysis

- **Correlation Heatmaps:** Given the number of numeric fields, identify if there are any interesting correlations between fields like `individualCount`, `decimalLatitude`, `decimalLongitude`, and others.
- **Scatter Plots:** These are particularly useful for spotting relationships and behavior between two numerical variables.
- Tools: `seaborn` (for heatmaps), `matplotlib` and `plotly` (for scatter plots).

FIGURA 3. Visualizaciones recomendadas por VisBot

Módulo 4: Generación de Visualizaciones (Plotly Express)

Este módulo traduce las recomendaciones abstractas en objetos gráficos concretos. Implementa métodos especializados como `create_scatter_map()` para contextos locales y `create_scatter_geo()` para proyecciones globales, permitiendo codificar hasta cinco dimensiones simultáneamente.

Módulo 5: Customización Interactiva y Reactividad

Proporciona un panel de control basado en widgets de Streamlit. El usuario puede ajustar variables de ejes, paletas de colores accesibles (*colorblind-safe*) y parámetros estadísticos (como el número de *bins* en histogramas) con re-renderizado inmediato gracias al estado de sesión de la plataforma.

4.3.3 Implementación de Visualizaciones Especializadas

El sistema aplica técnicas avanzadas para manejar la densidad y complejidad de los datos científicos, tales como:

- **Clustering Adaptativo:** Para evitar la saturación en mapas con miles de registros.

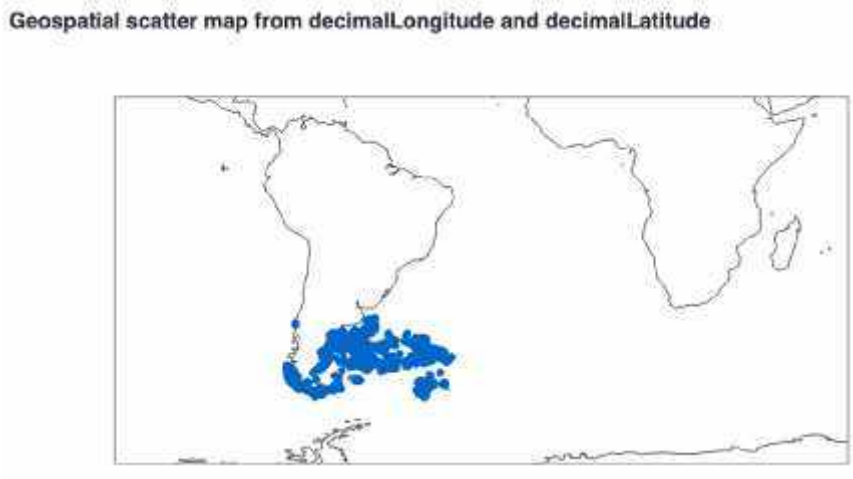


Figura 4. Mapa de dispersión geoespacial a partir de latitud y longitud

- **Optimización Perceptual de Color:** Uso de paletas de Brewer (1994) para distinguir taxones biológicos de forma efectiva.

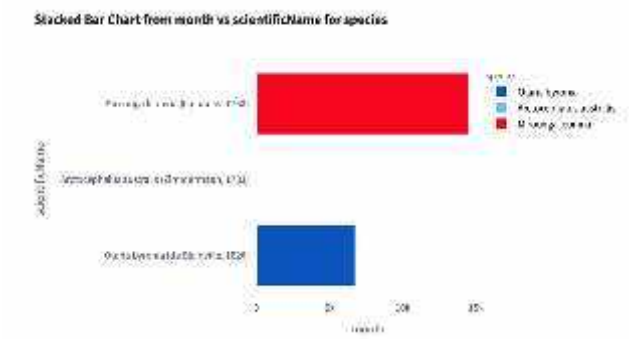


Figura 5. Gráfico de barras apiladas, de ocurrencias de Especies (usando nombre científico) por mes, caracterizadas por color según especie.

- **Análisis Estadístico de Distribución:** Aplicación de la regla de Sturges para *binning* y superposición de curvas de densidad (KDE).

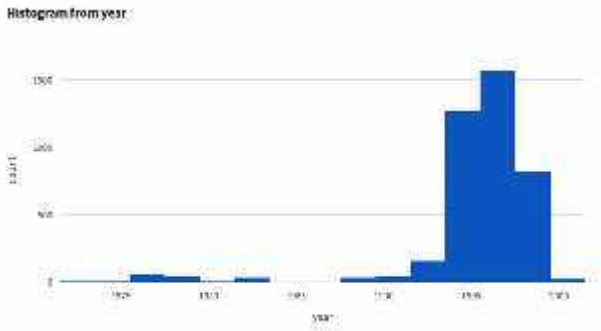


Figura 6. Histograma por Año

- **Detección de Patrones Ecológicos:** Cálculo de correlaciones de Pearson y clustering jerárquico (Algoritmo de Ward) para identificar gremios biológicos.

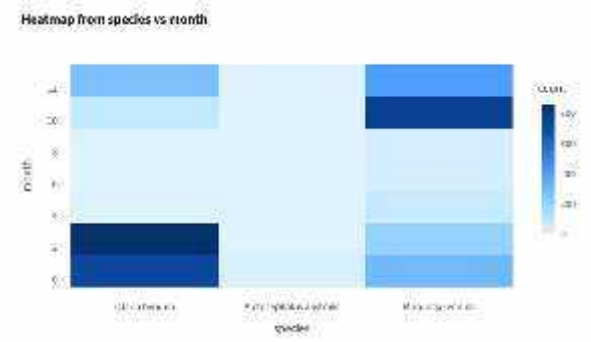


Figura 7. Mapa de calor de especies por mes

4.3.1. Selección del Dataset para el Caso de Estudio

Para validar el sistema en el dominio de las ciencias marinas, se utilizó el dataset *"Locations of Seals in Patagonian Large Marine Ecosystem"* (Campagna et al., 2015), disponible en GBIF. Este conjunto de datos real del **CENPAT-CONICET** incluye registros georreferenciados de alta precisión, datos temporales y clasificación taxonómica de especies como el león marino sudamericano (*Otaria flavescens*) y el elefante marino del sur (*Mirounga leonina*).

La evaluación del sistema arrojó resultados satisfactorios:

- **Relevancia:** 4.3/5 (Escala Likert).
- **Diversidad:** Promedio de 4 tipos de visualización por dataset.
- **Usabilidad (SUS):** 78.5/100 (Calificación "Bueno").

4.4. Resultados

La evaluación de **VisBot** se estructuró en tres dimensiones: rendimiento cuantitativo, experiencia de usuario y análisis comparativo.

4.4.1. Métricas y Criterios de Evaluación

Como se detalla en la **Tabla 1**, las tareas simples (Task 1) presentan una tasa de éxito del 98%, mientras que tareas complejas de interpretación de IA (Task 2 y 3) rondan el 90%.

Tarea	Descripción	Tiempo Promedio	Tasa de Éxito
Task 1	Carga del conjunto de datos *California Housing*	1.5 seg	98%
Task 2	Revisión de recomendaciones generadas por IA	2.1 min	92%
Task 3	Generación de 3 gráficos siguiendo recomendaciones	2.5 min	88%
Task 4	Creación de gráfico de barras comparativo	1.2 min	100%
Task 5	Elaboración de diagrama de dispersión relacional	1.5 min	98%
Task 6	Generación de visualización geoespacial	1.8 min	86%

Tabla 1. Métricas promedio de tiempo de completitud y tasa de éxito por tarea experimental

La **Tabla 2** profundiza en la eficiencia operativa, destacando una baja tasa de error (1.2 por sesión).

Dimensión	Métrica	Puntuación Promedio	Desviación Estándar	Escala
Relevancia y Precisión	Relevancia	4.3	0.4	1-5
	Exactitud Semántica	4.6	0.3	1-5
	Correctitud del Mapeo	87%	6%	0-100%
	NDCG@5	0.82	0.07	0-1
Diversidad	Tipos de Gráficos Sugeridos	4/5	-	conteo
Eficiencia	Tiempo por Tarea	42 seg	6 seg	segundos
	Tasa de Éxito en Tareas	92%	5%	0-100%
	Interacciones por Tarea	6.3	1.1	conteo
	Tasa de Errores	1.2	0.5	por sesión
Usabilidad	Puntuación SUS	78.5	4.2	0-100
	CSAT	4.5	0.3	1-5
	NPS	+64	-	-100–+100
Funcionalidad	Manejo de Ambigüedad	3.9	0.5	1-5
	Explicabilidad	75%	10%	0-100%

Tabla 2. Síntesis de métricas de evaluación y resultados promedio del sistema VisBot

Finalmente, la satisfacción se validó con un **SUS de 78.5**, un **CSAT de 4.5/5** y un **Net Promoter Score (NPS) de +64**, situando a VisBot como una herramienta excelente en el ámbito de software especializado, a pesar de tener un nivel de personalización intermedio frente a herramientas comerciales como Tableau.

5. DISCUSIÓN Y LIMITACIONES

5.1. Capacidades de VisBot

Se discute cómo la IA actúa como un mediador entre el dato crudo y la percepción humana. El sistema no solo dibuja gráficos, sino que ofrece una capa de interpretación que asiste al investigador en la fase exploratoria.

5.2. Limitaciones Identificadas

Se identifican desafíos como la dependencia de APIs propietarias, el costo por token en datasets extremadamente grandes y la posibilidad de alucinaciones en los LLM si el contexto de los datos es ambiguo.

6. CONCLUSIONES

La tesis concluye que la integración de IA y tecnologías semánticas es el camino para la próxima generación de herramientas de analítica visual. VisBot cumple con los objetivos de accesibilidad y rigor técnico, proporcionando una solución viable para los desafíos de datos en ciencia marina.

7. Líneas Futuras

A partir de las conclusiones y los límites identificados en su investigación sobre **VisBot** (Ceballos, 2025), se proponen las siguientes líneas de investigación futuras, organizadas por dimensiones técnicas, disciplinarias y éticas:

7.1. Evolución Técnica y Arquitectónica

- **Integración de Modelos de Lenguaje Locales (Edge LLMs):** Para mitigar la dependencia de APIs propietarias y reducir costos operativos, se propone investigar el rendimiento de modelos de código abierto (ej. **Llama 3** o **Mistral**) ejecutados en servidores locales. Esto garantizaría la privacidad de datos científicos sensibles y permitiría el ajuste fino (*fine-tuning*) con vocabularios específicos de la oceanografía.
- **Análisis Visual Multimodal:** Extender las capacidades de VisBot para que no solo procese datos tabulares o RDF, sino que pueda interpretar imágenes (satelitales, mapas de calor preexistentes o fotografías de especies) integrando modelos de visión por computadora para una recomendación de 360 grados.
- **Arquitectura de Procesamiento Distribuido:** Ante el desafío de datasets masivos, es necesario escalar la arquitectura modular hacia entornos de **Big Data** (ej. integración con Apache Spark), permitiendo que el perfilado de datos y la generación visual mantengan tiempos de respuesta interactivos.

7.2. Enriquecimiento Semántico y Calidad de Datos

- **Sistemas de Diálogo Predictivo (Conversational Analytics):** Implementar un módulo de clarificación activa donde el sistema, ante datos ambiguos o metadatos incompletos, inicie un diálogo con el investigador para refinar la intención analítica antes de generar el gráfico.
- **Enriquecimiento Automático mediante Knowledge Graphs:** Investigar la conexión en tiempo real con grafos de conocimiento globales (como **DBpedia** o **WoRMS**) para completar automáticamente información taxonómica o ambiental faltante en el dataset de entrada, mejorando la precisión de las recomendaciones.

7.3. Aplicación y Validación en Ciencia Marina

- **Visualización de Series Temporales 4D:** Desarrollar módulos específicos para la representación de procesos oceanográficos que ocurren en cuatro dimensiones (latitud,

longitud, profundidad y tiempo), optimizando la renderización de perfiles batimétricos dinámicos.

- **Colaboración Humano-IA en Tiempo Real:** Evaluar el impacto de VisBot en entornos de toma de decisiones críticas, como cruceros oceanográficos en curso, donde la velocidad de visualización puede orientar el cambio de rumbo de una expedición científica.

7.4. Dimensión Ética y Cognitiva

- **Auditoría de Sesgos en Recomendaciones:** Desarrollar métricas para detectar y corregir sesgos en las sugerencias del LLM (por ejemplo, una tendencia sistemática a recomendar ciertos tipos de gráficos sobre otros debido a los datos de entrenamiento del modelo).
- **Estudio de la Carga Cognitiva a Largo Plazo:** Realizar estudios longitudinales para determinar si el uso prolongado de asistentes de IA como VisBot mejora las capacidades analíticas de los científicos o si, por el contrario, genera una dependencia excesiva que limite el pensamiento crítico sobre la integridad de los datos.

7.5. Transferibilidad a Otros Dominios

- **Generalización a Dominios Sensibles:** Adaptar la metodología de "razonamiento visual" de VisBot a áreas como la epidemiología o la gestión de políticas públicas, donde la transparencia y la trazabilidad de por qué se eligió una visualización específica son fundamentales para la comunicación de riesgos.

REFERENCIAS BIBLIOGRÁFICAS

Bertin, J. (1967). *Sémiologie Graphique*. Gauthier-Villars.

Dibia, V. (2023). LIDA: A tool for automatic generation of grammar-agnostic visualizations and infographics using large language models. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 113–126.

Maddigan, P., & Susnjak, T. (2023). *Chat2Vis: Fine-tuning data visualisations using multilingual natural language text and pre-trained large language models*. arXiv. https://doi.org/10.48550/arXiv.2303.14292

Wu, Y., Wan, Y., Zhang, H., Sui, Y., Wei, W., Zhao, W., Xu, G., & Jin, H. (2024). Automated data visualization from natural language via large language models: An exploratory study. *Proceedings of the ACM on Management of Data, 2*(3), Article 149.

Streamlit Inc. (2023). *Streamlit* (Versión 1.28) [Software de computación]. https://streamlit.io

Plotly Technologies Inc. (2015). *Collaborative data science* [Software]. https://plot.ly

Ceballos, D. (2025a). *VisBot* [Aplicación web]. https://visbot.streamlit.app/

Ceballos, D., Zárate, M., & Delrieux, C. (2025). *Enhancing data analytics with AI-powered visualization recommendations: A large language model approach*. TechRxiv. https://doi.org/10.36227/techrxiv.175303443.39749915/v1

OpenAI. (2023). *GPT-4 technical report*. arXiv. https://doi.org/10.48550/arXiv.2303.08774

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*(387), 531–554.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. *IEEE Symposium on Visual Languages*, 336–343.

Tufte, E. R. (2001). *The visual display of quantitative information* (2.ª ed.). Graphics Press.

Brewer, C. A. (1994). Color use guidelines for mapping and visualization. *Visualization in Modern Cartography*, 123–147.

Sturges, H. A. (1926). The choice of a class interval. *Journal of the American Statistical Association*, *21*(153), 65–66.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244.

Campagna, C., Lewis, M. N., & Marin, M. R. (2015). *Locations of seals in Patagonian Large Marine Ecosystem* (v1.6). ArOBIS Centro Nacional Patagónico. https://doi.org/10.15468/qcwa73

Brooke, J. (1996). SUS: A quick and dirty usability scale. En P. W. Jordan, B. Thomas, B. A. Weerdmeester, e I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Behavior-Driven Microservice Architecture: un marco metodológico para la identificación iterativa de microservicios en contextos ágiles greenfield

Autor: Dr. Nicolás Battaglia (nicolas.battaglia@uai.edu.ar)
Directores: Dr. Gustavo Rossi (Gustavo.Rossi@UAI.edu.ar), Dr. Alejandro Fernández (casco@lifia.info.unlp.edu.ar)
Fecha de defensa: 22-12-2025
Universidad Abierta Interamericana – Facultad de Tecnología Informática
Buenos Aires, Argentina

Resumen

Las arquitecturas de microservicios se han consolidado en el desarrollo de software ágil; sin embargo, en contextos greenfield, la identificación de límites de servicio continúa siendo un desafío significativo. En estos escenarios, las decisiones arquitectónicas suelen tomarse bajo alta incertidumbre y con conocimiento parcial del dominio, apoyándose en criterios implícitos o experiencia individual. A pesar de la disponibilidad de prácticas como Domain-Driven Design (DDD) y Behavior-Driven Development (BDD), la literatura evidencia la ausencia de enfoques metodológicos que conecten de manera sistemática el comportamiento funcional con las decisiones arquitectónicas. Esta tesis aborda este problema desde un enfoque de Design Science Research (DSR). El problema se caracteriza mediante un ciclo de relevancia que integra evidencia empírica y análisis del estado del arte, y se valida a través de un ciclo de rigor basado en el análisis crítico de trabajos relacionados, identificando un vacío metodológico en la identificación de microservicios en entornos ágiles greenfield. Como respuesta, se propone Behavior-Driven Microservice Architecture (BDMA), un marco metodológico que integra BDD y DDD en un proceso iterativo orientado a derivar decisiones arquitectónicas a partir del comportamiento funcional. La evaluación exploratoria del marco, realizada mediante un experimento comparativo, aporta evidencia inicial de mejoras en la coherencia arquitectónica, la alineación con el dominio y la explicitación del razonamiento arquitectónico, junto con una curva de aprendizaje inicial. En conjunto, la tesis contribuye con un marco metodológico para la identificación de microservicios en contextos ágiles greenfield, abordando una brecha relevante en la literatura y sentando bases para futuras validaciones empíricas. Estos resultados abren la puerta a futuras investigaciones orientadas a la validación del marco en contextos industriales y al desarrollo de herramientas de soporte metodológico.

Introducción y contexto

Las arquitecturas de microservicios se han consolidado como un enfoque predominante para el desarrollo de sistemas software complejos; sin embargo, la correcta identificación de los límites de los microservicios continúa siendo un desafío central, particularmente en contextos ágiles greenfield, donde las decisiones arquitectónicas tempranas se toman bajo alta incertidumbre y sin referencias estructurales previas (Battaglia, 2025; Battaglia et al., 2024; Zhong et al., 2024). En ausencia de guías metodológicas explícitas, estas decisiones suelen basarse en criterios intuitivos o experiencia individual, lo que incrementa el riesgo de arquitecturas frágiles, fronteras inconsistentes y acumulación temprana de deuda técnica (Booch, 2006; Ford et al., 2017; Zimmermann, 2017). Prácticas consolidadas como Domain-Driven Design (DDD) y Behavior-Driven Development (BDD) aportan conceptos relevantes para el modelado del dominio y el descubrimiento funcional, respectivamente; sin embargo, la literatura evidencia que su aplicación suele ser fragmentada y carente de un proceso metodológico que articule explícitamente el comportamiento funcional con las decisiones arquitectónicas involucradas en el diseño de microservicios (Battaglia et al., 2025;

Oumoussa & Saidi, 2024; Taibi & Lenarduzzi, 2018). De forma consistente, los enfoques existentes para la identificación de microservicios abordan el problema de manera parcial, sin ofrecer un proceso integral, trazable y reproducible que conecte sistemáticamente el comportamiento del sistema con la arquitectura resultante. En este contexto, la literatura reciente enfatiza la necesidad de concebir la arquitectura como una hipótesis evolutiva, sujeta a validación y adaptación continua (Ford et al., 2017; Zimmermann, 2017), pero persiste un vacío metodológico respecto de cómo operacionalizar este principio en proyectos ágiles greenfield. En respuesta a esta limitación, la presente investigación plantea la necesidad de un marco metodológico sistemático, iterativo y reproducible que integre prácticas de descubrimiento funcional y modelado del dominio, proporcionando mecanismos explícitos de trazabilidad, validación y evolución arquitectónica frente a cambios de requisitos.

Problema de investigación y objetivos

A partir del análisis anterior, se identifican tres obstáculos principales que limitan la correcta identificación de microservicios en contextos ágiles *greenfield*: 1) La falta de integración metodológica entre DDD y enfoques ágiles, lo cual dificulta su aplicación en ciclos iterativos y colaborativos (Zhong et al., 2024). 2) La ausencia de mecanismos que aseguren la trazabilidad entre requerimientos funcionales y decisiones arquitectónicas, debilitando la validación continua del diseño (Kapferer & Zimmermann, 2020; Su et al., 2024). 3) La carencia de estrategias sistemáticas para delimitar microservicios en proyectos sin sistemas preexistentes, donde el conocimiento del dominio es aún parcial (Bajaj et al., 2022; Vural & Koyuncu, 2021)

Objetivo principal

- Diseñar y validar un marco metodológico denominado *Behavior-Driven Microservice Architecture*, que integre BDD y DDD, para facilitar la identificación iterativa de microservicios en proyectos ágiles *greenfield*, favoreciendo la trazabilidad funcional, la coherencia arquitectónica y la adaptación a cambios del dominio.

Objetivos secundarios

- Caracterizar los desafíos metodológicos actuales en la identificación de microservicios en entornos ágiles *greenfield*, para comprender las limitaciones de DDD en ciclos iterativos y la escasa trazabilidad entre requisitos funcionales y decisiones arquitectónicas.
- Diseñar la metodología BDMA como una secuencia iterativa de fases que combine prácticas de BDD con artefactos estratégicos de DDD para favorecer la identificación progresiva de microservicios.
- Implementar y documentar la aplicación de BDMA en un caso de estudio representativo del dominio.
- Evaluar empíricamente el impacto de BDMA

Marco conceptual

El marco conceptual se introduce para proporcionar las bases teóricas necesarias que permitan analizar críticamente los enfoques existentes y validar el problema identificado en el ciclo de relevancia, en coherencia con el enfoque DSR adoptado. El diseño de arquitecturas de microservicios se fundamenta en la descomposición de sistemas en servicios autónomos y alineados con capacidades del negocio; sin embargo, la literatura señala que este paradigma desplaza la complejidad hacia la

correcta definición de límites y responsabilidades. En este contexto, Domain-Driven Design aporta conceptos clave como lenguaje ubicuo y bounded contexts para estructurar dominios complejos y alinear decisiones arquitectónicas con el dominio, aunque requiere integrarse en un proceso ágil más amplio para resultar operativo. Complementariamente, Behavior-Driven Development facilita el descubrimiento y la validación temprana del comportamiento funcional mediante escenarios colaborativos, pero no ofrece mecanismos explícitos para derivar decisiones arquitectónicas. Finalmente, la arquitectura ágil y evolutiva concibe la arquitectura como una hipótesis sujeta a validación continua, enfatizando la necesidad de procesos que permitan su adaptación progresiva a medida que evoluciona el conocimiento del dominio y los requisitos.

Marco Metodológico

Esta investigación adopta Design Science Research (DSR) como enfoque metodológico, dado que busca diseñar y evaluar un artefacto prescriptivo (un método) para resolver un problema relevante de la práctica en ingeniería de software (Hevner et al., 2004; Peffers et al., 2007). En el marco de DSR, la construcción del problema de investigación se fundamenta mediante la interacción entre ciclo de relevancia (contexto y necesidad real) y ciclo de rigor (sustento teórico y análisis crítico de conocimiento existente).

Ciclo de relevancia: identificación empírica del problema

El ciclo de relevancia de la investigación se orientó a comprender las dificultades que enfrentan los equipos al diseñar arquitecturas de microservicios en contextos ágiles greenfield. Para ello, se realizó un análisis del contexto que combinó un mapeo sistemático de la literatura y un experimento exploratorio. El mapeo permitió identificar la coexistencia de enfoques parciales y fragmentados para la identificación de microservicios, sin un proceso integral que articule el comportamiento funcional con las decisiones arquitectónicas. Complementariamente, el experimento evidenció una alta variabilidad en las arquitecturas propuestas, dificultades para justificar los límites de servicio y una baja trazabilidad entre requisitos y arquitectura. En conjunto, ambos insumos permitieron caracterizar empíricamente el problema, mostrando que la principal dificultad radica en la ausencia de un proceso metodológico sistemático que guíe la identificación y evolución de microservicios desde las etapas tempranas del desarrollo.

Análisis crítico preliminar y formulación del problema

A partir de la evidencia obtenida en el ciclo de relevancia, se realizó un análisis crítico preliminar orientado a validar la consistencia y relevancia del problema identificado. Este análisis permitió identificar una brecha recurrente: si bien existen prácticas consolidadas para el descubrimiento funcional y para el modelado estratégico del dominio, no se dispone de un proceso metodológico que las articule de manera sistemática, trazable y reproducible para la identificación de microservicios en contextos greenfield. Como consecuencia, las decisiones arquitectónicas tempranas carecen de fundamentos explícitos, lo que dificulta su validación y adaptación frente a cambios de requisitos. Si bien esta caracterización se apoya en evidencia empírica y análisis contextual, su consolidación requiere contrastar dichas observaciones con los enfoques existentes, tarea que se aborda posteriormente mediante el ciclo de rigor.

Ciclo de rigor: análisis de trabajos relacionados

En el marco de Design Science Research, el ciclo de rigor tiene como objetivo asegurar que el problema identificado en el ciclo de relevancia se encuentre correctamente fundamentado y posicionado dentro del cuerpo de conocimiento existente. En esta investigación, dicho ciclo se orienta a analizar críticamente los enfoques propuestos en la literatura para el diseño e identificación de microservicios, con el fin de contrastar las observaciones empíricas iniciales y validar la existencia de un vacío metodológico.

Análisis de trabajos relacionados

El análisis de trabajos relacionados evidencia una amplia diversidad de enfoques para la identificación de microservicios, que abarcan heurísticas de dominio, modelos conceptuales, análisis estructural y técnicas algorítmicas basadas en minería de requisitos o clustering. Si bien estas propuestas aportan técnicas valiosas, la mayoría se focaliza en dimensiones particulares del problema sin ofrecer un proceso metodológico integral que articule de manera coherente estructura, comportamiento y dominio. Asimismo, muchos enfoques asumen la existencia de modelos previos o conocimiento consolidado del dominio, lo que limita su aplicabilidad en contextos greenfield. La literatura muestra además que la integración entre prácticas de descubrimiento funcional y modelado estratégico del dominio suele abordarse de forma implícita, sin definir mecanismos explícitos para derivar decisiones arquitectónicas trazables ni para validarlas y ajustarlas frente a la evolución de los requisitos. Como consecuencia, las propuestas analizadas presentan dificultades para justificar, comparar, reproducir o adaptar las decisiones de identificación de microservicios en distintos contextos.

Síntesis crítica y validación del problema

La revisión crítica de los trabajos relacionados permite validar las observaciones surgidas en el ciclo de relevancia y consolidar la formulación del problema de investigación. En particular, el análisis confirma que, si bien existen numerosas técnicas y enfoques parciales para la identificación de microservicios, persiste un vacío metodológico en cuanto a la existencia de un proceso iterativo, sistemático y reproducible que conecte el descubrimiento funcional del sistema con las decisiones arquitectónicas necesarias para definir y evolucionar microservicios en contextos ágiles *greenfield*. Esta síntesis refuerza la idea de que el problema no radica en la ausencia de prácticas o conceptos, sino en la falta de un marco que permita orquestar dichas prácticas dentro de un proceso coherente, manteniendo la trazabilidad entre requisitos, modelos de dominio y arquitectura. En consecuencia, se justifica la necesidad de diseñar un artefacto metodológico que integre estos enfoques y permita abordar de manera explícita la identificación y evolución de microservicios desde las primeras etapas del desarrollo. La convergencia entre el ciclo de relevancia y el ciclo de rigor establece así una base sólida para el ciclo de diseño, en el cual se propone y valida el marco metodológico BDMA como respuesta al problema identificado.

Ciclo de Diseño: Behavior-Driven Microservice Architecture (BDMA)

Como resultado del ciclo de diseño, esta tesis propone Behavior-Driven Microservice Architecture (BDMA) como un marco metodológico para diseñar microservicios en contextos ágiles *greenfield*, respondiendo a las brechas detectadas. BDMA integra principios estratégicos (DDD), tácticas funcionales (BDD) y prácticas de arquitectura evolutiva, operacionalizándolo en un conjunto de fases iterativas con actividades y artefactos concretos.

Alcance y premisa de entrada

BDMA parte de la premisa de que el proyecto ya dispone de un backlog funcional compuesto por escenarios BDD validados. El aporte del método se centra en cómo utilizar esos escenarios como insumo para el diseño arquitectónico, manteniendo trazabilidad y validación.

Estructura del método: cinco fases operativas

BDMA se compone de cinco fases operativas y verificables.

Fase 1: Clasificación funcional de escenarios:

Esta fase tiene como objetivo organizar el backlog funcional del sistema —expresado en escenarios BDD redactados en lenguaje Gherkin— en agrupaciones semánticamente coherentes que representen unidades funcionales relevantes como features, épicas o procesos de negocio. Esta clasificación permite descubrir límites funcionales naturales que pueden derivar en futuros bounded contexts. Inspirada en los principios de arquitectura evolutiva (Ford et al., 2017) y continuous architecture (Erder & Pureur, 2015; Woods et al., 2021). Su aplicación temprana garantiza que la arquitectura se alinee con las necesidades del negocio.

Fase 2: Extracción colaborativa de elementos del dominio:

Una vez agrupados los escenarios, esta fase se enfoca en identificar, mediante prácticas colaborativas, los elementos fundamentales del dominio expresados en cada escenario. El equipo técnico, junto con los expertos del negocio, lleva adelante sesiones estructuradas de Event Storming (Brandolini, 2018) y Example Mapping (Wynne, 2015) para revelar actores, comandos, eventos de dominio y entidades involucradas. Estas sesiones permiten descomponer el comportamiento deseado en piezas semánticas reutilizables, estableciendo una base conceptual sólida para el modelado estratégico. El lenguaje ubicuo emerge como producto de esta colaboración, y los comandos y eventos extraídos alimentan directamente el diseño de contextos e interfaces. Esta fase habilita la trazabilidad entre escenarios funcionales y elementos arquitectónicos, dando lugar a un mapa del dominio centrado en el comportamiento.

Fase 3: Delimitación y formalización de bounded contexts:

Con los elementos del dominio ya identificados, se procede a su agrupamiento en bounded contexts. Estos contextos se definen como unidades semánticamente cohesivas que agrupan responsabilidades, reglas de negocio, entidades y eventos afines, promoviendo la alta cohesión interna y el bajo acoplamiento externo. Se aplican aquí patrones de diseño propuestos por DDD, como context mapping y relaciones de dependencia explícitas entre contextos. El resultado es un mapa de contextos actualizado que refleja tanto la estructura organizacional como los flujos de comunicación entre componentes. Esta delimitación formal constituye el corazón del diseño estratégico de microservicios, ya que define fronteras claras de autonomía y evolución. La claridad obtenida en esta etapa permite construir una arquitectura sostenible en el tiempo, capaz de adaptarse sin perder coherencia.

Fase 4: Derivación de contratos e interfaces:

A partir de los contextos definidos en la etapa anterior, esta fase traduce los comandos y eventos del dominio en contratos técnicos concretos: API, endpoints o mensajes que expresan cómo cada servicio se comunica con otros o con el exterior. Se adopta un enfoque API-first, donde los contratos se diseñan

antes de la implementación, y se asegura que cada interfaz esté motivada y validada por al menos un escenario BDD. Esta conexión funcional garantiza que el diseño técnico sea verificable desde el punto de vista del negocio. Las interfaces se modelan siguiendo buenas prácticas para microservicios (Jamshidi et al., 2018; Newman, 2021; Richardson, 2018). A su vez, se mantiene la trazabilidad entre cada escenario y su operación técnica correspondiente. Esto permite que la arquitectura sea testeable de forma continua.

Fase 5: Validación continua y evolución arquitectónica:

La última fase sostiene el principio de que la arquitectura debe evolucionar al ritmo del dominio. A medida que se incorporan nuevos escenarios o cambian los existentes. Cualquier desviación detectada impulsa la revisión del diseño, que puede implicar la modificación de contextos, la actualización de interfaces o la refactorización de componentes. Cada decisión arquitectónica significativa se documenta mediante registros ADR, asegurando transparencia y auditabilidad en la evolución del sistema (Ford et al., 2017; Fowler, 2015). Además, se versionan el mapa de contextos, los contratos y los artefactos clave del marco, manteniendo una trazabilidad integral. Esta fase convierte a la arquitectura en una práctica viva, ajusta su estructura y se mantiene siempre alineada con los objetivos funcionales del sistema (Woods et al., 2021).

Ciclo de Consolidación

La evaluación de BDMA se estructuró como un experimento pedagógico en un contexto académico controlado, utilizando un diseño comparativo. Este enfoque permitió observar el impacto del uso del marco en la identificación de microservicios y contrastarlo con enfoques no guiados o implícitos, en los que los participantes debían tomar decisiones arquitectónicas sin contar con una metodología explícita.

El experimento se diseñó con el objetivo de evaluar no solo el producto final (arquitectura resultante), sino también el proceso de razonamiento seguido por los participantes, en coherencia con el carácter metodológico del artefacto propuesto.

El experimento se llevó a cabo con estudiantes de posgrado en ingeniería de software, organizados en equipos de trabajo. A todos los participantes se les proporcionó el mismo conjunto de escenarios BDD, representativos de un dominio funcional no trivial, y se les solicitó diseñar una arquitectura de microservicios a partir de dicho backlog funcional.

El diseño experimental contempló dos enfoques de trabajo:

- un enfoque no guiado, en el que los equipos abordaron la identificación de microservicios basándose en su experiencia previa y criterios personales;
- un enfoque guiado mediante BDMA, en el que los equipos aplicaron explícitamente las fases y artefactos definidos por el marco metodológico.

Esta configuración permitió analizar diferencias en términos de coherencia arquitectónica, trazabilidad funcional y calidad del razonamiento arquitectónico entre ambos enfoques.

Instrumentos y criterios de evaluación

La evaluación de los resultados se realizó utilizando una combinación de instrumentos cualitativos y cuantitativos, diseñados para capturar distintas dimensiones del proceso y del resultado. Entre los principales instrumentos se incluyen:

- una rúbrica de evaluación, orientada a analizar aspectos como la alineación con el dominio, la claridad de los límites de servicio, la trazabilidad entre escenarios y microservicios, y la consistencia de las decisiones arquitectónicas;
- un análisis comparativo de las arquitecturas resultantes, considerando la estructura de servicios propuesta, sus responsabilidades y relaciones;
- una encuesta de percepción, destinada a recoger la experiencia de los participantes respecto del uso del marco, su utilidad percibida y la claridad del proceso.

Estos instrumentos permitieron evaluar el impacto de BDMA desde múltiples perspectivas, evitando reducir la validación a métricas puramente técnicas.

Resultados y observaciones

Los resultados del experimento evidencian diferencias significativas entre los enfoques comparados. En los equipos que aplicaron BDMA se observó una mayor coherencia en la delimitación de microservicios, una mejor alineación con el dominio funcional y una trazabilidad explícita entre escenarios BDD, bounded contexts y servicios identificados. Asimismo, el uso del marco favoreció un razonamiento arquitectónico más estructurado y justificable, reduciendo la dependencia de decisiones implícitas o intuitivas. Los participantes destacaron, además, que la aplicación secuencial de las fases de BDMA facilitó la discusión colaborativa y el entendimiento compartido del dominio.

Como limitaciones, se identificó una curva de aprendizaje inicial, particularmente asociada al uso sistemático de escenarios BDD y a la necesidad de adoptar un razonamiento más explícito sobre decisiones arquitectónicas. Estas observaciones resultan coherentes con el carácter metodológico del marco y refuerzan la necesidad de entrenamiento y práctica para su adopción efectiva.

Consolidación del artefacto

En términos de Design Science Research, la evaluación realizada permite considerar a BDMA como un artefacto válido en un contexto exploratorio, aportando evidencia inicial sobre su utilidad y aplicabilidad. Si bien los resultados no pretenden ser generalizables a contextos industriales complejos, constituyen una base sólida para futuras evaluaciones en escenarios más amplios y diversos. La evidencia obtenida en esta fase contribuye a consolidar el marco metodológico propuesto y a justificar su aporte como respuesta al problema de investigación formulado, cerrando así el ciclo de diseño y evaluación de la investigación.

Conclusiones y contribuciones de la tesis

Esta tesis abordó el problema de la identificación de microservicios en contextos ágiles *greenfield*, un desafío recurrente tanto en la práctica industrial como en la literatura académica. A partir de un enfoque de Design Science Research, el problema fue caracterizado mediante el ciclo de relevancia —integrando un mapeo sistemático y un experimento exploratorio— y validado críticamente a través

del ciclo de rigor, permitiendo identificar un vacío metodológico en la articulación entre descubrimiento funcional, modelado del dominio y decisiones arquitectónicas.

Como respuesta a este problema, la investigación propuso Behavior-Driven Microservice Architecture (BDMA), un marco metodológico que integra de manera explícita prácticas de Behavior-Driven Development y Domain-Driven Design dentro de un proceso iterativo y trazable para la identificación y evolución de microservicios. El aporte central de BDMA no reside en la introducción de nuevas técnicas aisladas, sino en la orquestación sistemática de prácticas existentes, proporcionando un proceso reproducible que permite derivar decisiones arquitectónicas a partir del comportamiento funcional del sistema. La evaluación exploratoria del marco, realizada mediante un experimento pedagógico controlado, aportó evidencia inicial sobre la utilidad de BDMA para mejorar la coherencia de las arquitecturas resultantes, fortalecer la alineación con el dominio y favorecer un razonamiento arquitectónico más explícito y justificable. Asimismo, los resultados permitieron identificar limitaciones y desafíos asociados a la adopción del enfoque, tales como la curva de aprendizaje inicial, que constituyen insumos relevantes para su evolución futura.

Contribuciones y líneas futuras de investigación

Desde una perspectiva académica, la tesis contribuye al cuerpo de conocimiento en ingeniería de software al proponer un marco metodológico integrador para el diseño de microservicios en contextos ágiles *greenfield*, abordando una brecha identificada en la literatura. Desde una perspectiva práctica, BDMA ofrece a los equipos una guía concreta para reducir la dependencia de decisiones intuitivas, mejorar la trazabilidad entre requisitos y arquitectura, y sostener la evolución arquitectónica frente a cambios de requisitos. Finalmente, esta investigación sienta las bases para futuras líneas de trabajo orientadas a la validación del marco en contextos industriales, la incorporación de herramientas de soporte y la exploración de técnicas automáticas que asistan determinadas fases del proceso, manteniendo el enfoque metodológico propuesto.

Bibliografía

Bajaj, D., Goel, A., & Gupta, S. C. (2022). GreenMicro: identifying microservices from use cases in greenfield development. *IEEE Access*, *10*, 67008–67018.

Battaglia, N. (2025). Identificación Ágil de Microservicios Utilizando DDD y BDD. *XXVIII Ibero-American Conference on Software Engineering (CIbSE 2025 Doctoral Symposium)*.

Battaglia, N., Garcia, A. N., & Congiusti, A. (2024). Descubrimiento de Microservicios en Metodologías Ágiles: un mapeo sistemático de la literatura. *XXX Congreso Argentino de Ciencias de La Computación*, 1506–1510.

Battaglia, N., Rossi, G., & Fernández., A. (2025). Behavior-Driven Microservice Architecture: Un Marco Metodológico para la Identificación Iterativa de Microservicios en Proyectos Ágiles Greenfield. *Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería Del Conocimiento JIISIC 2025*.

Booch, G. (2006). The accidental architecture. *IEEE Software*, *23*(3), 9–11.

Brandolini, A. (2018). *Introducing EventStorming*. Leanpub.

Erder, M., & Pureur, P. (2015). *Continuous architecture: sustainable architecture in an agile and cloud-centric world*. Morgan Kaufmann.

Ford, N., Parsons, R., & Kua, P. (2017). *Building Evolutionary Architectures*. O'Reilly Media.

Fowler, M. (2015). *Microservice Trade-Offs*. https://martinfowler.com/articles/microservice-trade-offs.html

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

Jamshidi, P., Pahl, C., Mendonça Nane Kratzke, & others. (2018). Microservices: The Journey So Far and Challenges Ahead. *IEEE Software*, *35*(3), 24–35.

Kapferer, S., & Zimmermann, O. (2020). Domain-specific Language and Tools for Strategic Domain-driven Design, Context Mapping and Bounded Context Modeling. *MODELSWARD*, 299–306.

Newman, S. (2021). *Building Microservices* (2nd ed.). O'Reilly Media.

Oumoussa, I., & Saidi, R. (2024). Evolution of microservices identification in monolith decomposition: A systematic review. *IEEE Access*, *12*, 23389–23405.

Peffers, K., Rothenberger, M., Kuechler, B., & Tuunanen, T. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77.

Richardson, C. (2018). *Microservices Patterns: With examples in Java*. Manning Publications.

Su, R., Li, X., & Taibi, D. (2024). From Microservice to Monolith: A Multivocal Literature Review. *Electronics*, *13*(8), 1452.

Taibi, D., & Lenarduzzi, V. (2018). On the definition of microservice bad smells. *IEEE Software*, *35*(3), 56–62.

Vural, H., & Koyuncu, M. (2021). Does domain-driven design lead to finding the optimal modularity of a microservice? *IEEE Access*, *9*, 32721–32733.

Woods, E., Erder, M., & Pureur, P. (2021). *Continuous architecture in practice: Software architecture in the age of agility and devops*. Addison-Wesley Professional.

Wynne, M. (2015). Introducing Example Mapping. *Cucumber Blog*. https://cucumber.io/blog/bdd/example-mapping-introduction/

Zhong, C., Li, S., Huang, H., Liu, X., Chen, Z., Zhang, Y., & Zhang, H. (2024). Domain-driven design for microservices: An evidence-based investigation. *IEEE Transactions on Software Engineering*, *50*(6), 1425–1449.

Zimmermann, O. (2017). Microservices tenets: Agile approach to service development and deployment. *Computer Science-Research and Development*, *32*, 301–310.

DEFINICIÓN DE UN MODELO DE CLASIFICACIÓN DE TRÁFICO EN REDES DEFINIDAS POR SOFTWARE

Doctorado en Ciencias Informáticas
Universidad Nacional de la Plata, Facultad de Informática.
Autor: Alberto Daniel Núñez Agurto
Fecha de exposición: 12 de junio de 2025
Directores:
Ing. Luis Marrone
Ing. Walter Fuertes, PhD

RESUMEN

El crecimiento exponencial de aplicaciones y servicios en Internet ha incrementado la demanda de ancho de banda y calidad de servicio, complejizando la gestión de redes. En este contexto, la clasificación eficiente del tráfico es fundamental para optimizar recursos y garantizar desempeño y seguridad. No obstante, los métodos tradicionales basados en inspección profunda de paquetes presentan limitaciones de escalabilidad, privacidad y eficiencia frente a entornos dinámicos y cifrados. Las redes definidas por software (SDN) emergen como una alternativa al separar los planos de control y de datos, proporcionando una visión global del tráfico y capacidades de programabilidad directa. La integración de técnicas de aprendizaje automático en este entorno permite incorporar inteligencia a la red, mejorando la detección y clasificación del tráfico. En este contexto, la presente investigación propone un modelo de clasificación basado en aprendizaje profundo que aborda tanto aplicaciones como ataques de red. Mediante la metodología SEMMA, se entrenaron cuatro modelos: LSTM, BiLSTM, GRU y BiGRU, utilizando características extraídas de dos conjuntos de datos públicos. Los resultados experimentales, validados en un entorno de simulación, evidencian el desempeño superior del modelo GRU, confirmando la viabilidad y efectividad de la propuesta y su contribución a la mejora de la calidad de servicio y la seguridad en entornos SDN.
Palabras clave: *Redes definidas por software; clasificación de tráfico; clasificación basada en aplicaciones; aprendizaje automático.*

INTRODUCCIÓN

El crecimiento exponencial de aplicaciones, dispositivos inteligentes y servicios digitales ha incrementado significativamente el volumen y la diversidad del tráfico en Internet [1]. Esta evolución ha complejizado la gestión de las infraestructuras de red, las cuales deben garantizar calidad de servicio, eficiencia en el uso de recursos y seguridad frente a amenazas cada vez más sofisticadas [2]. En este contexto, la clasificación de tráfico se convierte en un mecanismo esencial para la optimización y administración efectiva de redes modernas.

Las redes definidas por software (SDN, por sus siglas en inglés *Software-Defined Networking*) emergen como una arquitectura que facilita esta gestión al separar el plano de control del plano de datos, proporcionando una visión global del estado de la red y habilitando la programabilidad centralizada [3]. Esta característica permite implementar políticas dinámicas de calidad de servicio,

seguridad y asignación de recursos [4]. Sin embargo, la creciente complejidad del tráfico, sumada al uso generalizado de cifrado y a la aparición constante de nuevas aplicaciones y ataques, limita la efectividad de los métodos tradicionales de clasificación basados en puertos o inspección profunda de paquetes [5], [6].

En este escenario, la incorporación de técnicas de aprendizaje automático (ML, por sus siglas en inglés *Machine Learning*) en entornos SDN ha despertado un gran interés, ya que permite identificar patrones complejos en los flujos de red y automatizar tareas de análisis, clasificación y detección de anomalías [7], [8]. No obstante, persisten desafíos relacionados con la selección de características, la generalización de los modelos y su integración eficiente en los controladores SDN.

Estas limitaciones evidencian la necesidad de desarrollar modelos de clasificación más robustos, escalables y adaptables, capaces de identificar simultáneamente tipos de aplicaciones y ataques, contribuyendo así a mejorar el rendimiento y la seguridad en redes definidas por software.

Este trabajo explora la integración de técnicas de aprendizaje automático en las redes definidas por software, para mejorar la exactitud y eficiencia de la clasificación de tráfico, abordando desafíos actuales y proponiendo un enfoque innovador para clasificar de manera efectiva los flujos de tráfico dentro de entornos de red dinámicos.

ALCANCE Y CONTRIBUCIONES DE LA TESIS

El objetivo de esta investigación es proponer y validar un modelo de clasificación de tráfico en redes definidas por software basado en técnicas de aprendizaje profundo, orientado a mejorar simultáneamente la calidad de servicio y la seguridad mediante la identificación de aplicaciones y ataques a nivel de flujo.

El alcance del trabajo comprende el análisis del estado del arte, el diseño del modelo de clasificación, su evaluación experimental con conjuntos de datos públicos y su validación en un entorno SDN emulado. A diferencia de propuestas previas, el modelo integra la clasificación de tipos de aplicación y detección de ataques en un esquema multiclase unificado.

Las principales contribuciones pueden resumirse en:

- Identificación de limitaciones actuales en la clasificación de tráfico en entornos SDN, particularmente en escalabilidad, desequilibrio de datos y generalización de modelos.
- Diseño y evaluación comparativa de cuatro arquitecturas recurrentes (GRU, BiGRU, LSTM y BiLSTM) aplicadas a clasificación multiclase basada en flujos.
- Propuesta de un esquema de selección de características obtenibles desde la interfaz norte del controlador SDN, reduciendo dimensionalidad y costo computacional sin degradar la exactitud.
- Integración y validación del modelo en un entorno SDN emulado, demostrando su viabilidad técnica para despliegue en redes reales.

FUNDAMENTOS DE SDN

Las redes definidas por software (SDN) constituyen un paradigma arquitectónico que desacopla el plano de control del plano de datos, permitiendo una gestión centralizada y programable de la red [9]. Este enfoque proporciona una visión global del estado de la infraestructura, facilitando la implementación dinámica de políticas de calidad de servicio, seguridad y optimización de recursos [3]. La arquitectura general de este paradigma se muestra en la Figura 1.

