



XXIII Edición del Workshop de Investigadores en Ciencias de la Computación

LIBRO de ACTAS

Abril 2021 - Chilecito – La Rioja – Argentina

Universidad Nacional de Chilecito - UNdeC

Red de Universidades con Carreras de Informática - RedUNCI

Frati, Fernando Emmanuel

XXIII Edición del Workshop de Investigadores en Ciencias de la Computación / Fernando Emmanuel Frati ; Fernanda Beatriz Carmona. - 1a ed adaptada. - Chilecito : UNdeC, 2021.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-24611-3-3

1. Computación. I. Carmona, Fernanda Beatriz. II. Título.

CDD 004.071

Comité Académico

Universidad de Buenos Aires (Ciencias Exactas)
GARBERVETSKY, Diego

Universidad de Buenos Aires (Ingeniería)
ECHEVERRIA, Adriana

Universidad Nacional de La Plata
PESADO, Patricia

Universidad Nacional del Sur
RUEDA, Sonia

Universidad Nacional de San Luis
PRINTISTA, Marcela

Universidad Nacional del Centro de la Provincia de Buenos Aires
ACITI, Claudio

Universidad Nacional del Comahue
GROSSO, Guillermo

Universidad Nacional de La Matanza
ETEROVIC, Jorge

Universidad Nacional de La Pampa
ALFONSO, Hugo

Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
KOREMBLIT, Gabriel

Universidad Nacional de Salta
GIL, Gustavo

Universidad Nacional de la Patagonia Austral
LASSO, Marta

Universidad Nacional de San Juan
RODRIGUEZ, Nelson

Universidad Nacional de Entre Ríos
NORIEGA, Jorge

Universidad Nacional de la Patagonia San Juan Bosco
BUCKLE, Carlos

Universidad Autónoma de Entre Ríos
TUGNARELLI, Mónica

Universidad Nacional del Nordeste
DAPOZO, Gladys

Universidad Nacional de Rosario
ZANARINI, Dante

Universidad Nacional de Misiones
KUNA, Horacio

Universidad Nacional del Noroeste de la Provincia de Buenos Aires
RUSSO, Claudia

Universidad Nacional de Chilecito
CARMONA, Fernanda

Universidad Nacional de Lanús
AZCURRA, Diego

Universidad Nacional de Santiago del Estero
DURAN, Elena

Escuela Superior de Ejército
ARROYO ARZUBI, Alejandro

Universidad Nacional del Litoral
LOYARTE, Horacio

Universidad Nacional de Río IV
ARROYO, Marcelo

Universidad Nacional de Córdoba
FRIDLENDER, Daniel

Universidad Nacional de Jujuy
HERRERA COGNETTA, Analía

Universidad Nacional de Río Negro
VIVAS, Luis

Universidad Nacional de Villa María
PRATO, Laura

Universidad Nacional de Luján
PANESSI, Wálter

Universidad Nacional de Catamarca
POLICHE, María Valeria

Universidad Nacional de La Rioja <i>MARTINEZ, Marcelo</i>	Pontificia Universidad Católica Argentina <i>GRIECO, Sebastián</i>
Universidad Nacional de Tres de Febrero <i>OLIVEROS, Alejandro</i>	Universidad del Salvador <i>ZANITTI, Marcelo</i>
Universidad Nacional de Tucumán <i>LUCCIONI, Griselda María</i>	Universidad del Aconcagua <i>GIMÉNEZ, Rosa</i>
Universidad Nacional Arturo Jauretche <i>MORALES, Martín</i>	Universidad Gastón Dachary <i>RUIDÍAS, Hector Javier</i>
Universidad Nacional del Chaco Austral <i>ZACHMAN, Patricia</i>	Universidad Argentina de la Empresa -
Universidad Nacional del Oeste <i>FOTI, Antonio</i>	Universidad del CEMA <i>GUGLIANONE, Ariadna</i>
Universidad Nacional de Cuyo <i>GARCÍA GARINO, Carlos</i>	Universidad Austral <i>COSENTINO, Juan Pablo</i>
Universidad Nacional de Mar del Plata <i>RÍOS, Carlos</i>	Universidad Atlántida Argentina <i>RATHMANN, Liliana</i>
Universidad Nacional de Quilmes -	Universidad Católica de La Plata <i>BERTONE, Rodolfo</i>
Universidad Nacional de Hurlingham <i>MEDRANO, Gustavo</i>	Instituto Tecnológico de Buenos Aires <i>BOLO, Mario</i>
Universidad Nacional de San Antonio de Areco <i>RAMÓN, Hugo</i>	Universidad Champagnat <i>PINCIROLI, Fernando</i>
Universidad Nacional de San Martín <i>ESTAYNO, Marcelo</i>	
Universidad de Morón <i>PADOVANI, Hugo</i>	
Universidad Abierta Iberoamericana <i>DE VINCENZI, Marcelo</i>	
Universidad de Belgrano <i>GUERCI, Alberto</i>	
Universidad Argentina John F. Kennedy <i>PANIZZI, Marisa</i>	
Universidad Adventista del Plata <i>BOURNISSEN, Juan</i>	
Universidad de la Cámara Argentina de Comercio y Servicios <i>MALBERNAT, Lucía</i>	
Universidad de Palermo <i>ALVAREZ, Adriana</i>	

Coordinadores

Agentes y Sistemas Inteligentes

Marcelo FALAPPA

Universidad Nacional del Sur

Marcelo ERRECALDE

Universidad Nacional de San Luis

Daniel PANDOLFI

Universidad Nacional de la Patagonia Austral

Procesamiento Distribuido y Paralelo

Marcelo NAIOUF

Universidad Nacional de La Plata

Verónica GIL COSTA

Universidad Nacional de San Luis

Javier BALLADINI

Universidad Nacional del Comahue

Arquitectura, Redes y Sistemas Operativos

Luis MARRONE

Universidad Nacional de La Plata

Daniel ARIAS FIGUEROA

Universidad Nacional de Salta

Orlando MICOLINI

Universidad Nacional de Córdoba

Tecnología Informática aplicada en Educación

Zulma CATALDI

Universidad de Buenos Aires

Alejandra ZANGARA

Universidad Nacional de La Plata

Gustavo GIL

Universidad Nacional de Salta

Computación Gráfica, Imágenes y Visualización

Martín LARREA

Universidad Nacional del Sur

María J. ABÁSOLO

Universidad Nacional del Centro de la Provincia de Buenos Aires

Roberto GUERRERO

Universidad Nacional de San Luis

Procesamiento de señales y Sistemas de Tiempo Real

Oscar BRÍA

INVAP

Fernando TINETTI

Universidad Nacional de La Plata

Nelson RODRIGUEZ

Universidad Nacional de San Juan

Ingeniería de Software

Pablo FILLOTTRANI

Universidad Nacional del Sur

Pablo THOMAS

Universidad Nacional de La Plata

Fernanda CARMONA

Universidad Nacional de Chilecito

Bases de Datos y Minería de Datos

Laura LANZARINI

Universidad Nacional de La Plata

Claudia DECO

Universidad Nacional de Rosario

Norma HERRERA

Universidad Nacional de San Luis

Innovación en Sistemas de Software

Monica TUGNARELLI

Universidad Nacional de Entre Rios

Marisa PANIZZI

Universidad Argentina John F. Kennedy

Guillermo FEIERHERD

Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur

Gladys DAPOZO

Universidad Nacional del Nordeste

Innovación en Educación Informática

Claudia RUSSO

Universidad Nacional del Noroeste de la Provincia de Buenos Aires

Elena DURÁN

Universidad Nacional de Santiago del Estero

Lucía MALBERNAT

Universidad de la Cámara Argentina de Comercio y Servicios

Seguridad Informática

Paula VENOSA

Universidad Nacional de La Plata

Jorge ETEROVIC

Universidad Nacional de La Matanza

Javier ECHAIZ

Universidad Nacional del Sur

Antonio CASTRO LECHTALLER

Universidad de Buenos Aires

Jurado Tesis Doctorado

Marcela PRINTISTA

Universidad Nacional de San Luis

Laura DE GIUSTI

Universidad Nacional de La Plata

Silvia CASTRO

Universidad Nacional del Sur

Alejandra CECHICH

Universidad Nacional del Comahue

Horacio KUNA

Universidad Nacional de Misiones

Comité Organizador Local

Rector:

Norberto CAMINO

Vicerrector Académico:

Germán ANTEQUERA

Secretario de Ciencia y Tecnología:

Walter ROBLEDO

Director de Departamento de Ciencias Básicas y Tecnológicas:

Manuel VELASCO

Director de Departamento de Ciencias de la Educación y la Salud:

Rosana QUIROGA

Director de la Escuela de Ingeniería:

Alberto Eduardo RIBA

Comisión curricular Escuela de Ingeniería:

Fernando Emmanuel FRATI

José Daniel TEXIER

Cristian RIOS

Gabriel QUIROGA SALOMÓN

Matías PÉREZ

Emmanuel PORTUGAL

Fernanda Beatriz CARMONA

Mara Lía ROVERO

María Elena LABAQUE

Índice de Artículos

Arquitectura, Redes y Sistemas Operativos	1
Análisis y Aplicaciones de Internet de las Cosas y Ciudades Inteligentes. Framework de trabajo	2
Análisis y Simulación de procesadores RISC-V en plataforma ISA abierta	6
Entorno de Contenedores de Emuladores que contienen Sistemas Embebidos	12
Estudio y Simulación de Redes Definidas por Software y Automatización de Red	18
Protocolo TCP: El RTT como un factor de evaluación del rendimiento	24
Redes de acceso de 5 generación	29
Salón de Clases Inteligente Eficiente Energéticamente	33
Simulación de enrutamiento BGP con GNS3	38
Simulación en la enseñanza de redes de computadoras en el nivel universitario en tiempos de pandemia	43
Tecnología de dispositivos de lógica reconfigurable aplicada en la implementación segura de sistemas de iot	50
Tecnologías de IoT y aprendizaje automático para la solución de problemas en el medio productivo y el cuidado del medioambiente	56
Una arquitectura de microservicios para dar soporte a la creación y ejecución de actividades de recolección de datos con intervención humana	62
Agentes y Sistemas Inteligentes	67
Análisis de textos con estructura	68
Aplicación de técnicas avanzadas de aprendizaje automático para identificar emociones en textos	73
Aprendizaje automático aplicado a la pandemia del virus Covid-19 en Argentina	78
Aprendizaje profundo en aplicaciones biomédicas, agronómicas y ambientales	82
Big Data Optimization con algoritmos metaheurísticos utilizando frameworks de computación distribuida	87
Ciudades inteligentes y sostenibles: iniciativas y desafíos	92
Deep Learning. Aplicaciones en Reconocimiento de Lengua de Señas, Generación de Lenguaje Natural e Imágenes Astronómicas	98
Desarrollo de sistemas de scheduling de producción en el contexto de la industria 4.0	103
Detección de patrones y tendencias en estudiantes universitarios de carreras de ingeniería para determinar el éxito académico aplicando Machine Learning	108
Experticia. Un modelo de Sistema Experto aplicado al Poder Judicial	113
Logística verde de residuos tecnológicos	119
Metaheurísticas, búsqueda estocástica y cómputo eficiente en optimización aplicada	124
Modelo prescriptivo dinámico para un sistema de eventos complejo	129
Modelos de Pronósticos para Incendios Forestales en la Provincia de Córdoba	134
Optimización de Carteras de Inversión: Un Benchmark con Modelos Clásico, de Computación Cuántica y de Hibridación AI /QC	139
Red neuronal multiescala para clasificación de la calidad vocal	144
Redes neuronales para la clasificación de partículas de arena	149