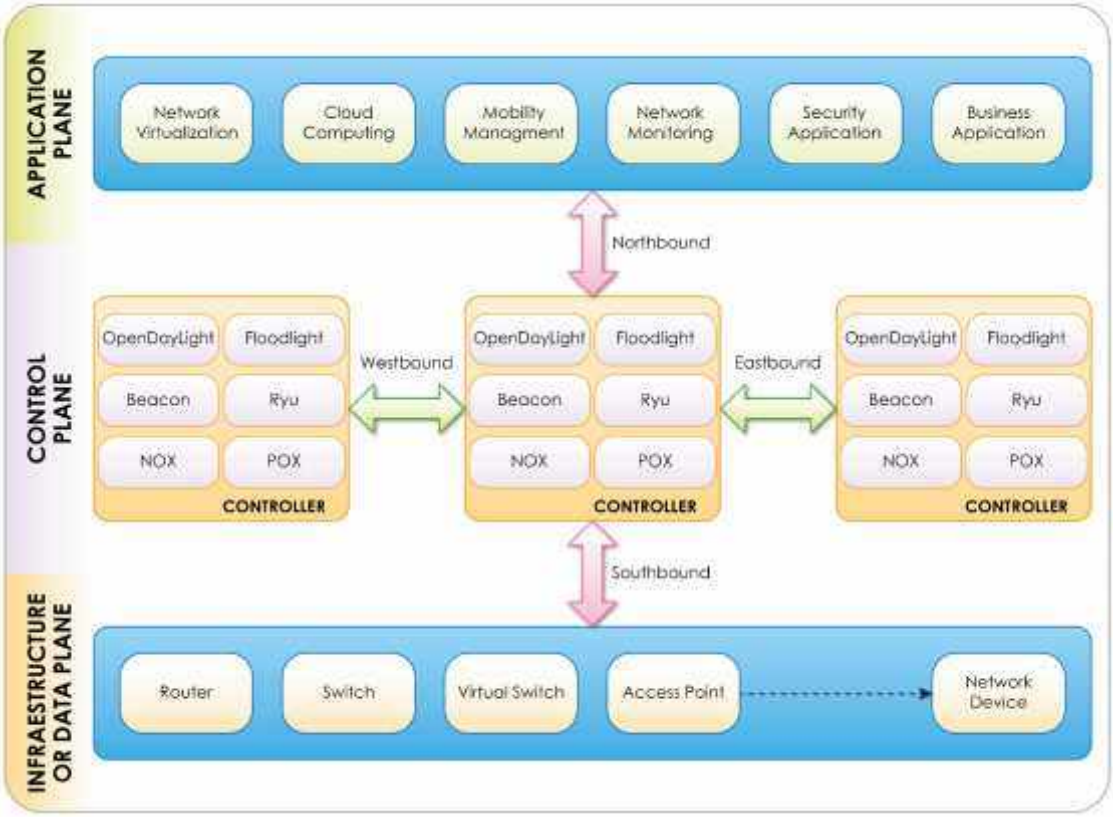


Figura 1. **Arquitectura de redes definidas por software.**

En la arquitectura SDN, el plano de datos está compuesto por dispositivos de reenvío que ejecutan reglas definidas por el controlador mediante protocolos como OpenFlow [10]. El plano de control, lógicamente centralizado, administra la topología, las reglas de flujo y la recopilación de estadísticas. A través de las interfaces hacia el norte, el controlador expone información del estado de la red a aplicaciones externas, lo que habilita la integración de técnicas de análisis y aprendizaje automático [4].

Esta capacidad de obtener métricas globales de tráfico y programar dinámicamente los flujos convierte a las SDN en un entorno idóneo para implementar modelos inteligentes de clasificación de tráfico. En particular, la disponibilidad de estadísticas de flujo desde el controlador permite construir clasificadores basados en características agregadas, evitando técnicas invasivas como la inspección profunda de paquetes.

Controlador SDN

El controlador SDN constituye el elemento central de la arquitectura, al concentrar la lógica de control y mantener una visión global del estado de la red [11]. A través de las interfaces hacia el norte (NBI),

expone estadísticas de flujo y métricas operativas que pueden ser utilizadas por aplicaciones externas para análisis avanzado y toma de decisiones.

Esta capacidad de recopilación y exportación de información resulta fundamental para la implementación de modelos de aprendizaje automático, ya que permite extraer características del tráfico sin recurrir a técnicas invasivas como la inspección profunda de paquetes. En este trabajo se aprovechó esta funcionalidad para obtener atributos de flujo directamente desde el controlador, facilitando la integración del modelo de clasificación en el entorno SDN.

APRENDIZAJE AUTOMÁTICO EN SDN

El aprendizaje automático es una aplicación de la inteligencia artificial que permite desarrollar sistemas que aprenden de forma autónoma identificando patrones complejos a partir de grandes conjuntos de datos. Desde un punto de vista operativo, el ML consta de dos fases: 1) la primera fase se refiere al *entrenamiento* y consiste en proporcionar a los algoritmos de ML un subconjunto del conjunto de datos empleado, denominado conjunto de entrenamiento, a partir del cual el modelo del sistema puede aprender, y 2) la segunda fase consiste en el *toma de decisiones*, donde el sistema puede estimar el resultado de una nueva entrada, basándose en el modelo entrenado. Los algoritmos de ML se clasifican a grandes rasgos en aprendizaje supervisado, no supervisado, semi-supervisado y de refuerzo [12], [13], [14], [15]. Los esfuerzos de investigación realizados durante siete décadas han dado lugar a numerosas técnicas de ML. En la Figura 2 se destacan los hitos más cruciales en la evolución del ML.

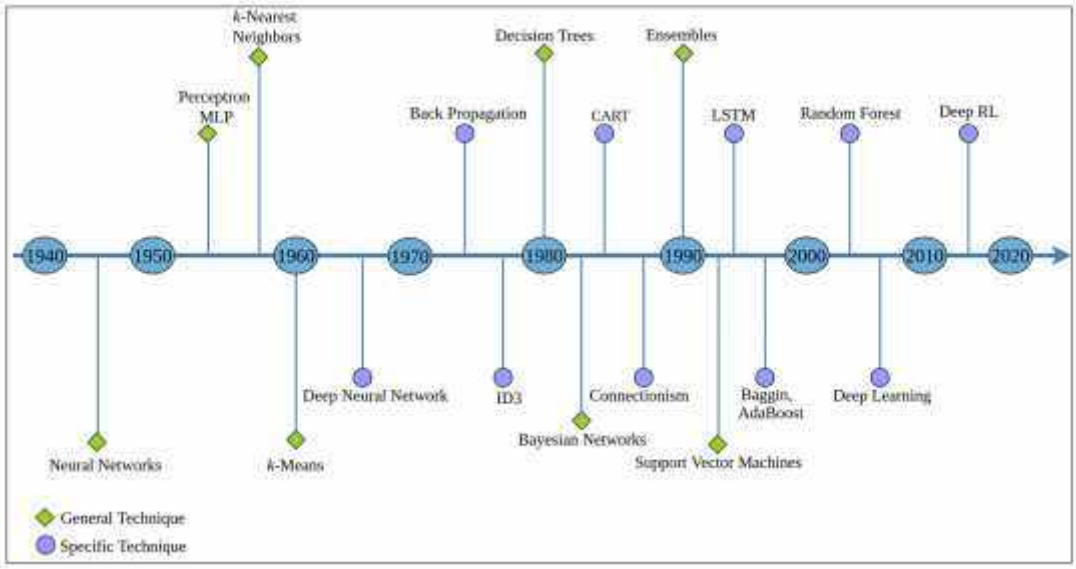


Figura 2. **Evolución de las técnicas de aprendizaje automático con hitos claves.**

En general, el ML es ideal para inferir soluciones a problemas con un gran conjunto de datos representativos. Las técnicas de ML están diseñadas para identificar datos ocultos y patrones en los datos. Por lo tanto, son muy adecuadas para resolver problemas en SDN. Por ejemplo, un problema de clasificación en SDN puede formularse para identificar tráfico anómalo.

MODELO DE CLASIFICACIÓN DE TRÁFICO

La investigación adoptó la metodología SEMMA (Sample, Explore, Modify, Model, Assess) para el desarrollo del modelo de clasificación de tráfico en entornos SDN [16]. Se emplearon dos conjuntos de datos públicos: InSDN e ISCX-VPN-nonVPN, integrados y reorganizados en cinco clases de tráfico: Multimedia, VoIP, Mensajería instantánea, Transferencia de archivos y Ataques. Con el fin de garantizar equilibrio entre clases, se realizó un proceso de muestreo uniforme, obteniendo 10 000 instancias por categoría.

En la fase de modificación se efectuaron procesos de limpieza de datos, codificación de etiquetas, normalización mediante estandarización Z-score y selección de características. Inicialmente se eliminaron atributos constantes y registros anómalos. Posteriormente, mediante análisis de información mutua y evaluación de viabilidad de extracción desde el controlador SDN (API norte del controlador Ryu), se seleccionaron cinco características relevantes, que se detallan en la Tabla 1. Esta reducción permite disminuir la dimensionalidad sin afectar el desempeño predictivo.

Tabla 1. **Descripción de las características.**

No.	Nombre característica	Descripción
1	Flow Duration	Duración del flujo en microsegundos
2	Flow Pkts/s	Número de paquetes de flujo por segundo
3	Flow Byts/s	Número de bytes de flujo por segundo
4	Fwd Pkts/s	Número de paquetes reenviados por segundo
5	Bwd Pkts/s	Número de paquetes de retorno por segundo

En la fase de modelización se implementaron cuatro arquitecturas de aprendizaje profundo basadas en redes neuronales recurrentes: LSTM, BiLSTM, GRU y BiGRU. El conjunto de datos se dividió en 70% para entrenamiento, 15% para validación y 15% para pruebas. Se utilizó la función de pérdida *categorical crossentropy*, optimizador Adam y 20 épocas de entrenamiento, tal como se detalla en la Tabla 2. Además, se evaluaron distintas longitudes de secuencia, determinándose como óptima una longitud de 115.

Tabla 2. **Configuración de hiperparámetros.**

Función de activación	Función de pérdida	Tamaño del lote	Algoritmo de optimización	Capa de entrada	No. de épocas
Sigmoid	Categorical crossentropy	16	Adam	64	20

Los resultados experimentales muestran que los modelos GRU y LSTM alcanzaron el mejor desempeño en la clasificación multiclase. Con cinco características seleccionadas, el modelo GRU obtuvo una exactitud de 99,65%, con valores elevados de precisión, recuperación y F1-score en todas las clases. Asimismo, presentó menor tiempo de entrenamiento y menor utilización de GPU en

comparación con los modelos bidireccionales. Cuando se utilizaron todas las características del conjunto de datos, la mejora en exactitud fue marginal, evidenciando la efectividad del proceso de selección de atributos, como se observa en la Figura 3.

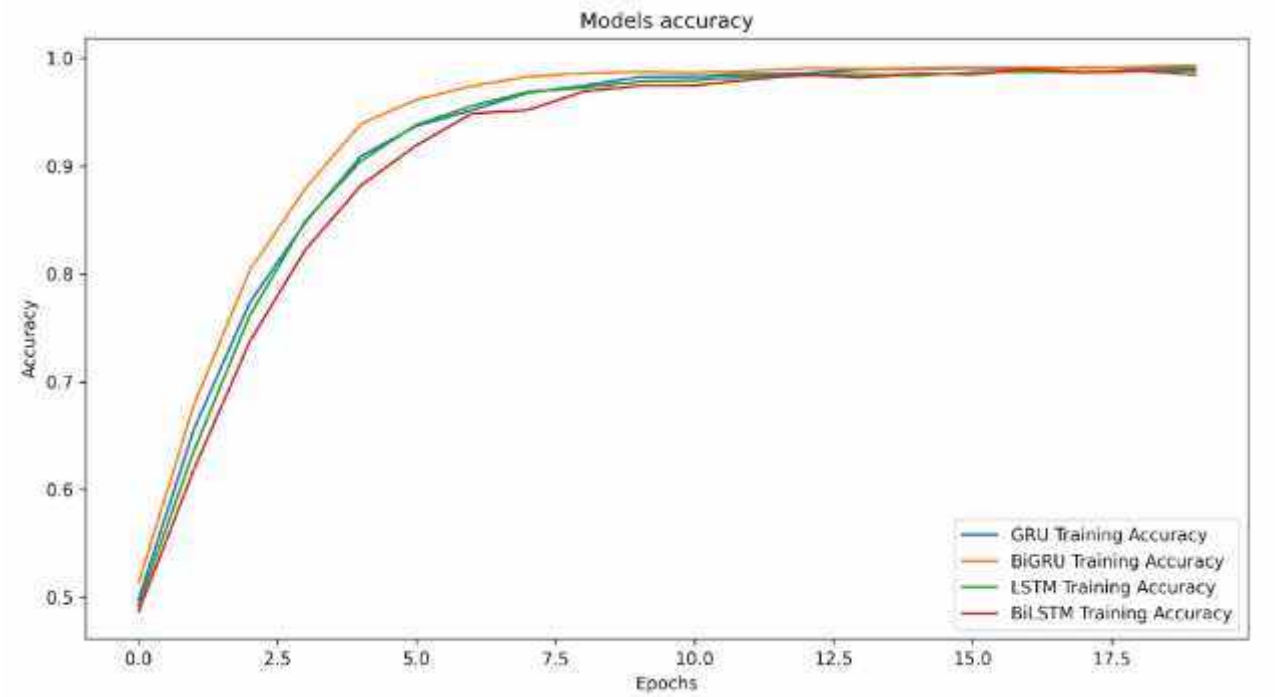


Figura 3. **Comparación de la exactitud de los algoritmos.**

IMPLEMENTACIÓN DEL ENTORNO DE PRUEBAS

El escenario de pruebas fue desarrollado en GNS3 [17], compuesto por un controlador Ryu, un Open vSwitch y un servidor de aplicaciones encargado de ejecutar los modelos entrenados, como se muestra en la Figura 4. La comunicación entre el controlador y el servidor se realizó mediante interfaces hacia el norte y mensajes JSON. Las pruebas de implementación demostraron la viabilidad técnica de integrar los modelos GRU y LSTM en un entorno SDN realista, validando su capacidad para clasificar tráfico en tiempo casi real y confirmando el potencial de despliegue en redes operativas.

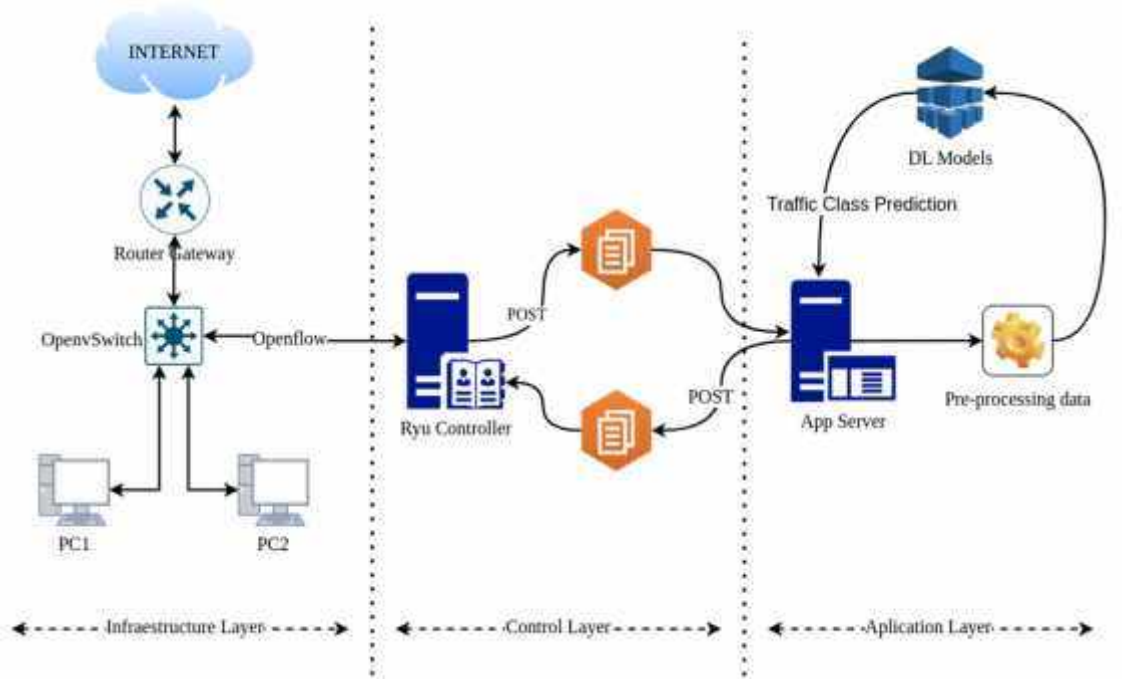


Figura 4. **Entorno para la implementación de los modelos de clasificación.**

CONCLUSIONES

La investigación realizada demuestra la viabilidad y efectividad del uso de técnicas de aprendizaje profundo para la clasificación de tráfico en redes definidas por software. Se propuso y validó un modelo de clasificación multiclase capaz de identificar tanto tipos de aplicaciones como ataques de red, contribuyendo simultáneamente a la mejora de la calidad de servicio y la seguridad en entornos SDN.

A partir del análisis del estado del arte se identificaron limitaciones asociadas a conjuntos de datos desequilibrados, problemas de escalabilidad y desafíos en la generalización de los modelos. En respuesta a ello, se desarrollaron cuatro arquitecturas basadas en redes neuronales recurrentes (GRU, BiGRU, LSTM y BiLSTM), entrenadas sobre conjuntos de datos públicos integrados y balanceados. La selección de cinco características relevantes, obtenibles desde la interfaz norte del controlador SDN, permitió reducir la complejidad computacional sin afectar significativamente la exactitud del modelo.

Los resultados experimentales evidencian que el modelo GRU alcanzó el mejor desempeño, con una exactitud del 99,65%, manteniendo altos valores de precisión y recuperación en todas las clases. Asimismo, se validó su integración en un entorno SDN emulado, confirmando su viabilidad técnica para despliegues en redes reales y su capacidad de clasificación en tiempo casi real. En conjunto, estos hallazgos demuestran que la integración de aprendizaje profundo en arquitecturas SDN constituye una estrategia eficaz para la gestión inteligente del tráfico.

LÍNEAS DE I+D FUTURAS

Como líneas futuras de investigación se plantea, en primer lugar, la validación de los clasificadores en entornos SDN de producción en tiempo real, permitiendo la clasificación instantánea de flujos

entrantes a partir de estadísticas de flujo y la aplicación dinámica de políticas de calidad de servicio y seguridad.

Asimismo, se propone la construcción de un conjunto de datos actualizado a partir de tráfico real en redes definidas por software, con el fin de evaluar la robustez y capacidad de generalización del modelo en escenarios más representativos y dinámicos.

Finalmente, se prevé explorar enfoques basados en aprendizaje por refuerzo que permitan desarrollar mecanismos adaptativos de clasificación y gestión de tráfico, capaces de responder de manera autónoma a la aparición de nuevas aplicaciones, patrones de uso y tipos de ataques.

PUBLICACIONES DERIVADAS DE LA TESIS

Los resultados de esta tesis doctoral han sido publicados en los siguientes trabajos:

- Nuñez-Agurto, D., Fuertes, W., Marrone, L., Benavides-Astudillo, E., Coronel-Guerrero, C., & Perez, F. (2024). A Novel Traffic Classification Approach by Employing Deep Learning on Software-Defined Networking. Future Internet, 16(5), 153.
- Nunez-Agurto, D., Fuertes, W., Marrone, L., & Macas, M. (2022). Machine Learning-Based Traffic Classification in Software-Defined Networking: A Systematic Literature Review, Challenges, and Future Research Directions. IAENG International Journal of Computer Science, 49(4).
- Nuñez-Agurto, D., Fuertes, W., Marrone, L., Benavides-Astudillo, E., & Vásquez-Bermúdez, M. (2023). Traffic Classification in Software-Defined Networking by Employing Deep Learning Techniques: A Systematic Literature Review. In International Conference on Technologies and Innovation (pp. 67-80). Springer, Cham.
- Nuñez-Agurto, D., Fuertes, W., Marrone, L., Castillo-Camacho, M., Benavides-Astudillo, E., & Perez, F. (2023, November). Attack Classification Using Machine Learning Techniques in Software-Defined Networking. In International Conference on Applied Technologies (pp. 241-256). Cham: Springer Nature Switzerland.
- Nuñez-Agurto, D., Fuertes, W., Marrone, L., Pisco, L., Ordoñez, X.,& Rodriguez, G. (2024, November). Application Based Traffic Classification by Employing Machine Learning Models on Software-Defined Networking. International Conference on Computer Science, Electronics and Industrial Engineering (pp. 355-370). Cham: Springer Nature Switzerland.

REFERENCIAS

[1] Cisco, "Cisco annual internet report (2018–2023)," Cisco, 2018--2023. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf

[2] O. Belkadi, A. Vulpe, Y. Laaziz, and S. Halunga, "ML-based traffic classification in an SDN-enabled cloud environment," *Electronics*, vol. 12, no. 2, 2023, doi: 10.3390/electronics12020269.

[3] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015, doi: 10.1109/JPROC.2014.2371999.

[4] J. Xie *et al.*, "A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393–430, 2019, doi: 10.1109/COMST.2018.2866942.

[5] A. W. Moore and K. Papagiannaki, "Toward the accurate identification of network applications," in *Passive and active network measurement*, C. Dovrolis, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 41–54.

[6] M. Finsterbusch, C. Richter, E. Rocha, J.-A. Muller, and K. Hanssgen, "A survey of payload-based traffic classification approaches," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 1135–1156, 2014, doi: 10.1109/SURV.2013.100613.00161.

[7] S. Ayoubi *et al.*, "Machine learning for cognitive network management," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 158–165, 2018, doi: 10.1109/MCOM.2018.1700560.

[8] M. Hayes, B. Ng, A. Pekar, and W. K. G. Seah, "Scalable architecture for SDN traffic classification," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3203–3214, 2018, doi: 10.1109/JSYST.2017.2690259.

[9] O. N. Foundation, 2014.

[10] C. Trois, M. D. Del Fabro, L. C. E. de Bona, and M. Martinello, "A survey on SDN programming languages: Toward a taxonomy," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2687–2712, 2016, doi: 10.1109/COMST.2016.2553778.

[11] H. Sufiev, Y. Haddad, L. Barenboim, and J. Soler, "Dynamic SDN controller load balancing," *FUTURE INTERNET*, vol. 11, no. 3, 2019, doi: 10.3390/fi11030075.

[12] T. T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 4, pp. 56–76, 2008, doi: 10.1109/SURV.2008.080406.

[13] M. Mohammed, M. B. Khan, and E. B. M. Bashier, *Machine learning: Algorithms and applications*. Crc Press, 2016.

[14] X. J. Zhu, *Semi-supervised learning literature survey*. University of Wisconsin-Madison Department of Computer Sciences, 2005.

[15] H. Wu and S. Prasad, "Semi-supervised deep learning using pseudo labels for hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1259–1270, 2018, doi: 10.1109/TIP.2017.2772836.

[16] A. Azevedo and M. Santos, "KDD, semma and CRISP-DM: A parallel overview," in *IADIS european conf. Data mining*, Jan. 2008, pp. 182–185.

[17] R. Mohtasin, P. W. C. Prasad, A. Alsadoon, G. Zajko, A. Elchouemi, and A. K. Singh, "Development of a virtualized networking lab using GNS3 and VMware workstation," in *2016 international conference on wireless communications, signal processing and networking (WiSPNET)*, 2016, pp. 603–609. doi: 10.1109/WiSPNET.2016.7566205.

UNIVERSIDAD NACIONAL DE LA PLATA
Facultad de Informática

Título de la tesis: Desarrollo de un modelo de elicitación de emociones a partir de las características de la música. Generación de un sistema recomendador.
Autor: Yesid Ospitia Medina.
Directores: Sandra Baldassarri (Directora), José Ramón Beltrán (Co-Director).
Asesora: Cecilia Sanz.
Fecha de defensa: Marzo de 2024.

Síntesis del Contenido (Abstract)

La presente investigación aborda una problemática crítica en la industria musical contemporánea: la invisibilización de talentos emergentes provocada por la hegemonía de los sistemas de recomendación basados en la popularidad. En un mercado saturado donde el volumen de lanzamientos diarios imposibilita el descubrimiento orgánico, los algoritmos actuales suelen reforzar sesgos que favorecen a los "artistas más famosos", dificultando la entrada de nuevas propuestas. El aporte sustancial de esta tesis radica en la transición de un modelo de recomendación basado en metadatos y popularidad hacia un paradigma híbrido fundamentado en el reconocimiento dinámico de emociones. Para ello, se presenta la creación del **Emotional Non-Superstar Artist-Dataset (ENSA) y el diseño de un sistema recomendador**. Los resultados demuestran la viabilidad de utilizar modelos de *machine learning* para la identificación de valencia y excitación (*arousal*), permitiendo una vinculación técnica entre la gramática musical y la respuesta afectiva. Las líneas de investigación futuras se proyectan en tres ejes estratégicos: la expansión de ENSA, la extensión funcional del recomendador a micro-estructuras musicales (puentes, solos) y la mitigación proactiva de sesgos emergentes.

1. CAPÍTULO 1: INTRODUCCIÓN

La industria musical global atraviesa una transformación estructural sin precedentes. La migración masiva de soportes físicos a servicios de *streaming* en la nube ha consolidado un ecosistema donde el éxito comercial depende de la visibilidad algorítmica. Según datos de la IFPI, los ingresos por *streaming* han crecido un 67% recientemente, sin embargo, esta expansión económica no ha garantizado la equidad competitiva. La importancia estratégica de este estudio reside en superar la dependencia de recomendaciones estadísticas de popularidad para centrarse en la percepción emocional, permitiendo que la música cumpla su función primordial de gestión del bienestar psicológico del oyente.

1.1. Resumen General

La investigación se estructura a partir de cinco preguntas orientadoras (P1-P5) diseñadas para desentrañar la representación computacional de la relación música-emoción. Mediante una metodología multidisciplinar que integra la Computación Afectiva, la Recuperación de Información Musical (MIR) y el aprendizaje supervisado, se busca determinar si es posible diseñar sistemas de recomendación que operen bajo la subjetividad de la percepción humana en lugar de métricas de mercado.

1.2. Conclusiones del Capítulo

La evidencia metodológica sugiere que la resolución de este problema requiere un enfoque de desarrollo ágil basado en prototipos evolutivos. Este marco permite que el sistema de recomendación madure a través de ciclos iterativos de experimentación, donde la medición de tasas de acierto y error garantiza un producto final sensible a las intenciones del compositor y las necesidades del oyente. *La efectividad de estas soluciones tecnológicas depende necesariamente de la construcción de un marco conceptual sólido que defina la naturaleza de la respuesta emocional humana frente al estímulo sonoro.*

2. CAPÍTULO 2: FUNDAMENTACIÓN TEÓRICA

La Computación Afectiva constituye el pilar tecnológico indispensable para decodificar la relación entre la gramática musical y la respuesta psicológica del individuo. A diferencia de los enfoques puramente estadísticos, este marco permite tratar la música como un lenguaje con intención expresiva, donde el procesamiento de señales digitales se traduce en estados de ánimo medibles. El rigor científico de esta base teórica diferencia este estudio de los análisis triviales, al fundamentar el reconocimiento emocional en principios físicos y psicológicos.

2.1. Resumen General

Se analizan las emociones mediante modelos categóricos (Ekman) y, de forma preferente por su valor estratégico, modelos dimensionales de Russell (Valencia/Excitación). Para el Reconocimiento de Emociones en la Música (MER), se evalúan librerías como Marsyas, MIR toolbox y PsySound 3, que permiten la extracción de características de bajo nivel. Asimismo, se fundamenta el uso de la lógica difusa para capturar la imprecisión del lenguaje humano y del aprendizaje de máquina para la clasificación determinística y predicción de valores.

2.2. Conclusiones del Capítulo

El análisis teórico ratifica la superioridad de los modelos dimensionales para las estrategias de recomendación, dada su capacidad para capturar la variabilidad emocional de una pieza a lo largo del tiempo. Se concluye que la integración de inteligencia artificial es vital para procesar la complejidad de descriptores acústicos y convertirlos en categorías afectivas de alto nivel. *Establecida la base teórica, resulta imperativo revisar el estado actual de la tecnología y las herramientas disponibles para implementar estos modelos.*

3. CAPÍTULO 3: ESTADO DEL ARTE

La madurez alcanzada por las librerías de extracción de características musicales es notable; no obstante, persiste una brecha crítica entre la capacidad técnica y la disponibilidad de datos etiquetados sin sesgos comerciales.

3.1. Resumen General

La literatura revela que los sistemas de predicción han evolucionado, pero el tratamiento de sesgos en los sistemas recomendadores musicales (MRS) es insuficiente. Los sistemas actuales favorecen a los "artistas famosos", relegando a los talentos emergentes a una posición con probabilidades bajas de recomendación.

3.2. Conclusiones del Capítulo

Se identifica una escasez de *datasets* con canciones originales completas y etiquetado emocional dinámico. Esta falta de recursos que integren la estructura musical justifica la implementación de los prototipos presentados en el Capítulo 4. *Los vacíos detectados en el estado del arte fundamentan la necesidad de implementar sistemas de reconocimiento emocional más precisos.*

4. CAPÍTULO 4: IMPLEMENTACIÓN DE SISTEMAS PARA EL RECONOCIMIENTO DE EMOCIONES EN LA MÚSICA

El diseño de los tres prototipos se concibió como un proceso iterativo de resolución de problemas técnicos. La investigación evalúa la importancia de transitar desde la predicción de valores numéricos hacia la clasificación categórica robusta para mejorar la aplicabilidad en entornos reales de recomendación.

4.1. Resumen de Prototipos

- **Sistema de Predicción (MLP):** Utilizó redes neuronales Perceptrón Multicapa. Logró un RMSE de 0.11, sin embargo, la extensión a clasificación por cuadrantes fracasó técnicamente: aunque alcanzó un 75% de precisión en el Q1, obtuvo un marginal 5% en el Q2 y un 21% en el Q4, invalidando su uso para sistemas multiclase precisos.
- **Sistema de Clasificación No Determinística (Fuzzy):** Implementó lógica difusa para manejar la ambigüedad del tempo y la excitación. Útil para la subjetividad, pero con una complejidad operativa que impide su escalabilidad en sistemas globales.
- **Sistema de Clasificación Determinística:** Evaluó SVM, Random Forest y MLP. El uso de técnicas de balanceo **SMOTETomek** permitió que el MLP superara las limitaciones de los datos desbalanceados, logrando el mejor rendimiento multiclase (AVG F: 0.50).

4.2. Conclusiones del Capítulo

La superioridad del enfoque determinístico, especialmente al emplear redes neuronales con pre-procesamiento estratificado, subraya las limitaciones de la lógica difusa para escalar motores de recomendación. El rigor metodológico en la fase de pre-procesamiento demostró ser más determinante para el éxito que la complejidad del algoritmo *per se* . *La eficacia de estos sistemas de reconocimiento depende intrínsecamente de la calidad de los datos de entrada, lo que condujo a la creación del dataset ENSA.*

5. CAPÍTULO 5: DISEÑO DEL DATASET MUSICAL: EMOTIONAL NON-SUPERSTAR ARTIST-DATASET (ENSA)

La creación de ENSA constituye una decisión estratégica para democratizar la industria musical. Al enfocarse en artistas no comerciales bajo licencias **Creative Commons** , este recurso actúa como una plataforma de visibilización. El etiquetado realizado por el propio artista añade una capa que captura la intención emocional original, un atributo inexistente en los *datasets* comerciales tradicionales.

5.1. Especificaciones y Análisis

ENSA integra 60 canciones originales completas con atributos clave:

- Inclusión de artistas noveles y obras completas (no fragmentos).
- Etiquetado de estructura musical (verso, coro, solos).
- Anotaciones dinámicas de valencia y excitación generadas por un modelo previamente entrenado.
- Metadata enriquecida con género y percepción de preferencia (*like/dislike*).

5.2. Conclusiones del Capítulo

La evidencia sugiere que ENSA se consolida como un recurso clave para mitigar sesgos de popularidad. El alto grado de coincidencia entre la etiqueta del artista y la percepción del oyente valida su utilidad para entrenar sistemas de recomendación equitativos. *Este recurso constituye el motor de entrenamiento para el sistema recomendador híbrido desarrollado en la etapa final de la tesis.*

6. CAPÍTULO 6: DESARROLLO DE UN SISTEMA DE RECOMENDACIÓN MUSICAL

El diseño del recomendador híbrido busca optimizar el emparejamiento entre obra y oyente mediante filtrado emocional y de contenido. La innovación reside en tratar la música como una **serie temporal** , superando la limitación de las "etiquetas estáticas" presentes en plataformas como Spotify.

6.1. Resumen General

El sistema emplea **Dynamic Time Warping (DTW)** para medir la similitud entre las trayectorias emocionales de las canciones. A diferencia de un promedio estático, el DTW permite comparar cómo evoluciona la excitación o la valencia. Experimentos de agrupamiento con **HDBSCAN y UMAP** permitieron identificar piezas de artistas noveles con perfiles emocionales análogos a canciones comerciales, basándose en la trayectoria afectiva real.

6.2. Conclusiones del Capítulo

Los resultados confirman que el agrupamiento por valencia y excitación a lo largo de la estructura de la canción es altamente efectivo. Tratar la música como una serie temporal permite recomendaciones más profundas, mitigando el sesgo técnico al priorizar la similitud afectiva sobre la fama comercial. *Estos hallazgos cierran el ciclo de investigación, vinculando el reconocimiento técnico con la aplicación práctica en la industria.*

7. CAPÍTULO 7: CONCLUSIONES Y TRABAJOS FUTUROS

El éxito de este proyecto se fundamenta en el cumplimiento del objetivo general de diseñar un sistema capaz de reconocer y recomendar música basándose en la percepción emocional. La trascendencia de esta tesis para la disciplina de la Informática reside en la aplicación de un marco de "justicia algorítmica" que equilibra la balanza competitiva entre artistas comerciales y talentos emergentes.

7.1. Cumplimiento de Objetivos

- **O1 (Estado de la Computación Afectiva):** Se determinó que el reconocimiento emocional es viable pero dependiente de la calidad del etiquetado.
- **O2 (Características Intrínsecas):** Se identificaron el tempo, el pitch y el modo como disparadores emocionales críticos.
- **O3 (Estrategias MER):** Se validó experimentalmente que el enfoque dinámico supera al estático.
- **O4 (Relación Música-Emoción):** Se construyó un modelo técnico vinculante entre gramática musical y dimensiones V/A.
- **O5 (Prototipo de Clasificación):** Se implementó exitosamente un clasificador multiclase robusto.
- **O6 (Estudio de Sesgos):** Se categorizaron los sesgos preexistentes, técnicos y emergentes en la industria actual.
- **O7 (Experimento de Apreciación):** Se recolectó y analizó el *feedback* de oyentes reales frente a estímulos controlados.
- **O8 (Dataset ENSA):** Se diseñó y publicó un repositorio de datos original con licencias abiertas.
- **O9 (Sistema Recomendador):** Se desarrolló un prototipo híbrido funcional basado en series temporales.

7.2. Contribuciones Detalladas

- **Dataset ENSA:** Esencial para la investigación ética, permitiendo el estudio de estructuras musicales completas en artistas no comerciales.
- **Modelos MER:** Se demostró la eficacia de redes MLP con SMOTETomek para clasificar emociones dinámicas con precisión científica.
- **Sistema Recomendador Híbrido:** Proporciona un mecanismo real para mitigar el sesgo de popularidad mediante el uso de series temporales y DTW.

7.3. Implicaciones Prácticas y Futuras Líneas

A pesar de los logros, la necesidad de expandir el volumen de datos para mejorar la generalización de los modelos de *deep learning* es evidente. Se proponen tres líneas futuras:

1. **Expansión de ENSA:** Diversificar géneros y contextos de escucha.
2. **Extensión del Recomendador:** Incluir procesamiento de lenguaje natural (NLP) para el análisis de letras.
3. **Estudio de Sesgos:** Monitorear sesgos emergentes en entornos de uso real prolongado para garantizar una recomendación ética y sostenible.

Desarrollo de un modelo de *mirroring* con indicadores y herramientas para el seguimiento y visualización de procesos educativos colaborativos mediados por tecnología digital en educación superior

Autor: Mitchell John Vásquez Bermúdez

Directora: Dra. Cecilia Sanz

Co-directora: Dra. Ma. Alejandra Zangara

Institución: Facultad de Informática de la Universidad Nacional de La Plata (UNLP)

Fecha de exposición: 25 noviembre 2025

Motivación

El trabajo colaborativo en el ámbito de la educación ha constituido un modelo de aprendizaje interactivo que invita a docentes y estudiantes a contribuir juntos, conjugar esfuerzos, competencias de talentos y generar conocimientos (Maldonado, 2007). Al mismo tiempo el trabajo colaborativo, como estrategia de enseñanza y aprendizaje, puede mejorar la calidad académica, además de ser importante para el desarrollo individual e integrado de los docentes y estudiantes en el ámbito educativo (Moreno et al., 2013).En este contexto, los estudiantes necesitan generar mecanismos de interacción grupal y también se requieren herramientas que reflejen acciones individuales y colectivas, reuniendo y resumiendo datos con una visualización, como por ejemplo un gráfico circular o un gráfico de línea de tiempo que muestre las contribuciones de los participantes (Arciniegas-Méndez et al., 2015). Estas herramientas se conciben como tecnologías de apoyo al aprendizaje colaborativo y permiten la visualización y seguimiento. De esta manera, el proceso colaborativo puede ser evaluado durante y al final de las actividades y es posible medir su impacto. Zangara (2018) define al *mirroring* como la visibilizarían y seguimiento de los indicadores del proceso colaborativo a disposición de todos los integrantes de un grupo de trabajo. Así, conforme la estrategia de mirroring aumenta, se profundiza la información sobre cómo los estudiantes interactúan. Las herramientas de visualización de *mirroring*, por su dinámica de trabajo, utilizan indicadores para las diferentes interacciones de cada uno de los participantes en el proceso colaborativo. De esta manera los autores Boroujeni et al. (2017) integraron la dinámica en los foros de discusión, utilizando la metodología de modelado de técnicas de análisis de redes sociales. Los indicadores evaluados fueron el tiempo, el contenido y la dimensión social. Djouad (2011) en su investigación realiza aportes de soluciones novedosas a dos aspectos esenciales de indicadores: 1) Integración de un Sistema Basado en Rastros como un nuevo componente para el diseño de sistemas (integración en un framework como Moodle); 2) Construcción de un conjunto de modelos de Trace con un proceso de transformación para calcular explícitamente los indicadores en herramientas como foros, wikis y chats. En otra investigación (Jermann & Dillenbourg, 2008) se propuso visualizar las tasas de participación de colaboradores, a través de *mirroring*, por medio de gráficos de barras y circulares que muestran de forma dinámica el número de palabras, así como el número de acciones de resolución de problemas. Además, las autoras Zangara & Sanz (2017) presentan una estrategia de seguimiento de trabajo colaborativo bajo la forma de *mirroring*, revelando indicadores en una experiencia educativa de postgrado durante 2016. Pese al valioso aporte de las diferentes investigaciones, se encuentran escasos estudios que centren su atención en desarrollar e

integrar herramientas de *mirroring* con un análisis de indicadores de seguimiento en el contexto de la educación superior. En consecuencia, esta temática resulta de interés y actualidad. Para ello, en esta tesis se estudian y proponen indicadores para el desarrollo de una herramienta de seguimiento y visualización del trabajo colaborativo individual y colectivo en el ámbito de la educación superior. Como parte del estudio se toman las bases del modelo de Comunidad de Indagación (CoI) para enriquecer la estrategia de *mirroring* con el análisis de los componentes afectivo, cognitivo y la presencia del docente en los procesos de colaboración.

Aportes de la Tesis

El aporte de la tesis, es un modelo de *mirroring* que avanza sobre los modelos existentes, porque aporta tanto indicadores como herramientas concretas para el seguimiento y la visualización. Tiene tres componentes. Estos componentes son: Diseñar una Actividad Colaborativa, Metodología de seguimiento orientada al *mirroring* y la implementación de un *dashboard*.

El aporte de diseñar la actividad colaborativa es considerar al diseño de la actividad colaborativa con las presencias del modelo CoI. Se propone en tres etapas: diseño de actividad colaborativa, colaboración y seguimiento. El diseño de actividad colaborativa, orienta al docente a construir una consigna adecuada, considerando los aspectos de la presencia cognitiva. Se debe presentar una situación desencadenante del tema a abordar, esto anima a los estudiantes a participar e involucrarse. La presencia docente con este diseño de actividad busca despertar las presencias cognitiva y social. En la etapa de colaboración, el docente tiene que poner el foco en facilitar la interacción y el intercambio. En este modelo *mirroring* se propone el uso de la herramienta foro, donde se va a integrar la idea de uso de metadatos y meta-anotaciones. La meta-anotación es una característica metacognitiva relacionada con la información que el estudiante proporciona. Para poder hacer meta-anotaciones, el estudiante tiene que reflexionar sobre su pensamiento, que es una característica de la metacognición. En el foro, el propio estudiante informa el tipo de mensaje que quiere enviar y para qué al grupo. Esto lo hace con meta-anotaciones al mensaje, agregándole información. Finalmente, el seguimiento a las actividades colaborativas conlleva el uso de instrumentos que permitan al docente conocer el proceso de colaboración. En esta etapa se sugiere el uso de herramientas de software que permitan visibilizar, a través de gráficos, los resultados de las interacciones de los estudiantes y docentes. Este diseño e implementación guía a cualquier docente que desea implementar actividades colaborativas. Se le propone una estructura en tres etapas, la cual permite que considere las tres presencias del modelo CoI desde la gestación de la actividad.

El segundo aporte la Metodología de seguimiento orientada al *mirroring*, avanza sobre la metodología de MetScín[1], sobre los indicadores que propone MetScín y suma las Presencias del modelo CoI a través de nuevos indicadores. Esto significa la vinculación de las presencias con los tipos de mensajes. La Presencia Social, es considerar la comunicación afectiva, la comunicación abierta y las emociones que expresan los participantes en los mensajes, ya sea empatía, solidaridad o comprensión. La Presencia Cognitiva está relacionada con el adquirir conocimientos y se la vincula con los mensajes de contenido que se refieren al contenido de la actividad. Estos mensajes están categorizados por perspectivas de análisis, consistencia de la

[1] **MetSCIn** (Metodología de Seguimiento de la Colaboración e Interacción). Tesis: Interacción e interactividad en el trabajo colaborativo mediado por tecnología informática Metodología de seguimiento en escenarios educativos.

actividad, resolución de problemas, así como mejora de productos, construcción del discurso grupal, entre otros. La Presencia Docente está relacionada con el tipo de mensajes de organización que consiste en la forma que se organiza la actividad a realizar: identificación de roles, distribución de tareas por roles, estrategias de trabajo, gestión de tiempo, entre otras. Esto permitió refinar y derivar en una serie de indicadores que se presentan a continuación:

- En la Presencia Social se consideró los indicadores de Mensajes afectivos relacionados a las expresiones de emociones positivas y negativas que se expresan en los aportes. En particular, se toman los aportes en un foro de Moodle.
- La Presencia docente, están relacionados a los indicadores de Mensajes de organización con los aspectos de identificación de roles, distribución de tareas por roles, manejo del tiempo y organización de tareas. Se decide categorizar los indicadores de esta forma debido a que la presencia docente está relacionando al diseño, la organización y la estructura de tareas durante una actividad colaborativa. Es decir, es un derrame del diseño de la actividad colaborativa realizada por el docente.
- La Presencia cognitiva está relacionada con los indicadores de Mensajes de contenido, y éstos tienen que ser definidos por los docentes antes de iniciar un trabajo colaborativo de acuerdo con una consigna específica.

Dentro de la presencia cognitiva, el foro ofrece la funcionalidad de meta-anotaciones, que permiten a los estudiantes etiquetar e identificar sus propios aportes sobre el contenido con resolución de problemas, con recuperación de información, distintos niveles, en relación con un tema específico de discusión. Así, se propone un modelo de indicadores basado en las tres presencias: social, cognitiva y docente.

El tercer aporte es la implementación de un *dashboard*. Esto es un avance porque ninguno de los *dashboards* contaba con los otros dos componentes: una interrelación entre un diseño de consigna con las presencias de CoI, más unos indicadores de seguimiento con una metodología de seguimiento. Esta herramienta refleja ambos componentes.

Líneas de I+D futuras

1. En cuanto a los aspectos teóricos, se abre la posibilidad de profundizar en los marcos de diseño de aplicaciones basadas en *mirroring*, en el contexto educativo. Además, se plantea considerar estos marcos teóricos en el desarrollo de nuevas aplicaciones.
2. La propuesta metodológica y VIST fueron desarrollados para Moodle. Se sugiere la adaptación y evaluación de su aplicabilidad en otros entornos virtuales de enseñanza y aprendizaje.
3. Se deja abierta una línea de investigación para abordar estudios similares en diferentes contextos educativos.
4. Resulta de interés llevar a cabo un estudio comparativo de diversas plataformas tecnológicas utilizadas en la educación superior. Este análisis se centraría en la influencia de estas plataformas en la implementación del *mirroring* y su impacto en los procesos educativos colaborativos.
5. Sería valioso realizar un estudio de caso comparativo que incluya un grupo de control. Esto permitiría medir el impacto de la estrategia de *mirroring* y del *dashboard* VIST, contrastando

los resultados de un grupo que utiliza el modelo propuesto frente a otro que no lo hace, para así validar la efectividad de la intervención.

6. Se propone continuar estudiando y aplicando los aportes del modelo de Comunidad de Indagación en el diseño y desarrollo de modelos de seguimiento para el trabajo colaborativo.
7. Se propone integrar las posibilidades de la Inteligencia Artificial (IA) para la interpretación de los mensajes de los estudiantes en el modelo de *mirroring*.

Descubrimiento y diseño de microservicios asistido por IA a partir de requisitos textuales

Resumen extendido de tesis doctoral

Daniel Narváez
CAETI – Universidad Abierta Interamericana (UAI)
josedaniel.narvaezf@alumnos.uai.edu.ar

Director: Gustavo Rossi
CAETI – Universidad Abierta Interamericana (UAI); LIFIA – Universidad Nacional de La Plata (UNLP)
gustavo@lifia.info.unlp.edu.ar

Co-director: Alejandro Fernández
LIFIA – Universidad Nacional de La Plata (UNLP)
alejandro.fernandez@lifia.info.unlp.edu.ar

Universidad: Universidad Abierta Interamericana (UAI), CAETI
Aprobada para defensa *Fecha de exposición:* 24/02/2026

Resumen

Esta tesis aborda el problema de descubrir y diseñar arquitecturas de microservicios a partir de requisitos textuales en escenarios *greenfield*, donde la ambigüedad semántica y la ausencia de modelos formales incrementan el riesgo de particiones plausibles pero deficientes. Se propone `ArchiGenMS`, un pipeline neuro-simbólico *generar–verificar* que integra (i) generación inducida por modelos de lenguaje grandes (LLMs), guiada por *prompt engineering* y búsqueda evolutiva; y (ii) una compuerta de verificación formal ejecutable en Lean 4 que impone invariantes de *well-formedness* (tipado fuerte del genotipo JSON, consistencia referencial y prohibición de auto-llamadas) antes de computar métricas estructurales.

La investigación se enmarca en *Design Science Research* (DSR), articulando un ciclo de relevancia (diseño desde texto), un ciclo de rigor (síntesis del estado del arte y formalización de métricas/invariantes) y un ciclo de diseño (construcción y evaluación del artefacto). La evaluación empírica se realizó sobre 22 escenarios de historias de usuario, ejecutando cinco generaciones por escenario, con registro completo de genotipos, veredictos `OK/ERROR`, métricas y tiempos en un paquete de replicación. Los resultados agregados evidencian convergencia sostenida: la aptitud promedio $f(G)$ disminuye 44,3 % (Gen. 1 a Gen. 5), con mejoras en cohesión (LCOM_{avg} −20,4 %) y acoplamiento (Coupling_{max} −13,9 %), mientras el tiempo de verificación en Lean se mantiene estable (2.9–3.7 s) y el tiempo de generación con LLM disminuye 19,1 %. En conjunto, ArchiGenMS restringe la incertidumbre generativa a un espacio de arquitecturas bien formadas y, sobre ese subconjunto, guía la optimización hacia diseños con mejores propiedades estructurales.

Palabras clave: IA neuro-simbólica, microservicios, LLMs, búsqueda evolutiva, verificación formal, Lean 4, métricas arquitectónicas, DSR.

1. Introducción, motivación y planteamiento del problema

La arquitectura de microservicios se ha consolidado como un estilo dominante para construir sistemas distribuidos evolutivos, al promover despliegue independiente, escalabilidad selectiva y alineación con capacidades de negocio [1], [2], [3]. Sin embargo, sus ventajas dependen críticamente de decisiones de partición no triviales: un *wrong cut* suele manifestarse como baja cohesión, acoplamiento excesivo,

dependencias indeseables y deuda arquitectónica que degrada mantenibilidad y agilidad [4], [5], [6].

El problema se intensifica en escenarios *greenfield*, donde el diseño debe emerger desde artefactos tempranos predominantemente textuales (p. ej., historias de usuario). En este contexto, la incertidumbre semántica y la ausencia de modelos estructurados dificultan inferir límites de servicio, operaciones y dependencias. Enfoques previos han explorado el descubrimiento desde historias de usuario mediante aprendizaje y agrupamiento semántico; si bien son promisorios, siguen siendo sensibles a la calidad del texto y, en particular, carecen de garantías estructurales explícitas [7], [8], [9], [10]. En consecuencia, una descomposición puede ser *plausible* desde el lenguaje natural y aun así ser inconsistente como arquitectura (p. ej., referencias a servicios inexistentes, auto-llamadas o esquemas JSON incoherentes).

La irrupción de los LLMs amplía el espacio de automatización: pueden proponer descomposiciones y relaciones a partir de requisitos [11], [12], [13]. No obstante, su naturaleza estocástica introduce una brecha de confianza: la salida puede ser persuasiva pero no necesariamente correcta ni computable. En paralelo, la *Search-Based Software Engineering* (SBSE) ofrece un marco sólido para explorar espacios de diseño mediante búsqueda y optimización, siempre que exista una noción clara de *validez* y una función cuantificable de *calidad* [14], [15].

En esta tesis, la entrada es un conjunto de requisitos textuales U. Se busca producir una arquitectura candidata como genotipo serializable (JSON) que define: (i) microservicios, (ii) operaciones/parámetros, y (iii) un grafo dirigido de llamadas. La salida debe satisfacer simultáneamente: (a) **validez estructural mínima**, mediante invariantes duros verificables (tipado fuerte del esquema, consistencia referencial y prohibición de auto-llamadas); y (b) **calidad arquitectónica**, mediante métricas computables de cohesión, granularidad y acoplamiento [16], [17]. La hipótesis central es que un pipeline *generar–verificar* puede acotar la incertidumbre generativa mediante verificación formal y, sobre el subconjunto de soluciones válidas, conducir la optimización hacia arquitecturas con mejores propiedades estructurales.

Estructura del documento. La Sección 2 presenta el aporte central y el encuadre metodológico en DSR, explicitando cómo el ciclo de relevancia, rigor y diseño guía la construcción del artefacto. La Sección 3 posiciona el trabajo frente al estado del arte, delimitando la brecha entre plausibilidad semántica y correctitud estructural. La Sección 4 describe la propuesta `ArchiGenMS`: representación del genotipo, invariantes verificables en Lean 4, métricas y ciclo operativo *generar–verificar*. La Sección 5 reporta la evaluación experimental y la evidencia de convergencia, incluyendo resultados agregados y un caso ilustrativo. Finalmente, la Sección 6 discute implicaciones, limitaciones y amenazas a la validez, y la Sección 7 sintetiza conclusiones y líneas futuras.

2. Aporte central, contribuciones y enfoque metodológico

El aporte central es `ArchiGenMS`, un marco neuro-simbólico para diseño *greenfield* desde requisitos textuales. Su rasgo distintivo es materializar un bucle *generar–verificar* en el que los LLMs actúan como motor de variación semántica (guiado por *prompt engineering* y muestreo por inspiraciones), mientras que un *kernel* simbólico en Lean 4 impone invariantes duros de *well-formedness* y entrega veredictos deterministas (`OK/ERROR`) junto con métricas computables para guiar la selección evolutiva [14], [18]. De este modo, la búsqueda se restringe a arquitecturas estructuralmente válidas y se optimiza únicamente dentro de ese subconjunto.

En términos técnicos, el *kernel* formal: (i) impone tipado fuerte del esquema JSON (rechazando genotipos incoherentes); (ii) valida consistencia referencial (todo *caller*/*callee* existe en el registro de servicios); y (iii) prohíbe auto-llamadas (*self-calls*). Solo si estas condiciones se satisfacen, se computan métricas de cohesión, granularidad y acoplamiento, utilizadas como señal de aptitud [16]. Además, `ArchiGenMS` incorpora persistencia experimental y memoria de élites, reforzando trazabilidad y reduciendo la tasa de candidatos inválidos.

Metodológicamente, la investigación sigue *Design Science Research* (DSR) [19], [20]. En el ciclo de relevancia, se define el problema de diseñar microservicios desde texto; en el ciclo de rigor, se

fundamentan decisiones con evidencia del estado del arte y principios SBSE, formalizando métricas e invariantes verificables [21], [22]; en el ciclo de diseño, se construye el pipeline extremo a extremo y se evalúa empíricamente con trazabilidad y paquete de replicación [23], [24].

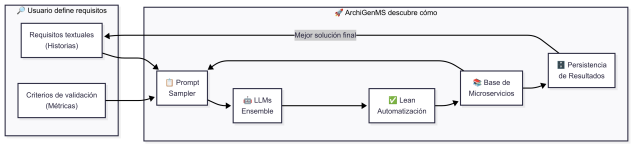
3. Posicionamiento frente al estado del arte

El descubrimiento de microservicios desde requisitos textuales ha sido abordado principalmente mediante técnicas de agrupamiento semántico y aprendizaje aplicadas a historias de usuario, buscando aproximar fronteras de servicio a partir de similitud de dominio [7], [8], [9], [10]. Sin embargo, incluso cuando la partición es plausible desde el lenguaje natural, suele persistir una brecha entre *plausibilidad semántica* y *correctitud estructural*: referencias inconsistentes, dependencias indeseables o artefactos no computables pueden degradar la validez del diseño resultante.

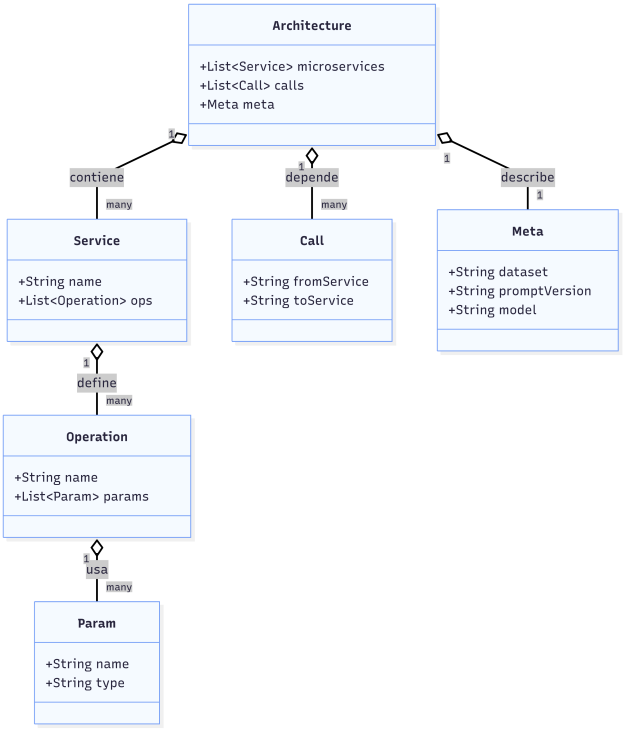
En paralelo, la generación asistida por LLMs ha emergido como mecanismo para proponer alternativas de arquitectura y apoyar decisiones de diseño a partir de texto [11], [12], [13]. No obstante, la estocasticidad del proceso generativo impone un problema de confianza: la salida puede ser persuasiva pero no necesariamente verificable. ArchiGenMS se ubica en el cruce de estas líneas: integra LLMs con búsqueda (SBSE) para explorar alternativas [14], [15], pero impone una compuerta formal ejecutable (Lean 4) para restringir el espacio de búsqueda a arquitecturas bien formadas [18]. De este modo, el enfoque no solo optimiza métricas de calidad [16], [17], sino que preserva una noción explícita de validez estructural mínima como precondición de evaluación y selección.

4. Propuesta: ArchiGenMS como artefacto verificable

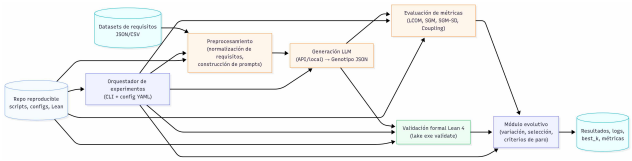
`ArchiGenMS` implementa un pipeline *generar–verificar* para descubrir y diseñar arquitecturas de microservicios a partir de requisitos textuales. La Fig. 1 sintetiza el flujo general, el ciclo evolutivo, la orquestación de alto nivel y la intuición de métricas empleadas para guiar la evaluación.



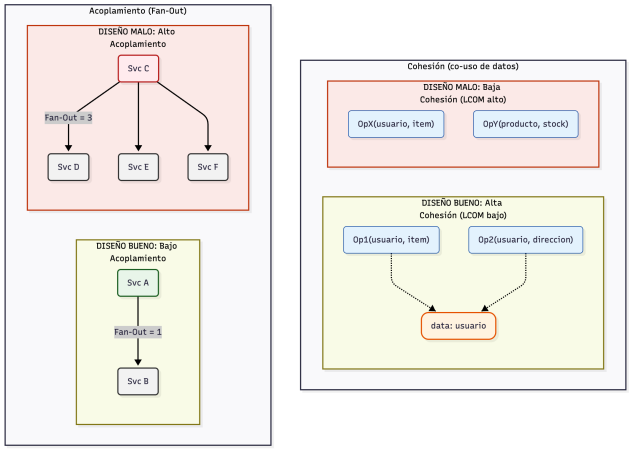
(a) Pipeline generar–verificar.



(b) Ciclo evolutivo.



(c) Arquitectura de alto nivel.



(d) Señales métricas (LCOM / fan-out).

Figura 1: Resumen visual de `ArchiGenMS`: flujo general, búsqueda evolutiva, arquitectura de orquestación e intuición de métricas.

Lectura de la Fig. 1. El panel (a) resume el principio arquitectónico del enfoque: la generación inducida por LLMs produce candidatos, pero la aceptación depende de una compuerta formal que impone validez mínima antes de evaluar calidad. El panel (b) muestra cómo esa compuerta se inserta en un ciclo inspirado en SBSE: la selección se realiza únicamente sobre individuos `OK`, evitando que mejoras numéricas surjan de artefactos inválidos. El panel (c) evidencia que la propuesta no es un prototipo aislado, sino un artefacto extremo a extremo: orquestación experimental, generación, verificación, cómputo de métricas y persistencia para trazabilidad. Finalmente, el panel (d) ancla la intuición de las señales métricas: se penalizan patrones típicos de degradación (baja cohesión vía LCOM alto y alto acoplamiento vía fan-out), mientras las restricciones duras permanecen bajo verificación formal.

Representación del genotipo arquitectónico

El espacio de búsqueda se define sobre una representación explícita serializable como JSON, compuesta por: (i) microservicios con operaciones parametrizadas; y (ii) llamadas dirigidas entre microservicios. Esta representación actúa como *genotipo* evolutivo: es el artefacto que el sistema genera, valida, evalúa y selecciona. En términos prácticos, permite expresar granularidad (número de operaciones por servicio), cohesión (co-uso de parámetros entre operaciones) y acoplamiento (fan-out de llamadas) [16].

Formalización mínima: genotipo, restricciones e indicadores

Sea un conjunto de requisitos textuales U. `ArchiGenMS` genera una arquitectura candidata como genotipo

$$G = (S, O, E),$$

donde S es el conjunto de microservicios, $O(s)$ es el conjunto de operaciones del servicio $s \in S$, y $E \subseteq S \times S$ es el conjunto de llamadas dirigidas (*caller* $\rightarrow$ *callee*).
La capa simbólica (Lean 4) impone invariantes duros de *well-formedness*:

$$\forall (u, v) \in E : \; u \in S \land v \in S \land u \neq v,$$

lo que garantiza (i) consistencia referencial y (ii) prohibición de auto-llamadas, además de tipado fuerte del esquema (rechazo de JSON incoherente) [18].
Sobre genotipos válidos, se computan métricas estructurales. El acoplamiento máximo (fan-out) se define como:

$$\text{Coupling}_{\text{máx}}(G) = \max_{s \in S} |\{ t \in S \mid (s, t) \in E \}| .$$

La granularidad máxima se aproxima por el tamaño del servicio:

$$\text{SGM}_{\text{máx}}(G) = \max_{s \in S} |O(s)|.$$

La cohesión se estima mediante una variante computable de LCOM basada en co-uso de parámetros entre operaciones: para $o_i, o_j \in O(s)$, sea $P(o)$ el conjunto de parámetros. Si P es el número de pares que **no** comparten parámetros y Q el número de pares que **sí** comparten, entonces:

$$\text{LCOM}(s) = \begin{cases} 0, & \text{si } |O(s)| < 2 \text{ o si } P = 0 \text{ o } Q = 0, \\ \dfrac{P}{P + Q}, & \text{en otro caso.} \end{cases}$$

Nota: Se utilizó la definición operacional de LCOM implementada en el experimento; en casos degenerados ($|O(s)| < 2$, $P = 0$ o $Q = 0$) la métrica se fija en 0 para estabilidad computacional.
Finalmente, la aptitud agregada (menor es mejor) se define como:

$$f(G) = \overline{\text{LCOM}} + \text{SGM}_{\text{máx}} + \text{SGM-SD}_{\text{sum}} + \text{Coupling}_{\text{máx}},$$

donde $\overline{\text{LCOM}}$ promedia LCOM por servicio y $\text{SGM-SD}_{\text{sum}}$ agrega dispersión de granularidad. Esta función guía la selección evolutiva dentro del subconjunto de genotipos formalmente válidos [16].

Cuadro 2: Ciclo *generar–verificar* en `ArchiGenMS` (vista operativa).

Entrada: requisitos U, generaciones T, población N; **Salida:** élites $\mathcal{E}$ con métricas y trazas.

Inicializar $\mathcal{P}_1 \leftarrow \text{Generate}(U)$ mediante LLM(s) restringiendo a JSON.
para $t = 1$ **hasta** T **hacer**
  para cada $G \in \mathcal{P}_t$ **hacer**
    $(ok, m) \leftarrow$ Verify_and_Measure(G) en Lean 4.
    si ok **entonces** computar $f(G) \leftarrow f(m)$ **si no** rechazar/penalizar.
  fin para
  Seleccionar élite $\mathcal{E}_t$ por menor $f(G)$ y persistir (G, m, f) como traza.
  Generar descendencia $C_t \leftarrow \text{Mutate}(\mathcal{E}_t, U)$ (LLM como variación semántica).
  Formar $\mathcal{P}_{t+1} \leftarrow \text{Survive}(\mathcal{E}_t, C_t)$ (con memoria de élites).
fin para
Retornar $\mathcal{E}_T$ y registros experimentales (JSONL/CSV).

Cuadro 1: Invariantes de *well-formedness* verificados en Lean 4 (compuerta de validez).

Invariante	**Verificación (idea formal)**	**Error típico**
Tipado fuerte del esquema	Parseo + chequeo de estructura tipada del genotipo $G = (S, O, E)$; campos obligatorios y tipos esperados	JSON incoherente
Consistencia referencial	$\forall (u, v) \in E : \; u \in S \land v \in S$	Servicio inexistente
Prohibición de autollamadas	$\forall (u, v) \in E : \; u \neq v$	*self-call*

Generación y verificación

La capa generativa utiliza LLMs como operadores de variación semántica para proponer particiones, operaciones y dependencias a partir del texto. Para reducir fragilidad, `ArchiGenMS` combina *prompt sampling* (condicionando con élites) y un *ensemble* que produce múltiples candidatos y selecciona el primero que supera verificación formal [11]. La verificación en Lean 4 actúa como compuerta: filtra genotipos inválidos y retorna métricas deterministas para guiar la búsqueda [18].

5. Evaluación experimental

La validación empírica se diseñó para determinar si `ArchiGenMS` (i) converge hacia arquitecturas con mejor calidad estructural, (ii) sostiene resultados consistentes sobre escenarios heterogéneos, y (iii) integra verificación formal sin sobrecarga prohibitiva. El diseño sigue el principio evaluativo de DSR [19], [20].

Dataset, variables y protocolo

Los experimentos se ejecutaron sobre el dataset público de historias de usuario de Dalpiaz [7] (22 archivos, un escenario por archivo), manteniendo fija la configuración del pipeline. Cada escenario induce una población inicial desde texto; cada individuo se valida en Lean 4 y, si es `OK`, se evalúa por métricas computables. Se iteró por cinco generaciones preservando élites y generando descendencia vía variación semántica inducida por LLMs. Genotipos, métricas y trazas se registran en `JSONL/CSV` para trazabilidad y replicación [24].

Las métricas evaluadas fueron: LCOM (cohesión), SGM/SGM-SD (granularidad/consistencia) y Coupling (fan-out) [16], [17].

Configuración experimental del pipeline

Los experimentos se ejecutaron con una configuración fija del pipeline: población $\mu = 10$, generaciones $T = 5$ y élite $E = 5$. La variación semántica se induce mediante el muestreo del modelo `gpt-4o-mini` con temperatura $\tau = 0{,}5$; no se fijó explícitamente una semilla global ni para el generador pseudoaleatorio de Python ni para el proceso generativo del LLM, por lo que el procedimiento es estocástico.

Cada escenario (archivo `.txt`) se procesó una vez de forma secuencial. En cada evaluación, la compuerta formal en Lean 4 valida el genotipo (`OK/ERROR`) y computa métricas deterministas sobre los individuos válidos. Para trazabilidad, se persisten el genotipo, veredicto, métricas y tiempos en `program_db.jsonl` y `results.csv`, además del mejor genotipo por escenario (`best_<story>.json`).

Adicionalmente, el sistema utiliza una memoria global de élites: antes de evaluar cada individuo, se incorporan como "inspiraciones" los n mejores genotipos históricos almacenados en `program_db.jsonl`. Esta estrategia reduce la tasa de candidatos inválidos y favorece transferencia de patrones estructurales entre escenarios, aunque introduce dependencia entre ejecuciones a través de dicha memoria.

Reproducibilidad y control experimental

Para reducir varianza no atribuible al artefacto, el protocolo fija (i) el conjunto de escenarios (22 archivos), (ii) el número de generaciones (cinco), y (iii) la definición de validez estructural y de métricas computables. Cada evaluación persiste: el genotipo G (JSON), el veredicto del kernel (OK/ERROR), el reporte de métricas y la aptitud $f(G)$, junto con metadatos de ejecución (tiempos de LLM y Lean, escenario y generación). Esta persistencia habilita auditoría posterior: permite reconstruir la secuencia de búsqueda, inspeccionar casos atípicos y replicar resultados bajo condiciones equivalentes [24].

Adicionalmente, la separación entre restricciones duras (Lean) y objetivos blandos (métricas) establece una trazabilidad causal: si un individuo mejora $f(G)$, la mejora se explica por cambios en métricas sobre un genotipo ya validado; si un individuo es rechazado, la causa se registra como violación de invariantes (Cuadro 1). Este diseño reduce el riesgo de optimizar sobre soluciones no arquitectónicas, un problema frecuente cuando la generación produce artefactos plausibles pero no computables.

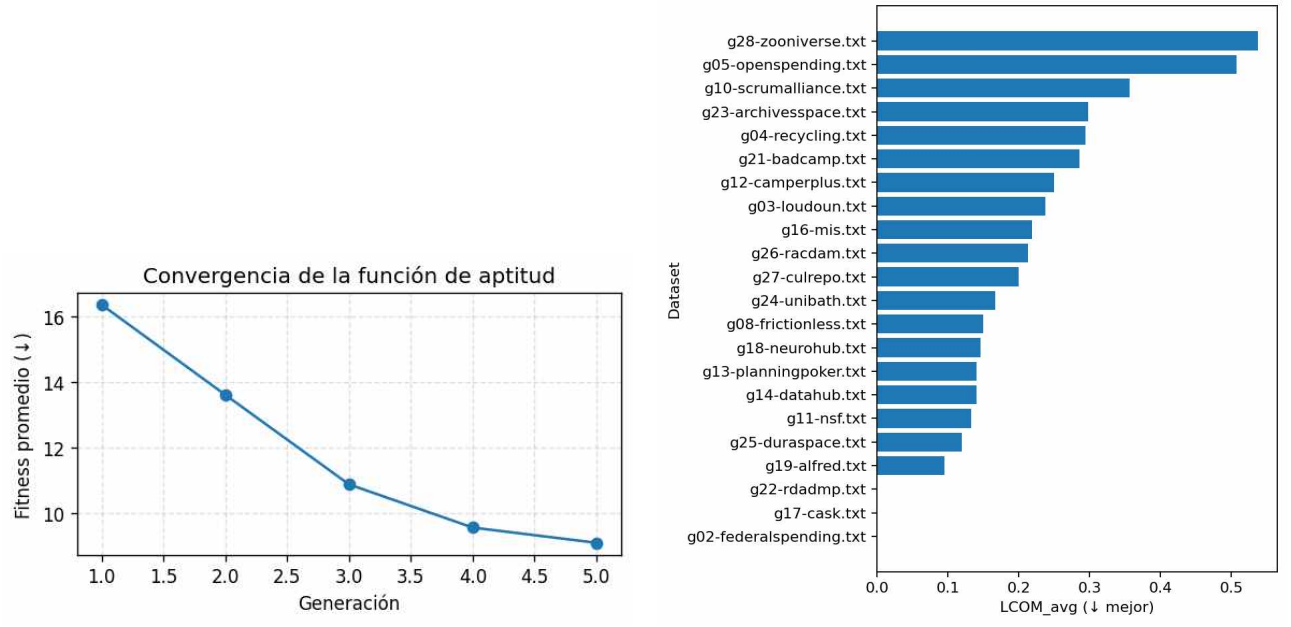
Lectura composicional: una intuición categórica del diseño

Como perspectiva conceptual complementaria, el diseño de microservicios puede interpretarse de forma composicional: los servicios se consideran *objetos* y las interacciones (llamadas/contratos) se modelan como *morfismos*. La orquestación emerge como composición (secuencial) y la coordinación de flujos independientes puede modelarse mediante una estructura monoidal (composición paralela). Bajo esta lectura, una refactorización que preserva invariantes puede entenderse como una transformación que mantiene propiedades por construcción, reforzando la idea de que ciertas garantías (validez estructural) deben imponerse antes de optimizar métricas. Esta línea abre una vía futura para formalizar reglas de composición y equivalencias de arquitectura usando herramientas estándar de teoría de categorías [25], [26].

Resultados agregados de convergencia

La Fig. 2 muestra: (a) la convergencia promedio de $f(G)$; y (b) una lectura por escenario basada en LCOM del mejor individuo válido. Complementariamente, el Cuadro 3 resume medias por generación (Gen. 1 vs Gen. 5).

Interpretación de la Fig. 2. En (a) se observa una disminución sostenida de $f(G)$, lo que sugiere que el ciclo *generar–verificar* no solo filtra, sino que también induce mejora intergeneracional en las métricas que componen la aptitud. En (b), la distribución por escenario evidencia heterogeneidad:



(a) Convergencia promedio de $f(G)$ (↓). (b) LCOM_{avg} del mejor individuo por escenario (↓).

Figura 2: Evidencia empírica: convergencia agregada y heterogeneidad por escenario.

Cuadro 3: Resumen agregado de convergencia (22 escenarios): medias por generación. Flechas: menor es mejor (↓).

Indicador	**Gen. 1**	**Gen. 5**	**Δ (%)**
$f(G)$ (↓)	16.346	9.104	−44,3
LCOM_{avg} (↓)	0.289	0.230	−20,4
Coupling_{max} (↓)	1.341	1.155	−13,9
Tiempo LLM (s) (↓)	17.868	14.452	−19,1
Tiempo Lean (s)	2.9–3.7	2.9–3.7	estable

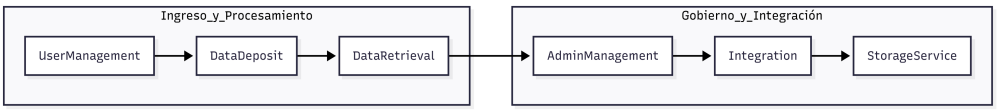


Figura 3: Arquitectura descubierta automáticamente para `g24-unibath` (validada por Lean).

existen dominios en los que se alcanza cohesión muy alta (LCOM_{avg} bajo) y casos atípicos donde la cohesión permanece menor. Esto es relevante porque indica que el enfoque es sensible a la estructura semántica del dominio y que los outliers pueden utilizarse como señales diagnósticas para refinar políticas de granularidad o añadir restricciones adicionales.

Los resultados completos por escenario (LCOM, SGM, SGM-SD, Coupling y $f(G)$) se reportan en el paquete de replicación [24].

Caso ilustrativo: `g24-unibath`

Para evidenciar interpretabilidad, se inspeccionó el escenario `g24-unibath` [7]. La Fig. 3 muestra la arquitectura descubierta: además de validarse en Lean (tipado, referencias consistentes y ausencia de *self-calls*), exhibe bajo acoplamiento observable y una estructura coherente con criterios métricos de microservicios [16], [17].

Interpretación del caso `g24-unibath` (Fig. 3). Este caso ilustra la interpretabilidad del artefacto: el diseño no se reporta como un conjunto de números, sino como una arquitectura inspeccionable que

ha superado invariantes de validez en Lean (estructura tipada, referencias consistentes y ausencia de auto-llamadas). Desde la lectura estructural, el diagrama permite discutir acoplamiento observable (fan-out limitado) y separación de responsabilidades, conectando la evidencia visual con las métricas empleadas en la aptitud y con el criterio de aceptación (`OK/ERROR`) del kernel formal.

6. Discusión, limitaciones y amenazas a la validez

Los resultados respaldan que `ArchiGenMS` operacionaliza un enfoque neuro-simbólico *generar–verificar*: Lean 4 actúa como compuerta de validez (`OK/ERROR`) y evaluador determinista de métricas, restringiendo la incertidumbre generativa a un espacio de soluciones bien formadas [18]. A nivel metodológico, el pipeline se alinea con SBSE al explorar un espacio de arquitecturas mediante una función de aptitud cuantificable [14]. A nivel arquitectónico, el enfoque no persigue solo plausibilidad semántica, sino señales estructurales medibles (LCOM/SGM/Coupling) interpretadas bajo criterios establecidos [16], [17].

La evaluación se realizó sobre un único dataset público [7]; aunque su diversidad aporta heterogeneidad, la generalización a contextos industriales, multilingües o con artefactos adicionales (UML/ADRs) requiere validación adicional. Asimismo, las métricas capturan calidad estructural estática; no se evaluaron propiedades dinámicas (rendimiento, resiliencia, escalabilidad). Finalmente, existe amenaza tecnológica: el comportamiento depende de modelos generativos específicos y de versiones del verificador formal; el paquete de replicación mitiga este riesgo al fijar condiciones experimentales [24].

7. Conclusiones y líneas futuras

Esta tesis presenta evidencia de que `ArchiGenMS` puede descubrir y diseñar arquitecturas de microservicios a partir de requisitos textuales mediante un pipeline neuro-simbólico que combina LLMs, búsqueda evolutiva y verificación formal en Lean 4. La evaluación muestra convergencia hacia mejores valores de aptitud y patrones consistentes de mejora estructural, manteniendo interpretabilidad mediante casos concretos y trazabilidad a través de registros experimentales reproducibles.

Como líneas futuras, se propone: (i) incorporar objetivos multi-criterio (p. ej., frentes de Pareto) para hacer explícitos *trade-offs* entre cohesión, granularidad y acoplamiento; (ii) ampliar el conjunto de invariantes verificables en Lean (p. ej., restricción de ciclos, contratos de interacción y consistencia referencial extendida); (iii) validar el enfoque en datasets multilingües y en dominios industriales; (iv) integrar métricas dinámicas y experimentos de *runtime*; y (v) explorar una formalización basada en álgebra de categorías para modelar microservicios e interacción como construcciones composicionales, de modo que comunicación/orquestación se representen mediante composición y las refactorizaciones se caractericen como transformaciones que preservan invariantes por construcción [25], [26].

Referencias

[1] J. Lewis y M. Fowler, "Microservices: a definition of this new architectural term (2014)," *URL: http://martinfowler. com/articles/microservices. html (cit. on p. 26)*, 2014.

[2] N. Dragoni et al., "Microservices: yesterday, today, and tomorrow," *Present and ulterior software engineering*, págs. 195-216, 2017.

[3] S. Newman, *Building microservices: designing fine-grained systems*. .º'Reilly Media, Inc.", 2021.

[4] D. Taibi, V. Lenarduzzi y C. Pahl, "Processes, motivations, and issues for migrating to microservices architectures: An empirical investigation," *IEEE Cloud Computing*, vol. 4, n.º 5, págs. 22-32, 2017.

[5] D. Neri, J. Soldani, O. Zimmermann y A. Brogi, "Design principles, architectural smells and refactorings for microservices: a multivocal review," *SICS Software-Intensive Cyber-Physical Systems*, vol. 35, n.º 1, págs. 3-15, 2020.

[6] H. Ünlü, D. E. Kennouche, G. K. Soylu y O. Demirörs, "Microservice-based projects in agile world: A structured interview," *Information and Software Technology*, vol. 165, pág. 107 334, 2024.

[7] F. Dalpiaz, *Requirements data sets (user stories)*, Mendeley Data, jul. de 2018. DOI: 10.17632/7zbk8zsd8y.1. dirección: https://doi.org/10.17632/7zbk8zsd8y.1.

[8] D. Bajaj, A. Goel y S. C. Gupta, "GreenMicro: identifying microservices from use cases in greenfield development," *IEEE Access*, vol. 10, págs. 67 008-67 018, 2022.

[9] F. H. Vera-Rivera, E. G. P. Cuadros, B. Perez, H. Astudillo y C. Gaona, "SEMGROMI—a semantic grouping algorithm to identifying microservices using semantic similarity of user stories," *PeerJ Computer Science*, vol. 9, e1380, 2023.

[10] G. Lucassen, F. Dalpiaz, J. M. E. van der Werf y S. Brinkkemper, "Improving agile requirements: the quality user story framework and tool," *Requirements engineering*, vol. 21, n.º 3, págs. 383-403, 2016.

[11] M. Esposito et al., "Generative AI for Software Architecture. Applications, Trends, Challenges, and Future Directions," *arXiv preprint arXiv:2503.13310*, 2025.

[12] J. R. A. Pereira et al., "Toward Generating Microservice Architectures from Textual Requirements with Large Language Models," en *Simpósio Brasileiro de Componentes, Arquiteturas e Reutilização de Software (SBCARS)*, SBC, 2025, págs. 79-89.

[13] R. Dhar, K. Vaidhyanathan y V. Varma, "Can llms generate architectural design decisions?-an exploratory empirical study," en *2024 IEEE 21st International Conference on Software Architecture (ICSA)*, IEEE, 2024, págs. 79-89.

[14] M. Harman, S. A. Mansouri e Y. Zhang, "Search-based software engineering: Trends, techniques and applications," *ACM Computing Surveys (CSUR)*, vol. 45, n.º 1, págs. 1-61, 2012.

[15] I. Saidani, A. Ouni, M. W. Mkaouer y A. Saied, "Towards automated microservices extraction using muti-objective evolutionary search," en *International conference on service-oriented computing*, Springer, 2019, págs. 58-63.

[16] O. Al-Debagy y P. Martinek, "A metrics framework for evaluating microservices architecture designs," *Journal of Web Engineering*, vol. 19, n.º 3–4, págs. 341-370, 2020.

[17] D. Taibi y K. Systä, *A Decomposition and Metric-Based Evaluation Framework for Microservices*, 2019. arXiv: 1908.08513 [cs.SE]. dirección: https://arxiv.org/abs/1908.08513.

[18] L. d. Moura y S. Ullrich, "The lean 4 theorem prover and programming language," en *International Conference on Automated Deduction*, Springer, 2021, págs. 625-635.

[19] A. R. Hevner, S. T. March, J. Park y S. Ram, "Design science in information systems research," *MIS quarterly*, págs. 75-105, 2004.

[20] A. Dresch, D. P. Lacerda y J. A. V. Antunes Jr, "Design science research," en *Design science research: A method for science and technology advancement*, Springer, 2014, págs. 67-102.

[21] D. Narváez, N. Battaglia, A. Fernández y G. Rossi, "Designing microservices using ai: A systematic literature review," *Software*, vol. 4, n.º 1, pág. 6, 2025.

[22] D. Narváez, G. H. Rossi y N. Battaglia, "Aplicación de Inteligencia Artificial en el Diseño de Microservicios," en *XXX Congreso Argentino de Ciencias de la Computación (CACIC)(La Plata, 7 al 11 de octubre de 2024)*, 2024.

[23] D. Narváez, N. Battaglia, A. Fernández y G. Rossi, "Descubrimiento automático de microservicios mediante modelos generativos y verificación formal," en *XXXI Congreso Argentino de Ciencias de la Computación (CACIC) (Viedma, 6 al 10 de octubre de 2025)*, Universidad Nacional de Río Negro (UNRN), Viedma, Argentina, 2025.

[24] D. Narváez, *ArchiGenMS: Reproducible Package*, https://github.com/jdanieln/archi-ai, Accedido: 22 jul. 2025, 2025.

[25] S. Mac Lane y S. MacLane, *Categories for the working mathematician*. Springer, 1971, vol. 5.

[26] B. Fong y D. I. Spivak, "Seven sketches in compositionality: An invitation to applied category theory," *arXiv preprint arXiv:1803.05316*, 2018.

DIAGNÓSTICO DE COMPETENCIAS BLANDAS MEDIANTE HERRAMIENTAS DE BUSINESS INTELLIGENCE

Doctorado en Informática

Alumna: Silvia V. Poncio

Director de tesis: Carlos Neil

Universidad Abierta Interamericana. Facultad de Tecnología Informática.

Centro de Altos Estudios en Tecnología Informática. Rosario. Argentina

{silvia.poncio, carlos.neil}@uai.edu.ar

Fecha de exposición: 19/12/2025

Síntesis de su contenido

El proyecto de investigación se desarrolla en la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI), sede Rosario, Argentina, en el marco del Centro de Altos Estudios en Tecnología Informática (CAETI). La investigación se inscribe dentro de la línea "Sociedad del Conocimiento y Tecnologías aplicadas a la educación", abordando específicamente el diagnóstico de competencias genéricas universitarias mediante herramientas de Business Intelligence (BI) [1].

El marco referencial se basa en las competencias genéricas definidas por el Consejo Federal de Decanos de Ingeniería (CONFEDI) en su Libro Rojo, documento que establece los estándares de segunda generación para la acreditación de carreras de Ingeniería en Argentina [2]. Estas competencias, adoptadas también por la Asociación Iberoamericana de Entidades de Enseñanza de la Ingeniería (ASIBEI) en la Declaración de Valparaíso, constituyen el faro orientador para la formación del ingeniero iberoamericano [3].

Marco Conceptual de las Competencias Profesionales

La formación universitaria contemporánea en Argentina se encuentra inmersa en un cambio de paradigma impulsado por la economía del conocimiento. Organismos como CONFEDI y ASIBEI reconocen que la educación superior debe trascender el saber teórico para centrarse en el desarrollo de competencias profesionales, lo que implica el saber hacer en contextos reales [2][3]. Desde la perspectiva de CONFEDI, una competencia se define como la capacidad de articular eficazmente un conjunto de esquemas (estructuras mentales) y valores, lo que permite movilizar distintos saberes en un contexto determinado con el fin de resolver situaciones profesionales [2]. Esta conceptualización integra diversos componentes esenciales: los saberes (conocimiento teórico, contextual y procedimental), el saber hacer (formalizado, empírico y relacional), el contexto profesional (escenario donde el egresado debe desempeñarse), y el desempleo profesional (actuación técnicamente competente y socialmente comprometida) [4].

En la visión de la UNESCO, afianzada en el informe Delors, se sostiene que la actuación idónea demanda la integración del saber ser, el saber conocer, el saber hacer y el saber convivir de manera compleja. Estos cuatro saberes fundamentales son la base de la educación para toda la vida [5].

Para asegurar un desarrollo curricular eficaz, las competencias en el ámbito de la ingeniería se clasifican en una tipología dual que distingue entre competencias específicas y genéricas:

Competencias Específicas (CE): Son aquellas propias de una determinada ocupación o profesión, caracterizadas por su alto grado de especialización. En el ámbito de la ingeniería en Argentina, derivan de las actividades reservadas al título, cuyo ejercicio tiene el potencial de comprometer el interés público (salud, seguridad o bienes). Por ejemplo, la Red de Ingeniería en Informática/Sistemas de Información (RIISIC) de CONFEDI estableció siete competencias específicas clave, incluyendo la capacidad de especificar, proyectar y desarrollar sistemas de información, así como certificar el funcionamiento, condición de uso o estado de sistemas informáticos [2].

Competencias Genéricas (CG): Son capacidades profesionales comunes a todos los ingenieros, independientemente de la rama específica. Representan habilidades transferibles esenciales para el desempeño profesional en cualquier contexto laboral. CONFEDI adoptó una síntesis de 10 competencias genéricas, consideradas complejas e integradas, que constituyen la base imprescindible para la consolidación de las competencias específicas [2][4].

Motivación

La formación universitaria contemporánea, especialmente en Argentina, se encuentra inmersa en un cambio de paradigma impulsado por la economía del conocimiento. Organismos como CONFEDI y ASIBEI reconocen que la educación superior debe trascender el saber teórico para centrarse en el desarrollo de competencias profesionales, lo que implica el saber hacer en contextos reales [2][3].

Tobón define las competencias como actuaciones integrales para identificar, interpretar, argumentar y resolver problemas del contexto, desarrollando de manera articulada el saber ser, saber convivir, saber hacer y saber conocer, con idoneidad, mejoramiento continuo y ética [4]. En este marco, las competencias blandas o soft skills adquieren relevancia creciente, especialmente en el contexto de la inteligencia artificial, donde las habilidades humanas como empatía, creatividad y liderazgo ético resultan irremplazables [6].

Habilidades Blandas en la Era de la Inteligencia Artificial

El concepto de soft skills surgió en la década de 1970 con el ejército de los Estados Unidos, refiriéndose originalmente a habilidades laborales relacionadas con acciones que afectan principalmente a las personas y al trabajo administrativo. Académicos contemporáneos las definen como competencias que, si bien no están directamente relacionadas con una tarea específica (como sí lo están las habilidades duras), son necesarias para cualquier puesto y se enfocan en las relaciones interpersonales [4].

Kubtová et al. plantean que existe consenso sobre la importancia crítica de las habilidades blandas para la empleabilidad, destacando comunicación efectiva, colaboración, trabajo en equipo, resolución de problemas complejos, pensamiento crítico y creatividad como capacidades esenciales para la sostenibilidad laboral. El valor de estas habilidades reside en que encapsulan la esencia atemporal de la naturaleza humana, aspectos que la inteligencia artificial no puede replicar. La necesidad de clarificar y estructurar las soft skills se ha intensificado debido al desarrollo tecnológico exponencial, especialmente la amplia disponibilidad de herramientas basadas en Inteligencia Artificial (IA), cuyo impacto en el trabajo basado en el conocimiento es significativo e inevitable. El valor de las habilidades blandas reside en que representan la base fundamental para la gestión de emociones humanas complejas, la toma de decisiones éticas y el fomento de interacciones significativas con terceros. La empatía, el liderazgo ético o la creatividad disruptiva (no replicativa) son capacidades que la IA no puede replicar [6].

Problemática y Justificación del Estudio

El desarrollo de modelos de Business Intelligence orientados al diagnóstico de competencias blandas en contextos universitarios constituye un campo escasamente explorado. La gestión del conocimiento (GC) como variable intangible sobre el desarrollo de competencias genéricas universitarias representa una problemática sobre la cual se enmarca esta investigación, aportando a la responsabilidad social de formar profesionales idóneos, cada vez más requeridos en el mercado laboral.

El escaso desarrollo de modelos creados desde BI que permitan llevar a cabo un diagnóstico de las competencias universitarias para generar espacios de reflexión respecto al desarrollo de las capacidades blandas requeridas para la formación de competencias transforma a este problema en una demanda altamente significativa. Las organizaciones empresariales dependen cada vez más de la información comprendida y procesada de un modo inteligente y de su capacidad de innovación, lo que les permite ser competitivas y establecer relación directa en la identificación y gestión del conocimiento. La inteligencia de negocios, según Belleza, Arias y Rico, se define como una herramienta tecnológica orientada a facilitar el acceso a información relevante en tiempo real, optimizando la elaboración de indicadores de gestión y mejorando la toma de decisiones estratégicas [7]. Aplicada al ámbito educativo, permite transformar datos cualitativos sobre competencias en conocimiento estratégico para la mejora institucional. El procedimiento de inteligencia de negocios permitirá obtener la GC con la realización de diagnósticos acerca de las competencias. Según The Data Warehouse Institute (TDWI), BI se define como la combinación de tecnología, herramientas y procesos que permiten transformar los datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia [8].

Aporte de la tesis a la disciplina

Las implementaciones realizadas entre 2018 y 2025 en difundida en distintos contextos México, Argentina, España, Cuba entre otros permitió validar y consolidar el Modelo Diagnóstico de Competencias Universitarias (MDCU) como herramienta estratégica para la gestión del conocimiento sobre habilidades blandas. El modelo que se presenta resultó de un proceso de investigación basado en BI, con el fin de obtener aproximaciones del desarrollo de las capacidades de las competencias genéricas del alumno universitario de manera visual. Inicialmente denominado Modelo de Evaluación de Competencias Universitarias (MECU) en 2016 [9], se redefinió como Modelo Diagnóstico de Competencias Universitarias (MDCU) en 2018, dado que el concepto de evaluación es una categoría altamente controvertida en el ámbito educativo y la intencionalidad del modelo trasciende las fronteras evaluativas [10].

El modelo se sustenta en la existencia de niveles de información que manejan los sistemas de BI, utilizando el test Myers-Briggs Type Indicator (MBTI) basado en la teoría de Carl Jung [11]. En el VI Congreso Internacional de Emprendimiento AFIDE (IX International Summer School of Entrepreneurship - ISSE) en Veracruz, México 2018, se presentó la aplicación del modelo. La población y muestra del estudio estuvo compuesta por 29 alumnos de 5to año de Ingeniería en Sistemas, con un promedio de edad de 30 años. Los resultados evidenciaron diferencias significativas entre primer y quinto año, sugiriendo un desarrollo progresivo de estas capacidades durante el trayecto universitario. Se observó la dificultad de los futuros egresados en la habilidad de relacionarse bien con otros, mientras que en la categoría "Actuar de manera autónoma" se evidenció poca dificultad en los proyectos personales (37,14%), pero la visión sistémica (actuar dentro de un gran esquema) presentaba dificultades en un 28,57%. Esta implementación permitió validar que el modelo posibilita realizar drill down (análisis en profundidad) hasta el nivel individual del alumno, habilitando evaluar el estado de situación individual como grupal para establecer estrategias de acción concreta, tales como cursos de formación que generen cambios en el desarrollo de estas competencias [10].

La implementación del diagnóstico para desempeñarse de manera efectiva en equipos de trabajo, se presentó en el Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI) en la Universidad Tecnológica Nacional, Facultad Regional San Francisco, Córdoba, Argentina. El universo de estudio comprendió estudiantes de 1er a 5to año de Ingeniería en Sistemas, con una población y muestra de 56 estudiantes. Los resultados mostraron que en la Categoría 2 "Interactuar en grupos heterogéneos", el 67,69% de los estudiantes manifestaron tener la habilidad de manejar y resolver conflictos, el 63,93% presentó habilidad de cooperar, y el 50,45% reconoció tener habilidad para relacionarse bien con otros. Se identificaron preferencias diferenciales en medios de comunicación según el año de cursado, con mayor preferencia por comunicación digital en años superiores. En primer año predominaba el medio verbal como preferido para la comunicación entre compañeros del grupo de referencia. En años más avanzados, esa comunicación se ampliaba a otros alumnos fuera del grupo de pertenencia. Dentro de las habilidades a mejorar detectadas se destacó un porcentaje muy bajo (21,43%) referido a la capacidad de tomar decisiones que permitan diferentes opiniones. Solo el 35,71% de los encuestados presentaba manejo efectivo de las emociones en la habilidad de relacionarse bien con otros [12].

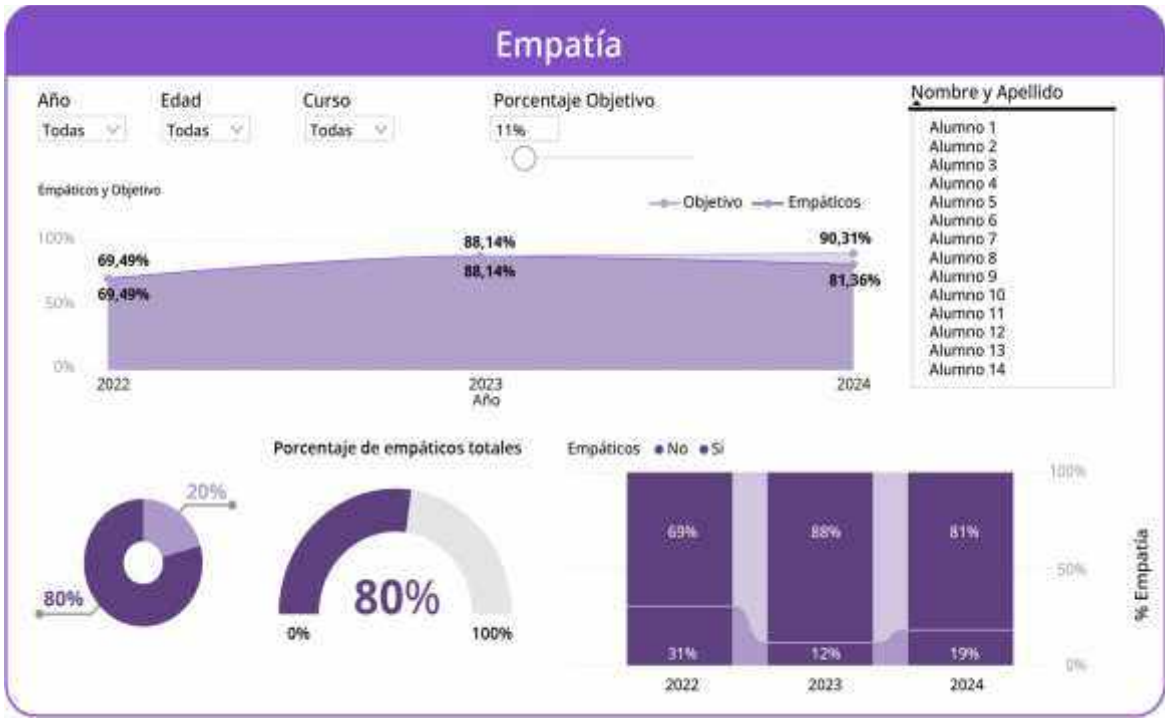
En España 2023, en colaboración con la Universidad de Salamanca, la investigación fue publicada en el libro "Persona, empresa, sociedad y emprendimiento en el contexto de los ODS" (Editorial Dykinson). El objetivo fue construir una experiencia educativa que permitiera observar qué sucede cuando se pone en práctica una acción a través de una actividad que contribuye al desarrollo de una competencia genérica, evidenciando que el uso del análisis de datos BI y prácticas del Management 3.0 pueden ser utilizadas en el ámbito educativo universitario [13].