Riesgos, dilemas éticos y buenas prácticas en inteligencia artificial	155
Subsistemas para Análisis de Textos	160
Utilizando Argumentación Rebatible en la Detección de Intrusión en Sistemas Biométricos	165
Bases de Datos y Minería de Datos	170
Análisis de calidad de vida de la población sanjuanina aplicando ciencia de datos	171
Análisis de Datos Simbólicos para Data Science	176
Aplicaciones de Bases de Datos Espaciales y Espacio Temporales en el ámbito agropecuario	181
Aplicaciones de Procesamiento de Lenguaje Natural y Ciencia de Datos	185
Bases de Datos no Convencionales: Índices y Operaciones	191
Ciencia de datos aplicada al estudio la fauna íctica en la zona del río Paraná	196
Ciencia de datos aplicada. estudio de casos en diversas áreas, focalizado en la industria agropecuaria de la provincia de rio negro	200
Ciencia de datos como herramienta de soporte en la gestión pública de calidad del agua	204
Clasificación automática de correos electrónicos	209
Indexación y Administración de Grandes Volúmenes de Datos	214
Inteligencia y analítica de negocios para la toma de decisiones en diferentes contextos	219
Mejoras Algorítmicas para Problemas de Búsquedas en Datos Masivos	224
Procesamiento inteligente de grandes volúmenes de información y de flujos de datos	229
Propuesta para la migración de datos desde SIGEVA-UNSJ hacia Dspace	235
Técnicas de Indexación para Bases de Datos Avanzadas	240
Web Mining y Text Mining: enfoques avanzados para analizar el contenido de grandes cantidades de información	245
Computación Gráfica, Imágenes y Visualización	250
Análisis de datos biométricos en huellas dactilares y su relación con la DM2 a través de ciencia de datos y procesamiento de imágenes.	251
Análisis Visual de Datos Multidimensionales	256
Análisis visual de movimientos microscápicos	260
Integración del Procesamiento Imágenes e Internet de las Cosas en la estimación temprana del rendimiento de cultivos frutales	265
Interacción humano-computadora: estrategias hacia la empatía computacional	270
La Televisión Digital Interactiva para el mejoramiento de los pueblos latinoamericanos	275
Métricas de Inmersión para Sistemas de Realidad Virtual mediante Modelos de Regresión	281
Realidad Extendida y Aplicaciones Móviles aplicadas a la Educación y a brindar Información al Ciudadano	286
Verificación y validación de representaciones visuales y sus interacciones	292
Visualizando la información en ciencia de datos	296
Innovación en Educación Informática	301
Adecuación de estrategias y métodos para fomentar el pensamiento computacional	302
Análisis del impacto de la clase invertida a través del uso de herramientas de la web 3.0 en la adquisición de competencias del perfil profesional del Ingeniero en Sistemas de Información.	308
Aplicación de la herramienta Quizizz como estrategia de Gamificación en la Educación Superior	313
Avances en la formación de docentes y alumnos como Investigadores Científicos Iniciales en Informática y Ciencias de la Computación	318
Diseño de estrategias didácticas para la enseñanza de ingeniería de requerimientos	324
Enseñanza introductoria de la programación: un estudio curricular por competencias	329

EscuelasTIC: estrategias para trabajar el pensamiento computacional en la escuela argentina	338
Experiencia de Enseñanza y Aprendizaje del Pensamiento Computacional en las carreras de Ingeniería	343
Herramientas y tecnologías para el desarrollo web desde el frontend al backend	347
Innovación educativa con tecnologías emergentes en el contexto de las prácticas educativas abiertas . . .	351
La articulación integrada entre academia, investigación y extensión como una experiencia de innovación en educación en Informática para los Cursos de Ingreso de la UNSJ en el marco de SIED UNSJ	355
Las redes neuronales artificiales y la deserción en el nivel universitario	362
Microcomputadoras para la enseñanza y el aprendizaje	368
Prácticas de formación del pensamiento computacional en la enseñanza superior mediada por tecnología	373
Prácticas de programación grupales en el aula estrategias didácticas para el desarrollo del pensamiento computacional en los primeros cursos de programación del ciclo inicial universitario	377
Rediseño de una propuesta didáctica para enseñar a programar	382
Rediseño Educativo para la Enseñanza y Aprendizaje del Cálculo Numérico	387

Ingeniería de Software 394

Aplicación de metodologías para contribuir con la calidad del software en sistemas de gestión y sistemas críticos	395
Aplicaciones Móviles 3D y Realidad Virtual	400
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	404
Atender aspectos lingüístico-cognitivos en la captura de términos del contexto	410
Clasificando información en sitios de CQA	415
Contextos digitales para asistencia de los ciudadanos: enfoques de experiencia de usuario	421
Desarrollo de incumbencias en el modelado de la vista estática	426
Despliegue de sistemas de software: estado de la práctica actual en PyMES de Argentina	431
Enfoque Metodológico Dirigido por Modelos para desarrollos de Internet/Web de las Cosas	435
Enfoques y Tendencias en el Desarrollo de Aplicaciones para Dispositivos Móviles.	440
Estrategia de evaluación de la reusabilidad de los conjuntos de datos de los portales de Infraestructuras de Datos Espaciales.	446
Estudios en torno a la Accesibilidad Web. Aporte regional en el NEA	451
Evaluación del atributo aprendizaje en aplicaciones web	456
Evaluación y mejora de narrativas digitales basadas en datos	461
Exploración de Técnicas de Machine Learning para Migración de Sistemas Legados hacia Microservicios	466
Hacia la definición temprana de requisitos de calidad de datos para productos de información en contexto ágil	471
Hacia un modelo integral de Accesibilidad en Sistemas e-Learning	477
i-QuAGI: Aproximación a un enfoque inteligente para la evaluación de calidad de procesos ágiles de software	481
Implementación de sistema de gestión de calidad utilizando herramientas case	486
Integración de herramientas de prueba automatizadas para evaluar la accesibilidad en aplicaciones web	491
La Sostenibilidad como Característica Transversal a la Calidad del Software	496
Marco de trabajo para la gestión del conocimiento en la administración de proyectos de desarrollo del software	502
Medición automática de aspectos humanos en equipos virtuales de desarrollo de software	507
Modelo de Diseño Instruccional en e-Learning	512

Modelo para curaduría de proyectos software de fuente abierta para estudios empíricos en ingeniería de software	515
Modelo Semántico de Repositorio de Procesos de Negocio para la Gestión de Procesos de Negocio Colaborativos	520
Reusabilidad en el contexto de Desarrollo de Sistemas para Big Data	525
Reuso de Servicios Heterogéneos basado en CBR	530
Sistemas de información organizacionales asistidos por analítica e inteligencia de negocios	535
Sistemas móviles accesibles e inteligentes para una sociedad inclusiva	540
Técnicas y Herramientas para la mejora de procesos, Gobernanza Digital y Calidad de Datos	545
Tendencia actual de la utilización de herramientas de Gestión de Proyectos Software en las instituciones de Educación de Superior de Jujuy	550
Toma de decisiones científica en la ingeniería de software mediante inteligencia computacional y análisis de datos	553
Verificación de Entregables del Proceso de Elicitación en PYMES Argentinas de la Región Centro	558
Innovación en Sistemas de Software	564
Captura multimodal para la inferencia de estados emocionales aplicados a contextos de computación afectiva	565
Diseño y desarrollo de Serious Games para la rehabilitación de problemas cognitivos en pacientes neurológicos implementando VRPN para las comunicaciones entre las interfaces y los dispositivos	570
Diseño y desarrollo de una aplicación móvil de monitoreo y seguimiento para escolares	574
Estrategias de análisis automático de datos urbanos: movilidad, turismo y desigualdad	578
Extracción de datos, análisis de información y predicción del comportamiento en el rubro económico chubutense en contexto de COVID-19	583
Framework para la generación de Procedimientos Semánticos aplicados a Navegadores de Realidad Aumentada	588
Investigación, desarrollo y publicación de un prototipo de segmento terreno satelital	594
Juegos serios de realidad aumentada orientados a entrenar y recuperar movimiento en niños con discapacidad	599
Mejorando la performance en Aplicaciones Web Progresivas mediante estrategias de utilización de la Cache	604
Procesamiento de datos meteorológicos para determinar la ocurrencia de heladas en la agricultura	609
Prototipo de Sistema para la Gestión de Controles de Tránsito Vehicular	614
Realidad aumentada y realidad virtual aplicadas a proyectos con fines sociales	618
Reconocimiento de emociones en la voz empleando redes neuronales y su integración en frameworks multimodales de educación emocional	623
Redes de conocimiento para la ayuda de cercanía - instrumento para la reactivación microeconómica	628
Seguimiento de Pacientes en Rehabilitación Robótica	633
Sistema multiplataforma para el registro y asistencia de pacientes con síndrome de intestino irritable	638
Sistemas de control de robots sobre plataforma de internet de las cosas	643
Supervisión y control de procesos	646
Tecnologías 4.0 para la Industria Argentina - Análisis y herramientas de IDi para la formulación de políticas y el desarrollo de la Industria 4.0	652
Tecnologías de Sistemas de QA aplicadas a la Web Semántica	657
Tecnologías emergentes aplicadas a problemas agrícolas en regiones áridas	662
Transformación Digital en tiempos de la Industria 4.0	667
Procesamiento Distribuido y Paralelo	673

Algoritmos paralelos y evaluación de rendimiento en plataformas de HPC	674
Arquitecturas Edge- Fog - Cloud en Procesamiento Distribuido	680
Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones	686
Cálculo Paralelo para la Multiplicación de Matrices	692
Cloud Computing, IaaS privados y públicos para el análisis de performance	695
Cómputo Paralelo y Distribuido: Métrica de Rendimiento, Aplicaciones de Datos Masivos e Inteligencia Artificial	700
Métodos de Inteligencia artificial para la reducción de incertidumbre en modelos de predicción de incendios	705
Optimización en el desarrollo de aplicaciones Serverless en entornos distribuidos	710
Procesamiento de problemas de Mecánica de Sólidos en entornos de Cloud Computing. El caso de estudios paramétricos	715
Procesamiento paralelo sobre arquitecturas serverless para tratamiento de datos provenientes del IoT .	720
Redes neuronales paralelas aplicadas a la visión computacional	725
Sintonización de un Método de Reducción de Incertidumbre: un proceso de medición-mejora para incrementar la eficiencia	729
Técnicas de modelado y simulación para arquitecturas HPC y salud	734
Tecnología HPC como motor de ciencia de la UNdeC	739
Tecnologías HPC Híbridas en la Solución de Problemas Complejos	744
Procesamiento de señales y Sistemas de Tiempo Real	749
Análisis de la revolución industrial 4.0 y Big Data. Proyecto de software: Oxímetro fotográfico	750
Modelos Matemáticos y Métodos Computacionales en ingeniería	755
Redes de Sensores, Robots Móviles y Simulación en Sistemas de Tiempo Real	759
Seguimiento del estado de una cámara de germinación hidropónica utilizando email basado en node-red	763
Técnicas de Optimización de Soluciones en Sistemas Embebidos	772
Redes de Cooperación Científica Internacionales	777
Consorcio Cloud Computing, Big Data y Emerging Topics	778
El Uso de Técnicas Computacionales para Mejorar el Cumplimiento de Técnicas del Recuerdo en Entornos Inteligentes (REMIND)	784
Proyecto CAP4CITY	787
Seguridad Informática	792
Aplicación de los Contratos Inteligentes en Internet de las Cosas	793
Aplicación de redes neuronales profundas para la detección automática de Nombres de Dominio generados de manera algorítmica	799
Avances en Aspectos de Seguridad en el Sistemas de Voto Electrónico OTP-Vote	804
Avances en Robustecimiento ante Ataques de Presentación y Falsificación para Sistemas Basados en el Análisis de Patrones de Tecleo	809
Criptografía Liviana para Internet de las Cosas e Internet de las Cosas Industrial	814
Criptografía Maliciosa y Ciberdefensa	818
Desarrollo de una dApp académica en la red Blockchain Federal Argentina	822
Desarrollo de una Guía para el abordaje de Incidentes de Ciberseguridad en Infraestructuras Críticas Industriales	827
Detección de vulnerabilidades en especificaciones de contratos inteligentes de la plataforma Ethereum	832
Indicadores de ciberseguridad de una red informática en un Laboratorio EDI	837