La experiencia de Cuba 2024 se centró en comportamientos de empleabilidad, identificando que el 75% de los estudiantes manifestaba iniciativa, creatividad e innovación, capacidades vinculadas al potencial emprendedor [14]. Esta implementación validó la aplicabilidad del modelo en contextos culturales diversos. Se presentó en el 14º Congreso Internacional de Educación Superior Universidad 2024 en La Habana, Cuba. El universo de estudio comprendió alumnos de 1º y 5º año de Ingeniería en Sistemas durante 2023, con una población y muestra de 110 alumnos (85 de primer año y 25 de quinto año). El enfoque se centró en diagnosticar el avance del aprendizaje autónomo en estudiantes de Ingeniería en Sistemas. Según ASIBEI [3], el ingeniero iberoamericano debe ser un profesional provisto de cultura general, con cosmovisión sistémica que le permita aplicar sus conocimientos en cualquier lugar del mundo, actuando con solidaridad social. En primer año, el 68% de los alumnos respondió afirmativamente a las preguntas sobre aprendizaje autónomo, mientras que en quinto año el porcentaje ascendió a 70,29%. Este incremento sugiere un desarrollo progresivo de la capacidad de aprender de forma continua y autónoma durante el trayecto universitario.

En Argentina 2025 se consolidó el modelo incorporando indicadores, logrando dashboards interactivos con capacidad de análisis multidimensional mediante cubos OLAP. Se implementaron exitosamente herramientas QlikView y Power BI, evidenciando la versatilidad del modelo para distintas plataformas tecnológicas. Se presentó en el XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025) en la Universidad Nacional de Cuyo, Mendoza, Argentina, y en el XX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET) en la Universidad Nacional de Avellaneda [15] [16].

Un recurso clave implementado fue el drill down, funcionalidad que permite profundizar una visión generalista hacia niveles de información más detallados. Desde un tablero que muestra el promedio general de competencias socioemocionales, resulta posible desglosar los datos por carrera, curso, cohorte o incluso por estudiante individual, facilitando diagnósticos más precisos y personalizados. Se incorporó el desarrollo de Key Performance Indicators (KPI) para las competencias blandas. Según Parmenter [17], los KPIs son herramientas esenciales para medir áreas específicas, enfatizando la

importancia de alinear estos indicadores con los objetivos estratégicos de una organización. Se desarrolló específicamente el KPI de "Empatía", observándose su progreso en ingresantes de primer año durante los períodos 2022 al 2024. Los resultados mostraron que los estudiantes ingresantes del primer año 2023 presentan niveles más elevados de empatía en comparación con cohortes anteriores, evidenciando la importancia de dinámicas diferenciadas dentro del ámbito universitario [16].



Dashborard Key Performance Indicator: Empatía.

Cubo de Datos OLAP y Data Warehouse

Se diseñó un cubo de datos OLAP (Online Analytical Processing) considerando los componentes según Kimball y Ross [8]. El Data Warehouse constituido contiene los datos ordenados y estructurados, lo cual permite una velocidad de acceso a la información sumamente rápida.

Reportes y Dashboards

Los reportes se generan en orden por cada competencia, subcompetencia y categorías detalladas. Brindan una visión desglosada con drill down hasta el nivel del alumno para que los usuarios puedan examinar y comprender mejor el diagnóstico, permitiendo en el dashboard interactivo explorar los datos y generar informes personalizados. Se incluyen indicadores clave de rendimiento (KPIs) de las competencias blandas desarrolladas a nivel de los alumnos, parametrizándolas. La visualización de la información permite observar el diagnóstico para que los analistas identifiquen oportunidades reflexivas en una determinada competencia e implementen estrategias tácticas entorno a dicho conocimiento visibilizado mediante gráficas representativas [10].

Líneas de investigación futuras

Los trabajos futuros se enmarcan en varias líneas de desarrollo e innovación que amplían el alcance y la aplicabilidad del modelo. La implementación del modelo en el ámbito empresarial permitirá diagnosticar el desarrollo de competencias sociales enmarcadas en el desarrollo profesional.

Representa una innovación metodológica que trasciende su origen académico para posicionarse como una herramienta estratégica en la gestión del talento humano. Al trasladar su lógica de visualización competencial al entorno organizacional, permite diagnosticar el desarrollo de competencias sociales —tales como trabajo colaborativo, comunicación efectiva, liderazgo y responsabilidad— enmarcadas en el proceso de profesionalización. Este modelo se articula con los principios de la gestión del conocimiento y la inteligencia organizacional, al ofrecer indicadores visuales que facilitan la toma de decisiones en áreas como selección, formación, evaluación del desempeño y planificación de trayectos formativos. El MDCU en contexto empresarial contribuye a fortalecer la empleabilidad interna, al identificar brechas formativas y orientar intervenciones específicas que alineen el desarrollo de competencias con los objetivos estratégicos de la organización. En este sentido, el modelo no solo diagnostica, sino que también potencia la capacidad de adaptación, innovación y resiliencia del capital humano, consolidando su aporte al rendimiento institucional y a la generación de ventajas competitivas sostenibles. Se propone llevar a cabo el proceso ETL (Extract, Transform, Load) con Inteligencia Artificial, proponiendo la comparación de los resultados a través de diferentes asistentes de IA y desarrollar luego el resto del Modelo con el sistema de BI ya previsto. Esta línea de trabajo permitirá observar una contribución a lo expresado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, cuando en su informe destacan la importancia de las habilidades socioemocionales en la educación y cómo estas competencias pueden ser evaluadas y mejoradas [18].

Habilidades Socioemocionales: Ampliación del Alcance

Es importante reflexionar respecto al amplio rango de habilidades que se encuentran en el ámbito socioemocional. En las últimas décadas, diversos modelos han buscado definirlas, agrupándolas en diferentes dimensiones en función de su enfoque teórico. Algunos modelos utilizan el concepto de habilidad, otros se refieren a competencias. Sin embargo, aluden en todos los casos al repertorio de capacidades que permiten vincularse afectivamente con uno mismo, con los demás y con el entorno. La UNESCO [18] se pregunta: ¿Cuáles son las habilidades socioemocionales? Esta pregunta fundamental guiará futuras investigaciones orientadas a expandir el modelo hacia nuevas dimensiones competenciales, incluyendo aspectos como inteligencia emocional, resiliencia, capacidad de adaptación al cambio, y competencias interculturales.

Transferencia y Escalabilidad

Para la transferencia de los resultados, se propone la realización de actividades colaborativas a través de talleres reflexivos que hagan uso del modelo de diagnóstico. El objetivo de estos espacios es fortalecer el desarrollo de las competencias blandas en los alumnos, contribuyendo de manera significativa a la consolidación de su perfil profesional-laboral. Este enfoque tiene la ventaja de ser altamente transferible y aplicable a cualquier organización. Dicha aplicabilidad se sustenta en la premisa de que toda organización es una entidad creada por individuos que comparten intereses y valores, y que buscan lograr ciertos objetivos. En este sentido, el modelo de diagnóstico de competencias blandas no se limita al entorno universitario, sino que puede ser adoptado por cualquier institución que se conciba como un producto social y que aspire a gestionar el talento de sus miembros de forma estratégica [16].

BIBLIOGRAFÍA

[1] S. Poncio, L. Semino, C. Cuña, N. Vozella, A. M. Trottini, P. Contesti, y Y. Montella, "Diagnóstico de competencias genéricas universitarias, a través de software Business Intelligence," en *Libro de actas – XXV Workshop de Investigadores en Ciencias de la Computación (WICC 2023)*, M. Sarobe, L. B. Cicerchia, H. D. Ramón, L. Esnaola, J. P. Tessore, y C. C. Russo, Eds. Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA), 2023, pp. 806-811.

[2] CONFEDI, *Libro Rojo de CONFEDI: Propuesta de Estándares de Segunda Generación para la Acreditación de Carreras de Ingeniería en la República Argentina*. Consejo Federal de Decanos de Ingeniería, 2014.

[3] ASIBEI, *Plan Estratégico 2016-2025*. Asociación Iberoamericana de Entidades de Enseñanza de la Ingeniería, 2016.

[4] S. Tobón, *Formación integral y competencias: Pensamiento complejo, currículo, didáctica y evaluación*, 4ta ed. Editorial ECOE, 2013.

[5] J. Delors, *La educación encierra un tesoro: Informe a la UNESCO de la Comisión Internacional sobre la Educación para el Siglo XXI*. París, Francia: UNESCO, 1996.

[6] J. Kubtová et al., "Soft Skills for the 21st Century: A comprehensive review of employability competencies in the age of artificial intelligence," *International Journal of Educational Research*, vol. 125, pp. 102-118, 2025.

[7] J. Belleza, M. Arias, y D. Rico, "Business Intelligence como herramienta para la toma de decisiones estratégicas en instituciones educativas," *Revista Iberoamericana de Tecnología en Educación*, vol. 14, no. 2, pp. 45-58, 2019.

[8] R. Kimball y M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3ra ed. Wiley, 2013.

[9] Poncio, S. V., Trottini, A. M., Ferrandi, M. A., & Zatta, M. S. (2016). Modelo de Evaluación de Competencias Universitarias (MECU). HCITISI - Sexto Congreso Argentino de La Interacción Persona Computador@, Telecomunicaciones, Informática e Información Científica. Huerta Grande, Córdoba. Argentina., ISBN 978.88.96.471.54.8-DOI 10.978.8896471/548. https://doi.org/https://doi.org/10.978.8896471/548.

[10] S. Poncio y A. M. Trottini, "Modelo Diagnóstico de Competencias universitarias (MDCU): Aprender a Emprender," presentado en VI Congreso internacional de Emprendimiento AFIDE. IX International Summer School of Entrepreneurship (ISSE), Veracruz, México, 2018.

[11] Myers, I., & Briggs, K. (1976). Myers-Briggs Type Indicator (MBTI)

[12] S. Poncio, V. Castañeira, D. Tedini, L. Almada, y A. Bravo, "Diagnóstico para desempeñarse de manera efectiva en equipos de trabajo en alumnos universitarios (MDCU)," presentado en el Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2020), Universidad Tecnológica Nacional, Facultad Regional San Francisco, Córdoba, Argentina, 2020.

[13] Poncio, S., Cuña, C., & Vozella, A. (2023). Análisis de datos y prácticas de gestión para el desarrollo de competencias de emprendimiento del estudiante universitario. Persona, empresa, sociedad y emprendimiento en el contexto de Los ODS, Dykinson, 94–99. https://www.torrossa.com/gs/resourceProxy?an=5477488&publisher=FZ1825#page=94

[14] S. Poncio Manassero, A. M. Trottini Vitullo, y B. Hernández Sánchez, "Aplicación de inteligencia de negocios para diagnosticar comportamientos de empleabilidad en estudiantes de Ingeniería en Sistemas," presentado en el 14º Congreso Internacional de Educación Superior Universidad 2024, La Habana, Cuba, 2024.

[15] C. G. Neil, N. Battaglia, M. De Vincenzi, N. Garrido, F. Maffei, S. Poncio, y M. Monferrato, "El impacto de inteligencia artificial en la definición, desarrollo y evaluación de competencias,"

presentado en el XXVII Workshop de Investigadores en Ciencias de la Computación (WICC 2025), Universidad Nacional de Cuyo, Mendoza, Argentina, 2025.

[16] S. Poncio, C. Cuña Giménez, N. Card, y G. Ruiz, "Data Warehouse Soft Skills: Modelo diagnóstico para la observación de la actitud empática," presentado en el XX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2025), Universidad Nacional de Avellaneda, Avellaneda, Argentina, 2025.

[17] Parmenter, D. (2015). Key Performance Indicators. Wiley. https://doi.org/10.1002/9781119019855.

[18] UNESCO. (2024). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084.

Carrera: Doctorado en Ciencias Informáticas

Título de la Tesis: "Formulación de Métricas para la Evaluación Integral de Metodologías de Gestión de Proyectos"

Autor: Patricia Raquel Cristaldo

Director/Codirector: Dra. Daniela López De Luise / Dr. Leandro Antonelli

e-mail de contacto del autor: cristaldop@frcu.utn.edu.ar

Universidad: Universidad Nacional de La Plata - Facultad de Informática

Fecha de exposición: 11 de Junio de 2024

Mini biografía académica/profesional

Ingeniera en Sistemas de Información de UTN. Posgrados: Especialista en Ingeniería en Sistemas de Información y Magíster en Ingeniería en Sistemas de Información. Experiencia profesional, Docente e Investigadora: Área Sistemas de Información y Gestión Ingenieril. Profesora de Enseñanza Superior en Sistemas de Información.

Palabras Claves: *indicar una lista de hasta 8 palabras clave separadas por punto y coma (;)*

Gestión de proyectos; Metodologías de gestión; Métricas de calidad; Procesamiento de Lenguaje Natural.

MOTIVACIÓN

La motivación principal de este trabajo es mejorar el proceso de definición del alcance de un proyecto y, por consiguiente, del producto que contiene el mismo, haciendo más atractiva dicha tarea, incorporando ludificacion. En función de ello este trabajo desarrolla un método que involucra un conjunto de métricas para la identificación del conocimiento implícito que se encuentra dentro de los documentos técnicos que describen el alcance del proyecto, aplicando minería de textos, para descubrir y extraer conocimiento relevante y generar un modelo que describa las características esenciales del proyecto. Esto es relevante dado que la complejidad del lenguaje natural dificulta el acceso a la información en los textos, por lo que la semántica expresada puede diferir del modelo pretendido de sistema.

Las principales contribuciones de este trabajo son:

- El diseño de métricas que permiten ponderar el grado de completitud de la descripción textual del alcance de un proyecto, plasmado en la documentación de dicho proyecto.
- El diseño de métricas que permiten establecer los criterios de medición y validación de los requisitos, a partir de los textos que describen los mismos.
- El diseño de métricas que permiten determinar si cada uno de los requerimientos, presentan algún criterio de trazabilidad en su descripción. Esto se basa en una extrapolación adecuada de las palabras claves en cada oración del documento que

describe el alcance del proyecto. El valor obtenido descansa especialmente en los estilos propuestos y en la ponderación inducida a nivel de sentencias del escrito. De esta manera se organizan los resultados midiendo la trazabilidad de cada requerimiento.

- Definición y diseño de un método general basado en un marco de medición que integra las métricas propuestas anteriormente y que son evaluadas en el contexto de esta tesis.
- Definir un método ludificado que aplique el marco de medición para la mejora de los proyectos. La ludificación busca mejorar el compromiso, la motivación y el desempeño de las personas encargadas de formular el alcance de un proyecto, incorporando elementos y técnicas de juego, haciendo que las actividades sean más atractivas.

Objetivo y Alcances

El objetivo de este trabajo es definir y desarrollar métricas e indicadores para la evaluación transversal de metodologías de gestión de proyectos, incluye desde el diseño y selección de las métricas, definición y ajuste matemático a partir de experimentos reales. Se integra además con un sistema ludificado. El proceso completo se realiza a partir de expresiones textuales de la descripción de un proyecto.

APORTES DE LA TESIS

El principal aporte de esta tesis es definir un método ludificado que aplica un marco de medición para la mejora en la gestión de los proyectos. En dicho marco, se integra el conjunto de métricas con resultados. Luego, se muestra al prototipo como herramienta que brinda ayuda para aplicar el método.

LÍNEAS DE I/D FUTURAS

- A partir de la implementación de los módulos de ludificación, aplicados con distintas plataformas, evaluar dichos comportamientos y generar las correspondientes estadísticas e informes. Es decir, evaluar estadísticamente los registros de la recopilación de datos gamificados y el análisis de esas estadísticas desde la perspectiva de los hallazgos publicados anteriormente.
- Establecer la relación entre las métricas y el conocimiento implícito en los documentos textuales que se utilizan para describir varias partes del proyecto, y un poco de minería de los resultados para descubrir la relación oculta con otras variables. Por la complejidad del lenguaje natural constituye un desafío que puede ayudar a resolver ciertos problemas en el desfase entre las definiciones originales y la producción final, y la cantidad de problemas en los enunciados iniciales de las propuestas elaboradas por las organizaciones.

- Relacionar las métricas para la identificación de conocimiento implícito, encontradas en textos que describen el alcance del proyecto, aplicando técnicas de minería de texto, relacionadas con el proceso de descubrir y extraer conocimiento relevante y no trivial de textos no estructurados. Sin embargo, la complejidad del lenguaje natural dificulta la extracción de información de los textos. En consecuencia, la construcción de representaciones de significado de propósito general a partir del texto aún depende de restricciones.
- Probar estadísticamente la relación entre más casos y métricas, las características de la documentación respetando el sesgo de administración y el impacto en los proyectos actuales y futuros de las métricas de la organización.
- Estas métricas pueden brindar un buen asesoramiento sobre cómo mejorar las prácticas actuales en la empresa de una manera muy precisa. Con una herramienta adecuada para recopilar esta información, se espera que los gerentes y propietarios puedan comprender mejor el proceso de la producción final.
- Ampliar la muestra de este estudio para generalizar los hallazgos y transferirlos a otros contextos.

PUBLICACIONES RELACIONADAS CON LA TESIS

---

a. **Patricia R. Cristaldo, Daniela López De Luise, Lucas La Pietra, Anabella De Battista, Jude Hemanth (2021)** - "*Metrics for the Systematic Evaluation of Project Management Methodologies"*. Global Research and Development Journal for Engineering, ISSN online 2455-5703, Vol.6 Issue 5, April 2021. (https://link.springer.com/chapter/10.1007/978-3-030-77637-4_3).

b. **Patricia R. Cristaldo, Daniela López De Luise, Lucas La Pietra, Anabella De Battista (2021) - "***Adjustment Level Metrics for Requirements in Project Management"*. Proceedings of 2021 8th Intl. Conference on Soft Computing & Machine Intelligence (ISCMI 2021), IEEE ISSN: 2640-0146, ISBN: 978-1-7281-8682-5, IEEE Catalog Number: CFP2144Z-USB, Cairo, Egypt. 2021.
(se adjunta Acta)

c. **Patricia R. Cristaldo, Daniela López De Luise, Lucas La Pietra, Anabella De Battista (2021) - "***Metrics for validation and traceability of Project Management Requirements"*. ID **GRDJEV06I050009** has been approved in GRD Journal for Engineering, ISSN: 2455—5703.
(https://issuu.com/grdjournals/docs/grdjev06i050009).

d. **Patricia R. Cristaldo, Daniela López de Luise, Lucas La Pietra, Anabella De Battista and Jude Hermanth (2021)** – Capítulo de libro: **—** "*Data Mining-Based*

Metrics for the Systematic Evaluation of Software Project Management Methodologies". Blockchain. Artificial Intelligence, and the Internet of Things., pag. (35-51) – Editorial Springer – DOI: 10.1007/978-3-030-77637-4_3. (https://www.springerprofessional.de/en/data-mining-based-metrics-for-the-systematic-evaluation-of-softw/19890306).

e. **Patricia Cristaldo, María Daniela López de Luise, Lucas La Pietra, Soledad Retamar, Anabella De Battista (2022) –** *"Medición para la evaluación transversal de metodologías de gestión de proyectos"*. XXIV WORKSHOP DE INVESTIGADORES EN CIENCIAS DE LA COMPUTACIÓN (WICC), Mendoza, Argentina. (http://sedici.unlp.edu.ar/bitstream/handle/10915/144323/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y)

f. **Patricia Cristaldo, María Daniela López de Luise, Lucas La Pietra (2022) –** ***"****Métricas para Metodologías de Gestión de Proyectos: planificación de calidad y equipo de gestión"*. VI IEEE Biennial Congress of Argentina (ARGENCON) Buenos Aires, Argentina.

g. **Patricia Cristaldo, María Daniela López de Luise, Lucas La Pietra (2022) –** ***"****Influencia de la Visión Organizacional en los riesgos de la Gestión de proyectos"*. 10º Congreso Nacional de Ingeniería Informática / Sistemas de Información - CONAIISI 2022, Entre Ríos. Argentina. ISBN 978-950-42-0218-9. (https://rtyc.utn.edu.ar/index.php/ajea/article/view/1146)

h. **Patricia R. Cristaldo, Daniela López de Luise, Lucas La Pietra (2023) – "***Metrics for Project Management Methodologies Elicitation"*. – Capítulo de libro: Perspectives and Considerations on the Evolution of Smart Systems. DOI: 10.4018/978-1-6684-7684-0.ch008 (https://www.igi-global.com/chapter/metrics-for-project-management-methodologies-elicitation/327531).

Presentación tesis doctorado - Concurso Raúl Gallard

Título:

Gemelos digitales híbridos para la enseñanza de robótica. Un entorno distribuido en tiempo real

Autor:

Mg. Néstor Adrián Balich

Directores:

Director: Dr. Gustavo Rossi

Co-Director: Dr. Carlos Neil

Universidad:

Universidad Abierta Interamericana

Fecha de exposición:

17 de diciembre 2025

1. Resumen

Esta tesis presenta el diseño, desarrollo, validación y evolución durante siete años de un entorno distribuido en tiempo real basado en gemelos digitales híbridos para la enseñanza de robótica y programación. El sistema integra simulación física, interacción remota multipresencia, ejecución distribuida y soporte multilenguaje, permitiendo que estudiantes de nivel primario, secundario, universitario y formación docente accedan a experiencias auténticas de programación y control robótico desde una interfaz unificada y accesible.

El aporte central consiste en la construcción de una arquitectura escalable que combina (software/hardware):

- Simulación distribuida en tiempo real.
- Sincronización multipresencia.
- Integración con hardware físico.
- Soporte para múltiples lenguajes de programación.
- Interfaz simplificada basada en Scratch/TurboWarp.
- Integración con entornos de computación en la nube (Collab).
- Incorporación progresiva de programación e interacción con IA.
- Implementación de rúbricas modernas para validación pedagógica.

El desarrollo no se limitó a un prototipo experimental: evolucionó durante siete años mediante ciclos incrementales de mejora (4 años y medio de cursada y 2 años y medio desde el comienzo de la línea de investigación), respaldado por 14 publicaciones científicas, implementación en múltiples instituciones y capacitación de más de 300 estudiantes y docentes.

El resultado es un entorno que reduce barreras de acceso a la robótica, permite validación experimental sin infraestructura costosa y posibilita experiencias de aprendizaje distribuidas, colaborativas y reproducibles.

2. Motivación y Problema Abordada

2.1. Problema de investigación preliminar

El aprendizaje de robótica en contextos educativos presenta desafíos persistentes: acceso desigual al hardware, tiempos limitados de uso en clase, mantenimiento técnico de los dispositivos, barreras de entrada para los docentes y restricciones presupuestarias. Si bien los simuladores han brindado una respuesta parcial a estos obstáculos, a menudo presentan una desconexión con el mundo físico, lo que genera experiencias de aprendizaje fragmentadas.

Por otra parte, aunque el concepto de gemelo digital ha demostrado gran efectividad en ámbitos industriales, su adopción en la educación en especial enfocado a la enseñanza de la robótica aún incipiente. Existen pocos modelos que articulen simultáneamente el control del robot real con su representación virtual en tiempo real

El modelo híbrido propuesto promueve la autonomía y la motivación intrínseca del estudiante, entendidas como factores esenciales para el aprendizaje significativo, dentro del contexto de transformación de la educación técnica argentina y del enfoque construccionista.

2.2. La pregunta que guía este trabajo es:

¿Cómo diseñar e implementar un modelo híbrido de gemelo digital que potencie el aprendizaje de robótica mediante la integración simultánea de entornos simulados y físicos?

La enseñanza de robótica y programación enfrenta múltiples desafíos:

- Costo y acceso limitado a hardware físico.
- Falta de escalabilidad en entornos presenciales.
- Brecha entre simulación académica y sistemas reales.
- Fragmentación de lenguajes y herramientas.
- Escasa integración entre simulación, evaluación y colaboración remota.

Los simuladores tradicionales presentan al menos una de estas limitaciones:

- No operan en tiempo real distribuido.
- No permiten multipresencia sincronizada.
- No integran hardware físico.
- No soportan múltiples lenguajes simultáneamente.
- No ofrecen interfaces accesibles para principiantes sin perder potencia para niveles avanzados.

La tesis parte de la hipótesis de que es posible diseñar un entorno híbrido que combine lo mejor de la simulación y el mundo físico, garantizando:

- Accesibilidad universal.
- Escalabilidad institucional.
- Rigurosidad técnica.
- Validación pedagógica.
- Evolución incremental sostenida.

2.3 Solución al Problema Identificado

Como respuesta a la problemática planteada, se desarrolla una arquitectura técnica de gemelo digital híbrido para robótica, implementada en el entorno denominado ESTER. Este entorno incorpora herramientas de control remoto, modelado virtual, y simulación distribuida "tecnologías alineadas con la visión de la Educación 4.0 y 5.0.

3. Aporte a la Disciplina

El aporte de esta tesis a la Informática y a la Tecnología Educativa puede organizarse en cinco dimensiones principales:

1. Aporte Arquitectónico

Se diseñó una arquitectura distribuida basada en:

- Separación de motor de simulación y capa de interacción.
- Sincronización en tiempo real cliente-servidor.
- Multipresencia con consistencia de estado (hasta 30 alumnos en tiempo real).
- Capacidad de ejecución paralela.

- Integración modular de dispositivos físicos.
- La arquitectura permite escalar desde una experiencia individual hasta escenarios colaborativos distribuidos geográficamente.

2. Aporte en Gemelos Digitales Híbridos

Se formaliza un modelo de gemelo digital híbrido educativo que:

- Replica comportamiento físico.
- Permite validación cruzada entre simulación y hardware real.
- Reduce riesgo experimental.
- Facilita iteración pedagógica.

No se trata solo de simular un robot, sino de mantener coherencia estructural entre modelo digital y sistema físico real.

3. Aporte Multilenguaje

El sistema permite programación en:

- Lenguajes textuales tradicionales.
- Lenguajes educativos.
- Integración con notebooks en la nube.
- Incorporación progresiva de herramientas de IA.

Esto elimina la dependencia de una única tecnología y favorece la continuidad formativa entre niveles educativos.

4. Aporte en Accesibilidad

La interfaz basada en Scratch/TurboWarp permite que:

- Estudiantes de primaria puedan interactuar.
- Personas sin experiencia previa programen.
- Docentes adopten la herramienta sin curva de aprendizaje prohibitiva.

Simultáneamente, el sistema no limita la complejidad para usuarios avanzados.

5. Aporte Empírico

El desarrollo fue validado mediante:

- 14 publicaciones científicas.
- Implementaciones reales.
- Más de 300 estudiantes y docentes capacitados.
- Iteraciones de mejora documentadas.
- Evaluaciones mediante rúbricas formales.

Se trata de una contribución respaldada por evidencia longitudinal y no por pruebas aisladas.

4. Metodología y Casos de Estudio

La investigación adoptó una metodología de diseño-desarrollo-prueba-validación incremental-evolutiva organizada en etapas:

Fase I – Prototipo Base

- Diseño conceptual.
- Desarrollo inicial del motor de simulación.
- Validación funcional básica.

Fase II – Distribución y Tiempo Real

- Implementación de sincronización remota.
- Pruebas de latencia.
- Validación multipresencia.

Fase III – Integración Física

- Conexión con hardware real.
- Validación de coherencia entre gemelo y sistema físico.
- Ajuste de parámetros dinámicos.

Fase IV – Multilenguaje y Nube

- Integración con Collab.
- Soporte para programación en múltiples lenguajes.
- Desarrollo de APIs de conexión.

Fase V – Validación Pedagógica

- Diseño e implementación de rúbricas modernas.
- Evaluación en entornos educativos reales.
- Medición de impacto formativo.



Figura 1: Evolución del desarrollo del simulador



Figura 2: Desarrollo de robot reales

4.6 Casos de estudio

Se implementaron múltiples escenarios:

- Formación inicial en programación (nivel primario).
- Introducción a robótica (secundario).
- Cursos universitarios (grado y posgrado)
- Capacitación docente.
- Trabajo colaborativo remoto.
- En todos los casos se midieron:
- Usabilidad.
- Curva de aprendizaje.
- Estabilidad del sistema.
- Participación estudiantil.
- Resultados académicos.

5. Resultados

Los principales resultados obtenidos incluyen:

- Estabilidad operativa en entornos distribuidos.
- Sincronización efectiva multipresencia.
- Reducción de costos de infraestructura.
- Aumento de participación estudiantil.
- Mejora en comprensión de conceptos de programación y control.
- Validación empírica del modelo híbrido.

Se comprobó que la combinación de simulación distribuida, accesibilidad y soporte multilenguaje amplía el alcance formativo sin sacrificar rigurosidad técnica.

6. Conclusiones y Contribución Final

La tesis demuestra que es posible construir un entorno de gemelos digitales híbridos que:

- Sea distribuido y multipresencia.
- Funcione en tiempo real.
- Integre hardware físico.
- Soporte múltiples lenguajes.
- Sea accesible a principiantes.
- Escale a niveles avanzados.
- Se valide empíricamente en contextos reales.

La contribución final no es únicamente tecnológica, sino sistémica: un modelo integral que articula arquitectura, pedagogía, validación empírica y evolución incremental sostenida.

La duración de siete años de desarrollo documentado evidencia madurez científica y tecnológica, consolidando una línea de investigación robusta.

7. Trabajo Futuro

Las líneas futuras incluyen:

- Integración más profunda con IA generativa.
- Expansión a entornos de realidad aumentada.
- Incorporación de analíticas de aprendizaje avanzadas.
- Escalamiento internacional.

- Publicación como plataforma abierta.
- Integración con estándares industriales de gemelos digitales.

8. Impacto Académico y Proyección

Las contribuciones principales pueden organizarse en tres dimensiones:

1. **Técnica**
   - Desarrollo de un sistema basado en Unity 3D con comunicación en tiempo real mediante WebSocket y API REST.
   - Sincronización fluida entre simulador y robots físicos con latencia inferior a 100 ms.
   - Integración de sensores virtuales y físicos que garantizan coherencia en la representación híbrida.
   - Arquitectura distribuida y escalable, ejecutable tanto en red local como en la nube.
2. **Pedagógica**
   - Diseño de un entorno que articula el construccionismo, el aprendizaje significativo, el alineamiento constructivo, y la educación basada en competencias.
   - Inclusión de estudiantes con distintos niveles de experiencia en programación gracias a la implementación de ATOM+, que facilita la transición de la programación visual a la textual.
   - Generación de prácticas accesibles, motivadoras y contextualizadas en problemas reales, que fortalecen competencias técnicas, cognitivas y actitudinales.
3. **Científica**
   - Aporte al campo emergente de los gemelos digitales en educación, un área donde la evidencia empírica todavía es incipiente y fragmentada .
   - Generación de producción académica validada en congresos y revistas indexadas, que documenta la aplicabilidad del modelo en contextos diversos.
   - Propuesta replicable y adaptable a otras disciplinas STEM y a contextos no técnicos, lo que amplía el alcance del paradigma de los gemelos digitales más allá de la ingeniería.

La presente tesis demuestra que los gemelos digitales híbridos pueden superar la brecha entre la simulación virtual y la práctica física, constituyéndose en un recurso pedagógico transformador. El entorno ESTER no solo aporta un avance técnico en términos de sincronización y multipresencia, sino que también ofrece un marco pedagógico robusto y validado, lo que lo convierte en una referencia para el diseño de futuros ecosistemas híbridos en educación.

8.1 Producción previa:

- Cloud Robotic – Revista Abierta de informática e Investigación Aplicada (2017) – Universidad Abierta Interamericana.
- Interfaz Natural Emotiva Niño-Robot – Workshop de Investigadores en Ciencias de la Computación (2014).
- Construcción de robots autónomos colaborativos – Tesis maestría en sistema informáticos, Universidad Abierta interamericana (2009).

8.2 Reconocimientos:

- AWS Cloud Credit for Research (2023): subsidio de USD 11.000 otorgado a ESTER para infraestructura cloud.

8.3 Presentaciones en congresos y revistas

Nro.	Año	Título / Trabajo	Evento / Publicación	Observaciones / Reconocimientos
1	2025	Sincronización multi-robot en tiempo real para el simulador ESTER y gemelo digital	Congreso Argentino de Ciencias de la Computación	Presenta sincronización multi-robot y bidireccionalidad en tiempo real.
2	2025	ATOM+: Integration of Visual Programming and Digital Twins in the ESTER Simulator	Jornada Argentina de Informática e Investigación Operativa Jornada Argentina de Informática e Investigación Operativa	Introduce ATOM+ como herramienta visual integrada al modelo híbrido.
3	2025	Implementación de Robot Físico con Gemelo Digital para el Simulador Ester	Congreso de Tecnología en Educación y Educación en Tecnología	Creación de robot físico para el simulador Ester con gemelos digitales.
4	2025	Gemelos Digitales en la Educación en Robótica: Un Mapeo Sistemático de la Literatura	Congreso Nacional de Ingeniería Informática / Sistemas de Información	Mapeo sistemático de la literatura con eje en gemelos digitales en educación.
5	2024	Simulador de robótica 3D multirobot para enseñanza de programación	Congreso Nacional de Ingeniería Informática / Sistemas de Información	Validación con estudiantes de grado; entorno multiusuario.
6	2024	Desarrollo de simulador de robótica cloud multipresencia para enseñanza de programación	Congreso Argentino de Ciencias de la Computación	Extiende ESTER al ámbito cloud; multipresencia y escalabilidad.
7	2024	Exposición del trabajo	Congreso Argentino de Ciencias de la Computación – Workshop Innovación Educativa	Premio a la presentación destacada.

8	2024	Dictado Materia programación robótica e impresión 3D, de la Maestría en Tecnología Educativa	Universidad Abierta Interamericana	Premio a Buenas Prácticas Educativas.
9	2023	Simulador de robótica en la nube para prácticas colaborativas	Congreso Nacional de Ingeniería Informática / Sistemas de Información	Extensión cloud y prácticas multiusuario.
10	2023	Simulador 3D de robótica distribuido en Unity para enseñanza de programación e IA	Revista Abierta de Informática Aplicada	Publicación científica indexada.
11	2023	Taller docente sobre simulación con ESTER	Congreso Nacional de Ingeniería Informática / Sistemas de Información	Transferencia didáctica y capacitación docente.
12	2017	Cloud Robotic	Revista Abierta de Informática Aplicada	Primer acercamiento al concepto de robótica en la nube.
13	2014	Interface Natural Emotiva Niño-Robot	Workshop de Investigadores en Ciencias de la Computación	Investigación en interacción humano-robot.
14	2009	Construcción de robots autónomos colaborativos	Tesis de Maestría – Universidad Abierta Interamericana	Antecedente directo de la línea de investigación en robótica educativa.

Tabla 1: presentaciones en congresos y revistas

9. Síntesis Final

Esta tesis consolida siete años de investigación aplicada, integrando teoría, arquitectura, implementación y validación pedagógica en un entorno distribuido de gemelos digitales híbridos para la enseñanza de robótica.

La combinación de:

- Desarrollo incremental sostenido.
- 14 publicaciones científicas.
- Más de 300 personas capacitadas.
- Evolución funcional documentada.
- Validación formal mediante rúbricas.
- Escalabilidad técnica.
- Accesibilidad universal.

Configura una contribución sólida, original y de alto impacto para la Informática argentina.

MARCO DE TRABAJO PARA LA EVALUACIÓN DE EXPERIENCIA DE USUARIO EN SISTEMAS INTERACTIVOS

Autora: Laura Nidia Aballay

Director: Dra. Lorena Bearzotti (Pontificia Universidad Católica de Valparaíso, Chile)

Codirector: Dr. Fernando Pinciroli (Universidad Nacional de San Juan, Argentina)

Doctorado en Ciencias de la Informática

Universidad Nacional de San Juan – Facultad de Ciencias Exactas, Físicas y Naturales

Fecha de exposición: 12 de noviembre de 2025.

Introducción

La presente tesis doctoral aborda la problemática de la evaluación de la Experiencia de Usuario (UX) en sistemas interactivos desde una perspectiva integral que incorpora explícitamente la dimensión emocional. Tradicionalmente, la evaluación de la UX se ha centrado en métricas funcionales tales como usabilidad, eficiencia, efectividad y satisfacción. Sin embargo, la literatura reciente en Interacción Humano-Computador (HCI) y Computación Afectiva demuestra que las emociones influyen significativamente en la percepción, adopción, retención y compromiso del usuario frente a la tecnología.

La motivación central de esta investigación surge de la ausencia de un procedimiento metodológico unificado y adaptable que permita integrar de manera sistemática la evaluación emocional en los procesos de análisis de UX. Si bien existen herramientas aisladas —cuestionarios afectivos, métricas fisiológicas o modelos conceptuales— no se dispone de un marco estructurado que articule estos componentes dentro de un proceso metodológico completo, replicable y validado empíricamente.

El objetivo general consistió en desarrollar y validar un marco de trabajo para la evaluación de la UX con enfoque en emociones, integrando métodos multimodales de recolección de datos de manera ética, simple, escalable y adaptable a distintos contextos interactivos.

Se planteó como hipótesis que es posible crear un marco de evaluación de la UX que incorpore las emociones de los usuarios, utilizando diferentes métodos de análisis y que pueda aplicarse de manera ética y adaptable en distintos tipos de sistemas interactivos.

La investigación adoptó un enfoque metodológico mixto, combinando revisión sistemática de literatura, diseño conceptual, desarrollo instrumental y validación empírica mediante estudio de caso. En una primera fase se realizó una revisión sistemática de publicaciones indexadas en bases como

IEEE, ACM y Springer, identificando métodos de evaluación emocional y marcos existentes. En una segunda fase se diseñó el marco EMO-UX, definiendo sus principios, estructura y componentes. En una tercera fase se desarrolló un cuestionario emocional contextualizado y validado por expertos. Finalmente, se implementó y validó el marco mediante un estudio piloto en el Campus Virtual del Sistema Institucional de Educación a Distancia (SIED) de la Universidad Nacional de San Juan.

Los resultados de la implementación evidenciaron que el instrumento diseñado obtuvo un coeficiente Alfa de Cronbach de 0.87, indicando alta consistencia interna entre dimensiones evaluadas y capacidad del marco para identificar patrones emocionales diferenciados según características del sistema. Se comprobó que la incorporación estructurada de métricas emocionales complementa las métricas tradicionales de usabilidad y permite detectar oportunidades de mejora no visibles mediante indicadores funcionales clásicos.

Justificación

En los últimos años, las emociones han adquirido relevancia como un aspecto fundamental en el diseño de experiencias digitales [1], [2]. Este marco se fundamenta en el principio del Diseño Centrado en el Usuario (DCU), que coloca las necesidades, expectativas y experiencias de los usuarios en el centro del proceso de diseño y evaluación [3]. Este enfoque permite que las plataformas se ajusten no solo a requisitos técnicos, sino también a las respuestas emocionales de los usuarios, promoviendo empatía en el diseño y mejorando la calidad de la interacción [4].

La consideración de las emociones permite identificar y mitigar factores que generan frustración o ansiedad, y potenciar elementos que producen satisfacción y bienestar [5], [6]. Diversos estudios demuestran que las emociones influyen en la percepción de valor, la fidelización y la satisfacción general del usuario con productos digitales [7]. En consecuencia, el marco EMO-UX integra la evaluación emocional como componente central de la UX para captar de manera más integral la interacción usuario-sistema.

Aporte de la Tesis

La tesis realiza un aporte metodológico original al campo de la Ingeniería de Software y la Interacción Humano-Computador, al proponer un marco formalmente estructurado y empíricamente validado que integra evaluación emocional y métricas tradicionales de calidad de software.

El marco EMO-UX se estructura en cinco fases iterativas: planificación, diseño del instrumento emocional, implementación de la evaluación, recopilación y análisis de resultados, y refinamiento. Integra métricas específicas tales como intensidad emocional promedio, frecuencia emocional, diversidad emocional y tasa de compensación emocional. Asimismo, contempla la posibilidad de integración multimodal futura (biometría, reconocimiento facial, EEG).

El marco EMO-UX aporta una herramienta aplicable en procesos de desarrollo de software, fortaleciendo prácticas de diseño centrado en el usuario en contextos académicos y productivos, con potencial transferencia tecnológica.

Alcance

El alcance de este marco de trabajo se centra en la evaluación de la UX en sistemas interactivos. Está diseñado para ser aplicado en diversos tipos de plataformas, como sitios web, aplicaciones móviles, sistemas de e-commerce, videojuegos, redes sociales y otros sistemas interactivos. Es adaptable a diferentes contextos y puede emplearse en escenarios tanto controlados (laboratorios) como en ambientes reales.

Su aplicación está orientada a la etapa de evaluación en el ciclo de vida de desarrollo de sistemas, brindando una metodología para analizar cómo las emociones de los usuarios impactan en su experiencia general, más allá de los aspectos funcionales y de usabilidad convencionales.

Está diseñado para ser implementado una vez que el sistema ha alcanzado un nivel de desarrollo suficientemente avanzado, como para ser usado y evaluado en condiciones cercanas a las reales, idealmente antes de su despliegue completo o durante fases piloto de prueba en usuarios.

El marco EMO-UX fue concebido con la intención de ofrecer una herramienta flexible y adaptable para la evaluación de la experiencia emocional del usuario en diversos entornos interactivos. Si bien los resultados obtenidos en este estudio han sido positivos y prometedores en el contexto específico del SIED de la Universidad Nacional de San Juan, es necesario reconocer que aún no se cuenta con evidencia empírica directa que confirme su aplicabilidad generalizada en otros contextos tecnológicos, institucionales o culturales.

A diferencia de marcos centrados exclusivamente en métricas cuantitativas o heurísticas, EMO-UX integra evaluación emocional estructurada con métricas validadas, aportando una articulación metodológica inexistente en modelos previos.

Futuras aplicaciones

Para avanzar en la validación externa y consolidar la afirmación de adaptabilidad, se propone desarrollar las siguientes líneas de investigación:

- Estudios de replicación en otras plataformas educativas virtuales (por ejemplo, Google Classroom, Blackboard) para verificar si las secciones y emociones evaluadas mantienen su coherencia y relevancia.
- Aplicación del marco en entornos no educativos, como sistemas de banca online, plataformas de salud digital o portales administrativos, evaluando la pertinencia de cada sección y adaptando la redacción de ítems si fuese necesario.
- Traducción y adaptación al idioma inglés, permitiendo su aplicación en un contexto internacional y facilitando su uso en investigaciones y evaluaciones de UX en una mayor diversidad de sistemas y culturas. Esta expansión no solo aumentaría su accesibilidad para investigadores y profesionales de UX en distintos países, sino que también posibilitaría estudios comparativos sobre la influencia de factores culturales en la experiencia emocional

del usuario. Además, al integrarse con marcos y estándares globales de evaluación de UX, EMO-UX podría consolidarse como una herramienta clave en la medición de emociones en contextos digitales, fomentando su adopción en proyectos de diseño centrado en el usuario a nivel mundial.

- Incorporar tecnologías emergentes como la integración de sensores y tecnologías biométricas que permitan capturar información emocional más allá del autorreporte. El uso de cámaras para reconocimiento facial de expresiones, sensores de ritmo cardíaco o dispositivos portables (wearables) facilitaría registrar respuestas fisiológicas asociadas a emociones durante el uso real de las plataformas.
- También se contempla la incorporación del análisis de sentimientos aplicado a comentarios y observaciones textuales de los usuarios. Esta línea complementaria permitiría captar aspectos cualitativos que pueden pasar desapercibidos en una escala estructurada, y contribuir a enriquecer el diagnóstico emocional con matices lingüísticos propios del contexto del usuario.

En síntesis, esta tesis contribuye a consolidar una evaluación UX más humana, integral y científicamente fundamentada, ampliando el alcance metodológico de la disciplina hacia la incorporación sistemática de la dimensión afectiva.

Referencias

[1] D. Norman, Emotional Design: Why We Love (or Hate) Everyday Things. New York, NY, USA: Basic Books, 2004.

[2] R. W. Picard, Affective Computing. Cambridge, MA, USA: MIT Press, 1997.

[3] ISO 9241-210:2019, Ergonomics of human-system interaction – Human-centred design for interactive systems.

[4] ISO/IEC 25010:2011, Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE).

[5] M. Hassenzahl, Experience Design: Technology for All the Right Reasons. Morgan & Claypool, 2010.

[6] J. A. Bargas-Avila and K. Hornbæk, "Old wine in new bottles or novel challenges? A critical analysis of empirical studies of user experience," Proc. CHI, 2011.

[7] P. Desmet, "Measuring emotion: Development and application of an instrument to measure emotional responses to products," in Funology, 2003.

Metodología integral de recomendación de revisores en revistas científicas basada en múltiples factores

Tesista doctoral: Dr. Guillermo Decoppet[1] - Universidad Abierta Interamericana
Directora de tesis: Dra. Patricia San Martín[2] (CONICET-UNR) - Codirector: Dr. Carlos Neil[3] (UAI)
Calificación: 10 (diez, sobresaliente) - Fecha de defensa: 31/10/2025

Motivación: El problema de selección de expertos para la Revisión por Pares

La revisión por pares constituye el pilar fundamental sobre el cual se asienta la validación del conocimiento científico contemporáneo. Para las revistas científicas es un desafío y una tarea muy compleja elegir los mejores revisores a sabiendas de que la calidad de la revisión tiene un impacto directo en su prestigio.

El crecimiento exponencial en la cantidad de manuscritos enviados a revistas científicas y eventos académicos, sumado a un requerimiento de especialización del conocimiento cada vez más diversificado, da cuenta de una serie de problemáticas críticas de esta labor.

En el contexto particular de las revistas científicas de la región, y específicamente en Argentina, la gestión editorial suele enfrentarse a una carencia de herramientas tecnológicas sofisticadas. Un número significativo de editores continúan realizando la búsqueda y selección de revisores de manera "artesanal", basándose en redes de contactos personales o búsquedas manuales en bases de datos. Este enfoque no solo es ineficiente en términos de tiempo, sino que también limita la diversidad de evaluadores y puede comprometer el prestigio de la revista si la calidad de la revisión no es adecuada.

Ante este escenario, la Inteligencia Artificial (AI), y más precisamente el Procesamiento de Lenguaje Natural (PLN), tiene un gran potencial en la búsqueda de soluciones a problemas de este tipo.

Objetivos

General: Proponer una metodología integral que utilice IA para el PLN a los fines de contribuir y apoyar la selección de revisores para artículos científicos incorporando una significativa multiplicidad de factores que contemplan los comités editoriales en dicha tarea.

Específicos

- Relevar analíticamente las variables intervinientes en la selección de revisores de artículos científicos definiendo cuáles de estas son tratadas en los enfoques metodológicos de las herramientas automatizadas de recomendación en la temática.
- Efectuar un reestudio del PLN con la finalidad de seleccionar los enfoques metodológicos más adecuados a la problemática de recomendación de revisores científicos.
- Elaborar una metodología configurando variables múltiples aplicadas al caso de revisores para artículos científicos en el campo de las Ciencias de la Educación a nivel de primer prototipo.
- Evaluar el primer prototipo de la metodología de recomendación propuesta con miembros de comités editoriales de revistas científicas del campo de las Ciencias de la Educación.

Metodología de la investigación

La tesis doctoral adoptó un enfoque propositivo. El desarrollo teórico y metodológico de la investigación se basa en un mapeo sistemático bibliométrico y la elección de un caso contextualizado. En el trayecto de la tesis doctoral se han implementado instrumentos y técnicas cualitativas, adhiriendo a un marco

[1] decoppet@irice-conicet.gov.ar
[2] sanmartin@irice-conicet.gov.ar
[3] Carlos.Nail@uai.edu.ar

sociotécnico cultural. La elaboración del prototipo se basó en un enfoque iterativo, en el que se desarrollaron y probaron componentes de manera progresiva, corrigiendo y ajustando continuamente cada parte antes de integrar el sistema completo. Además, se realizaron pruebas integrales que permitieron identificar nuevas opciones y mejoras, lo que favoreció la evolución constante del prototipo.

El caso de referencia que tomó la presente investigación es la Revista IRICE[4] publicada por el Instituto Rosario en Ciencias de la Educación (IRICE), dependiente del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y la Universidad Nacional de Rosario (UNR). Esta Revista cuenta con más de 25 años de trayectoria y la búsqueda de revisores la realiza el consejo editorial sin ningún tipo de herramientas de apoyo para la toma de decisiones. Esta situación se replica en las revistas que conforman la "Red de Revistas Académicas en Educación de la Argentina"[5]. Por lo cual se estima que el caso es modélico y en consultas a la mencionada Red se ha detectado un alto interés en los posibles resultados de esta investigación en la comunidad de editores científicos de la disciplina.