Investigación en ciberseguridad en un año de pandemia	843
Sistema Inteligente de Detección de Anomalías para IoT	848
Un Sistema de Voto Electrónico para la FCEyN (UNLPam.)	853
Tecnología Informática aplicada en Educación	858
Agregación de la Tecnología Móvil en los Aprendizajes	859
Análisis de los recursos de visualización de datos usados en un marco de trabajo de inteligencia de negocios aplicados en la educación superior	864
Desarrollo de instrumentación virtual para uso en carreras de ingeniería	869
Desarrollo y evaluación de competencias en la ingeniería de software en un entorno virtual de aprendizaje colaborativo	875
Diseño y construcción de objetos de aprendizaje basada en criterios de calidad	881
Diseño y desarrollo de herramientas y entornos digitales para escenarios educativos híbridos	886
EduScrum un Marco de Trabajo que puede Propiciar Aprendizaje Significativo	893
Estudio de herramientas para la implementación de laboratorios remotos	898
Experiencia de accesibilizar un MOOC considerando la plataforma y recursos multimediales	903
Identificación de regularidades en matemática, mediante determinación de patrones y con uso de software libre	909
Impacto de las condiciones socioeconómicas en el rendimiento y la deserción estudiantil	914
Interacción natural, entornos inmersivos y otras tecnologías emergentes aplicadas a contextos educativos	919
Intranets educativas para escuelas rurales aisladas de Salta	926
Jabutí EDU, plataforma de Robótica Educativa IOT open Hardware	931
Juegos serios y realidad aumentada. Oportunidades y retos para su inclusión en propuestas educativas	935
La Gamificación como estrategia didáctica en Geometría	940
Marco de referencia para la formulación, gestión e implementación de carreras a distancia según el SIED UNSJ. Desarrollo de Manual de procedimientos	945
Materiales educativos digitales para la construcción de conocimientos en la universidad	953
NetOS-Lab: Laboratorio portátil de Redes y Sistemas Operativos	958
Realidad virtual, aprendizaje inmersivo y realidad aumentada: Casos de Estudio en Carreras de Ingeniería	963
Relación entre Analítica del Aprendizaje y Diseño del Aprendizaje	969
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	973
Tecnologías Digitales y Colaborativas en Educación	978
Tecnologías Emergentes	984
Una aproximación para el Diseño de una Aplicación destinada a la confección de Diagramas de clases - Caso estudiantes universitarios con disminución visual	993
Una propuesta para la mejora de funciones ejecutivas en niños con autismo a través del desarrollo de habilidades de Pensamiento Computacional	997
Tesis de Doctorado	1002
Coordinación de dispositivos en ambientes ubicuos mediante coreografías	1003
Framework para la evaluación de calidad de procesos ágiles	1013
Integración escalable de realidad aumentada basada en imágenes y rostros	1023
Interacción Tangible en escenarios educativos. Diseño de una herramienta de autor para la creación de actividades educativas basadas en Interacción Tangible	1032
Método de Reducción de Incertidumbre Basado en Algoritmos Evolutivos y Paralelismo Orientado a la Predicción y Prevención de Desastres Naturales	1040

Obtención de reglas de clasificación difusas utilizando técnicas de optimización Caso de estudio	
Riesgo Crediticio1050
Proceso marco orientado a aspectos en las etapas tempranas del ciclo de vida del desarrollo de software para una transición en la industria1061

Autores

A

ABALLAY, Alicia: 535.
ABÁSULO, María José: 286, 275, 588, 886.
ACEVEDO, Joaquín: 461, 395.
ACEVEDO, Jonatan Uran: 822.
ACEVEDO, Karvin Díaz: 692, 725, 318.
ACOSTA, Nelson: 633, 599, 435.
ACUÑA, Cesar: 481.
ADAGIO, Matías: 12.
ADO, Mariana: 984.
ADORNO, Sebastián: 984.
AGAMENNONI, Osvaldo: 260.
AGÜERO, Gustavo: 200.
AGUILERA, Cynthia: 535.
AGUIRRE, Federico: 2.
AGUIRRE, Jesús Francisco: 351.
AGUIRRE, Sofía: 859.
AGUIRRE, Verónica: 440, 404.
AHMAD, Tamara: 984.
AIDELMAN, Y.: 98.
ALANIZ, Mariano: 219.
ALANIZ, Mario: 869.
ALBA, M. Daniela: 191.
ALBORNOZ, Enrique M.: 33.
ALDERETE, Claudia: 604.
ALDERETE, Romina Y.: 451.
ALE, Zamira: 720.
ALFONSO, Hugo: 87.
ALFONZO, Pedro L.: 451.
ALLAN, Claudia: 978.
ALLENDES OLAVE, Paola Andrea: 351.
ALLISIARDI, Tomás: 286.
ALONSO, Daniel: 56.
ALONSO DE ARMIÑO, Ana: 978.
ALVARADO, Yoselie: 270.
ALVAREZ, Jonatan: 739.
ALVEZ, Carlos: 165.
AMADEO, Ana Paola: 903.
AMARO, Marcos: 643.
AMAYA, Fabrizio: 296.

AMOR, Matías: 73, 78.
AMORE, Marcelo D: 869.
ANABALON, Diego: 530.
ANDRADA, José L.: 456.
ANDRADE, Daniel: 324.
ANTONINI, Antonela: 256.
AÑAIS, Leonardo: 931.
ARAMAYO, Ubaldo J. M.: 646.
ARANDA, Gabriela: 415.
ARANDA, Marcos: 869.
ARANDA ROMERA, Juan Antonio: 667.
ARCHUBY, Federico: 919, 886.
ARCIDIACONO, José: 62.
ARGAÑARAS, Pablo E.: 200.
ARGAÑARAZ, Jorgelina: 646.
ARGAÑARAZ, Mauro: 832.
ARGÜELLO, Daniel: 6.
ARIAS FIGUEROA, Daniel: 43, 486, 38, 710.
ARIZA, Alejandro: 235.
ARMANDO, Silvana: 471.
ARROYUELO, Jorge: 191.
ARTOLA, Verónica: 1032, 919.
ASTUDILLO, Gustavo Javier: 382, 940, 973, 919.
ATENCIO, Hernán: 710.
ATÍA, Julissa: 82.
AUDISIO, Melina: 550.
AZAR, Paola: 240.
AZCURRA, Diego: 643.

B

BACIGALUPE, María de los Ángeles: 914.
BAEZA, Natalia: 978.
BALAGUER, Federico: 265.
BALDASSARRI, Sandra: 919.
BALLADINI, Javier: 700, 686.
BALLESTEROS, Carlos: 859.
BALMACEDA, Silvina: 355.
BALMACEDA CASTRO, Iván: 477, 512.
BANCHOFF TZANCOFF, Claudia M.: 618, 338.

BARBERIS, Ángel R.: 387.
BAREIRO, Hernán: 33.
BARRERA, María A.: 446.
BARRERA NICHOLSON, Paula: 540.
BARRIENTO, Valeria: 245, 108.
BARRIONUEVO, Carlos: 565.
BARRIONUEVO, Mercedes: 958, 744.
BARRIOS, Teresita Haydee: 308.
BASCIANO, Ivan: 935.
BASGALL, María José: 674.
BASPINEIRO, Rodolfo: 43.
BAST, Silvia: 804, 382, 973.
BATTAGLIA, Nicolás: 875.
BAVERA, Francisco: 997.
BAZÁN, Patricia: 62.
BECERRA, María del C.: 103, 535.
BECERRA, Martín Ezequiel: 588.
BEGUERÍ, Graciela: 296.
BELCASTRO, Ángela: 893.
BELLONI, Edgardo: 435.
BELTRAMINI, Paola: 869.
BENDER, Cristina: 343.
BENQUERENCA MENDEZ, Nicolás: 695.
BERMÚDEZ, Carlos: 87.
BERNAL, Edna: 784.
BERÓN, Mario: 832, 804.
BERTACCINI, Daniel: 124.
BERTERO, Regina: 149.
BERTONE, Rodolfo: 859, 404, 893, 431.
BIANCHINI, Germán: 318, 729, 725, 705, 692.
BILBAO, Martín: 119.
BJERG, Ernesto: 256.
BLANC, Rafael: 558.
BLANCHET, Raul: 18.
BLARIZA, Juan Pablo: 33.
BOLANO, J. J.: 204.
BOLATTI, Diego: 848, 2.
BOND, Román: 695, 734.
BORACCHIA, Marcos: 545.
BORDÓN, José M.: 451.
BOSSERO, julio: 113.
BOUCIGUEZ, María José: 286.
BOURDETTA, Viviana: 558.
BRANDAN, F.: 119.
BRAÑA, Juan Pablo: 139.
BRIA, Oscar: 755.
BRITOS, Luis: 214.

BRITOS, Pamela: 200.
BRITOS, Paola: 200, 196.
BUCCELLA, Agustina: 525.
BUFFARINI, Abril: 919.
BUSNELLI, Luis: 113.
BUSUM FRADERA, Matías: 56.
BUTELER, Luciana: 355.

C

CABEZAS, Fernanda: 550.
CADENA, Carlos: 926.
CAFFARATTI, Gabriel: 799.
CAFFETTI, Yanina Andrea: 750.
CALABRESE, Julieta: 545.
CALOT, Enrique P.: 809.
CALVO, Inés: 945.
CAMELE, Genaro: 229.
CAMPOS, Matías: 570.
CANO, Eduardo: 869.
CANTO, Isabel: 355.
CAÑIBANO, Rodrigo: 700.
CAPELLO, Ivan: 426.
CAPPELLETTI, Marcelo: 82, 56.
CAPRARO, Flavio: 609.
CARBONARI, Daniela: 963.
CARDENAS, Marina E.: 68, 160, 134.
CARDOSO, Alejandra Carolina: 73, 78.
CARDOZO, Claudia: 421.
CAREGLIO, Claudio: 715.
CARMONA, Emilse: 355, 945.
CARMONA, Fernanda Beatriz: 881, 662, 739.
CARNUCCIO, Esteban: 12.
CARO, Angélica: 471.
CARRILAO ÁVILA, Federico: 903.
CARRIZO, Alejandra: 421.
CARRIZO, Claudio: 471.
CARRIZO, Jorgelina: 176.
CARRUTHERS, Juan A.: 515.
CASALINI, María Clara: 292.
CASANOVA, Carlos: 553, 558, 129, 181, 185.
CASFUZ, Damian: 410.
CASTILLO, Julio: 160, 134, 68.
CASTRO, Leandro: 382.
CASTRO, Silvia: 281, 260, 251, 292, 256.
CASTRO LECHTALER, Antonio: 881.
CATANIA, Carlos: 799.

CAVASIN, Nicolas: 209.
CAYMES SCUTARI, Paola: 729, 318, 725, 705, 692.
CECHICH, Alejandra: 525, 530, 415.
CEDARO, Karina: 558.
CELESTE, Carolina: 978.
CELLERI, Milagros: 638.
CENCI, Karina M.: 787.
CHALABE, Susana: 646.
CHAPPERON, Gabriela: 565.
CHARNE, Javier: 984.
CHÁVEZ, Edgar: 191.
CHECARELLI, Diego: 219.
CHEDIAK, Ernesto: 50.
CHESÑEVAR, Carlos Iván: 362.
CHIARANI, Marcela: 351.
CHIAROTTO, Agustín: 700.
CHICHI, Manuel: 553.
CHICHIZOLA, Franco: 674, 686, 680.
CHIODI, Andrea: 633.
CHIRINO, Pamela: 705, 725, 692, 318.
CHUK, Oscar Daniel: 149.
CIANCIO, María I: 909.
CICERCHIA, Benjamín: 984.
CICINELLI, Diana: 329.
CIFUENTES, Franco: 604.
CIPRIANO, Marcelo: 818, 793.
CLARA, Bibiana Beatriz Luz: 155.
COCCONI, Diego: 520.
COLACIOPPO, Nicolás: 864.
COLLAZOS ORDÓÑEZ, César Alberto: 286.
COLLOTI, Guillermo: 33.
COLUSSI, Natalia: 377.
COMPANY, Ana: 395, 302.
CONSTABLE, Leticia: 864.
CONSTANTINI, Francisco: 744.
CONSTANZO, Manuel: 680.
CONTI, Laura: 113.
CORBALÁN, Leonardo: 440.
CORGATELLI, Franco: 530.
CORO, Octavio Daniel: 108.
CORSO, Cynthia Lorena: 864.
CORTEZ, Joaquín: 296.
CORTEZ, Lucía: 799.
CORTEZ OVIEDO, Sebastián: 652.
COSTANZO, Manuel: 674, 686.
CRISTALDO, Patricia: 185.
CRISTINA, Federico: 400.

CRUZ, Alejandro: 599, 633.
CRUZ, Marcos: 525.
CUADROS, Patricia: 953.
CUBA L., Cinthia A.: 196.
CUNEO, Juan Manuel: 103.

D

DA ROCHA ARAUJO, Leonardo: 466.
DAL BIANCO, Pedro: 98.
DANIELE, Marcela: 491, 997.
DAPOTO, Sebastián: 400.
DAPOZO, Gladys N.: 302, 395.
DASSO, Aristides: 456.
DAVID, María A.: 265.
DAZA, Monica: 351.
DE ANTUENO, Joaquín: 680, 686.
DE BATTISTA, Anabella: 185, 240, 553.
DE DIOS BENÍTEZ, Juan: 435.
DE GIUSTI, Armando: 686, 680, 674, 759, 286, 275, 700, 787.
DE GIUSTI, Laura: 674, 680, 686.
DE LA TORRE, María de los Angeles: 240.
DE LUCA, Graciela: 12.
DE RENZIS, Alan: 530.
DE VINCENZI, Marcelo: 875.
DECO, Claudia: 343, 214.
DEJEAN, Gustavo: 265.
DEL MORAL SACHETTI, Lorena E.: 387.
DEL PEZO IZAGUIRRE, Evelyn: 286.
DEL VALLE, Marta: 251.
DEL VALLE, Nicolás: 993.
DELGADO, Cesar Daniel: 574.
DELIA, Lisandro: 440.
DELVECHIO, Tomás: 224.
DENON, Nicole: 82, 56.
DESTEFANIS, Eduardo A.: 144.
DI BONIS, Gastón: 200.
DI CICCIO, Carlos: 984.
DI GENARO, María E.: 191.
DI GIONANTONIO, Alejandra: 993.
DÍAZ, Alejandro: 633.
DÍAZ, Daniel: 103.
DIAZ, Gabriela: 787.
DÍAZ, Javier: 843, 903.
DIAZ ARAYA, Carlos Daniel: 667.
DÍAZ LABRADOR, Marycarmen: 24.

DIAZ-ACEVEDO, Karvin: 705.
DIBEZ, Pablo: 893.
DIESTE, Oscar: 1003.
DIVI, Pablo: 540.
DOGLIOTTI, Mariano: 604.
DOMINGUEZ, Ana Patricia: 355, 945.
DOMINGUEZ, Nicolás: 638.
DOMINGUEZ, Samuel Franco: 646.
DONADIO, Maximiliano: 200.
DOORN, Jorge: 410, 265.
DORZÁN, María Gisela: 181.
DOTI, Santiago: 56.
DRAZIC, Leandro: 219.
DUARTE, David: 695.
DUCID, Matías: 583.
DUFOR, Fernando: 29.
DURANTE, Mateo: 843.

E

ELIZALDE, María Celia: 410.
ELKFURY, Fernando: 623, 565.
ENCINAS, Diego: 695, 734, 759, 124, 686, 56, 680.
ERRECALDE, Marcelo: 92, 209.
ESCALANTE, Julián: 744.
ESCUDERO, C.: 98.
ESNAOLA, Leonardo: 984.
ESPÍNDOLA, María Cecilia: 302.
ESPINILLA, Macarena: 784.
ESPONDA, Silvia: 545.
ESTEVEZ, Elsa: 652, 787.
ESTREBOU, César: 229, 686.
ETCHART, Graciela Raquel: 165.
ETEROVIC, Jorge E.: 793, 822, 814.

F

FACCHINI, Higinio: 6, 18.
FACHAL, Adriana: 886.
FARIÁS, Andrés: 853.
FASSIO, Esteban: 124.
FAVA, Laura: 618.
FAVRET, Fabián: 33.
FELDMAN, Patricio: 652.
FENNEMA, Marta C.: 540.
FERNANDEZ, Alejandro: 139.
FERNANDEZ, Gustavo: 643.

FERNANDEZ, Jacqueline: 270.
FERNÁNDEZ, Juan Manuel: 209.
FERNÁNDEZ, Mariela: 512, 477.
FERNANDEZ BARIVIERA, Aurelio: 229.
FERNANDEZ SOSA, Juan: 440.
FERRACUTTI, Gabriela: 256.
FERREYRA, Juan Pablo: 520.
FIGUEROA, Karina: 191, 214.
FIGUEROLA, Patricia: 739.
FILIPPI, Jose Luis: 859.
FILLOTTRANI, Pablo Rubén: 652, 787.
FINQUELIEVICH, Susana: 652.
FLAQUE, Valeria: 945.
FLORES, Andrés Pablo: 530.
FOIS, Giuliana: 200.
FONTANA, Daniel: 50.
FORESI, Pedro D.: 869.
FRACCHIA, Carina: 978.
FRANZANTE, Juan Manuel: 185.
FRATI, Fernando Emmanuel: 739, 674.
FRATI, Francisco: 739.
FREGA, Gerardo: 113.
FUNES, Ana: 456.

G

GAETAN, Gabriela: 421.
GAGLIARDI, Marisa: 881.
GAGLIARDI, Olinda Edilma: 181.
GALARZA, Brian: 695.
GALDAMEZ, Mariela: 725, 705, 692, 318.
GALLARDO, Facundo: 507.
GALLARDO, Vanesa: 909, 953.
GALLO, Silvana: 674.
GAMARRA, Alvaro: 38, 43.
GAMEN, R.: 98.
GANGA, Leonel: 535, 176, 171.
GANUZA, María Luján: 256, 292.
GARAY, Cristian: 638.
GARCIA, Adrian: 853.
GARCIA, Edith: 818, 814.
GARCIA, Mario Alejandro: 144.
GARCÍA, A.: 535.
GARCÍA, Berta Elena: 351.
GARCÍA, Pablo: 853.
GARCÍA, Sergio: 113.
GARCIA GARINO, Carlos: 715, 50.

GARCÍA RAVLIC, Ignacio: 410, 265.
GARGIULO, Florencia: 256.
GARINO, Carlos Garcia: 898.
GARRIDO, Gladys: 196.
GARRIGA, Martín: 530.
GASTAMINZA, Manuel: 935.
GAUDIANI, Adriana Angélica: 124, 734, 674.
GAVILANES, Wilma: 286.
GERLERO, Patricia: 502.
GIACOMANTONE, Javier: 755.
GIBELLINI, Fabian: 837.
GIGANTE, Nora: 822.
GIL, Evangelina: 935.
GIL, Gustavo D.: 43.
GIMSON, Loraine: 43.
GIORDANO, Andrés: 224.
GIROLIMO, Ulises: 652.
GODOY, Diego Alberto: 33, 435.
GODOY, Pablo: 898.
GODOY, Pablo Daniel: 50.
GODOY, Sebastián: 355.
GOIN, Martin M. J.: 935.
GOMEZ, Julian: 82.
GOMEZ, Martin: 710.
GOMEZ, Mauro: 734.
GOMEZ, Sofía: 313.
GOMEZ, Soledad: 338.
GÓMEZ, Adriana María: 219.
GÓMEZ, Cintia: 351.
GÓMEZ, Verónica: 351.
GONZALEZ, Agustín: 224.
GONZALEZ, Alejandro: 886.
GONZALEZ, Gerardo: 827.
GONZALEZ, Liliana Mirna: 953, 945.
GONZALEZ, Nahuel: 809.
GONZALEZ, Rodrigo: 570.
GONZALEZ, Simón Pedro: 171.
GONZALEZ, Valeria A.: 486.
GOÑI, Oscar: 772.
GORDILLO, María Luisa: 535.
GORGA, Gladys: 919, 886.
GOZALEZ, Rodrigo: 124.
GRAFFIGNA, Carlos Esteban: 739.
GRAMAJO, Sergio: 2, 784, 848.
GRANADA, A.: 98.
GREINER, Cristina: 302, 395.
GROLLER, Wien Eduard: 256.

GRONDA, María Laura: 451.
GROSSO, Alejandro: 191.
GUERRA, Jorge: 799.
GUERRERO, Roberto: 270.
GUEVARA, Diego: 219.
GUIDET, Sebastian: 739.
GUIGUET, Marcelo: 984.
GUTIERREZ, Juan P.: 646.
GUTIERREZ, Mario: 355.
GUZMÁN, Jéscica V.: 82.