El trabajo de campo con editores se desarrolló en dos etapas. En la primera, para la construcción de los fundamentos teóricos y la metodología propuesta, se efectivizaron entrevistas y encuestas semiestructuradas a investigadores con experiencia en comités editoriales. La segunda etapa consistió en la evaluación y ajustes del prototipo recomendador, en la cual se realizaron consultas y evaluaciones contextualizadas a dichos expertos. Cabe mencionar que todos estos insumos fueron fundamentales en el desarrollo socio-técnico cultural aquí propuesto.

La estructura de la tesis consiste en una introducción general, luego el primer capítulo aborda un mapeo sistemático del estado del conocimiento. El segundo capítulo desarrolla los fundamentos teóricos-metodológicos que sustentan los aspectos socio-técnicos culturales de la metodología innovadora propuesta. La misma se desarrolla en el tercer capítulo. En el cuarto, que integra la segunda etapa del trabajo de campo, se presenta la aplicación de la metodología en un prototipo y se detallan los testeos y evaluaciones de laboratorio. Finalmente, se abordan las conclusiones, posibles líneas de trabajo a futuro, la bibliografía referenciada y los anexos documentales.

Los avances de la tesis fueron publicados en:

Decoppet, G. O., & San Martin, P. S. (2024). Systematic mapping of automated reviewer recommendation solutions. *Journal of Computer Science and Technology*, *24*(2), e16. https://doi.org/10.24215/16666038.24.e16

Decoppet, G. O., & San Martin, P. S. (2025). Hacia una metodología flexible al contexto para la recomendación automatizada de revisores. *E-Ciencias De La Información*, *15*(1). https://doi.org/10.15517/eci.v15i1.62426

Mapeo sistemático

Se realizó un mapeo sistemático de la bibliografía en donde se seleccionaron 524 publicaciones y en base a los criterios de inclusión y exclusión se terminaron seleccionando 25 artículos científicos que trataban la temática dentro de los años 2018 y 2023.

En base a las preguntas de investigación se puede destacar:

- En cuanto a factores detectados e intervinientes en la temática se encontraron los siguientes, ordenados en forma descendente por inclusión: "Experiencia" (en todos los artículos), "Conflicto de intereses", "Autoridad académica", "Interés del investigador", "Diversidad", "*Seniority*" (asociado al anterior) y "Revisiones anteriores" que fue utilizada en solo un artículo.

[4] Sitio oficial Revista IRICE: https://ojs.rosario- conicet.gov.ar/index.php/revistairice/

[5] RED de RED: https://saieinvestigacionedu.wixsite.com/saieargentina/reddered

- En cuanto a modelos de PLN, hubo una diversidad de técnicas utilizadas. En mayor medida modelos de temas y luego de espacio vectorial, llamando la atención de que la mayoría de las técnicas no incluyen información semántica. Pocas soluciones con información semántica (Word2vec, Doc2Vec, Bert). Otras soluciones con redes convolucionales, diferentes redes neuronales recurrentes y *prompt tuning*. Solo hubo dos casos sin técnicas.
- Gran diversidad de integración de factores, pero nulos o muy pocos parámetros de adaptabilidad y contextualización en las fórmulas. Generalmente fórmulas cerradas.
- Representación del artículo: lo más usado fue una combinación de títulos y resúmenes.
- Fuentes de datos: lo más utilizado, y no en general, fue *Digital Bibliography & Library Project* (DBLP) que es especializada en Cs. de la Computación, pero cada modelo usaba sus propias fuentes de datos según el ámbito de interés que motivó la investigación.

Fundamentos teórico-metodológicos

Del análisis del mapeo sistemático, bajo una perspectiva socio-técnica y cultural, y tomando los resultados de la encuesta al equipo editorial de la revista IRICE y de otro evaluador sénior del campo de las Ciencias Exactas, se identifican los siguientes aspectos clave:

- **Rol Asistivo de la IA:** La recomendación automática es una información de referencia. Siendo la decisión final responsabilidad del Comité Editorial (CE).
- **Flexibilidad y consenso:** debe ser adaptable a las políticas y prácticas de cada CE y campo disciplinar, incorporando un diálogo colaborativo para consensuar criterios y evitar sesgos.
- **Parametrización:** debe permitir la elección y ponderación de factores según los consensos del CE.

Desarrollo de la metodología de recomendación de revisores: Marco general

La **estructura principal** consta de **tres globalizadores de factores** para una visión holística, a saber:

1 – Producción Científica: tiene en cuenta principalmente las publicaciones científicas de los candidatos a revisores. Dentro de ella, se integran diferentes factores:

- Experiencia (Factor Disparador): Evalúa el grado de conocimiento del candidato a revisor en relación con el manuscrito y es de especial interés hacer las siguientes definiciones:
    - Representación del revisor: Mediante modelos de espacio vectorial, se extraen las características globales del texto representativo para generar un vector único por cada artículo. Este enfoque garantiza una mayor fidelidad representativa frente al modelado por temas, cuya abstracción suele omitir matices críticos del contenido y supera a las técnicas basadas en citas, las cuales carecen de la capacidad para resolver la similitud profunda entre los textos.
    - Secciones representativas de cada artículo: sobre cuál representación del artículo se hace la extracción de características.
    - Técnicas de PLN: teniendo en cuenta los modelos del espacio vectorial, las diferentes técnicas posibles para su representación.
    - Idiomas: No todas las técnicas de PLN trabajan en múltiples idiomas; se debe definir un idioma base de trabajo.
    - Desambiguación de nombres: Es el proceso que consiste en distinguir entre dos o más autores que tienen el mismo nombre o determinar si diferentes variantes de un nombre pertenecen a la misma persona física. Debido a que el campo nombres y apellidos varía generalmente entre distintas revistas, no puede tomarse como un campo clave. Requiere de técnicas de desambiguación.
    - Umbral de conocimiento específico: Pone un límite a la selección de revisores basado en un valor mínimo de representación de conocimiento asociado al manuscrito a revisar. Todo lo que no llegue al umbral es descartado.

- Interés del Investigador (Antigüedad): Pondera publicaciones recientes sobre antiguas, asumiendo que reflejan intereses actuales del investigador.
- Aporte de los Autores (Orden): Busca conocer la contribución individual según el orden de aparición de los autores en un artículo científico (considera tendencias disciplinares).
- Conflicto de Intereses (que afectan la objetividad): Detección de coautoría y misma afiliación son las más utilizadas en metodologías automáticas. Los otros requieren carga manual.

2 – Autoridad académica: Tiene en cuenta el prestigio del investigador, otorgando mayor ventaja comparativa basado en:

- Índices bibliométricos: miden el impacto de la producción científica (p. ej., citas, índice h).
- Titulaciones: Nivel más alto de estudio alcanzado por un investigador.

3- Revisiones anteriores: Considera el desempeño previo del revisor para una distribución equitativa de la carga de trabajo. Tiene en cuenta:

- Controles: de revisiones simultáneas, de tiempo de descanso, de revisiones por período.
- Penalizaciones: por cantidad de rechazos y por deficiencias en la labor calificada de revisión.
- Puntaje: basado en las calificaciones de revisiones anteriores.

Flexibilidad de la fórmula integral: a través de los siguientes puntos a considerar:

- Adaptabilidad plena: la metodología permite la modificación total de la fórmula integral.
- Configuración Personalizada: se puede considerar la elección de factores a utlizar, la definición de opciones internas para cada factor y la disponibilidad de parámetros (valores y operadores).
- Múltiples perfiles de configuración: Posibilidad de definir distintos perfiles de búsqueda (ej., iniciar con configuración estricta e ir flexibilizando).

Construcción de metodología integral

1 - Adquisición, preprocesamiento y guardado de datos: proceso fundamental que requiere cantidad y calidad de datos para llevar adelante la metodología automática propuesta. Sus componentes son:

- Identificación y elección de la fuente de datos: es clave elegir la fuente de datos apropiada según la temática y la revista científica vinculada para poder contar con los datos adecuados.
- Extracción de datos de las fuentes de información: según el tipo de acceso que se tenga a los datos, se utilizarán diferentes técnicas de extracción. Pudiendo ser entre otras posibilidades: contar con acceso directo a la base de datos, acceso a través de APIs, acceso por Web Scraping.
- Preprocesamiento: se generan los *pipelines* de preprocesamiento para las diferentes partes representativas o el texto completo. Es una tarea específica a cada tipo de dato y da como resultado la depuración de los datos para poder ser usados en las siguientes etapas de la metodología.
- Guardado de datos: se evaluará según el tipo de acceso la conveniencia o no del guardado y, en caso de conveniencia, se determinará cuál es el formato adecuado.

2- Globalizador de Producción científica: Datos y cálculos necesarios para aplicar la metodología.

En experiencia:

La elección de la parte representativa del artículo se guarda en el **Parámetro de parte representativa del artículo científico: PPR_{ac}:** la metodología ofrece 6 opciones: Títulos, Palabras clave, Resúmenes, Resúmenes y Títulos, Resúmenes, títulos y palabras claves, Contenido completo del artículo desde introducción o similar y hasta bibliografía sin incluir.

La elección de la técnica que se va a utilizar para la generación de vectores representativos sobre técnicas de PLN se guarda en **Parámetro de técnica de PLN: P_{tpln}**: donde se ingresan algunas técnicas. Como ejemplos de aquellas que se pueden considerar:

- Técnicas de modelos de vectores dispersos: TF-IDF, BM-25
- Técnicas de Word Embeddings: Word2Vec, Glove, fastText
- Técnicas de Document Embeddings: Doc2Vec
- Técnicas de transformers: BERT, GPT

La elección de la técnica de cálculo de similitud se guarda en el **Parámetro de técnica de similitud: (P_{ts})**: se utiliza similitud de coseno porque es la mejor técnica para problemas de este tipo para encontrar la similitud con independencia del tamaño del texto, focalizándose en la orientación y no la magnitud. Sirve tanto para vectores con información semántica como en vectores dispersos y el ordenamiento es descendente por similitud.

La aplicación de punto de corte se guarda en el **Parámetro de punto de corte: P_{pc}**: este valor define la cantidad de artículos priorizados que se tendrán en cuenta para la búsqueda de revisores.

Desambiguación de nombres: si no se cuenta con datos codificados de autores, es necesario desambiguar. Se propone un método de contrastación en bases científicas de perfiles de investigadores y bases de publicaciones, sobre coincidencia de artículos como primera medida. Al identificarse el perfil con publicaciones coincidentes en ambas bases, se puede determinar el autor y así efectuar la desambiguación.

La definición de tipo de umbral de conocimiento a utilizar se guarda en el **Parámetro de cálculo de umbral: Pc_{us}.** Este proceso puede ser **automático** realizando el cálculo basado en la similitud del artículo priorizado con menor puntaje o **manual** ingresando un valor correspondiente en el **Parámetro de cálculo manual de umbral: PCM_{us}**. Se recomienda que el valor ingresado sea menor a 1 y mayor a 0.

C**álculo de puntaje de experiencia genérica** para un candidato a revisor: Del primer artículo se toma el valor de similitud y en caso de tener más de uno como priorizados, de los siguientes se toma el valor de similitud menos el valor de umbral. Luego se pueden realizar tres tipos de cálculos, según opción de parametrización de la metodología en el **Parámetro de tipo de cálculo de similitud: PTC_s**

- Suma simple: SS = $\sum_{i=1}^{n} V_{d\ (i)}$
- Promedio total: PT = $\frac{\sum_{i=1}^{n}(\sum_{x=1}^{i} V_{d\ (x)})}{n}$
- Promedio ponderado: $\begin{cases} P_{(1)} = V_{d\ (1)} \\ P_{(i)} = \frac{(2(P_{(i-1)})+V_{d\ (i)})}{2} \end{cases}$ para i > 1

<u>En interés del investigador:</u>

Se utiliza el **Parámetro de cálculo de antigüedad: Pc_a** para determinar si se considera (0 – no se realiza / 1 – se realiza). En caso afirmativo se guarda el valor en el **Parámetro de antigüedad: P_a**

Procedimiento de cálculo: Se toma el año de publicación, se lo compara con el año actual. Se utiliza **tabla de cálculo de antigüedad T_{ca}** para buscar **P_a** en base al rango de años de antigüedad. Además, puede ser utilizado como una restricción en el caso de que un rango se consigne con puntaje cero.

<u>En Aportes de los autores:</u>

Se usa el **Parámetro de cálculo del aporte de los autores: Pc_{au}** para determinar si se considera (0 – no se realiza / 1 – se realiza). En caso afirmativo se guarda el valor en el **Parámetro de aporte de los autores: P_{au}**

Procedimiento del cálculo: Se toma el orden de aparición del autor. Se utiliza **tabla de cálculo de aporte de los autores T_{cau}** para buscar **P_{au}** en base a rango de ordenamiento de autores.

<u>En Control de conflicto de intereses:</u>

Se realiza el control sin excepción a todos los candidatos a revisores en base a:

- **Verificación de coautoría en tres niveles:** en nivel 1 se descartan los autores directos, luego se informan los coautores en nivel 2 y los coautores de coautores en nivel 3.
- **Verificación de afiliación:** se informa coincidencia en Instituto de Investigación y/o Universidad

Cálculo del valor del globalizador de producción científica:

Sobre el Cálculo de **Puntaje de producción científica P_{pc}** según sea el tipo de cálculo en **PTC_s**:

- Suma simple: SS = $\sum_{i=1}^{n} V_{d\ (i)} * P_{a(i)} * P_{a\ (i)}$
- Promedio total: PT = $\frac{\sum_{i=1}^{n}(\sum_{x=1}^{n} V_{d\ (x)} * P_{a(i)} * P_{a\ (i)})}{n}$
- Promedio ponderado: $\begin{cases} P_{\ (1)} = V_{d\ (1)} * P_{a(1)} * P_{a\ (1)} \\ P_{\ (i)} = \frac{(2(P_{\ (i-1)})+V_{d\ (i)} * P_{a(i)} * P_{a\ (i)})}{2} \end{cases}$ para i >1

3 - Globalizador de Autoridad Académica: Datos y cálculos necesarios para aplicar la metodología.

Se puede decidir la inclusión o exclusión de este globalizador con el **Parámetro de inclusión de autoridad académica: Pi_{aa}** (S-Habilita/N-Deshabilita)

En el caso de utilizarse se pueden utilizar los diferentes contextos de cálculo:

En Índices bibliométricos:

Se puede decidir si utilizarlo o no con el **Parámetro de cálculo sobre índices bibliométricos: PC_{ib}** (S – Habilita/ N – No habilita)

En el caso de habilitarlo, se utiliza el **Parámetro de tipo de índice bibliométrico: Pt_{ib}** donde se define qué tipo de índice se va a utilizar, como por ejemplo Citas totales, h-index, i10-index. El valor a considerar en la fórmula se guarda en el **Parámetro del índice bibliométrico: P_{ib}**

Procedimiento de cálculo: Se toma el valor del índice seleccionado. Se utiliza la **tabla de cálculo bibliométrico T_{cb}** para buscar **P_{ib}** dentro del rango de valores de dicho índice.

Se puede definir un valor por defecto del **Parámetro del índice bibliométrico VDP_{ib}** si no se tiene datos.

En Titulaciones:

Se puede decidir si utilizarlo o no con el **Parámetro de cálculo sobre titulaciones: PC_t** (S – Habilita/ N – No habilita). El valor a considerar en la fórmula se guarda en el **Parámetro de Título: P_t**

Procedimiento de cálculo: Se utiliza **tabla de cálculo de título T_{ct}** para buscar **P_t** en base al título.

Se puede definir un valor por defecto del **Parámetro de título VDP_t** en caso de no tener datos.

La **tabla de cálculo de título T_{ct}** tiene un orden jerárquico entre títulos que posibilita además poder utilizarla como filtrado de revisores. Por ejemplo, filtrar solo a los revisores con título de Doctores.

Cálculo del valor del globalizador de Autoridad Académica:

Para el **cálculo de Puntaje de autoridad académica** se utiliza el **Parámetro de cálculo entre índices bibliométricos y titulación: Pc_{ibt}**, en el caso de que estén habilitados ambos cálculos, se define cómo se van a integrar en un puntaje final. Pudiendo ser: suma, promedio o producto.

El resultado se guarda en el **Puntaje de autoridad académica: P_{aa}** en caso de corresponder.

4 - Globalizador de Revisiones anteriores: Datos y cálculos necesarios para aplicar la metodología.

Se puede decidir la inclusión o exclusión de este globalizador con el **Parámetro de inclusión de revisiones: Pi_r** (S-Habilita/N-Deshabilita)

En el caso de habilitarse, se debe realizar la registración de todos los eventos importantes que intervienen en los cálculos, en la **tabla de revisiones T_r**

Otra base interviniente es la **tabla de calificaciones T_c** acordada por consenso por todo el equipo editorial. La misma se utiliza para calificar la consistencia de las revisiones y definir el coeficiente a utilizar en cálculos del globalizador (en caso de estar habilitados estos cálculos).

La operatoria se divide en:

Cortes:

- Controles temporales y tiempos de no consideración, en caso de control de:
  - Pedidos simultáneos, se utiliza el **Parámetro de revisiones sobre pedidos simultáneos: P_{rps}** (0-no se aplica, 1 o más pedidos).
  - Tiempo de descanso entre revisiones, se utiliza el **Parámetro de revisiones sobre tiempo de descanso: P_{rtd}** (0-no se aplica, 1 o más meses).
  - Revisiones por período, se utiliza el **Parámetro de revisiones sobre pedidos por período: P_{rpp}** (0-no se aplica, 1 o más revisiones).

Penalizaciones:

- Por rechazo de revisiones: Se puede definir si utilizarlo o no en base al **Parámetro de habilitación de revisiones sobre penalización de rechazadas: PH_{rpr}** (S – Habilita/ N – No habilita).
  Para las penalizaciones se usa el **Parámetro de revisiones sobre penalización de rechazadas: P_{rpr}**
  Procedimiento de cálculo: se buscan y suma la cantidad de rechazos. Se utiliza **tabla de cálculo para rechazos T_{cr}** para buscar **P_{rpr}** en base al rango de valores y cantidad de rechazos. Se aplica sobre la fecha del último rechazo. Permite penalizar por cierta cantidad de tiempo en meses o por tiempo indeterminado.
- Por revisiones de baja calidad: Se puede definir si utilizarlo o no en base al **Parámetro de habilitación de control de revisiones de mala calidad: PH_{rmc}** (S – Habilita/ N – No habilita). Para las penalizaciones se usa el **Parámetro de revisiones sobre penalización por mala calidad: P_{rpmc}**
  Procedimiento de cálculo: se buscan y suman las veces que recibió malas calificaciones. Se utiliza **tabla de cálculo para malas calificaciones T_{cmc}** para buscar **P_{rpmc}** en base al rango de valores de cantidad de malas calificaciones. Se aplica sobre la fecha de la última mala calificación. Permite penalizar por cierta cantidad de tiempo en meses o por tiempo indeterminado

Puntaje:

En base a las calificaciones de la consistencia de las revisiones anteriores, se establece el **Puntaje de revisiones anteriores: Pa_{ra}** y en el caso de que un candidato a revisor tenga más de una revisión se establece el **Parámetro de cálculo entre valores del coeficiente de evaluación de revisiones: PC_{cer}**. Pudiendo ser: Suma, Promedio y Producto. En caso de que un candidato no tenga revisiones anteriores, se puede configurar el puntaje de revisiones por defecto.

Definición de cortes y/o cálculos:

La metodología permite elegir qué tareas se van a llevar a cabo en este globalizador con el uso del **Parámetro de cálculo de revisiones con cálculo y cortes: PC_{rcc}** siendo las opciones: cálculo-cortes, cálculo o cortes

5 - Fórmula de la metodología integral: Datos y cálculos necesarios para aplicar la metodología.

Se utilizan dos parámetros integrales:

Parámetro Integral de peso: para cada uno de los globalizadores.

- **P.I. de peso de producción científica PIP_{pc}**
- **P.I. de peso de autoridad académica PIP_{aa}**
- **P.I. de peso de revisiones anteriores PIP_{ra}**

La suma total de los tres pesos integrales debe ser igual a 100. $PIP_{pc} + PIP_{aa} + PIP_{ra} = 100$

Parámetro integral de peso de unidad: para cada uno de los globalizadores.

- **P.I. de peso de unidad de producción científica $PIPU_{pc}$**
- **P.I. de peso de unidad de autoridad académica $PIPU_{aa}$**
- **P.I. de peso de unidad de revisiones anteriores $PIPU_{ra}$**

Los valores de cada uno de ellos se definen entre 0 y 1 (vistos como fracciones con numerador en 1. Ej: 1/2,1/3,1/4)

Múltiples perfiles de configuración de la metodología:

Se muestra la fórmula que permite definir la cantidad de perfiles que cada revista considere necesarios:

$$R_{(i)}P_k = PIP_{pc,k} * PIPU_{pc,k} * P_{pc,k} + PIP_{aa,k} * PIPU_{aa,k} * P_{aa,k} + PIP_{ra,k} * PIPU_{ra,k} * P_{ra,k}$$

donde i son los candidatos a revisores seleccionados y $k=1,\ldots,n$ son los diferentes perfiles de búsqueda.

Prototipo y pruebas de laboratorio

El proceso investigativo requirió pruebas de laboratorio y un prototipo para el desarrollo evaluativo de la metodología propuesta. Se utilizó una PC con Linux Debian 12 Bookworm y Python 3.11 con las librerías necesarias. Entre las definiciones que requiere el proceso metodológico se adoptó el Repositorio Institucional (RI) del CONICET, similar a una fuente heterogénea. En el mismo, se seleccionaron los artículos científicos del campo de las Ciencias de la Educación. Se extrajeron: títulos, autores, resúmenes, palabras clave, fecha de publicación y textos completos (excepto embargados) y se agregó el idioma declarado. Cabe señalar que en esta base un grupo significativo de autores ya están desambiguados por CONICET con datos de afiliación y titulaciones. Además, para la desambiguación y datos bibliométricos se utilizaron los perfiles de Google Scholar. En el proceso de detección de idiomas se eligió fastText sobre otras librerías por mejor rendimiento y para traducción de títulos, resúmenes y palabras clave, se usaron modelos disponibles en Hugging Face. Las extracciones del RI de CONICET y de los perfiles de Google Scholar se realizaron por Web scraping. Finalmente se seleccionaron más de 3000 artículos y más de 3000 autores. Esto permitió obtener una cantidad suficiente para probar el prototipo.

En cuanto a "Experiencia", se aplicaron diferentes técnicas de PLN: TF-IDF, Word2Vec, fastText, Doc2Vec y Bert, todas con un alto grado de representación de diferentes etapas evolutivas del PLN. Dentro de BERT se utilizó el proyecto sentence-BERT con modelos que utilizan representaciones de al menos 512 tokens y que están entrenados en español. Se seleccionaron los proyectos rm8488/multilingual-e5-large-ft-sts-spanish-matryoshka-768-16-5e (Transformer 1) y hiiamsid/sentence_similarity_spanish_es (Transformer 2). Siempre que fuera posible se aplicó limpieza de stop-words y lematización, pero no se utilizó en el caso de Transformers.

Para la **Desambiguación,** se utilizaron los perfiles de Google Scholar y se pudo desambiguar 781 investigadores sobre más de 1200 que obtuvieron al menos un perfil en búsquedas de bases de perfiles. Se pudo detectar que 761 correspondían a un solo autor y había diez que tenían perfiles duplicados; en estos casos se utilizaron los que tenían más artículos cargados y mayor valor de indicadores. También se detectó que 675 asociaban un solo nombre de autor del RI, 50 asociaban dos nombres de autores y dos asociaban tres nombres de autores diferentes.

En cuanto a **Indicadores**, de los 781 investigadores, 751 tenían información de indicadores. Los indicadores posibles son citas totales, h-index total, i10-index total y, en todos los casos, además, existía la opción de los últimos 5 años.

Comparativas de partes representativas: según el mapeo sistemático, lo más utilizado fue título con resumen combinado sin explicar el motivo. Se probaron todas las partes representativas del artículo con el texto completo depurado. En la mayoría de las técnicas, el resultado arrojó el resumen como la mejor

representación. Solo en TF-IDF, la mejor representación fue: "resumen + título + palabras clave". En Doc2Vec los porcentajes se emparejaron entre "resumen + título + palabras clave" y el resumen, por lo cual se puede usar el resumen.

Evaluación de expertos: se solicitó a 10 expertos, en sus áreas específicas de conocimiento en Ciencias de la Educación, que seleccionaran artículos cercanos a su trayectoria dentro del grupo de artículos presentes en el Ri CONICET - Cs. de la Educación. En base a dicha selección se le enviaron informes de priorización de artículos científicos a los nombrados, presentando diferentes técnicas de procesamiento de lenguaje natural (PLN) aplicadas al artículo seleccionado dentro de su campo de especialización. Cada uno de ellos evaluó y priorizó las técnicas, asignando valores más altos a las más efectivas y viceversa, con el fin de identificar qué técnicas coincidían mejor con sus criterios de evaluación. El resultado fue que la técnica que mejor se adaptó fue una basada en Transformers. Llamó la atención que TF-IDF (sin información semántica) ocupe el segundo lugar junto a Doc2vec, luego otro modelo basado en Transformers y, por último, fastText seguido de Word2vec.

Análisis de informes de recomendación de revisores: una vez que se obtuvo la información de la parte representativa a utilizar (texto completo o en su defecto, el resumen) y la técnica de PLN que mejor funcionó con evaluación de expertos, se le pidió al comité editorial de la revista IRICE una configuración de parametrización completa de al menos un perfil de la metodología. Teniendo la información cargada, se enviaron a 7 expertos de la revista IRICE, 19 informes de recomendación con 4 preguntas abiertas sobre similitud temática, listado de revisores, parametrización y mejores posibles. Como conclusión se puede establecer que hubo una aceptación sólida y constructiva del método. Se hicieron sugerencias puntuales que mejoran la operatividad sin afectar la validez conceptual. Y se pudo concluir que el prototipo fortalece la selección de revisores con criterios objetivos y configurables.

Conclusiones y aportes de la tesis a la disciplina

Se validó una **metodología adaptativa y flexible** para la recomendación de revisores que permite a los comités editoriales configurar criterios según sus propias políticas. El sistema integra desde técnicas tradicionales de PLN hasta **Transformers**, asegurando su vigencia tecnológica.

Los puntos clave de este aporte son:

- **Multifactorialidad parametrizable**: Integra experiencia, interés del investigador, aporte de los autores, conflicto de intereses, autoridad académica y antecedentes de revisión. Posibilita ajustar los valores representativos de cada factor y de la fórmula integral. El factor de diversidad presente en otras propuestas fue reemplazado y mejorado por la integridad de la metodología que permite la adaptabilidad a la diversidad de revistas científicas que quieran utilizarlo.
- **Visión sociotécnica-cultural**: Supera modelos previos al tratar la diversidad de criterios editoriales como una solución y no como un obstáculo.
- **Optimización de procesos**: Determina la parte más representativa a utilizar del artículo y la técnica de procesamiento más adecuada para cada caso.

Todos los puntos incluidos a continuación son mejoras originales de esta metodología que no se hallaron en los antecedentes relevados.

En **Producción científica** con respecto a:

"Experiencia": Introducción del conocimiento específico relativo con umbral dinámico de conocimiento. Acumulación de conocimiento según artículos priorizados en una propuesta de cálculo innovadora. Selección flexible de partes representativas: Texto completo, Título, Resumen, Palabras clave y distintas combinaciones. Elección de técnicas de PLN según contexto editorial y posibilidad de incluir nuevas. Diferentes tipos de cálculos para integrar conocimiento de revisores con más de un artículo priorizado.

“Interés del investigador”: Inclusión/exclusión flexible de este factor. Configuración de rangos de antigüedad para determinar representatividad en vez de valores predefinidos. Posibilidad de utilizarlo como restricción de consideración.

“Aporte de los autores”: Inclusión/exclusión flexible de este factor. Única metodología que incorpora este factor (basado en orden de publicación). Configuración flexible de representatividad en base a rangos de orden según criterio editorial.

De manera general en el “Control de Umbral de Consideración”: restringe la inclusión de investigadores con valores por debajo del umbral.

En **Autoridad académica:** Se puede incluir/excluir este globalizador.

En “Indicadores bibliométricos”: Inclusión/exclusión flexible. Selección del indicador a usar. Configuración de rangos según criterio editorial. Valor por defecto si no tiene información

“Titulaciones”: Inclusión/exclusión flexible. Configuración de representatividad según criterio editorial. Puede usarse como restricción (mínimo nivel). Valor por defecto si no tiene información. Influye en el cálculo (no solo como regla de inclusión).

De manera general define cómo combinar bibliografía y titulaciones.

En **Revisiones anteriores:** Se puede incluir/excluir este globalizador.

En cortes: Control de revisiones simultáneas máximas. Control de tiempo de espera/descanso. Controles de revisiones anuales máximas. Rechazos (penalizaciones). Parametrizable (temporales e indefinidas). Calidad de revisiones (penalizaciones). Parametrizable (temporales e indefinidas).

En cálculo: Evaluación de consistencia de revisiones con tabla de calificaciones y representatividad por cada CE. Configura valores por defecto si faltan datos. Cálculos diferentes entre todas las revisiones registradas del mismo candidato a revisor.

En el modelo integral: Pesos diferenciados por globalizador (suma=100). Concepto de peso unitario por factor. Configuración de múltiples perfiles para búsqueda de revisores.

El preprocesamiento propuesto alcanzó una efectividad superior al 90 %. Los expertos confirmaron la solidez metodológica.

La propuesta constituye una herramienta innovadora y flexible que optimiza la selección de revisores, aportando transparencia, equidad y eficiencia al proceso. Se alinea con los principios de la UNESCO sobre el uso de IA en investigación, garantizando el respeto a los datos públicos, el uso responsable de algoritmos y el apoyo a la toma de decisiones humanas.

En cuanto a **líneas futuras de investigación** se estima: 1) Desarrollo de métodos más precisos de corte de artículos y aplicación de técnicas de aprendizaje automático; 2) Optimización mediante APIs y exploración de otras fuentes de información; 3) Integración de diferentes modelos (incluyendo mixtos); 4) Incorporación de un agente de IA generativa para llevar adelante algunas tareas de preprocesamiento y un análisis semántico más profundo que la similitud estadística/vectorial entre artículos seleccionados y el manuscrito a revisar; 5) Desarrollo de un modelo funcional accesible para revistas científicas.

En prospectiva, se busca optimizar una metodología con alto potencial de aplicación más allá de contextos regionales, dada su flexibilidad y adaptabilidad. Esto puede fortalecer significativamente la labor de los Consejos Editoriales de publicaciones científicas, contribuyendo a la calidad y eficiencia del proceso de revisión por pares.

Modelo combinado de Co-training y aprendizaje por transferencia para clasificación de documentos, a partir de un análisis comparativo de modelos de aprendizaje semi-supervisados

Autor

Alex Santiago Cevallos Culqui

Director

Dra. Claudia Pons

Codirector

Dr. Gustavo Rodríguez

Universidad

Universidad Nacional de La Plata

Fecha de exposición

12/11/2025

Resumen

La escasez de documentos etiquetados en la mayoría de los conjuntos de datos en distintos dominios dificulta la correcta clasificación de documentos y la generación de aprendizaje, además de implicar altos costos en términos de recursos para su proceso de etiquetado. En este contexto, los modelos de aprendizaje semi-supervisados (Semi-Supervised Learning, SSL) surgen como una alternativa para mitigar esta limitación, sin embargo, la falta de un análisis comparativo que evidencie las fortalezas y debilidades de los distintos tipos de modelos dificulta su selección. Así, se plantea una Revisión de Literatura Sistemática (Systematic Literature Review, SLR) que identifica que las principales limitaciones de los modelos están relacionadas con los límites de decisión y la adaptación de dominio, factores que afectan sus niveles de rendimiento medidos en precisión. Es por esta razón que en la presente tesis se propone diseñar un modelo combinado de clasificación SSL que optimiza el proceso de etiquetado y la clasificación de documentos, mejorando su eficacia y niveles de precisión. Para ello, se desarrolla un marco comparativo que evalúa los distintos tipos de modelos y se implementa una estructura que integra las mejores prácticas identificadas. La metodología de trabajo para el análisis de los modelos se fundamenta en el enfoque PICOC para la estrategia de búsqueda y en la guía PRISMA para la definición de los criterios de exclusión. El modelo propuesto se estructura empleando una combinación de técnicas de co-entrenamiento y transferencia de aprendizaje (COTRA) para el procesamiento y entrenamiento de datos respectivamente, su entrenamiento se refuerza mediante el uso del conjunto de datos pre-entrenado de BERT. A diferencia de modelos previos, la estructura de COTRA fortalece el estado del arte al abordar de manera conjunta los desafíos de adaptación de dominio y límite de decisión. Esto se logra mediante una representación multivista optimizada que integra diversas representaciones de documentos con información complementaria proveniente de fuentes pre-entrenadas. Esta estrategia permite reducir la

incertidumbre en la asignación de etiquetas y mejorar la capacidad de generalización en escenarios con datos etiquetados limitados, proporcionando un modelo más robusto y adaptable para la clasificación de textos en contextos con restricciones de datos. Para la evaluación de COTRA, se llevaron a cabo experimentos con documentos científicos clasificados en cinco y once categorías correspondientes a sus áreas de estudio. El modelo se comparó con modelos SSL individuales basados en auto-entrenamiento, así también con modelos que incorporan co-entrenamiento, algoritmos genéticos y aprendizaje por transferencia a través de pipelines de clasificación como enfoques zero-shot. COTRA ha logrado los mejores niveles de rendimiento en comparación con el resto de modelos, alcanzando una precisión máxima de 0,87 entre los modelos de co-entrenamiento, frente a la mejor métrica de 0,78 obtenida por los modelos individuales de auto-entrenamiento en la clasificación de cinco categorías. Estos resultados indican que el co-entrenamiento representa una estrategia efectiva para mejorar el desempeño predictivo en la clasificación de documentos.

Introducción

La creciente cantidad de documentos digitales disponibles en repositorios académicos e institucionales, ha generado un desafío significativo en la clasificación y categorización de los documentos, afectando tanto la facilidad de búsqueda de documentos como la calidad de la toma de decisiones en diversos ámbitos, desde la investigación científica hasta la gestión empresarial [1]. La dificultad de realizar un análisis de texto efectivo del significado de un documento generalmente provoca errores en su clasificación, limitando así la capacidad de las organizaciones para extraer conocimiento valioso de su información y facilitar su intercambio [2].

En este contexto, la búsqueda e implementación de técnicas eficientes de procesamiento de lenguaje natural para la clasificación automatizada de documentos fortalecen este escenario. Los beneficios de esta propuesta incluyen mejoras en la organización de los documentos, lo que facilita el acceso y la recuperación de información relevante, promoviendo el intercambio de conocimientos entre diferentes sectores, potenciando la innovación y la colaboración interdisciplinaria, lo cual es crucial en un mundo interconectado y en constante evolución. Además, conseguir una clasificación con niveles de precisión más altos, permite una toma de decisiones más informada y basada en evidencia, reduciendo así la probabilidad de error.

En este ámbito, se han implementado algunas estrategias para abordar el desafío de la clasificación y categorización de documentos. Un importante enfoque es el uso de modelos de aprendizaje supervisado y no supervisado, para la clasificación de texto. Por ejemplo, estudios han demostrado que los modelos de clasificación automática, como los basados en entrenamientos supervisados [3] y no supervisados [4][5][6], pueden alcanzar niveles de precisión considerables en la categorización de documentos. Estos modelos analizan características por medio de técnicas de procesamiento de lenguaje natural (Natural Language Processing, NLP), lo que permite automatizar la clasificación.

Sin embargo, los modelos supervisados dependen en gran medida de conjuntos de datos etiquetados, cuya disponibilidad puede ser limitada, lo que restringe su aplicabilidad en contextos donde las etiquetas son escasas o costosas de obtener [7]. Por otro lado, los modelos no supervisados, aunque útiles para descubrir patrones en datos no etiquetados, a menudo carecen de la precisión necesaria para clasificaciones más finas [8]. Esta limitación ha llevado a un creciente interés en los modelos semi-supervisados, que combinan las ventajas de ambos enfoques al utilizar un pequeño número de etiquetas junto con grandes volúmenes de datos no etiquetados [9]. Estos modelos permiten una

mayor flexibilidad y eficacia en la clasificación, facilitando el aprendizaje en escenarios donde la disponibilidad de datos etiquetados es un obstáculo [10]. No obstante, son escasos los estudios focalizados en una valoración de los modelos semi-supervisado [11].

Por tal razón, en el presente trabajo se propone un análisis comparativo de los diferentes tipos de modelos semi-supervisados en la clasificación de documentos, así como el desarrollo de un modelo combinado que integra los enfoques más eficientes extraídos de dicho análisis. La propuesta del análisis comparativo de diferentes modelos semi-supervisados en la clasificación de documentos permite identificar cuáles son los enfoques más eficaces para contextos específicos. Esto facilita a investigadores y profesionales tomar decisiones informadas sobre las técnicas que deben implementar en sus sistemas de gestión documental. Al mejorar la precisión y efectividad de la clasificación, la propuesta no solo optimiza el proceso de categorización, sino que también contribuye a la creación de herramientas más accesibles y eficientes para la organización de la información. Esto se traduce en una mejor visibilidad y recuperabilidad de los documentos, lo que potencia el intercambio de conocimiento y la colaboración entre usuarios, optimizando así el acceso a información relevante y actualizada.

Problemas

Escasez de análisis de modelos semi-supervisados.- Existe dificultad en acceder a información comparativa sobre el desempeño de los modelos semi-supervisados en la clasificación de documentos [12]. Actualmente, hay pocos estudios que aborden esta área, y las revisiones existentes no se enfocan específicamente en la clasificación de documentos [13], lo que limita la capacidad de los usuarios para tomar decisiones informadas. Sin disponer de un análisis comparativo claro, investigadores y profesionales enfrentan dificultades para elegir el modelo más adecuado para sus necesidades. Esto puede llevar a la implementación de técnicas ineficaces, resultando en una clasificación subóptima de documentos que afecta la gestión de información en las organizaciones. Las consecuencias incluyen baja eficiencia, costos elevados y decisiones erróneas que impactan negativamente en el acceso a información relevante [14].

La solución propuesta presenta un marco comparativo que evalúa cómo funcionan los modelos semi-supervisados en diferentes contextos de clasificación de documentos, considerando variables relevantes como el tipo de documento, número de clases, cantidad de datos etiquetados y sus niveles de precisión. Esta solución permite a los usuarios identificar las ventajas y desventajas de cada modelo en contextos específicos, fortaleciendo su capacidad para tomar decisiones informadas. Al proporcionar un análisis contextualizado, se mejora la selección de modelos y se optimiza la clasificación de documentos. Sin embargo, la solución propuesta puede tener un alcance limitado, ya que el análisis comparativo podría dejar algunas técnicas relevantes sin evaluar. Así como contextos que no se consideran en el análisis, lo que provocaría que la solución pueda ser menos efectiva.

Limitaciones de los modelos Semi-supervisados.- A través de la revisión de los diferentes modelos semi-supervisados, se ha identificado que algunos métodos no logran satisfacer adecuadamente las necesidades de clasificación [15], especialmente en contextos donde los conjuntos de datos etiquetados son escasos o inexistentes, lo que resulta en limitaciones del modelo en precisión y eficacia [16] [17]. Este problema es significativo, ya que la clasificación de documentos es fundamental en diversas aplicaciones, desde la gestión de información hasta la automatización de procesos. Una clasificación ineficaz puede comprometer la calidad de la información y afectar la

toma de decisiones, impactando directamente en la operación y competitividad de las organizaciones [18]. La falta de un modelo robusto que combine las mejores prácticas identificadas en la revisión de modelos semi-supervisados, limita la capacidad de las organizaciones para clasificar documentos de manera efectiva [19]. Esto puede traducirse para las empresas en pérdidas operativas, decisiones basadas en información incorrecta y oportunidades no aprovechadas.

Se propone el desarrollo de un modelo combinado que integra las mejores prácticas y enfoques de los diferentes modelos semi-supervisados analizados. Este modelo está diseñado para optimizar la clasificación de documentos, mejorando tanto la precisión como la eficacia, en contextos con conjuntos de datos etiquetados limitados. La solución propuesta aborda el problema combinando fortalezas de diferentes enfoques semi-supervisados, permitiendo que el modelo aproveche tanto los datos etiquetados como los no etiquetados y pre-entrenados. Esto aumenta la capacidad del sistema para clasificar documentos de manera más efectiva, lo que se traduce en una mejora en la calidad de la información y en la toma de decisiones.

Sin embargo, la solución también presenta algunas limitaciones. Una de ellas es la complejidad y robustez de ajuste del modelo, se puede requerir recursos técnicos significativos, así como posibles requerimientos computacionales más altos que podrían limitar su aplicabilidad en ciertas infraestructuras. A partir de estas limitaciones, el estudio de este trabajo puede ampliarse en la realización de, pruebas en diferentes dominios y contextos, se podría evaluar la versatilidad y adaptabilidad del modelo combinado, ajustando su diseño según las necesidades específicas de cada aplicación. También, se podría investigar técnicas de aprendizaje activo que permitan al modelo identificar y solicitar etiquetas para ejemplos específicos, mejorando así la calidad de los datos disponibles.

Desafíos de la adaptación de dominio en modelos semi-supervisados.- Existe dificultad para adaptar modelos semi-supervisados a nuevos dominios de datos. Estos modelos, que combinan información etiquetada, no etiquetada y pre-entrenada de otros dominios a menudo tienen falencias de generalización cuando se aplican a contextos diferentes, lo que resulta en un rendimiento subóptimo [20]. La falta de técnicas efectivas para transferir el conocimiento entre dominios limita la capacidad de estos modelos para aprender de manera eficiente en diversas situaciones [21].

Este problema es de importancia, ya que, en muchos casos, los datos etiquetados son escasos o costosos de obtener. La efectividad de los modelos semi-supervisados radica en su capacidad para aprovechar grandes volúmenes de datos no etiquetados o pre-entrenados. Sin embargo, si no pueden adaptarse a variaciones en los dominios, se pierde el valor potencial de estos datos. La incapacidad de los modelos para la adaptabilidad y generalización puede resultar en una disminución de su utilidad y en la insatisfacción de los usuarios.

La incapacidad de adaptar modelos a nuevos dominios puede conducir a un rendimiento deficiente, afectando la calidad de los resultados y las decisiones basadas en estos modelos. Además, la inversión en el desarrollo de soluciones que no se adaptan correctamente puede traducirse en un uso ineficiente de recursos financieros y humanos. Esta falta de adaptabilidad también puede limitar la capacidad de la organización para entrar en nuevos mercados, restringiendo así su crecimiento.

La propuesta para abordar los desafíos de adaptación de dominio en modelos semi-supervisados es la implementación de un modelo que utiliza técnicas de transferencia de aprendizaje. Esta técnica permite que el modelo aproveche el conocimiento adquirido en un dominio fuente para mejorar su rendimiento en un dominio objetivo. Este modelo incluye el uso de redes neuronales pre-entrenadas,

se ha planteado una arquitectura del modelo para que sea más flexible y facilite el ajuste fino de parámetros específicos del nuevo dominio.

La implementación de técnicas de transferencia de aprendizaje ayuda a mitigar las falencias de generalización de los modelos semi-supervisados, permitiendo que estos se beneficien de datos y características de otros dominios. Esto mejora la capacidad de los modelos para aprender de manera eficiente en contextos diferentes, incrementando así su precisión y efectividad. La adaptabilidad del modelo mejora, lo que puede conducir a una mayor satisfacción del usuario al ofrecer resultados más precisos y relevantes.

Una de las limitaciones del modelo propuesto radica en el proceso de adaptabilidad entre el dominio fuente y el dominio objetivo, el proceso de entrenamiento de ajuste fino puede ser complejo y requerir tiempo y recursos significativos, especialmente si se necesitan ajustar múltiples parámetros. Existe el riesgo de que el modelo se sobreajuste a los datos del nuevo dominio si no se implementan técnicas adecuadas de regularización, todo esto depende de la calidad de los datos del dominio fuente. Para ampliar el trabajo presentado, se podría investigar la aplicación de enfoques de aprendizaje activo, donde el modelo selecciona de manera proactiva los ejemplos más informativos para etiquetar en el nuevo dominio, mejorando así la calidad de los datos etiquetados.

Límite de decisión y clasificación efectiva.- Existe dificultad en el etiquetado óptimo de documentos sin etiquetar, causada por la presencia de documentos ubicados en los bordes de las agrupaciones a lo que se denomina Límite de decisión [22]. Estos documentos, que presentan etiquetas difusas, afectan negativamente la precisión del proceso de clasificación. A pesar de las técnicas de etiquetado disponibles, el fenómeno del límite de decisión provoca incertidumbre en la asignación de categorías, lo que resulta en etiquetados incorrectos y un deterioro del rendimiento del modelo, especialmente cuando se manejan múltiples categorías [23].

La importancia de este problema radica en que el etiquetado adecuado de documentos es crucial para el éxito de los modelos de clasificación [24] y, en consecuencia, para la efectividad de los procesos organizacionales que dependen del análisis de datos [25]. Un etiquetado inexacto no solo puede llevar a decisiones erróneas, sino que también afecta la calidad del servicio al usuario. En un entorno donde la información se genera y consume a gran velocidad, la capacidad de clasificar documentos de manera precisa es fundamental para mantener la competitividad y responder a las necesidades del mercado.

La necesidad de realizar ajustes o reentrenamientos frecuentes de los modelos debido a errores en el etiquetado puede traducirse también en costos adicionales en términos de tiempo y recursos humanos[26]. Asimismo, si los productos o servicios de la organización dependen de clasificaciones precisas y estas fallan, la satisfacción del usuario puede verse comprometida, afectando la reputación de la organización.

Para abordar el problema de clasificación efectiva por límite de decisión, se propone la estructura de un modelo combinado, que fusiona dos modelos destinados a mejorar la robustez de las decisiones de clasificación, así se consigue una redundancia y mejora en la identificación y el tratamiento de documentos en los bordes de las agrupaciones. Esta solución resuelve el problema planteado al permitir que el modelo disponga de redundancia en la decisión de clasificación, así gestiona mejor los documentos que se encuentran en el límite de decisión. Así se minimiza la ambigüedad en la clasificación, aumentando la precisión en la asignación de categorías.

La limitante del modelo combinado es que su estructura aumenta la complejidad del modelo y potencialmente los tiempos de procesamiento, debido a su redundancia en el proceso de entrenamiento. Además, si las técnicas de agrupamiento no están bien diseñadas, existe el riesgo de continuar generando errores en la clasificación, lo que provocaría disponer una robusta arquitectura sin capacidad de gestionar documentos en límite de decisión. Para ampliar la propuesta a partir de las limitaciones establecidas, se puede profundizar el estudio en técnicas de regularización y enfoques de entrenamiento incremental. Crear un marco de evaluación para medir el rendimiento del modelo bajo diferentes configuraciones, aplicar estas técnicas en proyectos del mundo real, y documentar los resultados permitiría obtener información valiosa, para identificar los más efectivos.

Objetivo del trabajo

La clasificación de documentos en entornos semi-supervisados enfrenta varios retos, como es la escasez de documentos etiquetados para fortalecer su entrenamiento o la dificultad para adaptar modelos a nuevos dominios. Para abordar estos problemas, es fundamental establecer un enfoque sistemático que permita desarrollar un modelo eficaz y robusto. Los siguientes objetivos han sido definidos para guiar el proceso del presente estudio, asegurando que se aborden las limitaciones identificadas, para la obtención de resultados que mejoran significativamente la calidad de la clasificación de documentos.

General

Diseñar un modelo combinado de clasificación semi-supervisada, que optimice tanto el etiquetado como la clasificación de documentos, mejorando su precisión y eficacia, fundamentado en una estructura que se derive de un análisis de la evaluación de los distintos tipos de modelos semi-supervisados.

Específico #1

Diseñar un marco comparativo que evalúe el desempeño de los diferentes modelos semi-supervisados en la clasificación de documentos, considerando variables como tipo de documento, número de clases y niveles de precisión.