H

HADAD, Graciela: 410.
HARARI, Ivana: 903.
HASPERUÉ, Waldo: 229, 809.
HENRIQUES, Pedro Rangel: 832.
HERNANDEZ, Nicolás: 160, 134.
HERRERA, Aureliano: 869.
HERRERA, M.: 535.
HERRERA, Myriam: 171.
HERRERA, Norma Edith: 240, 185.
HERRERA, Susana: 540.
HODES, Agustín: 431.
HOET, Leonardo: 553.

I

IBAÑEZ, Eduardo: 404, 886.
IBAÑEZ, Francisco: 103, 667.
IERACHE, Jorge: 565, 623, 809, 638, 588, 594.
IERACHE, Rocío: 565.
IGLESIAS, Luciano: 919, 886.
IOST, Daniel: 324.
IRASTORZA, Ramiro M.: 82.
IRRAZÁBAL, Emanuel: 515, 461, 395.
ISAJA, Melisa: 200.
ISTVAN, Romina Mariel: 914.
ITUARTE, Luis E.: 646.

J

JARA, Jimena: 734.
JERÉZ, Marcela: 324.
JIMBO SANTANA, Patricia: 1050, 229.
JIMENEZ, Adriana: 540.
JOFRÉ, Mónica: 191.

JOFRÉ, Nicolás: 270.
JURÁN, Tomás: 224.

K

KAMLOFSKY, Jorge: 827.
KARANIK, Marcelo J.: 848.
KASIÁN, Fernando: 214.
KIMURA, Isabel: 338.
KLENZI, Raúl Oscar: 296, 609.
KOUTSOVITIS, Florencia: 638.
KUCUK, Lucas G.: 931.
KULESZA, Raoni: 275.
KUNA, Horacio: 204.

L

LA PIETRA, Lucas: 185.
LAFUENTE, Guillermo Javier: 859.
LAFUENTE, Gustavo: 859.
LAMAS, Daniel A.: 347.
LANCIOTTI, Julieta: 680, 686.
LANFRANCO, Einar: 843.
LANZARINI, Laura: 98, 229.
LARREA, Martín: 292, 281.
LARROSA, Norberto: 978.
LASAGNA, Valeria: 914.
LAURENT, Roberto: 978.
LAVALLÉN, Pablo: 224.
LAZARTE, Ivanna: 313.
LEDESMA, Viviana: 113.
LEGUIZAMÓN, Guillermo: 119.
LEIVA, Alfredo: 945.
LEIVA, Lucas: 772.
LEMONS DE SOUZA FILHO, Guido: 275.
LENCINA, Paula: 984.
LEO, Rosana: 496.
LÉPEZ, Héctor: 535.
LESCANO FARIAS, Lara V.: 646.
LEZCANO, Andrea: 461.
LIBUTTI, Leandro: 686, 680.
LIGORRIA, Laura: 993.
LITTERIO, Alejandra: 139.
LLITERAS, Alejandra B.: 62.
LOBOS, Martín: 973, 853.
LOMORO, Dalma Agostina: 793.
LOPEZ, M.: 92.

LÓPEZ, Paula: 229.
LOPRESTI, Mariela: 744.
LOVOS, Edith: 935.
LUCERO, Maximiliano: 744.
LUDUEÑA, Verónica: 214, 191.
LUENGO, Pablo: 984.
LUGONES, Leopoldo: 486.
LUJÁN GANUZA, María: 260.
LUNA, Adriana: 149.
LUND, María Inés: 219, 171.
LUQUE, Emilio: 700.
LUQUE, Leandro: 260.
LUZURIAGA, Juan Manuel: 525.

M

MACCALLINI, Lucas: 734.
MACEDO, Adriana: 540.
MACHACA, Darío Alejandro: 29.
MACIA, Nicolás: 843.
MADOZ, Maria Cristina: 886.
MAIORANO, Ariel: 818.
MALBERNAT, Lucia Rosario: 155.
MALBERTI RIVEROS, María Alejandra: 296.
MALDONADO, Marilena: 540.
MALLEA, Adriana: 176.
MALVACIO, Eduardo: 818.
MAMANI, Federico Nahuel: 978.
MANERA, Roxana: 993.
MANGIARUA, Nahuel Adiel: 1023, 588.
MANRESA-YEE, Cristina: 886.
MANRIQUES, Patricia: 881.
MANSILLA, Alejandra: 859.
MANSILLA CURAK, Roberto Fernando: 898.
MARCHETTA, Martín G.: 799.
MARIANETTI, Osvaldo: 50, 898.
MARIN, Maria Bianca: 308.
MARINELLI, Marcelo: 763.
MARIÑO, Sonia: 451.
MÁRQUEZ, Victorina: 945.
MARRERO, Luciano: 404.
MARTIN, Eliana Sofía: 338.
MARTÍN, Adriana: 720, 710.
MARTÍN, Adriana Elba: 421.
MARTÍN, Sofía: 618.
MARTINEZ, Cecilia: 176.
MARTINEZ, Julieta: 646.

MARTINEZ, M. Vanina: 165.
MARTINEZ, María Roxana: 604.
MARTÍNEZ, Rodolfo: 525.
MARTÍNEZ, Sergio L.: 646.
MARTINEZ CAROD, Nadina: 415.
MARTINS, Adair: 978.
MASANET, María Isabel: 609.
MASCHERONI, Agustín: 395.
MASSÉ PALERMO, María Laura: 926.
MATKOVIC, Kresimir: 256.
MAUREL, María del Carmen: 308.
MAZALÚ, Rafaela: 525.
MEDEL, Diego: 710, 720.
MEDINA, Javier: 784.
MEDINA, Santiago: 680, 759.
MEDINA, Yanina: 302.
MEDRANO, José Federico: 245, 108.
MENDEZ, Eusebio: 387.
MENDEZ-GARABETTI, Miguel: 1040, 729.
MENDOZA, Dante H.: 265.
MERCADO, Viviana: 92.
MERENDA, Franco: 214.
MERKEL, German: 594.
MERLINO, Hernan: 638.
MERLINO, Hernán: 809.
MICIELI, Gustavo: 29.
MIGANI, Silvina: 219.
MIGNOLA, Rodrigo: 185.
MILLON TELLO, Roberto: 739.
MINDLIN, Ivan: 98.
MINETTI, Gabriela: 87.
MINETTI, Yamila: 382.
MIRANDA, Natalia: 744.
MIRASSO, Aníbal: 715.
MIRET, Franco: 185.
MOLERO, Betina: 978.
MOLINA, Ana Laura: 720, 710.
MOLINA, Claudio: 935.
MOLINA, D.: 92.
MOLINA, Maria Laura: 219.
MONCHO, Matias: 324.
MONGE, Agustina: 73.
MONTANARI, Gisela: 997.
MONTEJANO, Germán: 804, 1003.
MONTEJANO, Rio Germán: 853.
MONTES DE OCA, Damian: 984.
MONTES DE OCA, Erica: 674, 686.

MONTES DE OCA, Federico : 734.
MONTEZANTI, Diego: 56, 686, 680.
MONTTOYA SUÁREZ, Lina María: 200.
MORALEJO, Raul: 739.
MORALES, María I.: 540.
MORALES, Martin: 695, 734, 82, 56.
MORÁN, Marina: 700.
MORENO, Juan Gabriel: 993.
MORERO, Franco: 87.
MOTTA, Elías: 426.
MOYANO, Ezequiel: 324.
MOYANO, Marcelo: 530, 525.
MULLICUNDO, Felipe: 347.
MUÑOZ, Ana: 963.
MUÑOZ, Martín: 609.
MUÑOZ, Roberto: 837.
MUÑOZ, Rocío: 545, 787.
MURAZZO, María: 710, 720.
MURILLO, Natalia: 265.

N

NAIOUF, Marcelo: 674, 686, 680, 286, 275.
NEIL, Carlos: 875.
NERVO, Facundo: 565.
NODA, Carina: 415.
NORDIO, Mauricio: 919.
NORIEGA, Jaquelina: 351.
NOTRENI, Juliana: 837.
NÚÑEZ, Álvaro: 646.
NÚÑEZ, Enrique A.: 149.

O

ODENA, María Belén: 652.
OLGUIN, Luis Alberto: 235, 296.
OLIVA, Elisa S.: 909.
OLIVARES, Diego: 214.
OLIVERA, Lucas: 82, 56.
OLMOS, Jorge: 638.
OLSOWY, Verena: 440, 404.
ONTIVEROS, Patricia: 318.
ORDINEZ, Leo: 583, 578.
ORELLANA, Alejandra: 219, 945.
ORLANDI, Cristina: 700.
OROZCO, Sergio: 119.
ORTEGA, Manuel: 296.

ORTIZ, Felipe: 431.
OSIO, Jorge R.: 56, 695, 82.
OSORIO, Alejandra: 903.
OSYCKA, Liam: 525.
OTAZÚ, Alejandra: 667.
OVIEDO, Sandra: 667, 103.
OYARZUN, Matías: 657.

P

PACINI, Elina: 715.
PAGNONI, Verónica Karina: 451.
PALAU, Franco: 799.
PALAVECINO, Rosa: 540.
PALMA, Marcelo: 426.
PALMERO, Pablo Rafael: 181.
PANDOLFI, Daniel: 119, 92.
PANESSI, Walter F.: 628.
PANIEGO, Juan Manuel: 759, 686, 680.
PANIZZI, Marisa: 969, 431, 614, 404.
PANTANO, Juan Cruz: 535.
PAPELEUX, Luc: 715.
PAREDES, Rodrigo: 191.
PARRA, Lorena Vanesa: 720, 710.
PARRA, Susana: 978.
PASCAL, Andrés: 240.
PASCUAL, Kevin Isaías: 978.
PASINI, Ariel: 545, 404, 787.
PATERNO, Facundo: 525.
PAULETTI, Clarisa: 940.
PAUTSCH, G.: 204.
PAZO ROBLES, María Eugenia: 818.
PELIZA, Carlos: 29.
PERALTA, Lorena: 993.
PERALTA, Mario: 477, 496, 471.
PEREIRA DOS SANTOS, Rodrigo: 466.
PEREYRA RAUSCH, Fernando: 553.
PEREZ, Daniel: 859.
PEREZ, Gabriel: 984.
PEREZ, Marisa: 520.
PEREZ, Matías: 662, 739.
PEREZ, Santiago C.: 963, 6.
PÉREZ, D.: 119.
PÉREZ, Santiago: 18.
PEREZ IBARRA, Susana: 347.
PEREZ VILLAR, Gustavo: 113.
PESADO, Patricia: 400, 545, 440, 404, 286, 275, 886,

787.
PETRIS, Raquel H.: 302.
PI PUIG, Martin: 759, 686, 680.
PICCIN, Ana M.: 329.
PICCOLI, María Fabiana: 129, 744, 720.
PICKELNY, M.: 355.
PIERGALLINI, Maria Rosana: 984.
PINA, Joaquín: 275.
PINCIROLI, Fernando: 1061, 426.
PINTO, Juan A.: 515.
PINTO, Noelia: 1013, 481.
PINTO FIGUEROA, Sonia Eleonora: 219.
PISKORZ, Maria: 638.
PIZZINGRILLI, Paola: 200.
POL'LA, Matías: 525.
POLLO-CATTANEO, María Florencia: 502.
PONCE CEVOLÍ, Ismael: 935.
PONCE DE LEÓN, Alejo: 692, 725, 705, 318.
PONTHOT, Jean Philippe: 715.
PONTORIERO, Francisco: 953.
PONZIO, Claudio Luis: 853.
PORTUGAL MURCIA, Emmanuel Alejandro: 662, 881, 739.
PÓSITO DE ROCA, Rosa María: 945, 355.
POUSA, Adrián: 674, 686, 680.
PRADO, Lucas: 553.
PRALONG, Lourdes: 558.
PREISEGGER, Juan Santiago: 545, 787.
PRETTO, Jeremias: 843.
PRINTISTA, Marcela: 214.
PROCOPIO, Gastón: 113.
PROPATO, Dario: 614.
PUIG VALLS, D.: 98.

Q

QUEIRUGA, Claudia: 338.
QUINTANA, Fabio: 113.
QUINTEROS, Darío: 149.
QUIROGA, Facundo: 98.
QUIROGA, Macarena: 893.
QUISPE, José Rolando: 347.

R

RAMIREZ, U.: 204.
RAMON, Hugo: 984.

RAMOS, Lautaro: 185.
RASJIDO, José: 119.
RATTALINO, Donna: 739.
REINAUDI, Claudia: 224.
RETAMAR, Lautaro: 185.
RETAMAR, Soledad: 185.
REXACHS, Dolores: 700.
REY, M.: 204.
REYES, Carina Jimena: 926.
REYES, Nora: 191, 214.
REYES ZAMBRANO, Gary: 229.
REZINOVSKY, Alfredo: 799.
RIBA, Alberto Eduardo: 662, 739.
RICCA, Monica: 935.
RICCI, Santiago: 224.
RICHARD, Cristhian: 185.
RIESCO, Daniel: 832.
RIGONI, Brian: 324.
RIOS, Cristian: 662.
RIOS, G.: 98.
RÍOS, Laura Elena: 558.
RIPERTO, Adriel: 662.
RISSOLA, Esteban A.: 224.
RIVERA, Paula Cecilia: 739.
RIVERA, Ramiro: 185.
RIVEROS, Alejandro: 219.
RIVEROS ZAPATA, Adolfo: 646.
ROBINS, Daniel: 739.
ROCABADO, Sergio: 926.
ROCHA, Fabio: 466.
RODRIGUEZ, Agustín: 209.
RODRIGUEZ, Ismael: 686, 680.
RODRIGUEZ, Marina: 984.
RODRIGUEZ, Nelson: 710, 720, 43.
RODRIGUEZ, Rafael: 550.
RODRIGUEZ, Sandra Isabel: 512.
RODRÍGUEZ, G.: 270.
RODRÍGUEZ, Guillermo: 466.
RODRÍGUEZ, Rocío Andrea: 604, 574.
RODRIGUEZ EGUREN, Sebastián: 686, 680.
RODRÍGUEZ HERLEIN, Diego R.: 24.
RODRÍGUEZ MEDINA, Carlos Gustavo: 149.
ROGER, Sandra: 657.
ROGGERO, Patricia: 214.
ROJAS, Nicolás: 387.
ROJAS PAREDES, Andres: 124.
ROMAGNANO, María Rosalía Gema: 535, 171.

ROMANO, Lucas: 324.
ROMERO, Fernando: 759, 680.
ROMERO, Juan Manuel: 799.
RONCHETTI, Franco: 98.
ROSADO, Magdalena: 275.
ROSALES, Vanessa Aybar: 338.
ROSATTO, Daniel: 734.
ROSENSTEIN, Javier: 570.
ROSENZVAIG, Federico: 540.
ROSETE-SUÁREZ, Alejandro: 98.
ROSSET, Ana Lorena: 144.
ROTELLA, Carina: 318.
ROTTOLI, Giovanni Daián: 553.
ROVERO, Mara Lia Ines: 739.
ROZAS, Claudia: 700.
RUANO, Darío: 240.
RUBIO, Damián: 843.
RUCCI, Enzo: 674, 686, 680.
RUEDA, Marcela Mabel: 550.
RUEDA, Sonia V.: 787.
RUHL, Analía L.: 837.
RUITTI, Alberto Javier: 739.
RUIZ, Susana Beatriz: 171.
RUSSO, Claudia: 984.
RUSTICCINI, Alejandro: 822.

S

SABOLANSKY, Alejandro: 843.
SAGRIPANTI, Mauro: 415.
SAGULA, Jorge: 628.
SAIZAR, Victoria: 113.
SALAS, Andrea: 909, 176.
SALAZAR MESÍA, Natalí: 919.
SALDAÑO, Viviana E.: 421.
SALDARINI, Javier: 471.
SALGADO, Carlos: 477, 496, 446, 471.
SALINA, Mauro: 82, 56.
SALINAS, Sergio Ariel: 318.
SALTO, Carolina: 87.
SALVATORE, Juan E.: 56.
SAMAT, Pablo: 214.
SANCHEZ, Alberto: 496, 477, 471.
SANCHEZ, Ernesto: 38, 43.
SANCHEZ, Federico: 720, 710.
SANCHEZ, Horacio: 945.
SANCHEZ, Mariano: 674.

SÁNCHEZ, Cecilia Beatriz: 837.
SANCHEZ BALZARETTI, Ignacio: 837.
SANDOBAL VERON, Valeria C.: 308.
SANTOS, Damián: 643.
SANTOS, Graciela: 286.
SANZ, Cecilia: 919, 886, 935, 286.
SANZ, Victoria: 674, 686, 680.
SARMIENTO, Adriana: 535.
SAROBE, Monica C.: 984.
SATTOLO, Iris: 969, 565.
SAVARRO, Mauricio: 578.
SCAPPINI, Reinaldo: 2, 848.
SCHAB, Esteban: 129, 185, 553.
SCHEFFER, Maria: 171.
SCHENONE, Carlos: 82.
SCHIAVONI, Alejandra: 903.
SEBASTIÁN, Velazquez: 343.
SELZER, Matias Nicolas: 281, 292.
SEMENZATO, Darío: 700.
SEMKEN, Martha: 368.
SERAFINO, Sandra: 984.
SERRA, Ariel: 29.
SERRANO, Eliana: 984.
SEVILLA, Gustavo: 507.
SILVA, Telmo: 275.
SILVERA, Jorge A.: 486, 387.
SIMARI, Gerardo I.: 165.
SOCORRO, Raisa: 275.
SOLIGO, Pablo: 594.
SORIA, José: 881.
SOSA, Alejandra: 351.
SOSA, Hernan: 421.
SOSA, Magali: 171.
SOSA, Marcelo Omar Diogenes: 362.
SOSA BRUCHMANN, Eugenia: 362.
SOSA ZITTO, Rossana: 558.
SPERANZA, Libertad: 240.
SPOSITTO, Osvaldo Mario: 113.
STEFANONI, María Eugenia: 963.
STRACCIA, Luciano: 502.
STRERI, Nicolás Emanuel: 491.
SUPPI, Remo: 700.

T

TACACHO, Emilse: 926.
TAFFERNABERRY, Carlos: 958.

TAGARELLI, Sandra: 318.
TALAME, M. Lorena: 73, 78.
TALAY, Carlos A.: 24.
TARANILLA, María Teresa: 181.
TARIFA, Enrique E.: 646.
TEJADA, Jorge Damián: 662, 739.
TERRENI, Luciana Gabriela: 373.
TESONE, Fernando: 440, 404.
TESTA, Oscar: 1003.
TETZLAFF, Tomás: 124.
TEXIER, Jose: 739, 881.
TEZE, Juan C.L.: 165.
THOMAS, Pablo: 400, 440, 404, 787.
TINETTI, Fernando G.: 686, 680, 759.
TOCONAS, Nancy: 550.
TODT, Carolina Mariana: 848.
TOLOSA, Gabriel: 224.
TOLOZA, Juan M.: 772.
TOMASELLI, Gabriela: 481.
TORRENTS-BARRENA, J.: 98.
TORRES, Estela: 507.
TORRES, Luis: 793, 814.
TORRES, M.: 92.
TORRES, Silvia V.: 351.
TORTOSA, Nicolás: 481.
TOSINI, Marcelo A.: 772.
TRASMONTANA, Julio: 471.
TRIGILA, Mariano: 124.
TRIGO, Santiago: 827.
TRIPODI, Gustavo: 628.
TRIPODI, Josefina: 628.

U

URQUIJO, Ruben Ricardo: 763.
URRIBARRI, Dana K.: 292, 251.
UVA, Marcelo: 997, 491.

V

VALDEZ, Jorge Ceferino: 92.
VALERGA, Lucas Sosa: 251.
VALIENTE, Waldo: 12.
VARANDA PEREIRA, María J.: 832.
VARAS, Valeria: 119.
VARGAS, Javier Leonardo: 87.
VARGAS, Lucas Gastón: 720.

VARGAS, Mariano: 368.
VAZQUEZ, Martín: 772.
VÁZQUEZ, Pablo: 431.
VEGA, Micaela: 421.
VENOSA, Paula: 843, 338.
VERA, Carlos: 512, 477.
VERA, Cristina: 219.
VERA, Pablo: 574, 604.
VERA LACEIRAS, María Silvia: 251.
VERINO, Claudia: 520.
VIALE, Pamela: 343, 377.
VIANO, Hugo J.: 351.
VIAVATTENE, Hernán A.: 574.
VILLAGARCIA WANZA, Horacio: 686, 759.
VILLAGRA, Andrea: 92, 119.
VILLAGRA, Silvia: 421.
VILLANUEVA, Tomás: 324.
VILLARREAL, Martin: 324.
VILLCA, Raúl: 12.
VILUGRÓN, M.: 200.
VIOLINI, Lucía: 886.
VOLKER, Mariano: 12.

W

WELCH, Daniel: 240.

Y

YOMMI, Alejandra: 265.

Z

ZACCARDI, Gonzalo: 695.
ZAMBRANO, Jusmeidy: 739.
ZAMUDIO, Eduardo: 204.
ZANELLATO, Claudio: 700.
ZANGARA, Maria Alejandra: 886.
ZANGLA, María Soledad: 351.
ZAPATA, Sergio G.: 507.
ZARAGOZA, Juan: 638.
ZARATE, Pedro: 667.
ZÁRATE ÁLVAREZ, Nicolás Ignacio: 570.
ZEA CARDENAS, Milagros: 837.
ZELIGUETA, Laura: 426.
ZORATTO, Valeria: 415.

Arquitectura, Redes y Sistemas Operativos

Análisis y Aplicaciones de Internet de las Cosas y Ciudades Inteligentes. Framework de trabajo

Gramajo Sergio, Scappini Reinaldo, Bolatti Diego, Aguirre Federico

Universidad Tecnológica Nacional, Facultad Regional Resistencia
Departamento de Ingeniería en Sistemas de Información
Centro de Investigación Aplicada en Tecnologías de la Información
y la Comunicación (CInApTIC)
French 414 – Resistencia (3500) Chaco - Argentina
{sergiogramajo, rscappini, diegobolatti, federodani}@gmail.com

RESUMEN

El desarrollo de las nuevas tecnologías como Internet de las Cosas (IoT) ha planteado nuevas formas de generar aplicaciones en pos de mejorar los servicios en las ciudades con un impacto directo en la calidad de vida de las personas y el medio ambiente. Este concepto está relacionado a ciudades inteligentes, seguridad y la gestión de las telecomunicaciones que interconectan múltiples dispositivos con una interacción humana mínima. Este trabajo se orienta a estudiar y proponer soluciones teniendo en consideración las arquitecturas necesarias, la cultura, los estudios de factibilidad técnico-económicos, el uso de las TICs, los aspectos climáticos, las normativas locales o nacionales de uso de telecomunicaciones y espectro, entre otros aspectos relevantes para proponer frameworks de IoT que puedan ser transferidos al medio local.