Específico #2

Implementar un modelo combinado que integre las mejores prácticas de los modelos semi-supervisados analizados, utilizando técnicas de transferencia de aprendizaje para mejorar la adaptabilidad y generalización en nuevos dominios de datos.

Específico #3

Desarrollar un modelo que combine diferentes entrenamientos de clasificación semi-supervisada, enfocándose en la redundancia y mejora del manejo de documentos situados en los límites de decisión, con el fin de aumentar la precisión del etiquetado y reducir la ambigüedad en la asignación de categorías.

Contribuciones principales

Este trabajo se enfoca en mejorar la clasificación de documentos en entornos semi-supervisados, abordando desafíos críticos como la escasez de datos etiquetados y la adaptación a nuevos dominios. Las contribuciones del estudio ofrecen soluciones que no solo optimizan el rendimiento de los

modelos existentes, sino que también sientan las bases para futuras investigaciones en el área. A través de un análisis comparativo y el desarrollo de un modelo combinado, este trabajo busca enriquecer la práctica de la clasificación de documentos y abrir nuevas oportunidades de aplicación. Estas contribuciones son presentadas en cuatro partes.

Evaluación de modelos semi-supervisados: La presente investigación aborda la escasez de análisis comparativos de modelos semi-supervisados en la clasificación de documentos, una problemática que ha limitado la capacidad de los usuarios para tomar decisiones informadas. La contribución central radica en el desarrollo de un marco comparativo que evalúa el desempeño de diversos modelos en contextos específicos, teniendo en cuenta variables como el dominio, el tipo de documento, el número de clases y la cantidad de datos etiquetados. Esta herramienta no solo proporciona una visión de las ventajas y desventajas de cada enfoque, sino que también brinda herramientas a investigadores y profesionales para seleccionar el modelo más adecuado según sus necesidades de optimización en su gestión documental.

Además, el trabajo contribuye a establecer un estándar para futuros estudios en el área, fomentando un enfoque sistemático en la evaluación de modelos semi-supervisados. Aunque el marco puede tener limitaciones en su alcance, su implementación representa un avance significativo en el campo. En última instancia, esta contribución promueve una mayor comprensión y un uso efectivo de los modelos semi-supervisados, lo que puede transformar la manera en que las organizaciones gestionan y clasifican su información.

Adaptabilidad de dominios: Cuando los documentos etiquetados son escasos o costosos de obtener, el enfoque de adaptación de dominio mejora la capacidad de un modelo para generalizar su rendimiento al aplicarse en un nuevo dominio diferente al de su entrenamiento, es decir permite que el conocimiento de un dominio fuente se transfiera al dominio objetivo [27]. En este contexto, en el presente trabajo se estructura un modelo que utiliza técnicas de transferencia de aprendizaje para abordar los desafíos de adaptación de dominio en modelos semi-supervisados para clasificación de documentos. El modelo propone generar diferentes representaciones de características de los documentos, de modo que estas representaciones puedan ser sometidas a un entrenamiento por transferencia utilizando diccionarios pre-entrenados, así se incrementa la capacidad de generalización en nuevos dominios, esto ha permitido que el modelo se adapte mejor a variaciones en los datos, mejorando así su rendimiento en contextos diferentes.

Al integrar diversas representaciones de las características del documento, se obtiene un mejor contexto de las palabras en el dominio, lo que incrementa su semántica. De esta manera, el modelo se adapta mejor a los diferentes estilos y lenguajes entre el dominio de origen y el de destino. La optimización del uso de datos pre-entrenados y no etiquetados, reduce la necesidad de inversión en la obtención de documentos etiquetados, lo que permite un uso más eficiente de los recursos. La adaptabilidad del modelo crea oportunidades para abordar nuevos dominios y entornos con mayor eficacia, lo que fomenta su crecimiento, ya que puede ajustarse a las necesidades específicas de diferentes contextos.

Optimización de límites de decisión: En el proceso de categorización existen fronteras que separan las diferentes clases, estas fronteras conocidas como límite de decisión determinan la asignación de etiquetas a nuevos documentos, su flexible y correcta identificación es crucial para la una adecuada clasificación. Los documentos cercanos a este límite, representan casos importantes ya que una ligera variación en sus características puede cambiar su clasificación. Manejar adecuadamente este límite es esencial, especialmente en situaciones donde la precisión es crítica y los datos pueden ser ruidosos.

Se propone un modelo combinado que utiliza diferentes perspectivas o vistas de las características del documento, el modelo de clasificación mejora la precisión al capturar información diversa, lo que resulta en una clasificación más exacta, especialmente cerca del límite de decisión. La redundancia en las representaciones reduce errores al permitir que el modelo verifique y valide decisiones, mientras que la combinación de diferentes perspectivas aumenta la robustez frente al ruido en los datos. Además, gestionar vistas facilita el manejo de casos límite y mejora la confianza en las clasificaciones.

Clasificando artículos universitarios por áreas de investigación: La clasificación de documentos por líneas de investigación en las universidades es fundamental para organizar el conocimiento y facilitar el acceso a información relevante, lo que permite a investigadores y estudiantes encontrar recursos en áreas específicas. Una adecuada categorización de documentos, promueve la colaboración entre académicos que trabajan en temas similares, apoya la toma de decisiones estratégicas en la planificación académica y mejora la visibilidad de la investigación institucional, lo que puede atraer a nuevas investigaciones y financiamiento. En muchos casos, los repositorios universitarios almacenan documentos científicos sin una organización que contemple las áreas de investigación específicas de la universidad. Como resultado, los documentos se dispersan, algunas áreas de conocimiento quedan incompletas y se restringen las oportunidades de colaboración, además de dificultar la identificación de los principales focos de investigación en cada disciplina.

En este contexto, se ha desarrollado un modelo combinado de aprendizaje semi-supervisado con el objetivo de clasificar documentos científicos de los investigadores de la Universidad Técnica de Cotopaxi. Este modelo ha sido diseñado para adaptarse a las áreas de investigación de la universidad, para su entrenamiento, se utilizan conjuntos de datos pre-entrenados, lo que facilita el aprendizaje de patrones de clasificación a partir de ejemplos previos y mejora la precisión en la asignación de los documentos. Además, el modelo procesa los documentos desde diferentes perspectivas, considerando diferentes elementos del documento como el título, resumen y palabras clave. Este enfoque multidimensional permite una clasificación más precisa y ajustada a las particularidades de cada disciplina, optimizando la organización del conocimiento. De este modo, se mejora la visibilidad de las áreas de investigación y se generan mayores oportunidades de colaboración entre los investigadores.

Trabajos Futuros

No obstante, el modelo COTRA presenta ciertos desafíos que podrían abordarse en futuras investigaciones. Una de sus principales limitaciones radica en la necesidad de un ajuste fino adecuado, ya que un entrenamiento ineficiente podría llevar al modelo a sobreajustarse a las características del dominio fuente, lo que afecta su capacidad de generalización en el dominio destino. Otro desafío radica en la complejidad computacional generada por la estructura combinada utilizada para gestionar los documentos etiquetados entre las distintas vistas y los conjuntos preentrenados. La integración del transformer y el co-entrenamiento simultáneo aumentan los requerimientos de memoria y procesamiento, lo cual puede dificultar la implementación en entornos con recursos limitados.

También, se podrían explorar distintos enfoques para analizar la eficiencia en la clasificación de documentos científicos mediante la estructura del modelo COTRA. Por ejemplo, en lugar de las técnicas de co-entrenamiento, sería posible evaluar el rendimiento utilizando técnicas de ensamblado de modelos, combinando múltiples clasificadores para gestionar las predicciones. Asimismo, se podría experimentar con técnicas de aprendizaje activo, permitiendo que el modelo seleccione proactivamente los ejemplos más informativos para su etiquetado.

Además, sería relevante extender la aplicación del modelo COTRA a otros dominios que presenten características similares a los documentos científicos, como artículos de prensa, literatura médica, documentos legales, informes técnicos u otros. La capacidad del modelo para manejar datos no estructurados y transferir conocimiento entre dominios podría resultar beneficiosa en tareas de análisis de contenido, recuperación de información y generación de resúmenes automáticos.

Bibliografía

[1] G. Macgregor, "Digital repositories and discoverability: definitions and typology," in *Discoverability in Digital Repositories*, Routledge, 2023, pp. 11–31.

[2] U. Kuckartz, "Qualitative text analysis: A systematic approach," *Compendium for early career researchers in mathematics education*, pp. 181–197, 2019.

[3] G. Ginde *et al.*, "ScientoBASE: a framework and model for computing scholastic indicators of non-local influence of journals via native data acquisition algorithms," *Scientometrics*, vol. 108, no. 3, pp. 1479–1529, 2016, doi: 10.1007/s11192-016-2006-2.

[4] E. C. McNie, A. Parris, and D. Sarewitz, "Improving the public value of science: A typology to inform discussion, design and implementation of research," *Res Policy*, vol. 45, no. 4, pp. 884–895, 2016, doi: 10.1016/j.respol.2016.01.004.

[5] A. Ibáñez, "Machine Learning in Scientometrics," Universidad Politécnica de Madrid, 2015. [Online]. Available: http://oa.upm.es/36488/

[6] M. Maridueña, M. Leyva, and A. Febles, "Modelado y análisis de indicadores de ciencia y tecnología mediante mapas cognitivos difusos," *Ciencias de la información*, vol. 47, no. 1, pp. 17–24, 2016, [Online]. Available: http://biblat.unam.mx/es/revista/ciencias-de-la-informacion/articulo/modelado-y-analisis-de-indicadores-de-ciencia-y-tecnologia-mediante-mapas-cognitivos-difusos.

[7] T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised machine learning: a brief primer," *Behav Ther*, vol. 51, no. 5, pp. 675–687, 2020.

[8] M. T. Almuqati, F. Sidi, S. N. Mohd Rum, M. Zolkepli, and I. Ishak, "Challenges in Supervised and Unsupervised Learning: A Comprehensive Overview.," *Int J Adv Sci Eng Inf Technol*, vol. 14, no. 4, 2024.

[9] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Mach Learn*, vol. 109, no. 2, pp. 373–440, 2020, doi: 10.1007/s10994-019-05855-6.

[10] Y. Padmanabha, P. Viswanath, and B. Eswara, "Semi-supervised learning: a brief review," *International Journal of Engineering & Technology*, vol. 7, no. 1.8, p. 81, 2018, doi: 10.14419/ijet.v7i1.8.9977.

[11] J.-C. Su, Z. Cheng, and S. Maji, "A realistic evaluation of semi-supervised learning for fine-grained classification," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 12966–12975.

[12] S. Ruder, M. E. Peters, S. Swayamdipta, and T. Wolf, “Transfer learning in natural language processing,” in *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Tutorials*, 2019, pp. 15–18.

[13] Y. Ouali, C. Hudelot, and M. Tami, “An overview of deep semi-supervised learning,” *arXiv preprint arXiv:2006.05278*, 2020.

[14] G. Alene, “The Role of Management Information System in Improving Organizational Performance and Effectiveness in Case of Debre Markos City Administration Revenue Authority, Ethiopia,” *ICTACT Journal on Management Studies*, vol. 4, no. 1, pp. 691–697, 2018.

[15] A. Cevallos-Culqui, C. Pons, and G. Rodriguez, “Semi-supervised learning models for document classification: A systematic review and meta-analysis,” *Inteligencia Artificial*, vol. 26, no. 72, pp. 81–111, Jun. 2023, doi: 10.4114/intartif.vol26iss72pp81-111.

[16] N. Chen, B. Ribeiro, C. Tang, and A. Chen, “Multi-label learning vector quantization for semi-supervised classification,” *Intelligent Data Analysis*, vol. 23, no. 4, pp. 839–853, 2019, doi: 10.3233/IDA-184195.

[17] L. Borrajo, A. Seara Vieira, and E. L. Iglesias, “An HMM-based synthetic view generator to improve the efficiency of ensemble systems,” *Log J IGPL*, vol. 28, no. 1, pp. 4–18, 2020, doi: 10.1093/jigpal/jzz067.

[18] I. Khong, N. A. Yusuf, A. Nuriman, and A. B. Yadila, “Exploring the impact of data quality on decision-making processes in information intensive organizations,” *APTISI Transactions on Management*, vol. 7, no. 3, pp. 253–260, 2023.

[19] A. Masmoudi, H. Bellaaj, K. Drira, and M. Jmaiel, “A co-training-based approach for the hierarchical multi-label classification of research papers,” *Expert Syst*, vol. 38, no. 4, pp. 1–19, 2021, doi: 10.1111/exsy.12613.

[20] A. Søgaard, *Semi-supervised learning and domain adaptation in natural language processing*. Springer Nature, 2022.

[21] M. Das, X. Chen, X. Yuan, and L. Zhang, “Federated semi-supervised domain adaptation via knowledge transfer,” *arXiv preprint arXiv:2207.10727*, 2022.

[22] M. Alemi, A. Bosaghzadeh, and F. Dornaika, “Graph-Based Semi-Supervised Learning with Bipartite Graph for Large-Scale Data and Prediction of Unseen Data,” 2024.

[23] M. N. Rizve, K. Duarte, Y. S. Rawat, and M. Shah, “In defense of pseudo-labeling: An uncertainty-aware pseudo-label selection framework for semi-supervised learning,” *arXiv preprint arXiv:2101.06329*, 2021.

[24] R. Liu *et al.*, “Recent Advances in Hierarchical Multi-label Text Classification: A Survey,” *arXiv preprint arXiv:2307.16265*, 2023.

[25] T. Fredriksson, J. Bosch, H. H. Olsson, and D. I. Mattos, “Machine learning algorithms for labeling: Where and how they are used?,” in *2022 IEEE International Systems Conference (SysCon)*, 2022, pp. 1–8.

[26] A. Mahadevan and M. Mathioudakis, “Cost-Effective Retraining of Machine Learning Models,” *arXiv preprint arXiv:2310.04216*, 2023.

[27] M. H. Tanveer, Z. Fatima, S. Zardari, and D. Guerra-Zubiaga, “An In-Depth Analysis of Domain Adaptation in Computer and Robotic Vision,” *Applied Sciences*, vol. 13, no. 23, p. 12823, 2023.

MODELO DE CALIDAD PARA EVALUACIÓN Y MEJORA DE VISUALIZACIONES DE INFORMACIÓN NARRATIVAS

Autora: Andrea Fernanda Lezcano Airaldi

Director: Emanuel Agustín Irrazábal

Codirector: Jorge Andrés Diaz-Pace

Fecha de exposición: 25 de agosto de 2025

Universidad Nacional del Nordeste

Resumen

Las visualizaciones narrativas combinan recursos gráficos y estructuras discursivas para facilitar la comprensión de datos en contextos diversos. Sin embargo, evaluar su calidad sigue siendo difícil, ya que muchos enfoques priorizan métricas de rendimiento (p. ej., precisión o tiempo) y dejan en segundo plano la narrativa, la utilidad percibida y la experiencia del usuario. Esta tesis propuso NaVi-Q, un modelo de calidad para evaluar y asistir la mejora de visualizaciones narrativas en interfaces de usuario. Basado en estándares de Ingeniería del Software y principios de diseño en visualización, el modelo organiza buenas prácticas en una estructura jerárquica de criterios, propiedades observables y aplicaciones evaluativas con instrumentos concretos, y permite ajustar ponderaciones mediante perfiles de calidad según objetivo y contexto. El modelo se operacionalizó en una herramienta software que genera diagnósticos y recomendaciones y que incorpora soporte semiautomatizado, integrando verificaciones automáticas para aplicaciones y asistencia basada en IA multimodal para apoyar la evaluación y sugerir mejoras. La validación en tres etapas (estudio exploratorio, experimento controlado y estudio de caso aplicado) aportó evidencia de consistencia y utilidad para detectar problemas frecuentes y orientar mejoras, favoreciendo su integración en procesos de diseño, revisión y mejora continua.

Palabras clave: Visualización de información, visualizaciones narrativas, evaluación de calidad, modelo de calidad.

Motivación

La visualización de información busca transformar datos abstractos en representaciones interpretables para apoyar procesos cognitivos, análisis y toma de decisiones [1]. En particular, las visualizaciones narrativas integran recursos gráficos con estructuras discursivas para comunicar un mensaje sobre los datos, facilitando la comprensión y la discusión en contextos organizacionales y analíticos [2], [3]. Su relevancia se ha reforzado en años recientes, al reportarse efectos sobre comprensión, confianza y actitudes frente a determinados temas [4], [5], por ejemplo, en contextos de crisis como la pandemia de COVID-19, donde visualizaciones narrativas se utilizaron para comunicar evolución de casos, riesgos y medidas [6]. En el desarrollo de software, estas visualizaciones se incorporan en interfaces, reportes y tableros, donde no solo deben codificar datos correctamente, sino también sostener objetivos comunicativos y atender a públicos diversos, por lo que su evaluación resulta necesaria desde fases tempranas [7], [8].

Si bien existen fundamentos consolidados para el diseño de visualizaciones [9], [10], las visualizaciones narrativas requieren consideraciones adicionales —por ejemplo, proveer contexto y estructurar un flujo narrativo— que introducen compensaciones de diseño y complejizan la evaluación [11]. Aun con buenas prácticas disponibles, la evaluación rigurosa sigue siendo un problema abierto: muchos enfoques se apoyan en métricas tradicionales (tiempo/error) y no capturan dimensiones como interpretación del mensaje, coherencia narrativa o respuesta del usuario, además de verse afectada por factores subjetivos y contextuales [12]. En consecuencia, se identifica la falta de marcos integrados y herramientas con criterios operacionalizables y comparables, especialmente para el caso narrativo [13].

En respuesta a este problema, la tesis propone NaVi-Q[1], un modelo de evaluación prescriptivo para visualizaciones narrativas en productos software. El mismo se operacionaliza en una herramienta que

[1] www.naviq.com.ar

guía la evaluación, calcula puntajes y produce un diagnóstico desagregado por criterio y propiedad, además de recomendaciones accionables orientadas a la mejora. El modelo estructura la evaluación en una jerarquía de criterios (objetivos de calidad de alto nivel), propiedades (atributos observables de la visualización) y aplicaciones evaluativas (verificaciones concretas con instrumentos de medición), y permite ajustar ponderaciones mediante perfiles de calidad según el objetivo y el contexto de uso. De este modo, permite revisiones sistemáticas y comparables, orientadas a decisiones de diseño, sin depender exclusivamente de juicios subjetivos o métricas aisladas.

El trabajo se desarrolló bajo el enfoque de Investigación en Ciencia del Diseño (DSR) [14], estructurado en cinco actividades: explicar el problema, definir objetivos y requerimientos, diseñar y desarrollar el artefacto, demostrar su viabilidad y evaluarlo. La evaluación del artefacto se realizó de forma progresiva mediante una evaluación a pequeña escala, un experimento controlado y un estudio de caso aplicado, generando evidencia sobre viabilidad, consistencia y aplicabilidad en un contexto real e incorporando iteraciones a partir de los hallazgos. La **Fig. 1** muestra las actividades realizadas para alcanzar el objetivo de la tesis en el marco de DSR, incluyendo la secuencia de diseño y refinamiento del artefacto y los estudios empíricos que sustentan su demostración y evaluación.

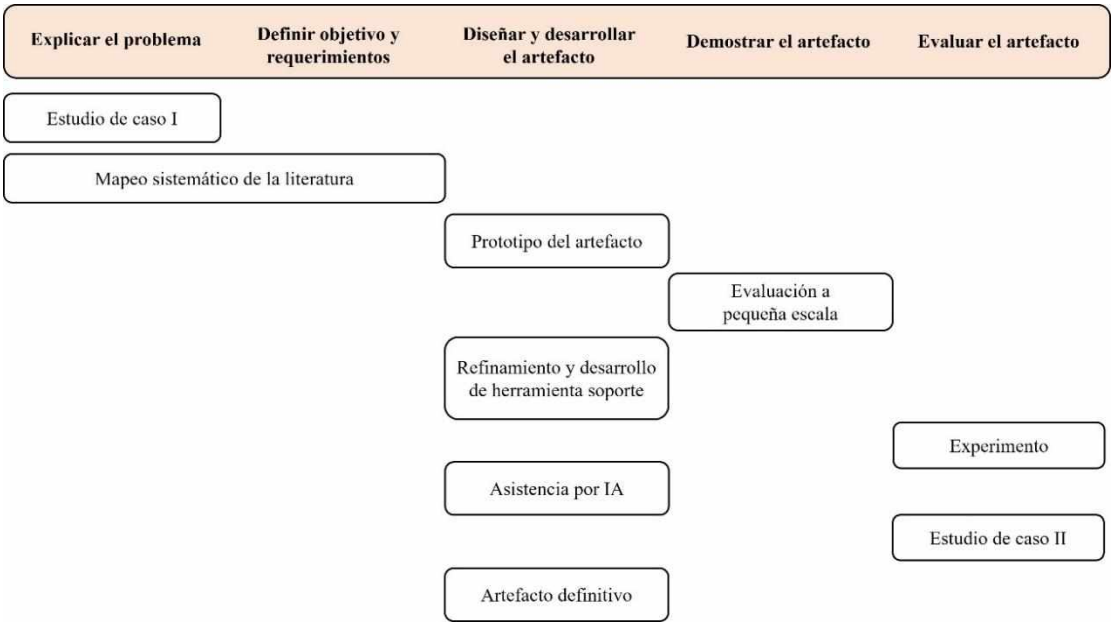


Fig. 1: Organización de actividades y tareas desarrolladas en el marco de DSR

Visualización Narrativa

La narrativa de datos (*data storytelling*) se presenta en la literatura bajo múltiples términos —historia con datos, visualización narrativa, historia dirigida por datos—, con un punto en común: el uso de estructuras narrativas para presentar datos de forma que facilite la comprensión y enriquezca la experiencia del público. Diversos autores describen este enfoque como la convergencia de datos, narrativa y elementos visuales [15], perspectiva retomada por trabajos recientes que también destacan la influencia del medio en la construcción y percepción del mensaje [16].

Además de definiciones basadas en componentes, se observa una conceptualización de la narrativa como secuencia ordenada de visualizaciones complementadas con texto e imágenes [17], donde el relato se sostiene mediante un recorrido que guía la atención del público. Esta guía se materializa a través de recursos como anotaciones, resaltados, señalización visual y explicaciones textuales que conectan la evidencia con el mensaje. En esa línea, se propone describir una visualización narrativa como una composición de “fragmentos” apoyados por datos y reforzados mediante mecanismos de comunicación que permiten construir sentido de manera incremental [18].

Un rasgo importante es que la narrativa no se limita a "embellecer" la codificación de datos: introduce decisiones explícitas sobre contexto, encuadre, progresión y énfasis, que influyen en lo que el público interpreta como relevante. Por ello, se plantea una visión iterativa del diseño de visualizaciones narrativas, que incluye requisitos, diseño, implementación y evaluación, y que reconoce la necesidad de refinar el artefacto en función del público y del escenario de uso [19]. Esta perspectiva resulta especialmente pertinente en productos software, donde las visualizaciones suelen formar parte de tableros o reportes y deben balancear claridad, consistencia y objetivos comunicativos, a menudo bajo restricciones de interfaz y diversidad de usuarios.

Evaluación en Visualización de Información

La literatura propone múltiples enfoques para evaluar visualizaciones —modelos teóricos, heurísticas, listas de verificación, cuestionarios, algoritmos y sistemas basados en IA— que aportan criterios útiles, pero suelen cubrir dimensiones parciales y difieren en su grado de formalización y soporte instrumental. En este marco, existen propuestas con jerarquías y modelos de medición explícitos, como Q-VIS, que introduce categorías ponderadas y ajuste de prioridades por contexto, aunque sin validación empírica [20]. También se reportan enfoques de inspección, donde predominan heurísticas adaptadas a diseño visual o toma de decisiones [21] y heurísticas orientadas a valorar el potencial de una visualización [22]. sin embargo, tienden a depender del juicio experto y no siempre contemplan adaptabilidad al contexto ni dimensiones relevantes, como señalan trabajos previos [23].

Además, se han propuesto métodos focalizados en criterios específicos, por ejemplo marcos para evaluar complejidad asociada a carga cognitiva [24] o enfoques centrados en comprensión que no definen métricas de cuantificación y mantienen alta dependencia del evaluador [25]. En visualizaciones narrativas, la práctica profesional evidencia falta de procedimientos sistemáticos y predominio de enfoques informales, lo que motivó marcos heurísticos derivados de la práctica para cubrir diseño visual y credibilidad de datos [26].

En años recientes se incorporaron enfoques automatizados como apoyo a tareas puntuales. Revisiones sobre IA aplicada a visualización distinguen métodos basados en reglas, aprendizaje automático e híbridos, e identifican la evaluación como una tarea central [27]. Se reportan algoritmos para estimar memorabilidad a partir de rasgos visuales [28] y propuestas con modelos de lenguaje para puntuar visualizaciones con criterios explícitos o detectar visualizaciones engañosas [29], [30], generalmente como complemento de la evaluación humana cuando se requiere interpretar contexto, público y objetivo comunicativo. En conjunto, este panorama respalda la necesidad de un marco de evaluación adaptable y operacionalizable para considerar múltiples dimensiones de calidad de forma sistemática, facilitando su adopción en escenarios reales, incluyendo productos software.

Desarrollo del Modelo NaviQ

NaVi-Q es un modelo prescriptivo para evaluar la calidad de visualizaciones de información narrativas. Su diseño se apoya en la noción de calidad como constructo medible y en marcos de referencia de evaluación y medición, adoptando como base ISO/IEC 25010 [31] e ISO/IEC 33003 [32], y un proceso de construcción en seis pasos: (1) definir alcance y contexto, (2) diseñar el metamodelo, (3) identificar requerimientos, (4) definir propiedades, (5) establecer reglas de evaluación y (6) instanciar el modelo para un caso de uso.

El modelo se enfoca en propiedades externas de calidad, es decir, características observables en la visualización como artefacto final, sin abordar su implementación técnica ni su impacto en uso. En términos de alcance, considera visualizaciones frecuentes en VI (p. ej., gráficos de líneas y barras) y su integración en documentos o tableros, tanto en modalidad estática como interactiva.

NaVi-Q organiza la evaluación mediante un metamodelo jerárquico que descompone la calidad en criterios, propiedades y aplicaciones evaluativas. Los criterios representan objetivos de calidad de alto nivel que agrupan dimensiones amplias del desempeño comunicativo y de uso de la visualización

(p. ej., comprensión o usabilidad). Las propiedades corresponden a atributos observables de la visualización que concretan esos objetivos en aspectos evaluables (por ejemplo, claridad, contextualización o consistencia). Las aplicaciones operacionalizan cada propiedad mediante verificaciones específicas, asociadas a un instrumento con opciones de respuesta; este instrumento incorpora la opción N/A cuando el ítem no resulta aplicable al caso evaluado (ver **Fig. 2**).

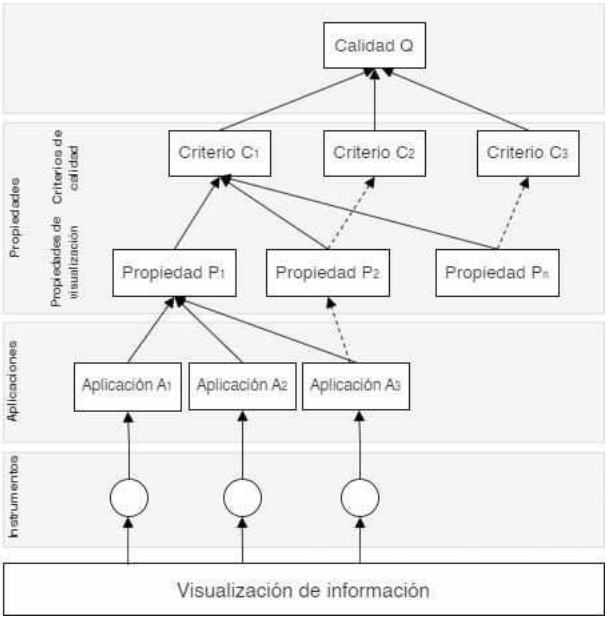


Fig. 2: Elementos y estructura del metamodelo.

El modelo define cinco criterios de calidad: Comprensión (C01), Compromiso (C02), Información (C03), Memorabilidad (C04) y Usabilidad (C05), y establece su relación con 17 propiedades (P01–P17) y con 55 aplicaciones (AP01–AP55), derivadas de buenas prácticas y criterios identificados mediante un mapeo sistemático de la literatura; el mismo fue publicado en la conferencia Evaluation and Novel Approaches for Software Engineering (ENASE) 2025 [33].

La medición adopta un modelo de medición formativo [32], en el que la calidad se construye como combinación de medidas observadas, lo que permite interpretar contribuciones por criterio y propiedad. En este marco, la agregación se alinea con técnicas de toma de decisiones multiatributo (MADM, por sus siglas en inglés) y utiliza una lógica compensatoria, adecuada para distintos perfiles de calidad según contexto. La calidad se calcula mediante el método de suma ponderada. Los pesos se derivan con el Proceso Analítico Jerárquico [34] (AHP por sus siglas en inglés) y, para reducir la carga de comparaciones del esquema clásico, se adopta AHP-Express [35], que disminuye las comparaciones necesarias a (n−1) manteniendo la jerarquía y la trazabilidad.

Las propiedades se miden en una escala de razón 0–1, interpretables como grado de cumplimiento, y se formalizan funciones de agregación por niveles (calidad global, criterios, propiedades, aplicaciones) coherentes con la estructura jerárquica. Como configuración inicial, los pesos de las aplicaciones dentro de una propiedad pueden asignarse de forma equitativa y redistribuirse cuando alguna aplicación no sea aplicable.

Finalmente, NaVi-Q incorpora el concepto de perfil de calidad para adaptar la evaluación a distintos escenarios de uso. Un perfil define un conjunto de ponderaciones que refleja prioridades contextuales (por ejemplo, educación, exploración o toma de decisiones), preservando una estructura común de criterios, propiedades y aplicaciones. De este modo, el mismo modelo puede aplicarse de forma consistente en dominios diferentes, ofreciendo resultados comparables y recomendaciones alineadas con los objetivos del escenario evaluado. El desarrollo del modelo, junto con la validación empírica resumida en la sección siguiente, fueron publicados en la revista IEEE Access [36].

Validación

La validación se estructuró como un proceso progresivo en tres etapas consecutivas, cada una con objetivos, preguntas e instrumentos propios: (i) una evaluación exploratoria a pequeña escala para analizar viabilidad inicial, (ii) un experimento controlado para estudiar consistencia y capacidad discriminativa en un entorno estructurado, y (iii) un estudio de caso aplicado para observar el uso del modelo en un contexto real de trabajo (ver **Fig. 3**).



Fig. 3: Esquema general de validación.

En la primera etapa de validación se realizó una evaluación a pequeña escala para comprobar la viabilidad inicial del modelo en un contexto controlado, siguiendo recomendaciones metodológicas para estudios preliminares con pocas fuentes de datos [37].

La evaluación se diseñó alrededor de tres preguntas: (PI-1) identificar fortalezas y limitaciones percibidas del modelo, (PI-2) estimar consistencia entre evaluadores, y (PI-3) caracterizar dificultad de aplicación. Para instrumentar el estudio, se configuró el perfil Estándar usando AHP-Express con 51 participantes en un taller de visualización, obteniendo pesos normalizados por criterio. El conjunto de datos incluyó 18 visualizaciones (12 estáticas y 6 interactivas). Participaron 12 evaluadores con niveles de experiencia heterogéneos; se preparó una planilla con criterios, propiedades y aplicaciones, con listas desplegables por aplicación y cálculo automático de puntajes, y luego se recolectó percepción de uso mediante encuesta Likert y preguntas abiertas.

Los resultados mostraron alta consistencia inter-evaluador (ICC = 0,984) y sugirieron que las instrucciones y la formulación de aplicaciones fueron interpretadas de forma similar. En cuanto la dificultad de aplicación, la mayoría calificó la experiencia como “algo fácil” o “ni fácil ni difícil”, con tiempos por visualización que variaron con la experiencia (menos de 10 minutos en evaluadores con mayor experiencia; ~12 minutos en experiencia intermedia; ~15 minutos en quienes no tenían experiencia específica).

Los comentarios cualitativos aportaron causas accionables: los perfiles intermedios y sin experiencia señalaron terminología poco común, necesidad de ejemplos o referencias visuales y mayor apoyo para responder algunas aplicaciones. Esto motivó una primera iteración del artefacto, orientada a facilitar su uso, donde se reorganizó el conjunto de aplicaciones reduciéndolo de 55 a 36 y se desarrolló una herramienta software para operacionalizar el proceso, incorporando ejemplos visuales y mecanismos de ayuda contextual (**Fig. 4**).

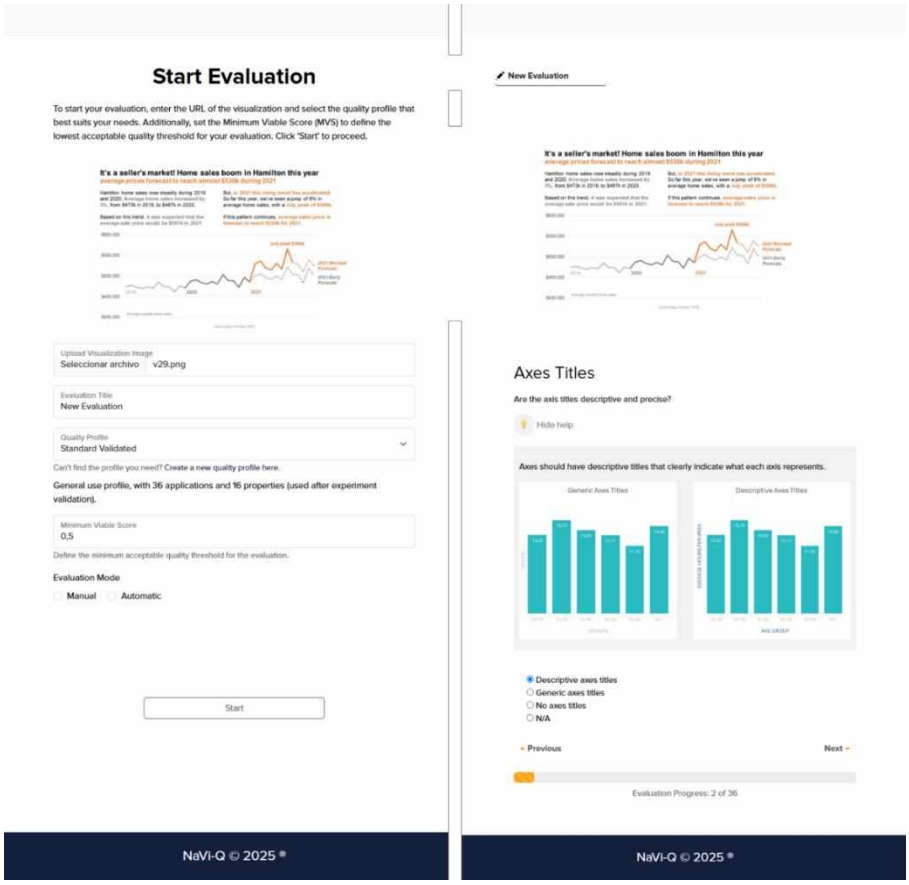


Fig. 4: Interfaz de la herramienta NaVi-Q

En la segunda etapa se ejecutó un cuasi experimento para analizar el comportamiento de NaVi-Q en condiciones controladas, evaluando su capacidad discriminativa y la consistencia de los puntajes. Se utilizó el perfil Estándar y, como referencia, un evaluador externo aplicó el modelo sobre 30 visualizaciones, que luego se agruparon en tres niveles de calidad (alta, media y baja; 10 por nivel). Los participantes (expertos y no expertos en el uso de visualizaciones) evaluaron visualizaciones de los tres niveles, siguiendo un diseño con factor inter-sujetos (experiencia) e intra-sujetos (nivel de calidad). Al finalizar, completaron una encuesta de experiencia de uso de la herramienta con ítems tipo Likert y un comentario abierto.

Respecto de la discriminación por nivel (H1), los puntajes aumentaron de forma consistente entre baja, media y alta calidad y las diferencias resultaron significativas (Kruskal–Wallis, $p < 0{,}001$), indicando que el modelo distingue niveles predefinidos. Para analizar la coherencia con la evaluación de referencia (H2), la correlación de Spearman entre puntajes de participantes y de referencia fue muy alta ($\rho = 0{,}931$; $p = 3{,}21e\text{-}106$), evidenciando fuerte concordancia.

Además, al comparar los puntajes de cada grupo con los valores de referencia se aplicó Wilcoxon para muestras relacionadas, sin diferencias significativas ni en expertos ($p = 0{,}655$) ni en no expertos ($p = 0{,}410$). En la comparación directa entre expertos y no expertos (H3), Mann–Whitney U tampoco mostró diferencias significativas ($U = 7344{,}5$; $p = 0{,}789$). Finalmente, un modelo de efectos mixtos confirmó un efecto significativo del nivel de calidad ($p < 0{,}001$), sin evidencia de efecto de la experiencia ni de la interacción experiencia–nivel, reforzando que NaVi-Q discrimina niveles de calidad sin depender de la experiencia del evaluador.

La encuesta final mostró valoraciones altas para la comprensión de las instrucciones y enunciados (promedio 4,35) y resultados también positivos para la claridad de la relación entre criterios y

aplicaciones (4,03) y la utilidad percibida (3,95). La satisfacción general se mantuvo en un nivel favorable (3,8), mientras que el ítem de apoyo recibido durante la evaluación obtuvo un valor moderado (3,7).

El análisis cualitativo de respuestas abiertas agrupó sugerencias en tópicos vinculados con claridad de explicaciones, mejoras de interfaz, proceso de evaluación, terminología y contextos de aplicación. Estos hallazgos motivaron una segunda iteración orientada a reducir carga cognitiva y brindar asistencia durante la evaluación, dando lugar a dos flujos de uso de la herramienta: (a) manual asistido y (b) automático (**Fig. 5**). En el flujo manual asistido se incorporaron (i) verificaciones automáticas para aplicaciones automatizables (contraste y accesibilidad de colores), presentadas como información de apoyo dentro de la evaluación, y (ii) asistencia basada en IA con un modelo multimodal para sugerir la opción más adecuada a partir del enunciado, el texto de ayuda y la visualización cargada. En el flujo automático, el sistema ejecuta la evaluación de forma integral y genera un diagnóstico con sugerencias específicas de mejora asociadas a recomendaciones, a partir de la visualización y del resultado de evaluación.

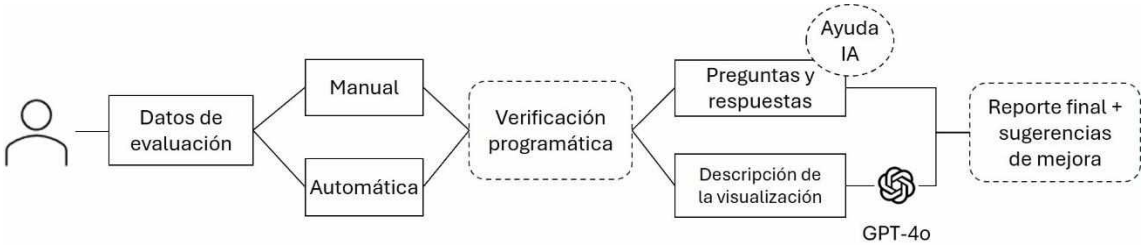


Fig. 5: Esquema del proceso de evaluación

La tercera etapa consistió en un estudio de caso aplicado para observar el funcionamiento del modelo en un entorno de trabajo real, siguiendo pautas de Yin [38] y formulando preguntas orientadas a utilidad de recomendaciones (PI-1), impacto de aplicar mejoras en comprensión, información y usabilidad (PI-2) y percepción general de uso e integración al flujo de trabajo (PI-3).

El estudio se realizó en colaboración con el grupo de investigación SWK de la Universidad de Hamburgo, en el marco de una estancia de investigación otorgada en conjunto por el Servicio Alemán de Intercambio Académico (DAAD) y la Secretaría de Educación Argentina. Se aplicó NaVi-Q a un tablero de un proyecto real de software, usando un perfil de calidad personalizado para priorizar comprensión, información y usabilidad. La evaluación se ejecutó en modalidad manual: un integrante del equipo aplicó la herramienta, resolvió dudas con ayuda contextual asistida por IA y consensuó respuestas con el grupo. Con base en los resultados, el sistema identificó aplicaciones con baja puntuación y generó sugerencias de mejora que el equipo discutió y priorizó según esfuerzo, permitiendo además agrupar patrones recurrentes de problemas. Estos resultados fueron consolidados en un manuscrito enviado a una revista y actualmente en proceso de revisión.

Para medir el impacto de las mejoras implementadas, se condujo un estudio con usuarios de diseño comparativo: 24 estudiantes avanzados de Informática fueron asignados aleatoriamente a un grupo control (visualizaciones originales) y un grupo experimental (versiones mejoradas), trabajando de forma individual y en condiciones equivalentes. Se aplicó un cuestionario de comprensión basado en tareas [39] y, al finalizar, una encuesta de Usabilidad del Sistema (SUS) adaptada para estimar usabilidad percibida [40].

Los resultados se resumen en la **Fig. 6**, donde el grupo experimental alcanzó mayores niveles de precisión en 4 de 5 preguntas y mostró confianza significativamente superior, y en la Tabla 1 para la encuesta SUS, con un puntaje promedio significativamente mayor en el grupo experimental.

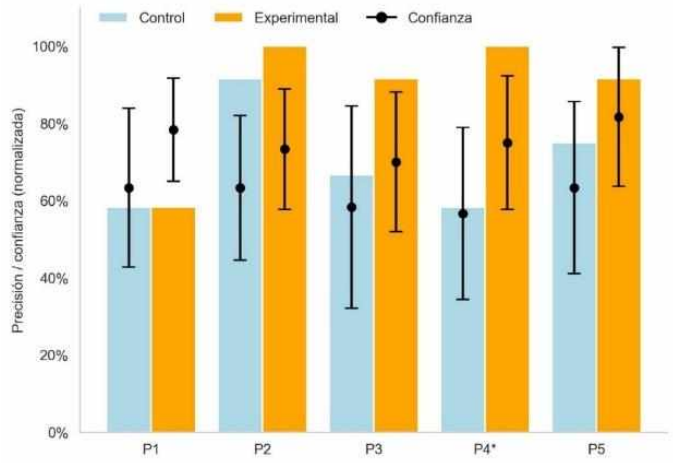


Fig. 6: Precisión y confianza normalizada por pregunta y grupo.

Tabla 1: Media, mediana, rango y desvío estándar de la encuesta SUS por grupo.

Grupo	Media	DE	Min.	Max.	Mediana
Control	63.12	6.22	52.5	70	65
Experimental	83.54	3.91	77.5	90	82.5

En cuanto a la percepción general, los participantes ubicaron el uso del modelo principalmente una vez que existe una primera versión de la visualización, valorándolo como herramienta de revisión para guiar iteraciones. También destacaron su utilidad para detectar problemas no triviales (omisiones narrativas, falta de atribuciones y deficiencias de legibilidad) y señalaron como principales limitaciones la sobrecarga en proyectos con muchas visualizaciones, la extensión del cuestionario para usuarios no expertos y la necesidad de integración al flujo de trabajo; en ese sentido, la modalidad asistida aparece como un recurso para reducir barreras de adopción.

Contribuciones

La contribución principal de esta tesis es NaVi-Q, un modelo prescriptivo para evaluar la calidad y asistir en la mejora de visualizaciones narrativas integradas en productos software. El modelo estructura la evaluación en una jerarquía de criterios, propiedades y aplicaciones evaluativas con instrumentos concretos, permite configurar perfiles de calidad mediante ponderaciones ajustables según el contexto de uso y se operacionaliza en una herramienta web que calcula puntajes, produce diagnósticos desagregados y emite recomendaciones accionables.

Como contribuciones adicionales, el trabajo aporta: (i) validación progresiva en tres etapas que habilitó iteraciones y mejoras orientadas a la adopción; (ii) un vocabulario evaluativo derivado de literatura y estándares, materializado en aplicaciones con escalas y criterios explícitos; (iii) soporte semiautomatizado con verificaciones automáticas (p. ej., contraste y accesibilidad de colores) y asistencia multimodal para apoyar respuestas y sugerencias; y (iv) un marco reutilizable para comparar enfoques y promover prácticas consistentes de evaluación y mejora continua.

Los resultados de esta tesis se difundieron en publicaciones que acompañan su evolución: un estudio de caso en J.UCS (2021) [41], un simposio doctoral en CIBSE (2022) [42], un mapeo sistemático en ENASE (2025) [33] y el artículo en IEEE Access (2025) [36], que en conjunto documentan la motivación, la base conceptual, el desarrollo del modelo y la evidencia empírica del enfoque propuesto. El modelo y la herramienta desarrollados fueron registrados y se encuentran protegidos por derechos de autor conforme a la legislación vigente en el país, bajo el expediente EX-2025-60931702-APN-DNDA#MJ. Este registro formaliza la autoría del artefacto y respalda su uso en actividades académicas y de transferencia tecnológica.

Trabajos Futuros

Para extender la línea de investigación, se propone validar NaVi-Q en dominios adicionales (p. ej., periodismo. educación), con audiencias y objetivos comunicativos distintos, a fin de analizar su flexibilidad y robustez en escenarios narrativos heterogéneos. En paralelo, resulta pertinente ampliar el metamodelo para contemplar formatos más dinámicos e interactivos (tableros en tiempo real, visualizaciones con desplazamiento secuencial), incorporando nuevas propiedades y aplicaciones cuando sea necesario.

Respecto de la integración con IA, una dirección relevante consiste en comparar modelos de lenguaje en tareas de evaluación y recomendación (p. ej., GPT-4o, Claude, Gemini), utilizando métricas cuantitativas y contrastes con respuestas de referencia. Complementariamente, se plantea explorar estrategias de ingeniería de prompts y, cuando sea viable, ajustes del modelo para mejorar precisión, adaptando el asistente según el dominio y nivel de detalle requerido.

Desde el punto de vista de la herramienta, se proponen extensiones que favorezcan adopción e integración: un módulo de entrenamiento para usuarios no expertos con casos ilustrativos y retroalimentación inmediata; una API para incorporar el motor de evaluación en otros sistemas (editores, plataformas colaborativas); y un modo de revisión iterativa que registre versiones sucesivas de una visualización para analizar su evolución en procesos de mejora continua.

Referencias

[1] S. Card, J. Mackinlay, and B. Shneiderman, *Readings in Information Visualisation. Using Vision to Think.: Using Vision to Think*, no. January. Morgan Kaufmann Publishers, 1999.

[2] E. Segel and J. Heer, "Narrative visualization: Telling stories with data," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 6, pp. 1139–1148, 2010, doi: 10.1109/TVCG.2010.179.

[3] W. Willett, J. Heer, J. M. Hellerstein, and M. Agrawala, "CommentSpace: Structured support for collaborative visual analysis," *Conference on Human Factors in Computing Systems - Proceedings*, pp. 3131–3140, 2011, doi: 10.1145/1978942.1979407.

[4] H. Shao, R. Martinez-Maldonado, V. Echeverria, L. Yan, and D. Gašević, "Data Storytelling in Data Visualisation: Does it Enhance the Eficiency and Efectiveness of Information Retrieval and Insights Comprehension?," *Conference on Human Factors in Computing Systems - Proceedings*, p. 21, May 2024, doi: 10.1145/3613904.3643022/SUPPL_FILE/3613904.3643022-TALK-VIDEO.VTT.

[5] M. Fansher *et al.*, "Narrative visualizations: Depicting accumulating risks and increasing trust in data," *Cogn. Res. Princ. Implic.*, vol. 10, no. 1, pp. 1–27, Dec. 2025, doi: 10.1186/S41235-025-00613-W/FIGURES/15.

[6] A. Lezcano Airaldi, J. A. Diaz-Pace, and E. Irrazábal, "Data-driven Storytelling to Support Decision Making in Crisis Settings: A Case Study," *JUCS - Journal of Universal Computer Science*, vol. 27, no. 10, pp. 1046–1068, Oct. 2021, doi: 10.3897/jucs.66714.