Palabras Clave: Internet de las cosas, Ciudades Inteligentes, Framework.

CONTEXTO

Este trabajo de investigación se desarrolla en el marco del proyecto “Análisis y Aplicaciones de Internet de las Cosas y Ciudades Inteligentes basadas en Telecomunicaciones y Seguridad” del Centro de Investigación Aplicada en TICS (CInApTIC) de la Universidad Tecnológica Nacional, Facultad Regional Resistencia.

1. INTRODUCCIÓN

En la actualidad la integración de las Tecnologías de Información y de Comunicación (TICs) y las ciudades ha promovido nuevas formas de abordar necesidades humanas en sociedad y surgen conceptos como ciudad de la información, ciudad digital, ciudad inteligente e IoT [1].

La ciudad inteligente surgió como una solución para abordar los desafíos que surgen con el crecimiento de la urbanización y la población. Sin embargo, el concepto de ciudad inteligente todavía está evolucionando y no se ha incorporado en todo el mundo debido a barreras tecnológicas, económicas y gubernamentales [1] [2].

En un sentido más amplio, IoT fue el resultado de la evolución de las telecomunicaciones y la electrónica que conectan millones de dispositivos, los avances tecnológicos, las redes inalámbricas de sensores (WSN) y las interacciones máquina a máquina (M2M) [3] [4] con mínima intervención humana es el principio de facto de IoT [5]. Además, los dispositivos conectados entre sí comparten información y tienen acceso a información autorizada de otros dispositivos para luego tomar de decisiones contextuales sobre determinada situación o problema [6].

Es así que este proyecto, como objetivo general, se orienta a desarrollar frameworks o modelos de Internet de las Cosas y Ciudades Inteligentes en base a estudios e identificación de atributos sobre

telecomunicaciones, seguridad y tecnología que puedan aplicarse a la región.

Objetivos específicos

1. Analizar las arquitecturas y protocolos usados en IoT y Ciudades Inteligentes.
2. Analizar las técnicas de seguridad para los modelos estudiados
3. Desarrollar escenarios de prueba y generación de aplicaciones.
4. Generar posibles transferencias al medio local y/o proponer modelos de estudio para nuestras investigaciones.
5. Desarrollar prototipos de IoT en base a estándares y adaptarlos al medio regional.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación que se abordan en el proyecto están vinculadas con:

- **Frameworks genérico de captación y presentación de datos:** Consiste en el desarrollo del framework que pueda ser adaptado a diferentes situaciones del medio local.
- **Arquitectura de redes de información para IoT:** Para el diseño de un detector de anomalías sobre IoT.
- **Inteligencia Artificial:** Se realizará un análisis de las técnicas de Machine Learning, con el objetivo de seleccionar la mejor opción para la implementación en el módulo de detección de anomalías.
- **Seguridad de IoT:** Se analizarán los ataques de seguridad y anomalías que se presentan en un entorno de IoT.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos en el proyecto podrán ser utilizados en las siguientes áreas del conocimiento:

A) Arquitecturas de redes de Información para IoT. El relevamiento y análisis de las

nuevas tendencias de redes de información de corto y amplio rango, conllevará a publicaciones científicas y transferencias al medio local o regional. Esto propiciará el contacto con investigadores de nivel internacional y nacional de otras instituciones para posibles intercambios de experiencias como el que se está llevando a cabo con el proyecto REMIND de la Unión Europea¹.

B) Programación y pruebas de diversos dispositivos usados para IoT y ciudades inteligentes como sensores y equipos de telecomunicación entre ellos sin intervención humana y que ayude a la toma de decisiones y mejore la gestión que lo utilice.

Una vez finalizado el proyecto generará nuevo conocimiento y aplicaciones que pueden ser transferidos tanto a entornos de investigación como diferentes entornos organizacionales o empresas del medio. En este sentido se pretende impulsar el intercambio de conocimiento con investigadores de otras instituciones, asistencia a eventos científicos, elaboración de publicaciones científicas, estancias de investigación en el exterior, convenios de transferencia, etc.

Hasta el momento se ha obtenido un prototipo del modelo o framework que se denomina SiMo (Sistema de Monitoreo). Este engloba tecnologías e infraestructura para hacer frente a soluciones para IoT. Esto permite la creación de un ecosistema interconectado en donde el uso de recursos convencionales aumenta en eficiencia gracias a la aplicabilidad de la tecnología. SiMo permite la integración de múltiples dispositivos hardware, entre ellos sensores y actuadores para la captación de datos de acuerdo con los parámetros que mejor se ajusten a las necesidades de los usuarios y escenarios de trabajo. Además, permite personalizar distintas configuraciones para cada uno de ellos y posteriormente obtener los datos para almacenarlos y a su vez controlar objetos (como sensores,

¹ (REMIND) de Horizonte2020 <https://remind-research.com/>

actuadores, equipos, etc.) de manera simple y amigable para que puedan cooperar entre ellos para alcanzar objetivos en común.

SiMo, se encuentra dividido en dos grandes módulos: SiMoW (SiMo Web) y SiMoRa (SiMo en Raspberry), dedicados a tareas puntuales que hacen en su totalidad a la solución final.

Ambos módulos están diseñados para trabajar en conjunto, como se observa en la Fig. 1. Todos los parámetros de configuración y funcionamiento de SiMoRa pueden ser seteados y consultados desde SiMoW, mediante el uso de REST API [7]. El desarrollo del Framework se basa en un diseño de desarrollo de software siguiendo una arquitectura de capas, permitiendo añadir nuevas funcionalidades según se necesite, para trabajar sobre diversos escenarios, siendo totalmente transparente al usuario. Esto facilita la escalabilidad y adaptabilidad a las necesidades de cualquier proyecto que se plantee como objetivo.

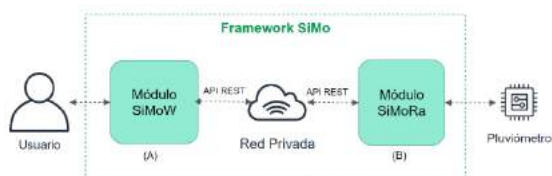


Fig. 1. Framework SiMo con módulos (A) SiMoW y (B) SiMoRa

En cuanto a las transferencias al medio existen dos líneas abiertas:

A) El primer sistema se basa en crear una red de estudio meteorológico y ambiental sobre 42 puntos distribuidos en toda la provincia. La medición se hará en los nodos de fibra óptica provinciales y los datos serán enviados a un centro de procesamiento en la ciudad de Resistencia.

B) Sistemas de Telecomunicación con LORA y dispositivos IoT mediante convenio a realizarse con la empresa YEAP. Se basa en un sistema de sensores para estudio de condiciones ambientales en espejos de agua o lagunas de Resistencia Chaco.

4. FORMACIÓN DE RECURSOS HUMANOS

Con el proyecto se pretende contribuir a la formación de recursos humanos desde diversas áreas:

- **Formación de becarios:** El proyecto cuenta con la participación de alumnos becarios del último año de la carrera de Ingeniería en Sistemas de Información que están realizando su práctica supervisada.
- **Alumnos de la carrera de Ingeniería en Sistemas de Información:** Se prevé realizar actividades de actualización y talleres con alumnos de las cátedras del área de redes de información, comunicaciones y seguridad informática. Además, por las propias características de los temas que involucra el proyecto se pueden realizar actividades en cátedras como inteligencia artificial.
- **Formación de jóvenes profesionales:** Se prevé la incorporación de jóvenes profesionales de Ingeniería en Sistemas de Información con la intención de seguir con una carrera en investigación universitaria. Los cuales pueden incorporarse en carácter ad-honorem al proyecto o a través de becas de iniciación en la investigación.
- **Formación de postgrado:** A partir de las líneas de investigación desarrolladas en el proyecto se prevé que el Ing. Diego Bolatti finalice su doctorado mediante una tesis vinculada a este proyecto.
- **Equipo de trabajo:**
 - **Director:**
 - Gramajo, Sergio.
 - **Investigadores de apoyo:**
 - Bolatti, Diego
 - Scappini, Reinaldo
 - Karanik, Diego
 - **Becario alumno:**
 - Todt, Carolina
 - Federico Aguirre

5. BIBLIOGRAFÍA

- [1] Sekhar N. Kondepudi, et. al. An overview of smart sustainable cities and the role of information and communication technologies. Set of ITU-T's Technical Reports and Specifications. 2016.
- [2] Bhagya Nathali Silva, Murad Khan, Kijun Han. Towards sustainable Smart Cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities and Society*. 38. 2018
- [3] Silva, B. N., Khan, M., & Han, K. Internet of things: A comprehensive review of enabling technologies, architecture, and challenges. *IETE Technical Review*, 1–16. 2017.
- [4] Silva, B. N., Khan, M., & Han, K. Big data analytics embedded Smart City architecture for performance enhancement through real-time data processing and decision-making. *Wireless Communications and Mobile Computing*. 2017.
- [5] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29, 1645–1660. 2013.
- [6] Vermesan, O., Friess, P., Guillemin, P., Giaffreda, R., Grindvoll, H., Eisenhauer, M., et al. Internet of things beyond the hype: Research innovation and deployment. *IERC Cluster SRIA*. 2015.
- [7] J. Mueller, SOAP and REST Basics and Differences <https://smartbear.com/blog/test-and-monitor/understanding-soap-and-rest-basics>, último acceso 07/03/2021

Análisis y Simulación de procesadores RISC-V en plataforma ISA abierta

Daniel Argüello, Higinio Facchini, Santiago Pérez
CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería)
Facultad Regional Mendoza, Universidad Tecnológica Nacional
Rodríguez 273, Mendoza, Argentina - 0261-5244576
(*darguell,higiniofac,santiagocp*)@frm.utn.edu.ar

Resumen

Cualquier implementación SoC (System on Chip) que incluya algún procesador embebido debe pagar regalías (royalties) a través de la compra de la propiedad intelectual o de la licencia arquitectural. Es de gran importancia, para superar las brechas en el diseño e implementación, la idea de introducir una interfaz Software/Hardware abierta (ISA abierta) RISC V, que sea una realidad, no sometida a regalías. En 2010 nació una iniciativa en la Universidad de California en Berkeley para desarrollar el procesador RISC-V de ISA abierto y público que elimina la mayor parte de las restricciones impuestas por los ISAs propietarios. El objetivo del proyecto de investigación es centrarse en los Procesadores Docentes (ProcDoc-RV), en línea con los autores Paterson y Hennesy, y en los Procesadores sencillos (Micro-RV), orientados a aplicaciones embebidas, para actividades de investigación, análisis y simulación arquitectónica, a fin de determinar métricas cualitativas y cuantitativas de rendimiento de dichos Procesadores. Este documento se corresponde con el proyecto PID sobre RISC V, de la UTN Mendoza, en articulación con la Universidad de Zaragoza (España), que se espera fomente la difusión de la temática, y una mayor innovación en el desarrollo e implementación de productos electrónicos, se puedan compartir diseños

y lograr accesibilidad a usuarios en general (y para aplicaciones específicas), solucionar problemas sin realizar grandes inversiones, y cualificar profesionales en el área.