[7] T. Munzner, *Visualization Analysis and Design*, 1st ed. A K Peters/CRC Press, 2014.

[8] C. Ware, *Information Visualization: Perception for Design (Interactive Technologies)*, 4th ed. Elsevier, 2020. doi: 10.1016/B978-0-12-812875-6.01001-X.

[9] W. S. Cleveland and R. McGill, "Graphical Perception and Graphical Methods for Analyzing Scientific Data," *Science (1979).*, vol. 229, no. 4716, pp. 828–833, Aug. 1985, doi: 10.1126/SCIENCE.229.4716.828.

[10] E. R. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. Graphics Press, 2001. [Online]. Available: https://www.jstor.org/stable/530384?origin=crossref

[11] J. Hullman, S. Drucker, N. H. Riche, Bongshin Lee, D. Fisher, and E. Adar, "A Deeper Understanding of Sequence in Narrative Visualization," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2406–2415, Dec. 2013, doi: 10.1109/TVCG.2013.119.

[12] S. Carpendale, "Evaluating Information Visualizations," in *Information Visualization*, vol. 4950 LNCS, no. January 1970, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 19–45. doi: 10.1007/978-3-540-70956-5_2.

[13] N. Elmqvist and J. S. Yi, "Patterns for visualization evaluation," *Inf. Vis.*, vol. 14, no. 3, pp. 250–269, Jul. 2015, doi: 10.1177/1473871613513228.

[14] P. Johannesson and E. Perjons, *An introduction to design science*, vol. 9783319106. 2014. doi: 10.1007/978-3-319-10632-8.

[15] C. Edmond and T. Bednarz, "Three trajectories for narrative visualisation," *Visual Informatics*, vol. 5, no. 2, pp. 26–40, Jun. 2021, doi: 10.1016/j.visinf.2021.04.001.

[16] Z. Zhao and N. Elmqvist, "The Stories We Tell About Data: Surveying Data-Driven Storytelling Using Visualization," *IEEE Comput. Graph. Appl.*, vol. 43, no. 04, pp. 97–110, Jul. 2023, doi: 10.1109/MCG.2023.3269850.

[17] R. Kosara and J. MacKinlay, “Storytelling: The next step for visualization,” *Computer (Long Beach. Calif).*, vol. 46, no. 5, pp. 44–50, 2013, doi: 10.1109/MC.2013.36.

[18] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale, “More Than Telling a Story: Transforming Data into Visually Shared Stories,” *IEEE Comput. Graph. Appl.*, vol. 35, no. 5, pp. 84–90, Sep. 2015, doi: 10.1109/MCG.2015.99.

[19] S. F. Sakarang, M. F. Asli, and M. Hamzah, “Narrative visualization: Taxonomy, methodologies, challenges, and opportunities,” *Inf. Vis.*, vol. 24, no. 2, pp. 199–223, Apr. 2025, doi: 10.1177/14738716251315916.

[20] H. Haase, “Evaluating the quality of scientific visualizations: the Q-VIS reference model,” in *Visual Data Exploration and Analysis V*, R. F. Erbacher and A. Pang, Eds., May 1998, pp. 123–131. doi: 10.1117/12.309534.

[21] C. Forsell and J. Johansson, “An heuristic set for evaluation in information visualization,” in *Proceedings of the International Conference on Advanced Visual Interfaces - AVI ’10*, New York, New York, USA: ACM Press, 2010, p. 199. doi: 10.1145/1842993.1843029.

[22] E. Wall *et al.*, “A Heuristic Approach to Value-Driven Evaluation of Visualizations,” *IEEE Trans. Vis. Comput. Graph.*, vol. 25, no. 1, pp. 491–500, Jan. 2019, doi: 10.1109/TVCG.2018.2865146.

[23] B. S. Santos, B. Q. Ferreira, and P. Dias, “Heuristic Evaluation in Information Visualization Using Three Sets of Heuristics: An Exploratory Study,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9169, 2015, pp. 259–270. doi: 10.1007/978-3-319-20901-2_24.

[24] Y. Zhu, X. Suo, and G. S. Owen, “Complexity Analysis for Information Visualization Design and Evaluation,” in *Advances in Visual Computing*, vol. 4841 LNCS, no. PART 1, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 576–585. doi: 10.1007/978-3-540-76858-6_56.

[25] H. Padda, S. Mudur, A. Seffah, and Y. Joshi, “Comprehension of Visualization Systems - Towards Quantitative Assessment,” in *First International Conference on Advances in Computer-Human Interaction*, IEEE, 2008, pp. 83–88. doi: 10.1109/ACHI.2008.19.

[26] N. Errey, J. Liang, T. W. Leong, and D. Zowghi, “Evaluating narrative visualization: a survey of practitioners,” *Int. J. Data Sci. Anal.*, vol. 18, no. 1, pp. 19–34, Mar. 2024, doi: 10.1007/s41060-023-00394-9.

[27] A. Wu, Y. Wang, X. Shu, D. Moritz, W. Cui, and H. Zhang, “AI4VIS : Survey on Artificial Intelligence Approaches for Data Visualization,” vol. XX, no. Xx, 2018.

[28] X. Fu, Y. Wang, H. Dong, W. Cui, and H. Zhang, “Visualization Assessment: A Machine Learning Approach,” *2019 IEEE Visualization Conference, VIS 2019*, pp. 126–130, 2019, doi: 10.1109/VISUAL.2019.8933570.

[29] L. Y.-H. Lo and H. Qu, “How Good (Or Bad) Are LLMs at Detecting Misleading Visualizations?,” Jul. 2024, [Online]. Available: http://arxiv.org/abs/2407.17291

[30] S. Shin, S. Hong, and N. Elmqvist, “Visualizationary: Automating Design Feedback for Visualization Designers using LLMs,” pp. 1–13, Sep. 2024, [Online]. Available: http://arxiv.org/abs/2409.13109

[31] ISO/IEC 25010, “Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Product quality model,” 2023. Accessed: May 26, 2025. [Online]. Available: https://www.iso.org/standard/78176.html

[32] ISO/IEC 33003, “Information technology — Process assessment — Requirements for process measurement frameworks,” 2015. Accessed: Dec. 04, 2022. [Online]. Available: https://www.iso.org/standard/54177.html

[33] A. Lezcano Airaldi, E. Irrazábal, and A. Diaz Pace, “Best Practices and Evaluation Methods for Narrative Information Visualizations: A Systematic Review,” in *Proceedings of the 20th International Conference on Evaluation of Novel Approaches to Software Engineering*, SCITEPRESS - Science and Technology Publications, 2025, pp. 378–389. doi: 10.5220/0013202000003928.

[34] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York: McGraw-Hill, 1980.

[35] J. E. Leal, “AHP-express: A simplified version of the analytical hierarchy process method,” *MethodsX*, vol. 7, p. 100748, 2020, doi: 10.1016/j.mex.2019.11.021.

[36] A. Lezcano Airaldi, J. A. Carruthers, E. Irrazabal, and A. Diaz-Pace, “Measuring Narrative Visualizations: a Model Based Evaluation Framework and Validation,” *IEEE Access*, vol. 13, pp. 87350–87368, May 2025, doi: 10.1109/ACCESS.2025.3570485.

[37] C. Robson, *Small-Scale Evaluation: Principles and Practice*. Sage Publications, 2017.

[38] R. K. Yin, *Case Study Research Design and Applications*. Sage Publications Ltd., 2017.

[39] S. Lee, S.-H. Kim, and B. C. Kwon, “VLAT: Development of a Visualization Literacy Assessment Test,” *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 1, pp. 551–560, Jan. 2017, doi: 10.1109/TVCG.2016.2598920.

[40] J. Brooke, “SUS: A ‘Quick and Dirty’ Usability Scale,” in *Usability Evaluation In Industry*, 1st ed., CRC Press, 1996, pp. 207–212. doi: 10.1201/9781498710411-35.

[41] A. L. Airaldi, J. A. Diaz-Pace, and E. Irrazábal, “Data-driven Storytelling to Support Decision Making in Crisis Settings: A Case Study,” in *Memorias de las JAIIO*, Dec. 2022, pp. 65–65. Accessed: Jul. 26, 2023. [Online]. Available: https://publicaciones.sadio.org.ar/index.php/JAIIO/article/view/299

[42] A. Lezcano Airaldi, E. Irrazábal, and J. A. Diaz-Pace, “Narrative Visualizations Best Practices and Evaluation: A Systematic Mapping Study,” Jun. 2022, doi: 10.21203/RS.3.RS-1735564/V1.

Predicción del Comportamiento, Clasificación Semántica y Ponderación de Riesgos de Incendios Rurales

Carlos Brys

Facultad de Ciencias Económicas

Universidad Nacional de Misiones

carlos.brys@fce.unam.edu.ar

Información de la Tesis

Director: Dr.Cs. **David Luis La Red Martínez**

Codirector: Dr.Cs. **Marcelo Julio Marinelli**

Facultad de Ciencias Exactas, Químicas y Naturales. Universidad Nacional de Misiones. Fecha de la defensa: **24 de junio de 2025**

Motivación

La tesis doctoral surge de una preocupación creciente por el efecto de los incendios rurales en la biodiversidad y el medioambiente, un problema que se agudiza ante el cambio climático y la actividad antropogénica.

La motivación principal es la necesidad apremiante de contar con herramientas tecnológicas para mejorar la detección de incendios, apoyar la toma de decisiones, disminuir el tiempo de respuesta y optimizar la gestión de los recursos para prevenir y mitigar el impacto de los incendios rurales, que presentan un aumento en la cantidad de incidentes y su consecuente efecto en el medio ambiente.

Los modelos matemáticos tradicionales usados comúnmente, como el *Índice Meteorológico de Peligro de Incendios Forestales (FWI)* presentan limitaciones, especialmente en regiones con vegetación diversa. Este índice, desarrollado en Canadá, no considera las características específi- cas de la combustibilidad de la vegetación heterogénea, lo que puede llevar a una subestimación del riesgo real de incendios en zonas rurales con transición de diferentes tipos de uso del suelo.

Es imperioso disponer de algoritmos computacionales más precisos, adaptados a las condiciones específicas de distintos ámbitos geográficos para aplicarlos en la gestión de incendios rurales,

con el fin de proteger la biodiversidad, el patrimonio cultural y la seguridad de las personas.

Se pretende superar las limitaciones de los modelos tradicionales existentes, como el FWI, y ofrecer una solución innovadora integral para la evaluación del riesgo y la predicción del comportamiento del fuego, integrando herramientas computacionales y tecnologías emergentes para la gestión eficiente de los recursos en la lucha contra los incendios rurales.

En este sentido, se destacan los conceptos que fundamentan la motivación de la tesis:

- **Necesidad de un índice adaptado:**
  Existe una clara necesidad de desarrollar un **Índice de Riesgos de Incendios Rurales (IRIR)** adaptado a distintos contextos geográficos, que mejore la precisión del FWI. Este nuevo índice debe considerar factores como la diversidad de la cobertura vegetal, datos topográficos, geoespaciales y meteorológicos. Impulsan esta motivación la escasa investigación global sobre el desarrollo de índices alternativos de peligro de incendio, donde la mayoría de los estudios se centran en adaptar el FWI en lugar de desarrollar nuevos índices más precisos y completos.
- **Importancia del área de estudio:**
  La investigación se centra en la "Ecorregión de la Selva Misionera", una zona de gran valor socioeconómico y alta biodiversidad. Los incendios en esta región, provocados por la actividad antropogénica, afectan gravemente a los ecosistemas locales. La necesidad de proteger la biodiversidad y el patrimonio cultural de las comunidades aborígenes que habitan en esta región añade una urgencia adicional a este estudio.
- **Limitaciones de los sistemas de alerta temprana:**
  Los sistemas de alerta temprana existentes tienen limitaciones, y esta investigación pre- tende abordarlas mediante el desarrollo de un modelo computacional predictivo. Es ne- cesario proporcionar información precisa e inmediata a las autoridades competentes para identificar, modelar, clasificar y evaluar la probabilidad de ocurrencia de los incendios, permitiendo una gestión más eficiente de los recursos y las estrategias de mitigación.
- **Brechas del conocimiento:**
  La falta de métodos predictivos adaptados a distintos contextos geográficos exige el desa- rrollo de *algoritmos computacionales predictivos para determinar el comportamiento del fuego y un índice de riesgo de incendios rurales alternativo al FWI*, que tengan en cuen- ta las características específicas de la región en estudio y las condiciones locales para proporcionar una evaluación más precisa del riesgo de incendios.
- **Marco teórico y metodológico:**
  La necesidad de un marco teórico sólido, que explore y modele conceptos clave como la propagación del fuego, los factores que influyen en su desarrollo y la relevancia de los sistemas de información geográfica. La formulación de una metodología cuantitati- va/cualitativa para examinar la bibliografía relacionada con la propagación de incendios, el riesgo potencial y la influencia del clima y la topografía.

- **Desarrollo de un modelo computacional innovador:**
  La propuesta central de la tesis es el desarrollo de un **modelo computacional predicti- vo** que utilice las técnicas más contemporáneas para obtener resultados más precisos que el FWI. El modelo propuesto considera las características específicas de la región geográ- fica en estudio, las condiciones ambientales locales y datos en tiempo real provenientes de fuentes abiertas. Se enfoca en la predicción del comportamiento del fuego, la clasificación semántica del peligro y la ponderación de los riesgos.

Aporte a la disciplina

La principal contribución de este trabajo es un algoritmo predictivo de la *Velocidad de Propa- gación del Fuego* (VPF), que considera variables estáticas y altamente dinámicas que denota una *Escala de Peligro de Incendios* (EPI), basada en la proximidad de los focos de calor a elementos críticos que se calcula en base a los siguientes parámetros:

- Una velocidad inicial de propagación del fuego que depende del tipo de combustible en el suelo,
- La humedad del suelo,
- La velocidad del viento,
- La radiación solar en una hora del día.

A diferencia de los otros modelos mencionados anteriormente que calculan la tasa de propa- gación (*Rate of Spread ROS*) y requieren conocer de antemano el valor de múltiples variables, el modelo sólo necesita un par de coordenadas geográficas para dar una predicción del com- portamiento del fuego, ya que todas las variables necesarias para el cálculo se obtienen online, en tiempo real y en cualquier momento, consultando a través de APIs de sistemas de código abierto. Esto hace que el modelo sea mucho más dinámico y no requiera la intervención de expertos ni el preprocesamiento de datos.

Para predecir el comportamiento de un incendio rural, el modelo incorpora y combina otros factores multivariantes no contemplados por la fórmula de *Rothermel*, como datos ambientales reales y pasados, uso del suelo, hora del día y topografía. Se propone un método de agregación multidimensional (tiempo, espacio y entorno) para calcular la velocidad de propagación del fuego (VPF) y predecir el flujo de desplazamiento.

El siguiente aporte de la tesis es un *Modelo Computacional* para calcular el potencial de que ocurra un incendio en el tiempo inmediato denominado **"Índice de Riesgo de Incendios Rurales" (IRIR)**, que tiene en cuenta variables estáticas como la inflamabilidad del combus- tible en el terreno y otras variables altamente dinámicas, como datos localizados en tiempo real, como la meteorología.

El IRIR es un algoritmo alternativo al FWI que representa el riesgo de ignición del combustible en el suelo, expresado en un *Indice de Inflamabilidad* (II) que tiene en cuenta las características específicas de la región y las condiciones locales para proporcionar una evaluación más precisa del riesgo de incendios rurales. Usando etiquetas lingüísticas se expresan los niveles de riesgo en una *Escala de Riesgo de Incendios Rurales* (ERI).

El IRIR es un indicador centrado en el ser humano para enfatizar el riesgo percibido de inicia- ción de incendios debido a la actividad antropogénica y representa el riesgo de inflamabilidad del combustible sobre el terreno, basado en el porcentaje de cobertura de diferentes tipos de combustible vegetal en la superficie de un área.

El IRIR es más preciso a la hora de clasificar el riesgo de ignición de un incendio en una región. Para aumentar sustancialmente la precisión, se utilizan métodos de aprendizaje automático para realizar una clasificación automática del uso del suelo a partir de imágenes satelitales. Por último, para facilitar la interpretación de los resultados se usa la lógica difusa para asignar etiquetas lingüísticas a los rangos de valores.

El IRIR se destaca por su enfoque dinámico y su capacidad para utilizar datos en línea sin la necesidad de intervención humana ni preprocesamiento de datos. Además, incorpora técnicas de inteligencia artificial como el aprendizaje automático para realizar una clasificación del uso del suelo a partir de imágenes satelitales. También mejora la interpretación de los resultados asignando etiquetas lingüísticas a los rangos de valores absolutos, lo que facilita una comprensión más intuitiva de los riesgos de incendio en las regiones analizadas.

La innovación y novedad de esta propuesta de solución frente a los modelos estudiados, donde se muestra que el modelo permite analizar situaciones en tiempo real, en línea, de cualquier foco de incendio detectado en la región bajo supervisión, directamente desde una página web y sin necesidad de recurrir a un centro de proceso de datos. Todos los cálculos se realizan bajo demanda utilizando el lenguaje Javascript en el navegador web del cliente. Como requisito tecnológico, este modelo sólo necesita un servidor web estándar y una conexión a las fuentes de datos abiertas para acceder a los recursos gratuitos en línea mediante consultas a las APIs.

Se creó una *Sala de Crisis Virtual* que se manifiesta en un *Mapa Interactivo para la Gestión de Emergencias Ambientales*: (https://sig.misiones.gob.ar/mapas/emergencia/) con actualizaciones periódicas, donde se puede consultar en tiempo real las características de los focos de calor detectados por los sensores satelitales, la clasificación del riesgo y los factores que inciden en su comportamiento.

Se creó además un mapa interactivo para visualizar del *Índice de Riesgo de Incendios Rurales* (https://sig.misiones.gob.ar/mapas/irir/) donde se presenta mediante una visualiza- ción de mapa de calor y un mapa de colores categóricos para los valores difusos de IRIR según la clasificación semántica. Esta representación emplea una paleta de colores para ilustrar la va-

Cuadro 1: Comparación General de los Trabajos Relacionados y el Modelo Propuesto

Modelos Revisados	Modelo Propuesto
Descriptivo	Predictivo
Análisis post-facto fuera de línea	Análisis bajo demanda en un mapa inter-activo en tiempo real
Imágenes de satélite	Datos descargados cada hora de FIRMS-NASA
Imágenes históricas	Datos en tiempo real
Procesamiento de imágenes	Análisis de datos
Información meteorológica histórica	Información meteorológica en tiempo real
No considera la relación de la dirección del viento con el aspecto del terreno	Considera la relación de la dirección del viento con el aspecto del terreno
No considera la hora del día	Considera como factor potencial la hora del día en que se analiza cada fuente de calor
Análisis de áreas quemadas	Análisis de puntos de cualquier coordenada LatLon en la región
Informes históricos	Informes en tiempo real bajo demanda

riación del índice de riesgo de incendio rural en distintas regiones geográficas. La visualización es útil para identificar áreas de mayor riesgo, facilitando a los usuarios una mejor comprensión del IRIR y permitiéndoles tomar medidas preventivas para mitigar el riesgo de incendios rurales y actuar de manera adecuada.

Esta contribución a la gestión del riesgo de incendios rurales proporciona a las autoridades una valiosa herramienta para tomar decisiones mejor informadas basadas en evidencia y mitigar los efectos dañinos de los incendios.

Trabajos Futuros

Sobre la base de los avances y contribuciones de este estudio, varias áreas serían objeto de futuras investigaciones. Entre ellas, la precisión del algoritmo de clasificación de focos de incendio podría mejorarse mediante técnicas avanzadas de aprendizaje automático, redes neuronales y análisis de datos. Esto permitiría una mejor identificación y priorización de los focos potencialmente peligrosos y una respuesta más rápida y eficaz por parte de los organismos responsables.

Integración con Soluciones Existentes y Futuras Aplicaciones

El Modelo Computacional de Comportamiento de Incendios Rurales se integra con soluciones existentes a través de la discusión de las implicaciones prácticas de los resultados y median- te las recomendaciones para su implementación en entornos reales. Además, se prevén varias aplicaciones futuras, como la mejora de la precisión de los modelos de predicción mediante la

incorporación de nuevas variables y técnicas avanzadas de análisis de datos. También se sugiere la exploración de enfoques innovadores para la gestión de incendios, incluyendo el uso de tecno- logías emergentes y la integración de sistemas de apoyo a la toma de decisiones. Se identifican áreas específicas donde se puede profundizar, como la evaluación de riesgos en contextos espe- cíficos y la adaptación de los modelos existentes a diferentes entornos geográficos y climáticos. Se enfatiza la importancia de la colaboración interdisciplinaria y el seguimiento continuo de avances en el campo para abordar los desafíos actuales en la predicción y gestión de incendios rurales.

Se plantean las posibles aplicaciones de los hallazgos encontrados en el estudio sobre el com- portamiento del fuego en entornos rurales que incluyen:

- Mejora de las políticas públicas y estrategias gubernamentales, para prevenir y mitigar los efectos nocivos de los incendios rurales en diferentes contextos socioeconómicos, lo cual podría implicar la evaluación de programas de educación y concientización, así como la mejora de infraestructuras y equipamiento para la prevención y el control de incendios.
- Utilización de los resultados para sensibilizar a la población sobre los riesgos de incendios rurales y promover prácticas seguras de manejo del fuego en áreas propensas a incendios, con el fin de aumentar la conciencia pública y fomentar la participación en acciones de prevención.
- Contribución a una gestión más efectiva del riesgo de incendios rurales y a la protección de comunidades y ecosistemas vulnerables mediante la aplicación de los hallazgos en el ámbito del manejo del fuego y en la toma de decisiones.
- Posible impacto en la evolución de teorías existentes y generación de nuevas perspectivas teóricas en el campo de estudio, enriqueciendo el marco teórico-conceptual y abriendo nuevas líneas de investigación.
- Mejora de la precisión del algoritmo de clasificación de focos de incendio mediante téc- nicas avanzadas de aprendizaje automático, redes neuronales y análisis de datos, lo que permitiría una mejor identificación y priorización de atención a los focos potencialmente peligrosos, y una respuesta más rápida y eficaz por parte de los organismos responsables.
- Exploración de cómo los cambios en el clima y el medio ambiente pueden afectar la propa- gación y el comportamiento del fuego en diferentes regiones geográficas, con el objetivo de predecir mejor los patrones de los incendios rurales y desarrollar estrategias más eficaces de prevención y mitigación.
- Investigación sobre cómo las ontologías pueden mejorar la interoperabilidad y la integra- ción de datos entre los diferentes sistemas y organizaciones implicados en la gestión del riesgo de incendios rurales, lo que permitiría una mejor coordinación entre las distintas partes involucradas y una respuesta más rápida ante situaciones críticas.

Tendría un impacto significativo investigar cómo los cambios en el clima y el medio ambiente pueden afectar a la propagación y el comportamiento del fuego en diferentes regiones geográficas.

Esto podría ayudar a predecir mejor los patrones futuros de los incendios rurales y a desarrollar estrategias más eficaces de prevención y mitigación.

Se podría explorar cómo las ontologías pueden mejorar la interoperabilidad y la integración de datos entre los diferentes sistemas y organizaciones implicados en la gestión del riesgo de incendios rurales. Esto permitiría una mejor coordinación entre las distintas partes involucradas y una respuesta más rápida ante situaciones críticas.

Además, es necesario un desarrollo adicional de la taxonomía y la ontología de incendios pro- puestas para mejorar la comunicación interinstitucional. La solución propuesta puede abordar eficazmente las limitaciones asociadas a la combinación de datos abiertos y conocimientos exper- tos sobre incendios mediante la actualización y el perfeccionamiento continuos de la ontología, integrando al mismo tiempo técnicas de aprendizaje automático. Estas actualizaciones, a su vez, conducirán a una mayor precisión y fiabilidad en la predicción del riesgo y el comportamiento de los incendios rurales.

Sería interesante evaluar cómo las tecnologías emergentes, como ser: los sensores remotos autó- nomos conectados a la Internet de las cosas (IoT), los drones y vehículos aéreos no tripulados (VANT) y los satélites, pueden mejorar la detección temprana y la vigilancia en tiempo real de los incendios rurales y determinar las zonas más susceptibles de incendiarse. Esto podría ayudar a detectar los incendios con mayor rapidez e intervenir en su extinción antes de que se propaguen.

Además, podría analizarse cómo se pueden mejorar las políticas públicas y las estrategias gu- bernamentales para prevenir y mitigar los efectos nocivos de los incendios rurales en diferentes contextos socioeconómicos. Esto podría incluir la evaluación de los programas de educación y concientización, la mejora de las infraestructuras y el equipamiento para la prevención y el control de incendios, y la identificación de las mejores prácticas para implicar a las comunidades locales en la gestión del riesgo de los incendios rurales.

A partir del trabajo de investigación propuesto sobre el Índice de Riesgo de Incendios Rurales (IRIR), existen varias vías de investigación futuras y áreas de mejora que explorar:

- Validación y generalización:
  Si bien el índice IRIR mostró resultados prometedores en la provincia de Misiones de Argentina, es esencial validar su desempeño en otras regiones con diversos tipos de ve- getación y condiciones ambientales. Probar el modelo en varias localizaciones geográficas demostrará su generalización y robustez.
- Incorporación de factores de cambio climático:
  Dado que el cambio climático influye en el comportamiento de los incendios, la integración de variables de cambio climático en el modelo IRIR puede mejorar su capacidad de pre- dicción. Parámetros como las tendencias de temperatura, los fenómenos meteorológicos

extremos y los cambios en los patrones de precipitación deberían tenerse en cuenta para evaluar cómo afectan al potencial de ignición en las zonas rurales.

- Integración dinámica de datos:
  La fuerza del IRIR reside en su capacidad para utilizar datos centrados en tiempo real, como la información meteorológica, para evaluar el riesgo de incendio en un futuro in- mediato. La ampliación del modelo para incluir otros datos relevantes en tiempo real, como patrones de actividad humana o datos de teledetección, podría mejorar aún más su precisión.
- Resolución a escala fina:
  Investigar los beneficios potenciales de aumentar la resolución espacial del índice IRIR. Esto podría implicar el análisis del riesgo de incendio a una escala más pequeña, lo que permitiría una identificación más precisa de las zonas de alto riesgo y ayudaría en los esfuerzos de prevención de incendios.
- Mejoras en el aprendizaje automático:
  Perfeccionar continuamente los algoritmos de aprendizaje automático utilizados en el mo- delo IRIR para mejorar la precisión de las predicciones. Investigar técnicas avanzadas de aprendizaje automático, como el aprendizaje profundo, para capturar patrones y relacio- nes complejas dentro de los datos.
- Big Data y Analítica de Datos:
  El análisis de grandes volúmenes de datos, conocido como Big Data, permite identificar patrones complejos que no son evidentes para los humanos, lo que resulta crucial para predecir incendios al integrar múltiples fuentes de información, como datos meteorológicos, satelitales, sociales y de sensores, mejorando así la precisión de las predicciones; además, el uso de modelos climáticos basados en datos históricos ayuda a anticipar cómo el cambio climático influirá en la frecuencia e intensidad de estos eventos, mientras que herramientas como *SparkR* (El servidor Apache Spark integrado con el lenguaje R) permiten realizar simulaciones en tiempo real para modelar rápidamente la propagación de incendios y respaldar decisiones estratégicas.
- Aplicar la Tecnología LIDAR (*Light Detection and Ranging*):
  LIDAR utiliza pulsos láser para generar mapas tridimensionales detallados de la vege- tación y el terreno, lo que permite aplicaciones clave en la gestión de incendios rurales, como el mapeo de combustible para identificar áreas con alta densidad de biomasa seca que podrían alimentar incendios, así como el modelado de la propagación del fuego pa- ra simular cómo éste podría desplazarse a través del terreno y la vegetación, brindando información crucial para la planificación y respuesta ante emergencias.
- Blockchain para Gestión de Datos:
  La Blockchain puede ser utilizada para garantizar la transparencia y seguridad en el intercambio de datos relacionados con incendios forestales, permitiendo aplicaciones como el seguimiento de recursos para monitorear el uso de agua, equipos y personal durante la

extinción de incendios, así como la verificación de datos para asegurar que la información recopilada por sensores y satélites sea confiable y no manipulada, además de que proyectos piloto están explorando su uso para coordinar esfuerzos internacionales en la lucha contra incendios.

- Desarrollo de la interfaz de usuario:
  Crear una interfaz fácil de usar para que las agencias de gestión de incendios y las auto- ridades locales puedan acceder e interpretar fácilmente los datos de IRIR. Esta interfaz podría proporcionar mapas de riesgo en tiempo real, tendencias históricas y umbrales de riesgo personalizables para la alerta temprana y la toma de decisiones estratégicas.
- Incorporación de factores sociales:
  Considerar la integración de factores sociales en el modelo IRIR para comprender y pre- decir mejor el riesgo de inicio de incendios debido a actividades humanas. Factores como la densidad de población, los patrones de uso del suelo y los incidentes históricos de incendios podrían proporcionar información valiosa.
- Colaboración y puesta en común de datos:
  Fomentar la colaboración con otras instituciones de investigación, agencias de gestión de incendios y partes interesadas para compartir datos y conocimientos. Este enfoque colaborativo puede conducir a una comprensión más completa de los factores de riesgo de incendios y fomentar el desarrollo de estrategias de gestión de incendios más eficaces.
- Actualización del Índice IRIR:
  Dado que el comportamiento del fuego y el potencial de ignición pueden cambiar con el tiempo debido a diversos factores, debe asegurarse de que el índice IRIR se actualiza y perfecciona periódicamente para reflejar las condiciones cambiantes y los nuevos hallazgos de la investigación.
- Estudios comparativos:
  Llevar a cabo estudios comparativos entre el índice IRIR y otros sistemas existentes de evaluación del riesgo de incendios rurales en varias regiones y bajo diferentes condiciones. Esto proporcionará una perspectiva más amplia sobre los puntos fuertes y débiles del modelo IRIR y su potencial de adopción en diferentes contextos.
- Incorporación de datos de redes sociales:
  Las plataformas de redes sociales se utilizan cada vez más como fuente de información en tiempo real durante las catástrofes naturales. La investigación futura debería centrarse en el análisis y la integración de datos de redes sociales en el algoritmo propuesto para mejorar la precisión de la evaluación y predicción del riesgo de incendio.
- Integrar la inteligencia artificial a un sistema de apoyo a la toma de decisiones:
  El modelo propuesto proporciona información que puede utilizarse para la toma de de- cisiones estratégicas y para predecir el comportamiento de los incendios rurales. La in- vestigación futura se puede orientar en integrar la inteligencia artificial generativa a un sistema de apoyo a la toma de decisiones (SSD), que pueda recomendar estrategias de mi-

tigación y ayudar a los responsables a interpretar los datos para tomar decisiones basadas en evidencia en tiempo real.

Título: *Semánticas Argumentativas Basadas en Diálogo: Fundamentos Teóricos, Representación y Análisis Formal*

Tesista: Lic. Soto, Yamil Osvaldo Omar
Correo electrónico: yamil.soto@cs.uns.edu.ar
Directores: Dr. Simari, Gerardo Ignacio y Dr. Falappa, Marcelo Alejandro
Fecha de defensa: 13 de febrero de 2026
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, San Andrés 800, Bahía Blanca, Buenos Aires, Argentina

Resumen

Este artículo resume la tesis de doctorado en Ciencias de la Computación titulada *Semánticas Argumentativas Basadas en Diálogo: Fundamentos Teóricos, Representación y Análisis Formal*, de ahora en más, la tesis, la cual comprende el estudio de las semánticas argumentativas basadas en diálogo en el contexto de la Argumentación Computacional abstracta basada en inferencia. La tesis presenta el enfoque basado en *bundle sets*, un formalismo general y modular para la caracterización de semánticas argumentativas basadas en diálogo capaz de representar tanto semánticas existentes como nuevas propuestas. Dicho formalismo es utilizado para caracterizar un conjunto de semánticas, entre ellas la aceptabilidad-*ac*, la cual es un aporte novedoso de la tesis. La aceptabilidad-*ac* introduce un comportamiento no presente en la literatura de las semánticas basadas en diálogo que consiste, intuitivamente, en rechazar a aquellos argumentos involucrados en una situación dudosa, a la que llamaremos indecisión, y confinar los efectos de la misma solo a los argumentos que la hayan originado. El desarrollo de la aceptabilidad-*ac* responde (principalmente) al interés por contar con semánticas que ofrezcan un tratamiento alternativo a los ciclos de derrotas de longitud impar, situaciones en las cuales las semánticas basadas en admisibilidad ofrecen resultados poco satisfactorios. La aceptabilidad-*ac* y el resto de semánticas basadas en diálogo abordadas en la tesis son formalmente analizadas a través de un conjunto de principios originales. Dicho análisis nos permite profundizar en las propiedades de las semánticas estudiadas y en las relaciones que existen entre las mismas, lo que a su vez nos permite guiar el desarrollo de nuevas semánticas y fomentar su adopción en sistemas de representación del conocimiento y razonamiento. Este resumen extendido se organiza como sigue: en la Sección 1 se introducen las bases de la Argumentación Computacional y de las semánticas argumentativas. Las contribuciones de la tesis y las motivaciones detrás de cada una de ellas se detallan en la Sección 2. En la Sección 3 se listan las publicaciones correspondientes a los distintos capítulos de la tesis. Finalmente, en la Sección 4 se delinean los trabajos futuros.

1. Preliminares

La tesis se enmarca en el área conocida como *Argumentación Computacional*, un área de la Inteligencia Artificial—abordada en la tesis exclusivamente desde la perspectiva de la *Inteligencia Artificial Simbólica* [19] (para más detalles respecto a aproximaciones no simbólicas a la Argumentación Computacional, recomendamos [10])—que se ocupa de múltiples tópicos vinculados con la noción de argumento y de las relaciones que pueden establecerse entre los mismos. En otras palabras, en este campo se propone representar y razonar acerca de la información de una manera fuertemente inspirada en cómo las personas construyen, intercambian y analizan argumentos en su día a día para desenvolverse en entornos donde la información puede ser controvertida, incompleta y/o inconsistente [2].

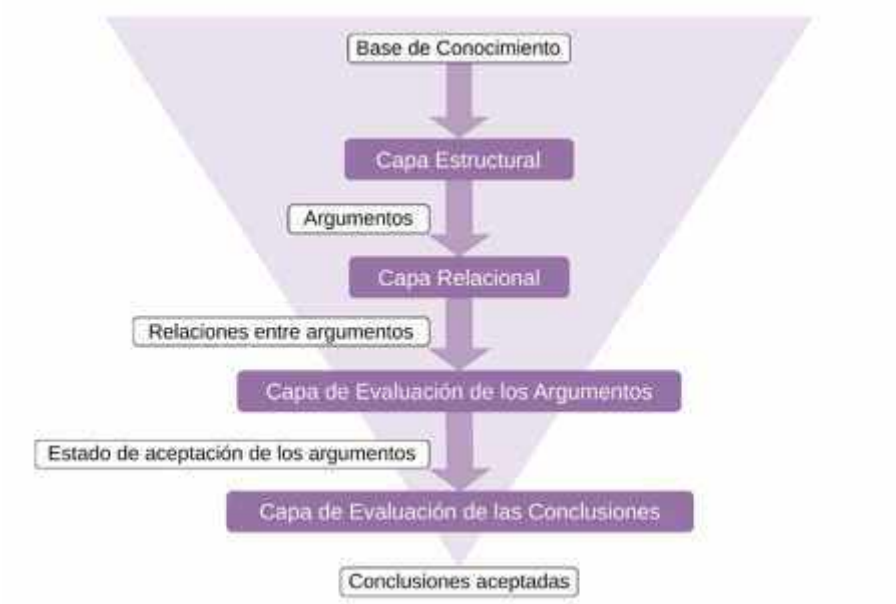


Figura 1: Capas del enfoque argumentativo basado en inferencia.

La tesis adhiere al *enfoque basado en inferencia* [24]—también llamado *monológico* y *relacional*—para la Argumentación Computacional. Este enfoque se centra en determinar qué conjunto(s) de piezas de información pueden extraerse de un único cuerpo de información posiblemente incompleto, inconsistente e incierto. A pesar de las diferencias que pueden existir entre los trabajos que adoptan esta perspectiva, podemos considerar un conjunto de pasos o capas lo suficientemente generales como para adaptarse a todos ellos. Una *capa estructural*, donde los argumentos son construidos a partir del cuerpo de información representado; una *capa relacional*, que identifica y representa las relaciones existentes entre los argumentos; una *capa de evaluación* de los argumentos, que ofrece un método formal para resolver los conflictos presentes entre los argumentos y, de esta manera, determinar su estado de aceptación o justificación; y finalmente una *capa de evaluación de las conclusiones*, donde en función de qué argumentos han sido aceptados—esto es, considerados como (tentativamente) verdaderos—se establece qué conclusiones deben aceptarse. Estas capas son ilustradas en la Figura 1. Aquellos formalismos que adoptan este enfoque utilizan las capas anteriormente mencionadas como nociones intermedias para representar una noción no monotónica (o rebatible) de consecuencia lógica. De esto se sigue que el propósito principal de la argumentación, en este contexto, es el de la inferencia, lo que da nombre al enfoque.

De lo anterior se sigue que la Argumentación Computacional constituye un marco general para el tratamiento de información contradictoria, que también puede ser incompleta e incierta. Además, su método para la toma de decisiones ofrece explicaciones sobre las mismas, aspecto que contribuye a aumentar el entendimiento, la transparencia y la confianza en los sistemas computacionales basados en Inteligencia Artificial, algo crucial a la luz de la rápida y generalizada adopción de los mismos en nuestra vida diaria, fenómeno que ha revitalizado fértiles discusiones sobre, por ejemplo, la *Inteligencia Artificial Explicable* [1]. De esto se sigue que el campo de la Argumentación Computacional es un paradigma atractivo y eficaz para la representación del conocimiento y el razonamiento en Inteligencia Artificial. Las diferentes formas en que los argumentos y sus relaciones pueden ser construidas y evaluadas ofrecen una amplia gama de posibilidades y aplicaciones, tales como la filosofía [31], el razonamiento jurídico [25], la teoría de

juegos [26], la democracia digital [5], la lingüística [21], las aplicaciones de software conversacionales [8], el razonamiento clínico [20], entre otras.

Además del enfoque basado en inferencia, la tesis adopta la perspectiva para la Argumentación Computacional conocida como *argumentación abstracta* [16]. Este paradigma cuenta con dos conceptos principales, el de *marco argumentativo* y el de *semántica argumentativa*. Un marco argumentativo es un grafo dirigido donde los nodos representan a los argumentos y los arcos representan una relación de derrota entre los mismos. La Figura 2 (a) ilustra un marco argumentativo donde los argumentos son representados como triángulos y los arcos dirigidos representan las relaciones de derrota de tal manera que un arco dirigido desde un argumento $\mathbf{X}$ a otro argumento $\mathbf{Y}$ representa que el primero derrota al segundo.

Si bien la argumentación abstracta omite varios elementos relevantes de la argumentación—como la estructura de los argumentos—ofrece un paradigma sencillo y con bases teóricas sólidas, enfocado en la evaluación de argumentos. Dicho de otro modo, este enfoque se abstrae de la construcción de los argumentos, de la determinación de las relaciones entre ellos y de la identificación de las conclusiones aceptadas, para centrarse en determinar qué (conjunto(s) de) argumentos pueden ser aceptados. Esto es representado por el segundo concepto fundamental de la argumentación abstracta. Esencialmente, una *semántica argumentativa* es una función que asigna a un marco argumentativo un conjunto de conjuntos de argumentos, cada uno de los cuales representa una postura razonable (de acuerdo a ciertos principios) que un agente racional puede adoptar.

Las semánticas argumentativas son uno de los tópicos más relevantes en la argumentación abstracta, motivo por el cual son objeto de múltiples y continuos estudios. En la literatura se han propuesto varias (familias de) semánticas argumentativas (por ejemplo, [16, 6, 13]), donde cada una de ellas caracteriza distintos comportamientos razonables que se adaptan de mejor manera a diferentes esquemas de razonamiento y/o dominios de aplicación. Además, existen diversas propuestas para definirlas, como el enfoque extensional [16], el basado en la recursividad de los componentes fuertemente conexos [4], el basado en etiquetados [23] y el basado en diálogo [22].

El último enfoque mencionado es al que adhiere la tesis y consiste en determinar el estado de aceptación de los argumentos mediante juegos de suma cero, denominados diálogos, en los que intervienen dos partes contrapuestas: el *proponente*, que intentará defender un argumento en discusión, i.e., mostrar que es (tentativamente) verdadero; y el *oponente*, que atacará dicho argumento con la intención de desmentir su veracidad. El primer movimiento del diálogo siempre corresponde al proponente, quien presenta un argumento inicial que es puesto a prueba. Luego, el oponente y el proponente se turnan para intercambiar argumentos que derroten el último argumento introducido por su contraparte. Este intercambio produce una secuencia de argumentos usualmente denominada *disputa*. Una disputa es ganada por uno de los participantes del diálogo en caso de que sea él quien logre concluirla, esto es, el último elemento de la disputa corresponde al argumento de un jugador tal que su contraparte no es capaz de presentar ningún contra-argumento. Esto, sin embargo, no finaliza el diálogo, debido a que la parte vencida puede retroceder a alguna parte previa de la disputa y extenderla, comenzando una nueva disputa en el contexto del mismo diálogo. Luego, el conjunto de todas las disputas que comparten el mismo argumento inicial representa todas las maneras posibles en que puede desarrollarse un diálogo. Estas disputas pueden representarse en una estructura de tipo árbol denominada *árbol de disputa*, donde la raíz es el argumento en discusión y cada rama corresponde a una disputa. No obstante, no toda disputa es razonable, por lo cual, necesitamos un mecanismo que nos permita enfocarnos en un subconjunto de las mismas. Este mecanismo se

denomina *protocolo* (o *función de movimiento legal*) y consiste en un conjunto de reglas que determinan cuáles son los movimientos legales disponibles para cada participante en cada momento del diálogo. En consecuencia, el protocolo rige el movimiento de los argumentos en el juego, y distintos protocolos permiten caracterizar diferentes semánticas argumentativas. Un argumento es aceptado de acuerdo a una semántica argumentativa basada en diálogo si el proponente, siguiendo el protocolo adjudicado por la semántica, es capaz de defenderlo de todos sus derrotadores. Esto es caracterizado formalmente a través de la noción de *estrategia ganadora*, la cual es un subárbol del árbol de disputa tal que toda rama finaliza en un argumento del proponente, y para todo argumento del oponente el proponente presenta un único contra-argumento.

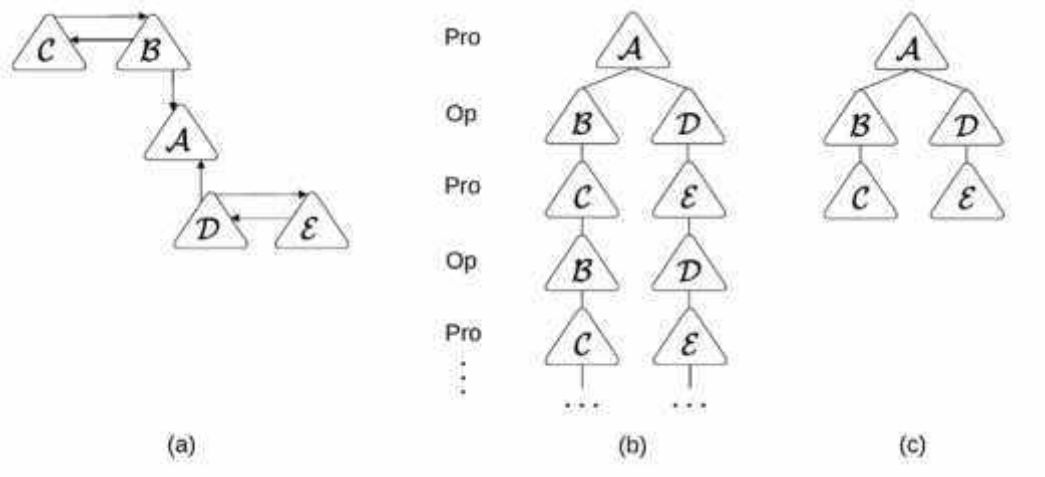


Figura 2: (a) un marco argumentativo. (b) un árbol de disputa infinito para el argumento **A**. (c) una posible estrategia ganadora para **A**, donde asumimos un protocolo que prohíbe repetir argumentos.

La Figura 2 (b) ilustra un árbol de disputa para el argumento **A** del marco argumentativo de la Figura 2 (a). El mismo es representado como un árbol en el cual un nodo **N**, representado por un argumento **X**, tiene un hijo **N'**, representado por un argumento **Y**, solo si **Y** derrota a **X**. Las disputas de nuestro árbol de disputa son infinitas debido a los ciclos de derrotas que existen entre los argumentos **B** y **C**, y **D** y **E**. La Figura 2 (c) muestra una estrategia ganadora para el argumento **A** en la cual asumimos un protocolo que prohíbe la repetición de argumentos en una disputa.

Las semánticas basadas en diálogo resultan atractivas por múltiples razones. Ofrecen un punto de vista complementario a otros enfoques—como los basados en extensiones o etiquetados—para analizar los fundamentos de las semánticas argumentativas. Al centrarse en la evaluación individual de los argumentos sin requerir el cálculo de conjuntos completos de extensiones o etiquetados, son una base adecuada para el desarrollo de algoritmos. Además, pueden aplicarse diversas técnicas a los diálogos para mejorar su eficiencia, lo que permite el desarrollo de algoritmos eficientes para la evaluación de argumentos [11]. En sistemas multi-agente, estas semánticas pueden mejorar la comunicación entre los agentes, lo que a su vez contribuye a la efectividad de la coordinación de los agentes para la realización de tareas. Para aquellos sistemas que traten con información delicada, pueden proporcionar explicaciones de por qué cierta pieza de información—un argumento o una conclusión—se considera (tentativamente) cierta. Esto puede contribuir a la explicabilidad, transparencia y confiabilidad de estos sistemas. Lo anterior resulta particularmente significativo para el desarrollo de herramientas de representación de conocimiento y razonamiento basadas en argumentación, aplicables a entornos del mundo real [18, 20].

2. Motivación y Contribuciones

A continuación discutiremos las principales contribuciones expuestas en la tesis. Para cada una de ellas, comentaremos las razones que motivaron su desarrollo.

2. 1. Enfoque Basado en Bundle Sets

La primera de nuestras contribuciones es el formalismo denominado *enfoque basado en bundle sets*[1], el cual tiene como objetivo representar formalmente, y de manera general y modular, a multitud de semánticas argumentativas basadas en diálogo, tanto existentes como aún por desarrollar. Otros formalismos existen para dicho fin, como los propuestos en [17, 22], no obstante, según nuestro leal saber y entender, ninguno de ellos nos permite cumplir con nuestro objetivo de representar satisfactoriamente el conjunto de semánticas argumentativas que nos interesa. Por tanto, basándonos en lo expuesto en [9], adaptamos e incorporamos algunas definiciones para satisfacer los requerimientos específicos de nuestra línea de investigación.

El concepto principal del enfoque basado en *bundle sets* es aquel que le da nombre: *bundle set*. Esta noción se refiere a un diálogo acerca de un determinado argumento y es el análisis de esta estructura lo que nos permite establecer el estado de aceptación del argumento bajo discusión. Esencialmente, un bundle set es un conjunto de *líneas argumentativas*, donde cada línea argumentativa es una secuencia de argumentos tal que cada elemento de la misma es un derrotador de su predecesor inmediato. Cada línea argumentativa en un bundle set representa una disputa diferente acerca del mismo argumento, por lo tanto, las líneas argumentativas siempre comparten al menos un argumento, aquel sobre el que se debate, pero también siempre difieren en al menos otro. En este enfoque la caracterización de distintas semánticas argumentativas se encuentra embebida en el concepto de *aceptabilidad*, el cual representa un conjunto de reglas que determina cuáles son los argumentos que pueden utilizarse en una discusión. Basta con modificar este conjunto de reglas para capturar diferentes semánticas argumentativas, por lo tanto, la definición de las mismas es modular. Por lo anterior, nuestra propuesta sirve como un marco teórico para la representación y análisis de múltiples semánticas y la obtención de conclusiones generales a todas ellas. Este punto está estrechamente vinculado con nuestra última contribución.

2. 2. Aceptabilidad Basada en la Anulación de Componentes Aumentados

Nuestra siguiente contribución consiste en la definición de una semántica argumentativa que exhibe un comportamiento ausente en la literatura sobre semánticas basadas en diálogo. Denominaremos a dicha semántica como *aceptabilidad basada en la anulación de componentes aumentados*, *aceptabilidad-ac* para abreviar. Esencialmente, la aceptabilidad-*ac* cuenta con un comportamiento muy específico respecto a dos puntos estrechamente vinculados con un tipo particular de conjunto de argumentos, al cual denominaremos *indecisión*. Una indecisión es un conjunto de argumentos donde cada miembro puede afectar el estado de aceptación de los demás miembros, a la vez que ninguno de ellos se ve afectado por argumentos externos al conjunto. En estos casos, no contamos con información suficiente para decidir a favor de uno (o más) de los miembros del conjunto por sobre los demás, de ahí el nombre de indecisión. Con lo anterior en mente, el comportamiento de la aceptabilidad-*ac* puede describirse respondiendo a las siguientes dos preguntas: *(i)* ¿qué estado de aceptación corresponde a los argumentos involucrados en una indecisión?, y *(ii)* ¿cómo debe verse

[1] El término *bundle* set podría traducirse como conjunto de haces; no obstante, ante la ausencia de una mejor traducción, mantendremos el término original en inglés.

afectado el estado de aceptación de un argumento que no pertenece a una indecisión si es derrotado únicamente por argumentos que participan en una? La aceptabilidad-*ac* rechaza a todos los argumentos pertenecientes a una indecisión, a la vez que garantiza que el estado de aceptación de aquellos argumentos derrotados por miembros de una indecisión no se vea afectado por dichas derrotas. Nos referiremos a esta última característica del comportamiento de la aceptabilidad-*ac* como *bloqueo de indecisión*.

Existen múltiples razones por las cuales el comportamiento descrito anteriormente resulta razonable y significativo. En primer lugar, y de forma intuitiva, parece razonable rechazar a todos los argumentos involucrados en una indecisión, puesto que todos se consideran dudosos y ninguno logra superar a los demás argumentos del conjunto. Es decir, no tenemos motivos para decantarnos por ningún subconjunto de la indecisión por sobre los demás. Con respecto al bloqueo de indecisión, el razonamiento detrás de aceptar a aquellos argumentos únicamente derrotados por miembros de una indecisión es el siguiente: como todos los derrotadores del argumento bajo análisis son rechazados, y como el argumento bajo análisis no se encuentra en una situación dudosa (i.e. no pertenece a una indecisión), luego no existen motivos para rechazarlo. Otra razón que motiva el desarrollo de la aceptabilidad-*ac* es que la misma permite representar un esquema de razonamiento ausente en la literatura de semánticas basadas en diálogo y similar al que involucra, por ejemplo, al principio penal *in dubio pro reo*—según el cual, ante la presencia de evidencia contradictoria, debe juzgarse a favor del acusado.

Finalmente, la aceptabilidad-*ac* ofrece una alternativa a un problema presente en las principales semánticas argumentativas basadas en el concepto de *admisibilidad* [16]. En estas semánticas, la presencia de ciertos ciclos de derrotas de longitud impar produce que se rechacen argumentos que podrían ser aceptados, lo que constituye un resultado poco satisfactorio. La aceptabilidad-*ac* resuelve este problema al enfocar su comportamiento en el manejo de las indecisiones y no en el concepto de admisibilidad, pero lo hace sin renunciar a varias propiedades positivas derivadas de este último concepto.

Ideas similares a las abordadas en la aceptabilidad-*ac* fueron propuestas por primera vez en [13, 14, 15]. En dichos trabajos, los autores utilizaron una versión modificada del esquema basado en la recursividad de los componentes fuertemente conexos (del inglés, *strongly connected component recursiveness*) [4] para definir una familia de semánticas argumentativas, denominadas *semánticas basadas en la anulación de los componentes fuertemente conexos* (*semánticas SCC-void* para abreviar), donde cada semántica embebe el bloqueo de indecisión. En la tesis probamos que la aceptabilidad-*ac* constituye una caracterización basada en diálogo de uno de los integrantes de la familia de semánticas *SCC-void*, la semántica *grounded SCC-void* [2]. Esto significa que ambas definiciones aceptan (y rechazan) el mismo conjunto de argumentos. El resultado anterior solo vale en los marcos argumentativos tradicionales (tal cual fueron definidos en [16]). Además, probamos que en los marcos argumentativos que incluyen la relación de *subargumento*—un tipo adicional de relación entre argumentos que consideramos en la tesis—la aceptabilidad-*ac* se comporta satisfactoriamente, resultado que no vale para la semántica *grounded SCC-void* puesto que la misma fue concebida exclusivamente para marcos tradicionales.

2. 3. Principios para el Análisis de Semánticas Argumentativas Basadas en Diálogo

[2] En general, mantendremos el nombre de las semánticas argumentativas en inglés.

Finalmente, proponemos un conjunto de principios para el análisis formal de las semánticas basadas en diálogo. Esta contribución resulta significativa debido a que el análisis formal de las semánticas argumentativas—independientemente del enfoque adoptado para definirlas—ha sido de interés desde la concepción de las mismas, puesto que a través del mismo podemos comprender en mayor detalle el comportamiento de cada semántica, las propiedades que subyacen al mismo y establecer relaciones de similitud y de diferencia entre las semánticas. Este tipo de análisis es un elemento clave para promover la adopción de semánticas argumentativas en sistemas del mundo real, debido a que nos ofrece herramientas para identificar aquellas propiedades que una semántica debe poseer para adaptarse satisfactoriamente a los requerimientos de un dominio específico. Asimismo, estas propiedades formales pueden guiar el desarrollo de nuevas semánticas cuyo comportamiento sea más adecuado para determinadas aplicaciones. Lo anterior ha motivado el desarrollo del enfoque basado en principios (véase, por ejemplo, [3, 30, 12]), el cual propone definir un conjunto de características, denominadas principios, y analizar las semánticas argumentativas a la luz de las mismas. Siguiendo esa misma línea, ampliamos el enfoque existente proponiendo un nuevo conjunto de principios. Cada uno de los mismos surge de concebir las semánticas argumentativas como diálogos, lo que resulta novedoso ya que el resto de principios disponibles en la literatura adopta otro enfoque para la concepción de las semánticas argumentativas.

Los principios introducidos son utilizados para analizar la aceptabilidad-*ac* y la caracterización basada en diálogo de cuatro semánticas adicionales, a saber: la semántica *grounded* [16], la semántica preferida (crédula) [16], la semántica de propósito específico presentada en [17]—a la que nos referiremos como semántica *default*—y la semántica *pairwise cogency* (crédula) [7]. Como resultado de este análisis, obtenemos una mayor comprensión de algunas de las propiedades de estas semánticas, de las relaciones entre ellas, de los dominios en los que podrían desempeñarse más satisfactoriamente y múltiples indicios acerca de líneas futuras de investigación relacionadas con este tipo de semánticas.

3. Publicaciones

A continuación se indican los artículos publicados como resultado del trabajo llevado a cabo durante la realización de la tesis. Por cada una de las publicaciones, se detalla brevemente su contribución y se indican los capítulos de la tesis con los que se vinculan sus resultados.

- En el artículo "*A mathematical conceptualization of bundle sets in defeasible logic programming*" [28] publicado en *Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing (SAC 2024)*, extendemos las bases del enfoque basado en *bundle sets*. Este artículo se corresponde con el Capítulo 3 de la tesis.
- En el artículo "*Towards a dialogue game-based semantics for extended abstract argumentation frameworks based on indecision-blocking*" [29] publicado en *Proceedings of the 16th International Conference on Scalable Uncertainty Management (SUM 2024) tomo 15350 de Lecture Notes in Computer Science*, realizamos un estudio preliminar de algunas semánticas también consideradas en la tesis, a saber, las semánticas *grounded* y *default*, además de una tercera que fue un primer intento por conseguir una semántica basada en diálogo que embebiese el bloqueo de indecisión. La necesidad de contar con un marco formal para analizar rigurosamente las semánticas argumentativas basadas en diálogos fue la principal conclusión que obtuvimos de este análisis. Además, los resultados de este artículo

sentaron las bases para el posterior refinamiento y desarrollo de las contribuciones presentadas en los Capítulos 3, 4 y 5 de la tesis.

- En el artículo "*A principle-based framework for analyzing dialogue game-based semantics*" [27] publicado en *Proceedings of the 22nd International Conference on Principles of Knowledge Representation and Reasoning (KR 2025)*, realizamos un refinamiento del enfoque basado en bundle sets, adaptando varias definiciones introducidas en las publicaciones previas para contemplar los requerimientos de representar un conjunto más amplio de semánticas basadas en diálogo. Además, propusimos un conjunto de principios formales para analizar de manera rigurosa las propiedades de las semánticas abordadas y las relaciones existentes entre las mismas. Los resultados de la presente publicación se hallan reflejados en los Capítulos 3 y 5 de la tesis.

Al momento de la redacción de este resumen extendido, contamos con un artículo en desarrollo en el cual brindamos la definición de aceptabilidad-*ac*. Para esto, desarrollamos los conceptos que sirven de base a dicha definición. Además, estudiamos las propiedades de la aceptabilidad-*ac* y su relación con las semánticas *SCC-void* y con las semánticas basadas en diálogo contempladas en la tesis. Este contenido se corresponde con lo expuesto en los capítulos 4 y 5 de la tesis.

4. Trabajo Futuro

A partir de los resultados obtenidos en la tesis, múltiples líneas futuras de investigación permanecen abiertas. En primer lugar, planeamos extender nuestro análisis de las semánticas argumentativas basadas en diálogo. En este sentido, queremos considerar un espectro más amplio de semánticas, incluyendo, por ejemplo, a la semántica estable y a los meta-diálogos para la semántica preferida escéptica y las semánticas *weak cogency* y *cyclic cogency*. También pretendemos explorar distintas variantes de cada semántica a fin de evaluar cómo diferentes versiones de los protocolos afectan el análisis basado en principios. Además, pretendemos emplear un mayor abanico de principios e investigar en detalle la relación entre los principios existentes para las semánticas basadas en extensiones o en etiquetados y aquellos introducidos en nuestro trabajo. Por ejemplo, pretendemos dilucidar relaciones del tipo "*si una semántica basada en extensiones σ satisface el principio X, luego el protocolo correspondiente a σ satisface el principio Y*".

En la tesis tratamos la incorporación de elementos no considerados en los marcos clásicos, en particular, la relación de subargumento. Si bien esto nos permitió extender el comportamiento de las semánticas, también exigió consideraciones adicionales a la hora de garantizar resultados satisfactorios en las mismas. Por lo tanto, estamos interesados en explorar y evaluar semánticas argumentativas basadas en diálogo definidas en formalismos que extienden a los marcos empleados en la tesis mediante la incorporación de elementos—tales como etiquetas o múltiples agentes—y relaciones adicionales—por ejemplo, derrotas indirectas o colectivas, y diferentes tipos de soporte.

El punto inmediatamente anterior nos lleva también a la última de nuestras líneas futuras de investigación, la cual consiste en el análisis basado en principios de aquellas semánticas argumentativas empleadas en contextos multi-agente. Este tópico ha recibido relativamente poca atención, en comparación con las semánticas que adoptan el enfoque basado en inferencia, y resulta significativo puesto que en los sistemas multi-agente suelen considerarse elementos que en general se dejan de lado en la argumentación abstracta, como la naturaleza dinámica de los diálogos, la distribución de la información entre los participantes, los distintos actos de hablas y las

responsabilidades que cada uno asigna a los participantes, y las distintas estrategias que las partes involucradas pueden emplear, tanto para iniciar como para finalizar un diálogo.

Referencias

[1] Adadi, A., Berrada, M.: Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access 6, 52138–52160 (2018).

[2] Atkinson, K., Baroni, P., Giacomin, M., Hunter, A., Prakken, H., Reed, C., Simari, G.R., Thimm, M., Villata, S.: Towards artificial argumentation. AI Mag. 38(3), 25–36 (2017).

[3] Baroni, P., Giacomin, M.: On principle-based evaluation of extension-based argumentation semantics. Artif. Intell. 171(10-15), 675–700 (2007)

[4] Baroni, P., Giacomin, M., Guida, G.: Scc-recursiveness: a general schema for argumentation semantics. Artif. Intell. 168(1-2), 162–210 (2005)

[5] Bench-Capon, T.J.M., Atkinson, K., Wyner, A.Z.: Using argumentation to structure e-participation in policy making. Transactions on Large-Scale Data and Knowledge-Centered Systems 18, 1–29 (2015)

[6] Bodanza, G.A., Tohmé, F.A.: Two approaches to the problems of self-attacking arguments and general odd-length cycles of attack. J. Appl. Log. 7(4), 403–420 (2009).

[7] Bodanza, G.A., Tohmé, F.A., Simari, G.R.: Beyond admissibility: accepting cycles in argumentation with game protocols for cogency criteria. J. Log. Comput. 26(4), 1235–1255 (2016).

[8] Castagna, F., Kökciyan, N., Sassoon, I., Parsons, S., Sklar, E.: Computational argumentation-based chatbots: A survey. J. Artif. Intell. Res. 80, 1271–1310 (2024).

[9] Chesñevar, C.I., Simari, G.R.: A lattice-based approach to computing warranted beliefs in skeptical argumentation frameworks. In: Veloso, M.M. (ed.) IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007. pp. 280–285 (2007)

[10] Cocarascu, O., Toni, F.: Argumentation for machine learning: A survey. In: Baroni, P., Gordon, T.F., Scheffler, T., Stede, M. (eds.) Computational Models of Argument - Proceedings of COMMA 2016, Potsdam, Germany, 12-16 September, 2016. Frontiers in Artificial Intelligence and Applications, vol. 287, pp. 219–230. IOS Press (2016).

[11] Cohen, A., Gottifredi, S., García, A.J.: A heuristic pruning technique for dialectical trees on argumentation-based query-answering systems. In: Cuzzocrea, A., Greco, S., Larsen, H.L., Saccà, D., Andreasen, T., Christiansen, H. (eds.) Flexible Query Answering Systems - 13th International Conference, FQAS 2019, Amantea, Italy, July 2-5, 2019, Proceedings. Lecture Notes in Computer Science, vol. 11529, pp. 101–113. Springer (2019).

[12] Dauphin, J., Rienstra, T., van der Torre, L.: A principle-based analysis of weakly admissible semantics. In: Prakken, H., Bistarelli, S., Santini, F., Taticchi, C. (eds.) Computational Models of Argument - Proceedings of COMMA 2020, Perugia, Italy, September 4-11, 2020. Frontiers in Artificial Intelligence and Applications, vol. 326, pp. 167–178. IOS Press (2020).

[13] Dondio, P.: A proposal to embed the in dubio pro reo principle into abstract argumentation semantics based on topological ordering and undecidedness propagation. In: Dondio, P., Longo, L. (eds.) Proceedings of the 2nd Workshop on Advances In Argumentation In Artificial Intelligence, co-located with XVII International Conference of the Italian Association for Artificial Intelligence, AI3 @AI*IA 2018, 20-23 November 2018, Trento, Italy. CEUR Workshop Proceedings, vol. 2296, pp. 42–56. CEUR-WS.org (2018)

[14] Dondio, P., Longo, L.: Beyond reasonable doubt: A proposal for undecidedness blocking in abstract argumentation. Intelligenza Artificiale 13(2), 123–135 (2019).

[15] Dondio, P., Longo, L.: Weakly complete semantics based on undecidedness blocking. CoRR abs/2103.10701 (2021)

[16] Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif. Intell. 77(2), 321–358 (1995).
[17] García, A.J., Simari, G.R.: Defeasible logic programming: An argumentative approach. Theory Pract. Log. Program. 4(1-2), 95–138 (2004).
[18] Garcia, A.C., Martinez, M.V., Deagustini, C.A.D., Simari, G.I.: A multi-agent system for addressing cybersecurity issues in social networks. In: Benferhat, S., Casini, G., Meyer, T., Tettamanzi, A. (eds.) Proceedings of 1st Workshop on AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering, co-located with 20th International Conference on Principles of Knowledge Representation and Reasoning (KR 2023), Rhodes, Greece, September 3-4, 2023. CEUR Workshop Proceedings, vol. 3495, pp. 43–54. CEUR-WS.org (2023)
[19] Honavar, V.: Symbolic artificial intelligence and numeric artificial neural networks: towards a resolution of the dichotomy. In: Computational architectures integrating neural and symbolic processes: a perspective on the state of the art, pp. 351–388. Springer (1995)
[20] Hong, S., Xiao, L., Zhang, X., Chen, J.: Argmed-agents: explainable clinical decision reasoning with llm disscusion via argumentation schemes. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 5486–5493. IEEE (2024)
[21] Lawrence, J., Reed, C.: Argument mining: A survey. Comput. Linguistics 45(4), 765–818 (2019).
[22] Modgil, S., Caminada, M.: Proof theories and algorithms for abstract argumentation frameworks. In: Simari, G.R., Rahwan, I. (eds.) Argumentation in Artificial Intelligence, pp. 105–129. Springer (2009)
[23] Pollock, J.L.: Justification and defeat. Artif. Intell. 67(2), 377–407 (1994)
[24] Prakken, H.: Historical overview of formal argumentation. FLAP 4(8) (2017)
[25] Prakken, H., Sartor, G.: Law and logic: A review from an argumentation perspective. Artif. Intell. 227, 214–245 (2015).
[26] Rahwan, I., Larson, K.: Argumentation and game theory. In: Simari, G.R., Rahwan, I. (eds.) Argumentation in Artificial Intelligence, pp. 321–339. Springer (2009).
[27] Soto, Y.O., Cohen, A., Deagustini, C.A.D., Martinez, M.V., Simari, G.I.: A principle-based framework for analyzing dialogue game-based semantics. In: Proceedings of the 22nd International Conference on Principles of Knowledge Representation and Reasoning, KR 2025, Melbourne, Australia, November 11-17, 2025. pp. 555–565 (2025)
[28] Soto, Y.O., Deagustini, C.A.D., Martinez, M.V., Simari, G.I.: A mathematical conceptualization of bundle sets in defeasible logic programming. In: Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing. p. 779–786. SAC '24, ACM (2024)
[29] Soto, Y.O., Deagustini, C.A.D., Martinez, M.V., Simari, G.I.: Towards a dialogue game-based semantics for extended abstract argumentation frameworks based on indecision-blocking. In: Destercke, S., Martinez, M.V., Sanfilippo, G. (eds.) Scalable Uncertainty Management - 16th International Conference, SUM 2024, Palermo, Italy, November 27-29, 2024, Proceedings. Lecture Notes in Computer Science, vol. 15350, pp. 445–460. Springer (2024)
[30] van der Torre, L., Vesic, S.: The principle-based approach to abstract argumentation semantics. FLAP 4(8) (2017)
[31] Walton, D.N.: What is reasoning? What is an argument? The journal of Philosophy 87(8), 399–419 (1990)

Televisión Digital Interactiva (TVDI) para reducir las alteraciones de la marcha en adultos mayores

Autor: María Magdalena Rosado Álvarez
Directora: Dra. María José Abásolo (UNLP)
Codirector: Dr. Telmo Silva (Universidad de Aveiro)

Tesis presentada para obtener el grado de Doctor en Ciencias Informáticas

Doctorado en Ciencias Informáticas, Facultad de Informática, Universidad Nacional de La Plata (UNLP)

Fecha de defensa: 03 de diciembre de 2025

Síntesis

Motivación

El envejecimiento poblacional constituye uno de los fenómenos más significativos del siglo XXI, con impacto directo en la sostenibilidad de los sistemas de salud y en la autonomía funcional de las personas mayores. Entre sus consecuencias más críticas se encuentran las alteraciones de la marcha y el incremento del riesgo de caídas, eventos que comprometen la estabilidad, la movilidad y la independencia. Más allá de su dimensión clínica, este escenario expone un problema estructural en términos de escalabilidad, sostenibilidad y formalización tecnológica de los modelos actuales de rehabilitación domiciliaria.

Si bien los programas de rehabilitación multicomponentes han demostrado eficacia en la mejora del equilibrio y la movilidad, los enfoques tradicionales presentan limitaciones significativas: dependencia de supervisión presencial especializada, altos costos operativos, barreras geográficas y dificultades de adherencia sostenida en entornos domiciliarios. Estas restricciones no solo afectan la continuidad terapéutica, sino que evidencian la ausencia de infraestructuras computacionales formales capaces de trasladar la prescripción clínica a entornos distribuidos con coherencia metodológica y validación estructurada.

En paralelo, el desarrollo de las Tecnologías de la Información y Comunicación ha transformado los modelos de interacción humano-computadora y ha impulsado el crecimiento de la salud digital. Dentro de este ecosistema, la Televisión Digital Interactiva (TVDi) representa una infraestructura tecnológica estratégica por su alta penetración domiciliaria, capacidad de interactividad y potencial para integrar personalización y monitoreo en sistemas distribuidos. Su familiaridad en población adulta mayor la posiciona como un entorno tecnológico inclusivo con posibilidades reales de escalabilidad.

No obstante, el estado del arte muestra que las aplicaciones sobre TVDi se han orientado principalmente a servicios informativos o de entretenimiento, con limitada formalización arquitectónica en contextos de rehabilitación funcional. No se identifican modelos computacionales que integren de manera sistemática evaluación funcional estructurada,

personalización adaptativa basada en niveles diferenciados de riesgo, monitoreo desacoplado de adherencia y validación mediante modelos formales de aceptación tecnológica dentro de un marco explícito de ingeniería de software distribuido.

El desarrollo de una solución en este dominio plantea desafíos propios de las Ciencias Informáticas: formalizar la integración entre evaluación clínica y sistemas interactivos distribuidos; diseñar arquitecturas robustas y escalables sobre infraestructura de TVDi; implementar mecanismos de personalización adaptativa estructurada; garantizar usabilidad y accesibilidad en población adulta mayor; y validar el sistema mediante marcos formales de aceptación tecnológica y evaluación experta. Asimismo, la interoperabilidad y replicabilidad del modelo se constituyen en condiciones necesarias para su transferencia a distintos escenarios de salud digital.

En este contexto, la presente investigación propone el diseño, desarrollo y validación de un modelo computacional basado en Televisión Digital Interactiva orientado a la rehabilitación funcional de la marcha en adultos mayores. La propuesta integra una arquitectura tecnológica formal, personalización adaptativa de ejercicios según evaluación funcional, monitoreo estructurado de adherencia y validación mediante instrumentos clínicos estandarizados y modelos formales de aceptación tecnológica. De este modo, se establece un marco metodológico replicable para la construcción de sistemas interactivos clínicamente estructurados, contribuyendo al avance disciplinar de las Ciencias Informáticas aplicadas a la salud digital.

Aporte de la tesis a la disciplina

La presente investigación introduce un modelo arquitectónico y computacional para la rehabilitación funcional de la marcha basado en TVDi, sustentado en una arquitectura formal de microservicios que integra evaluación clínica estructurada, personalización adaptativa y monitoreo desacoplado en entornos domiciliarios.

A diferencia de implementaciones previas en salud digital, generalmente limitadas a soluciones funcionales o plataformas audiovisuales sin formalización estructural explícita, el modelo propuesto establece principios arquitectónicos definidos para la integración sistemática de lógica clínica en sistemas interactivos distribuidos. Esta formalización transforma la TVDi en una infraestructura tecnológicamente estructurada y metodológicamente validada para su aplicación en rehabilitación funcional.

Componentes estructurales del modelo March TV

El modelo se organiza bajo principios formales de arquitectura distribuida, especialización funcional y modularidad evolutiva, garantizando estabilidad operativa, mantenibilidad y escalabilidad.

1. Arquitectura distribuida basada en microservicios

El sistema adopta un esquema de microservicios con responsabilidades claramente delimitadas, donde los actores (Administrador, Fisioterapeuta y Adulto Mayor) interactúan con servicios independientes. Esta configuración asegura separación entre lógica clínica, gestión de contenidos y visualización, aislamiento de fallos y evolución modular sin reconfiguración global.

La aplicación explícita de principios de ingeniería de software distribuido al ámbito de la rehabilitación domiciliaria constituye una contribución en el diseño de sistemas interactivos clínicamente estructurados.

2. Formalización computacional de la prescripción terapéutica

El núcleo innovador del modelo reside en la transformación de resultados de evaluación funcional en insumos estructurados que activan un mecanismo formal de clasificación por nivel de riesgo y asignación personalizada de rutinas terapéuticas.

La separación entre *back-end* clínico y *front-end* de visualización preserva el control profesional de la prescripción, mientras que el usuario accede exclusivamente a contenidos autorizados. Este esquema establece un patrón arquitectónico para la integración sistemática de lógica clínica dentro de sistemas interactivos distribuidos.

3. Gestión modular de contenidos y visualización

El almacenamiento y distribución de ejercicios se implementan como servicio autónomo mediante infraestructura NAS (*Network Attached Storage*; almacenamiento conectado a la

red) y Plex Media Server, permitiendo actualización independiente de contenidos y control de acceso basado en perfiles.

El módulo de visualización en Smart TV opera como consumidor desacoplado, sin contener lógica clínica, consolidando coherencia arquitectónica, seguridad operativa y separación de responsabilidades.

4. Monitoreo estructurado y trazabilidad terapéutica

El modelo incorpora un servicio independiente de registro de actividad que captura métricas de reproducción y adherencia sin interferir en la experiencia del usuario. Este módulo habilita análisis longitudinal de cumplimiento terapéutico e introduce un esquema formal de trazabilidad en plataformas de TVDi aplicadas a rehabilitación funcional.

Integración metodológica y validación estructurada

El aporte se consolida mediante un marco metodológico que articula validación por expertos (método Delphi), evaluación de aceptación tecnológica (Modelo TAM) y aplicación empírica en entorno real con instrumentos clínicos estandarizados. Esta integración no solo valida la eficacia funcional, sino que formaliza un procedimiento replicable para el diseño y evaluación de sistemas interactivos en salud digital.

Contribución disciplinar

El principal aporte de la tesis radica en la formalización de un patrón arquitectónico y metodológico para la integración de evaluación funcional estructurada, personalización adaptativa y monitoreo desacoplado dentro de sistemas interactivos distribuidos sobre TVDi.

La investigación no se limita al desarrollo de una solución tecnológica específica, sino que establece un marco arquitectónico transferible para la incorporación de lógica clínica en plataformas interactivas domiciliarias. En consecuencia, la tesis amplía el alcance disciplinar de los modelos distribuidos en el ámbito de la salud digital, consolidando una contribución sustantiva a las Ciencias Informáticas aplicadas a la rehabilitación funcional.

Resultados Obtenidos

El análisis de los resultados se estructuró en cuatro dimensiones integradas: impacto funcional y consolidación domiciliaria, experiencia humano-tecnología, validación estructural del modelo y robustez arquitectónica, con el propósito de evaluar de manera conjunta la efectividad clínica y la solidez computacional del sistema March TV.

Las evaluaciones se realizaron en tres momentos: T0 (medición inicial previa a la intervención), T1 (evaluación intermedia tras 15 sesiones supervisadas en laboratorio) y T2 (evaluación final al concluir las 29 sesiones totales, incluyendo la fase domiciliaria.

1. Impacto funcional y consolidación domiciliaria

La mejora progresiva observada en la Escala de Tinetti y en la prueba *Timed Up and Go* (TUG) confirma que la intervención estructurada produjo cambios estadísticamente significativos en equilibrio y movilidad.

El incremento acumulado del 16.96% en Tinetti y la reducción del 28.81% en TUG evidencian mejoras clínicamente relevantes en estabilidad postural y velocidad de desplazamiento, variables directamente asociadas al riesgo de caída.

Un aspecto crítico del análisis radica en que las mejoras alcanzadas en T1 no solo se mantuvieron en T2, sino que se incrementaron tras la implementación domiciliaria. Esto demuestra que la arquitectura distribuida permitió trasladar la prescripción clínica al entorno doméstico sin pérdida de coherencia metodológica.

La sostenibilidad de los resultados en T2 indica que la lógica de asignación personalizada y la organización jerárquica de rutinas, gestionadas mediante microservicios desacoplados, actuaron como estructura habilitante para la progresión terapéutica fuera del entorno controlado.

2. Experiencia humano-tecnología: dimensión emocional y aceptación

La experiencia de interacción fue evaluada desde dos perspectivas complementarias: respuesta emocional del usuario y aceptación tecnológica.

Los resultados del *Self-Assessment Manikin* (SAM) evidenciaron una evolución progresiva en las tres dimensiones emocionales evaluadas. El agrado aumentó de 7.20 en T0 a 8.60 en T2; la activación pasó de 7.00 a 8.70; y la percepción de control se incrementó de 7.80 a 9.00. Esta progresión indica no solo familiarización con el sistema, sino consolidación de una experiencia emocional positiva y creciente sensación de dominio tecnológico.

De manera complementaria, el Modelo de Aceptación Tecnológica (TAM) mostró predominancia de puntuaciones máximas en utilidad y facilidad de uso, con 89.61% de valoraciones en el nivel superior y varios indicadores con desviación estándar nula.

Desde una perspectiva de ingeniería de software, estos resultados confirman que la modularidad y separación de responsabilidades propias de la arquitectura de microservicios no introdujeron fricción cognitiva ni barreras de adopción. Por el contrario, favorecieron una interacción estable, predecible y emocionalmente confortable para el adulto mayor.

La convergencia entre mejora funcional, experiencia emocional positiva y alta aceptación tecnológica fortalece la integración entre diseño computacional y aplicabilidad clínica.

3. Validación estructural del modelo

El consenso superior al 85% alcanzado mediante el método Delphi respalda la consistencia conceptual y técnica del modelo propuesto. La estabilidad de las respuestas en la segunda ronda evidenció madurez estructural y coherencia interna de la arquitectura planteada.

La validación experta no solo ratifica la pertinencia clínica del sistema, sino que confirma la adecuación del esquema arquitectónico distribuido como marco organizador del proceso terapéutico.

4. Robustez arquitectónica y condiciones operativas

Durante la implementación no se registraron fallos estructurales del sistema. La arquitectura basada en microservicios demostró funcionamiento estable y desacoplado, manteniendo integridad operativa entre módulos clínicos, de contenido y de visualización.

La ausencia de degradación del servicio, incluso bajo condiciones externas adversas como los cortes de energía eléctrica, evidencia que el diseño modular favoreció resiliencia operativa y mantenibilidad estructural. La reducción muestral estuvo asociada exclusivamente a factores ambientales y no a limitaciones técnicas del modelo.

En conjunto, los resultados demuestran que el modelo March TV no solo produjo mejoras funcionales estadísticamente significativas, sino que validó empíricamente una arquitectura computacional formal capaz de sostener prescripción terapéutica personalizada, monitoreo estructurado y experiencia de usuario positiva en entorno domiciliario.

La integración desacoplada entre evaluación funcional, asignación adaptativa y visualización interactiva consolidó un marco replicable para la implementación de rehabilitación funcional distribuida, fortaleciendo la contribución del trabajo dentro del ámbito de las Ciencias Informáticas aplicadas a la salud digital.

Posibles líneas de investigación futuras.

Como proyección de la investigación desarrollada, se plantean líneas de trabajo orientadas a fortalecer la dimensión computacional, clínica y operativa del modelo March TV.

En primer lugar, se propone incorporar mecanismos de adaptación inteligente en tiempo real mediante la integración de algoritmos de aprendizaje automático que ajusten dinámicamente la dificultad terapéutica según la evolución funcional del usuario. Esta línea permitiría avanzar hacia un sistema predictivo capaz de anticipar progresiones y optimizar la personalización del entrenamiento.

Asimismo, resulta pertinente profundizar en la robustez arquitectónica del sistema mediante la implementación de estrategias de tolerancia a fallos, modos offline y almacenamiento local sincronizable, con el objetivo de garantizar continuidad operativa en entornos con infraestructura eléctrica o conectividad inestable.

Otra línea relevante consiste en la integración de sensores de movimiento y módulos de análisis biomecánico automatizado, que permitan enriquecer el monitoreo estructurado y fortalecer la retroalimentación objetiva en tiempo real, ampliando la capacidad analítica del modelo.

Desde una perspectiva de escalabilidad, se propone evaluar la implementación del sistema bajo infraestructuras en la nube, explorando configuraciones que permitan balanceo de carga, orquestación de servicios y mayor número de usuarios simultáneos sin degradación del desempeño.

En el ámbito clínico, se plantea extender la aplicación del modelo a poblaciones con accidente cerebrovascular, enfermedad de Parkinson, artrosis u otras condiciones asociadas a movilidad reducida, con el fin de analizar la transferibilidad del marco computacional en distintos perfiles funcionales.

Finalmente, se considera necesario desarrollar estudios formales de costo-beneficio y costo-efectividad que comparen la solución basada en Televisión Digital Interactiva con modelos tradicionales de rehabilitación presencial, evaluando su impacto económico y su potencial de implementación a gran escala.

En conjunto, estas líneas futuras consolidan una agenda de investigación orientada a la evolución del modelo como infraestructura computacional escalable y transferible para la rehabilitación funcional distribuida dentro del campo de las Ciencias Informáticas aplicadas a la salud digital.

RESUMEN EXTENDIDO DE TESIS DE DOCTORADO

Transhumamismos, experiencias e inteligencia artificial: Articulaciones pedagógicas en ambientes virtuales en la Facultad de Humanidades, UNMDP

Autora: Mg. Gladys Vanesa Fernández

Director: Dr. Francisco Ramallo

Co-director: Dr. Alejandro González

Universidad: Universidad Nacional de La Plata (UNLP) – Facultad de Informática

Grado otorgado: Doctor en Ciencias Informáticas

Fecha de exposición: 16 de diciembre de 2025

Resumen

Esta tesis analiza la integración de inteligencia artificial (IA) en la gestión de ambientes virtuales de aprendizaje en la Facultad de Humanidades de la Universidad Nacional de Mar del Plata (UNMdP), en diálogo con debates transhumanistas y posthumanistas sobre la relación entre tecnología y educación. Desde el rol de coordinadora del Área de Asistencia a los Procesos Pedagógicos (APP), se desarrolló una investigación cualitativa con diseño exploratorio-descriptivo que combinó observación participante, análisis de comunicaciones institucionales —4.200 correos electrónicos y 130 mensajes internos de la plataforma Moodle del período 2020-2024— y evaluación de modelos predictivos de aprendizaje. Los hallazgos evidencian que las consultas administrativas son recurrentes y susceptibles de automatización mediante agentes conversacionales, con picos de demanda al inicio de cada cuatrimestre y durante los períodos de examen. La investigación concluye que la implementación de IA en la gestión administrativa universitaria debe realizarse de forma crítica, priorizando la transparencia algorítmica, la formación docente continua y el diseño centrado en las necesidades situadas del contexto latinoamericano, complementando —y no reemplazando— el rol de los equipos técnico-pedagógicos.

Palabras clave: inteligencia artificial, educación superior, transhumanismo, posthumanismo, ambientes virtuales, Moodle, analítica del aprendizaje, chatbot administrativo.

1. Motivación y contexto

La expansión acelerada de las tecnologías digitales en el ámbito educativo ha puesto en el centro del debate académico la pregunta por los modos en que la inteligencia artificial puede transformar —o distorsionar— las prácticas de enseñanza y aprendizaje en la educación superior. La Facultad de Humanidades de la UNMdP no es ajena a este proceso: el crecimiento de la matrícula, la diversificación de las modalidades de cursada y la pandemia de COVID-19 aceleraron la adopción

masiva del campus virtual Moodle como entorno principal de gestión pedagógica y administrativa, generando una demanda sostenida de soporte técnico y atención de consultas que supera la capacidad operativa del equipo APP.

La motivación central de esta investigación surge de la experiencia concreta de la autora como integrante de dicho equipo, quien observó que una proporción significativa del tiempo de trabajo se destinaba a responder consultas reiterativas de carácter administrativo, en detrimento del acompañamiento pedagógico personalizado que constituye la función sustantiva del área. Esta tensión entre las demandas operativas y la vocación pedagógica del APP configura el problema práctico que la tesis busca abordar mediante la exploración de soluciones basadas en IA.

Al mismo tiempo, la investigación reconoce que la incorporación de IA en los ambientes educativos no es un proceso técnicamente neutro. Las plataformas digitales y sus algoritmos reproducen supuestos epistemológicos, relaciones de poder y modelos de sujeto que deben ser interrogados desde una perspectiva crítica. En este sentido, la tesis se inscribe en el campo de los estudios sobre tecnología educativa, pero lo hace desde una mirada filosófica y pedagógica que recupera los debates transhumanistas y posthumanistas para pensar la relación entre agentes humanos y no humanos en los procesos de enseñanza y aprendizaje.

2. Marco teórico y conceptual

El marco teórico de la tesis se articula en torno a tres grandes líneas conceptuales. La primera recupera los debates entre humanismo, transhumanismo y posthumanismo, con especial atención a la obra de Rosi Braidotti, cuya perspectiva posthumanista desafía la concepción del sujeto autónomo e individual del humanismo clásico y propone pensar la subjetividad como proceso relacional, híbrido y co-constituido por entidades humanas y no humanas. Esta perspectiva resulta fecunda para analizar las interacciones en los ambientes virtuales, donde docentes, estudiantes, algoritmos y plataformas conforman ensambles complejos que no pueden reducirse a la lógica instrumental de la herramienta al servicio del usuario.

La segunda línea conceptual articula el concepto de articulación pedagógica, recuperado del trabajo de Mesquita (2022) desde una perspectiva descolonizadora. A diferencia de la mediación pedagógica tradicional, que concibe al docente como intermediario neutral entre el contenido y el estudiante, la articulación pedagógica reconoce el carácter político, situado y relacional de los encuentros educativos. Esta noción permite analizar los ambientes virtuales no como simples contenedores de información, sino como espacios donde se configuran relaciones de poder, saberes situados y experiencias encarnadas.

La tercera línea conceptual sistematiza los debates en torno a la IA en educación, relevando antecedentes sobre el uso de sistemas de aprendizaje automático, analítica del aprendizaje y agentes conversacionales en la educación superior. Se destaca el trabajo crítico de Sadin (2020), quien advierte sobre los riesgos de la mercantilización del conocimiento y la reducción de la experiencia educativa a parámetros algorítmicos cuantificables.

3. Diseño metodológico

La investigación adopta un diseño cualitativo de carácter exploratorio-descriptivo, organizado en tres fases articuladas. La Fase 1 consistió en una revisión y reflexión crítica de la literatura sobre IA en educación superior, transhumanismo y posthumanismo, complementada con el relevamiento de

políticas y regulaciones sobre IA en los planos internacional, nacional e institucional. Esta fase permitió construir el marco conceptual y contextual de la tesis.

La Fase 2 se centró en la observación participante y el análisis sistemático de las comunicaciones institucionales del APP. El corpus analizado comprende 4.200 correos electrónicos y 130 mensajes internos de Moodle correspondientes al período 2020-2024. Se utilizaron técnicas de análisis de contenido para identificar los tipos de consulta más frecuentes, su distribución temporal, los actores involucrados y los tiempos de respuesta. Esta fase permitió construir un mapa detallado de las demandas operativas del APP y fundamentar empíricamente la propuesta de automatización.

La Fase 3 comprendió la planificación, análisis y evaluación de los modelos predictivos y analíticos disponibles en el software Moodle, así como el diseño de una propuesta de chatbot administrativo entrenado con las consultas reales relevadas en la Fase 2. Se desarrollaron estrategias de validación que incluyen la definición de indicadores de desempeño del bot y la elaboración de una propuesta normativa para la protección de datos personales.

Desde el punto de vista epistemológico, la tesis asume explícitamente la posición de la investigadora como participante del campo estudiado. Lejos de concebir esta situación como un sesgo a neutralizar, la investigación la reconoce como una oportunidad epistémica que permite acceder a dimensiones de la experiencia institucional que difícilmente pueden capturarse desde una perspectiva puramente externa. Esta elección metodológica, coherente con el enfoque posthumanista adoptado, rechaza la pretensión de objetividad positivista y asume el carácter situado, reflexivo y encarnado del conocimiento producido.

4. Resultados principales

El análisis de las comunicaciones institucionales reveló que las consultas recibidas por el APP presentan una alta concentración en períodos específicos del ciclo académico: inicio de cuatrimestre (inscripciones, claves de automaterícula, configuración de materias) y períodos de examen (habilitación de actividades, gestión de notas). Estas consultas exhiben un patrón recurrente y predecible que las hace técnicamente susceptibles de automatización mediante agentes conversacionales.

La distribución temática de las consultas muestra que las solicitudes relacionadas con el acceso al campus virtual y la inscripción a materias constituyen la categoría más frecuente, seguidas por consultas sobre la configuración de actividades pedagógicas y la gestión de usuarios. El análisis de los tiempos de respuesta evidenció demoras significativas en los períodos de mayor demanda, con impacto directo en la experiencia de los usuarios y en la carga laboral del equipo.

En cuanto a la evaluación de los modelos predictivos de Moodle, los resultados revelaron limitaciones importantes en el contexto educativo argentino. Los modelos predeterminados de la plataforma presentan altos índices de falsos positivos —es decir, identifican como estudiantes en riesgo de abandono a estudiantes que en realidad no enfrentan dificultades significativas—, lo que puede generar intervenciones pedagógicas innecesarias o, peor aún, estigmatizar a determinados grupos de estudiantes. Estas limitaciones se explican, en parte, porque los modelos fueron diseñados para contextos anglófonos y no contemplan las particularidades del sistema universitario público argentino, donde la irregularidad en el acceso a dispositivos y conectividad no necesariamente refleja desvinculación académica.

Como propuesta central de la tesis, se diseñó un chatbot administrativo entrenado con las consultas reales relevadas durante la investigación. El bot fue concebido como un sistema capaz de responder preguntas frecuentes de carácter administrativo, liberar tiempo del equipo técnico para el

acompañamiento pedagógico personalizado y operar de manera complementaria —no sustitutiva— con respecto a los profesionales del área. La propuesta incluye criterios de selección de plataforma tecnológica, indicadores de evaluación del desempeño del bot y lineamientos para una normativa institucional de protección de datos personales.

5. Aporte de la tesis a la disciplina

El aporte de esta tesis se despliega en tres dimensiones complementarias. En primer lugar, ofrece un aporte empírico al producir un diagnóstico detallado y sistemático de las demandas operativas de un área de asistencia pedagógica en la educación superior pública argentina, en un contexto donde este tipo de estudios es escaso. Los datos producidos sobre la tipología, frecuencia y distribución temporal de las consultas administrativas constituyen insumos valiosos para la toma de decisiones institucionales en universidades con características similares.

En segundo lugar, la tesis realiza un aporte metodológico al integrar perspectivas cualitativas y técnicas de un modo inusual en el campo de las Ciencias Informáticas. La combinación de observación participante, análisis de contenido institucional y evaluación de modelos algorítmicos, desde una posición explícitamente situada y reflexiva, enriquece el repertorio metodológico disponible para investigar la IA en contextos educativos. Esta integración metodológica desafía la dicotomía entre investigación técnica e investigación humanística, proponiendo un enfoque que reconoce la interdependencia entre ambas tradiciones.

En tercer lugar, la tesis realiza un aporte conceptual al articular los debates transhumanistas y posthumanistas con el análisis empírico de experiencias concretas de implementación de IA en ambientes virtuales. Esta articulación teórico-empírica permite ir más allá de los análisis puramente instrumentales o de los análisis filosóficos desconectados de la práctica, produciendo una comprensión más rica y crítica del lugar de la IA en los procesos educativos. En particular, la noción de articulación pedagógica empleada en la tesis ofrece un marco conceptual fértil para analizar las interacciones entre agentes humanos y no humanos en los ambientes virtuales, con potencial de aplicación en otros contextos institucionales.

6. Conclusiones

La investigación concluye que la IA puede constituir un recurso valioso para mejorar la gestión administrativa de los ambientes virtuales en la educación superior, siempre que su implementación se realice de manera crítica, contextualizada y orientada a fortalecer —y no debilitar— las capacidades pedagógicas de los equipos institucionales. La automatización de consultas repetitivas mediante chatbots puede liberar recursos humanos para tareas de mayor complejidad y valor pedagógico, pero requiere un diseño cuidadoso que contemple las particularidades del contexto, la protección de datos personales y la formación de los usuarios.

Los modelos predictivos disponibles en las plataformas comerciales presentan limitaciones significativas cuando se aplican en contextos distintos a aquellos para los que fueron diseñados. En el caso del sistema universitario público argentino, estas limitaciones se expresan en altos índices de falsos positivos que pueden tener consecuencias negativas para los estudiantes y erosionar la confianza en los sistemas de analítica del aprendizaje. La investigación propone transitar hacia modelos que incorporen variables más diversas y representativas de las experiencias estudiantiles, considerando la multiplicidad de formas en que los estudiantes interactúan con el conocimiento más allá de la actividad dentro de la plataforma.

Desde una perspectiva teórica, la tesis concluye que los enfoques posthumanistas y transhumanistas ofrecen marcos conceptuales pertinentes para pensar la integración de IA en la educación, en tanto descentran la mirada instrumental y reconocen la agencia de los elementos no humanos en la configuración de las experiencias educativas. Sin embargo, esta perspectiva no debe conducir a una aceptación acrítica de la tecnología: la transparencia algorítmica, la equidad en el acceso, la formación docente y la regulación ética son condiciones necesarias para que la IA contribuya a una educación más justa e inclusiva en el contexto latinoamericano.

7. Líneas de investigación futuras

La presente investigación abre varias líneas de trabajo que podrían desarrollarse en el futuro. En primer lugar, resulta prioritaria la implementación efectiva del chatbot administrativo diseñado y la evaluación de su desempeño en condiciones reales de uso, considerando tanto indicadores técnicos (precisión de las respuestas, tasa de escalamiento a operadores humanos) como pedagógicos (impacto en la satisfacción de los usuarios y en la calidad del acompañamiento pedagógico brindado por el equipo APP).

En segundo lugar, sería relevante extender el estudio comparativo a otras facultades de la UNMdP y a otras universidades nacionales, con el fin de identificar patrones comunes y especificidades institucionales en la gestión de los ambientes virtuales. Esta comparación podría contribuir al diseño de soluciones más robustas y transferibles entre instituciones con características similares.

En tercer lugar, la investigación abre la puerta al desarrollo de modelos predictivos propios, adaptados a las particularidades del sistema universitario público argentino. Esto implicaría la construcción de datasets etiquetados en contextos locales, la definición colaborativa de indicadores de riesgo culturalmente pertinentes y la evaluación sistemática del impacto de las intervenciones algorítmicas sobre distintos grupos de estudiantes.

En cuarto lugar, la dimensión ética y regulatoria identificada como una asignatura pendiente en el contexto argentino amerita estudios específicos orientados al diseño de marcos normativos institucionales que protejan los derechos de los estudiantes, garanticen la transparencia de los algoritmos y regulen el uso de datos personales en los entornos educativos digitales.

Por último, en el plano teórico, la articulación entre perspectivas posthumanistas y análisis técnicos de sistemas de IA merece ser profundizada, tanto en su dimensión filosófica como en sus implicaciones para el diseño y evaluación de tecnologías educativas. Esta línea de investigación podría contribuir al desarrollo de un enfoque propiamente latinoamericano para pensar la IA en educación, que tome en cuenta las condiciones materiales, las tradiciones pedagógicas y los horizontes políticos específicos de la región.

Nota

Este resumen extendido fue elaborado sobre la base de la tesis doctoral completa, presentada en la Facultad de Informática de la Universidad Nacional de La Plata (UNLP) en abril de 2025. La tesis desarrolla en detalle los fundamentos teóricos, el diseño metodológico, los resultados empíricos y las propuestas aquí sintetizados.

XXVIII WICC Bs.As 2026

XXVIII Workshop de Investigadores en Ciencias de la Computación