El personal principal relacionado a esta línea de investigación son docentes de la UTN Mendoza, y de la Universidad de Zaragoza, en las Cátedras afines a las Arquitecturas de Computadoras.

Palabras clave: RISC, CISC, RISC V, ISA Abierto

Introducción

En 2010 nació una iniciativa en la Universidad de California en Berkeley para desarrollar el procesador RISC-V [1] de ISA abierto y público que elimina la mayor parte de las restricciones impuestas por los ISAs propietarios.

Existe un proyecto interuniversitario y participativo empresarial en Europa consistente en: 'Red-RISC-V: Investigación, formación y prospectiva en sistemas RISC-V' [2], en el cual participa la Universidad de Zaragoza y ha servido para la articulación y presentación del presente proyecto a través de un convenio entre esa universidad y la UTN-FRM, aunque en este caso aplicado específicamente sobre los Procesadores docentes y los Procesadores sencillos de RISC-V.

RISC-V tiene respaldo de la comunidad académica internacional, con

numerosos grupos activos en EEUU y Europa.

Contexto

El software abierto es ya una realidad que ha facilitado enormemente, tanto su propia evolución, como su uso y adaptación para aplicaciones específicas sin mayores inversiones que las de contar con profesionales cualificados.

Desde la misma perspectiva, se puede afirmar que en la actualidad es difícil encontrar algún producto electrónico, en cualquier escala, que no incluya un procesador. Mientras que existen múltiples procesadores comerciales para el desarrollo de sistemas Hw/Sw, la gran mayoría de ellos son núcleos con el repertorio de instrucciones (ISA) cerrado, protegido y propietario, cuya utilización está sometida a caras licencias de explotación o limitada al uso de componentes y sistemas de desarrollo disponibles en el mercado. Cualquier SoC (System on Chip) que incluya algún núcleo se ve sometido al pago de royalties vía compra de la IP o de la correspondiente licencia arquitectural. Un ejemplo es la familia de procesadores RISC ARM con gran penetración de mercado, que cubre aplicaciones desde sensores para IoT, hasta multiprocesamiento paralelo aplicado a teléfonos inteligentes, o servidores para centros de datos (HTC) y supercomputación (HPC). Cualquier SoC (System on Chip) que incluya algún núcleo ARM se ve sometido al pago de royalties vía compra de la IP o de la correspondiente licencia arquitectural (que permite adaptar el diseño a necesidades específicas).

Si bien existen estándares y software abiertos (por ejemplo, protocolos TCP/IP, OS Linux), la introducción de una interface Software/ Hardware para el

desarrollo de procesadores con repertorio de instrucciones (ISA) abierto y público tuvo sus contratiempos. Por ejemplo, se pueden mencionar como precedentes de ISA abiertos el SPARC V8 (1994) u el OpenRISC (2000-2011), cuyas propuestas no se consolidaron. Sin embargo, en 2010 comenzó desde la Universidad de California, Berkeley, otra propuesta que prosperó y, en el año 2014 ya estaba introducida como la arquitectura de procesadores RISC V. RISC (Computadoras con un Conjunto de Instrucciones Reducido) ha superado en ventas altamente a CISC (Computadoras con un Conjunto de Instrucciones Complejo), y actualmente domina el mercado de arquitecturas. En la actualidad la mayoría de los procesadores trabajan con ISA cerrado y propietario, y se pagan caras licencias (RISC ARM).

Teniendo en cuenta el amplísimo campo de aplicación de los procesadores: IoT (Internet de las Cosas), Sistemas y Procesadores Embebidos en SoC (System on Chip), Computadoras de alta performance y rendimiento (HPC, HTC), Centros de Datos (DC), y computadoras Ware-House (WSC), surge la importancia de contar con un ISA abierto y público que facilite una mayor innovación vía competiciones libres y diseños compartidos, investigación y desarrollo sin pagar caras licencias por ISA propietario que, además, obligan y utilizan diseños propietarios que también se pagan.

Como se mencionó precedentemente, existe un proyecto en Europa llamado: Red-RISCV: "Investigación, formación y prospectiva en sistemas RISC-V". Según se puede apreciar en el proyecto mencionado Red-RISCV, existe una realidad que ha facilitado la evolución de RISC V, ya que la meta es lograr cualificar profesionales, teniendo en cuenta el diseño y fabricación de SoCs

con procesos accesibles técnica y económicamente, si se evitan núcleos propietarios. RISC V ha inducido a otras arquitecturas propietarias a abrir también su ISA, como MIPS [3]. RISC V tiene respaldo de la comunidad académica internacional, y está dominando las tendencias y estrategias de futuro y promete una rápida expansión [4,5,6,7]. La Comisión Europea ha identificado el ISA abierto RISC V, como el ISA del futuro acelerador europeo en el marco del European Processor Initiative (EPI) [8]. Esta iniciativa está siendo liderada por el Centro Nacional de Supercomputación de Barcelona (BSC – Barcelona Supercomputing Center), promotor de esta red junto al Centro Nacional de Microelectrónica (CNM).

La Universidad de Zaragoza participa de la Red Europea y, además, mediante un convenio entre la Universidad de Zaragoza y la UTN-FRM se ha permitido su incorporación a este proceso tecnológico. Por otro lado, dos doctores de dicha Universidad española, integra el PID mencionado de la UTN Mendoza, y son miembros del Doctorado en Ingeniería, mención Computación de la misma institución.

Procesadores y complejidad

Existen 3 familias de procesadores RISC V con distintos grados de complejidad, más una cuarta familia dedicada a aceleradores específicos:

- a) Procesadores Docentes (ProcDoc-RV): Diseños con repertorios de instrucciones mínimos, en la línea del libro de Patterson y Hennesy [9].
- b) Procesadores sencillos (Micro-RV): Diseños con soporte para un repertorio base de enteros y alguna extensión estándar, orientados por ejemplo a aplicaciones embebidas, Rocket (UC Berkeley) [10] o diferentes

implementaciones de la plataforma PULP (ETH Zurich) [11].

- c) Procesadores de altas prestaciones (ProcAP-RV). Algunos ejemplos: BOOM (Berkeley out-of-orderprocessor)[12] y ET Maxion de Esperanto Technologies[13].
- d) Aceleradores (Acel-RV): aceleradores para aplicaciones específicas como pueden ser medicina personalizada, encriptación, redes neurales, etc. Aquí se incluirán los aceleradores ya previstos en la European Processor Initiative y otros aceleradores en desarrollo por los miembros de la red.

En cada familia, o entre familias, aparecen elementos de diseño, experimentación y optimización comunes. Para este proyecto se trabajará con las dos primeras familias descritas, de acuerdo a los recursos disponibles.

Metodología

Para el desarrollo del proyecto se ha previsto las siguientes tareas:

Tarea 1: Recopilar de información y estudio de los Procesadores Docentes (ProcDoc-RV) y los Procesadores sencillos (Micro-RV).

Tarea 2: Compilar trabajos de investigación sobre la temática, determinando escenarios experimentales, herramientas de simulación, tráfico utilizados, métricas medidas, resultados y conclusiones.

Tarea 3: Establecer mecanismos comparativos de los trabajos de investigación compilados, usando cuadros, índices, ponderaciones, etc.

Tarea 4: Definir escenarios de experimentación que faciliten el contraste entre ambos tipos de Procesadores: Los Procesadores Docentes (ProcDoc-RV) y

los Procesadores sencillos (Micro-RV), en la jerarquía RISC-V5.

Tarea 5: Construir tablas comparativas de las métricas para cada escenario. Se construirán tablas y graficas comparativas de las prestaciones o métricas para cada caso, según las combinaciones que permitan los diversos escenarios.

Tarea 8: Documentar, publicar y difundir resultados.

Tarea 9: Redactar Informe Final.

Líneas de Investigación y Desarrollo

Los temas y líneas de investigación, que se tratarán durante el desarrollo del proyecto se encuentran comprendidas en las Ciencias de la Computación e Informática para Equipos de procesamiento, y se pueden identificar como:

- Investigación y análisis de Procesadores.
- Arquitectura de Procesadores.
- Tecnologías RISC.
- Arquitecturas abiertas (ISA).
- Simulación VHDL sobre FPGA.
- Procesadores ProcDoc-RV.
- Procesadores Micro-RV.

Objetivos

Objetivo General:

Determinar e identificar ámbitos de aplicación y métricas de eficiencia por modelación asociadas con estas nuevas tecnologías basadas en arquitecturas abiertas, para los Procesadores Docentes (ProcDoc-RV) y los Procesadores sencillos (Micro-RV).

Objetivos Específicos:

Para el presente proyecto se buscan los siguientes objetivos:

- Determinar el impacto de avance de esta tecnología abierta y pública en nuestro ámbito, y la diversidad de modelos arquitectónicos disponibles.
- Determinar las características a detalle de los procesadores seleccionados, los escenarios de experimentación, para su posterior simulación.
- Determinar las configuraciones óptimas de los procesadores seleccionados y sus ámbitos de aplicación, para cada uno de los escenarios de experimentación.
- Definir las necesidades actuales locales a fin de iniciar desarrollos para alcanzar una masa crítica de conocimientos y poder interactuar, en el ámbito RISC V, con investigadores y profesionales de la UE y EEUU.
- Aplicar los conocimientos alcanzados para implementar una aproximación básica RISC V de estudio, simulación VHDL sobre una Field Programmable Gate Array (FPGA).

Con el uso de las herramientas de simulación basadas en VHDL, se busca trabajar en distintos escenarios de trabajo, de menor a mayor complejidad, que faciliten el contraste entre ambos tipos de procesadores planteados (ProcDoc-RV y Micro-RV).

En general se definirán métricas a obtener para cuantificar los distintos parámetros de trabajo de los procesadores, en cuanto a rendimiento, velocidad, consumo de recursos, uso de memoria interna, etc para poder realizar tablas, cuadros y gráficos comparativos para cada escenario planteado.

Estado de Avance

En proyectos anteriores de investigación y análisis de procesadores se han

realizado estudios de arquitecturas y tecnologías de Procesadores. Específicamente con un proyecto anterior, denominado “IMPLEMENTACIÓN Y ANÁLISIS DE PRESTACIONES DE UN PROCESADOR ELEMENTAL DIDÁCTICO CON VHDL Y FPGA”, se trabajó con el uso de herramientas de simulación y análisis de un procesador básico.

El equipo de trabajo está integrado por docentes investigadores, y becarios graduados y alumnos que son miembros del Centro UTN CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería) de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional, dentro del Área de Electrónica Aplicada. El Director, CoDirector y el Asesor Técnico son Docentes de la Cátedra Arquitectura de Computadoras de la Carrera de Ingeniería en Sistemas de Información y han trabajado desde hace 15 años en las temáticas de las Arquitecturas de Computadoras y las Técnicas Digitales.

Formación de Recursos Humanos

Mediante el presente proyecto se busca la formación de Recursos Humanos con el fin de:

- Aglutinar RRHH calificados en torno a estas nuevas tecnologías RISC-V y crear las sinergias para facilitar su participación en proyectos e iniciativas nacionales (2021-2022) e internacionales (2023-2024).
- Incidir en los gestores de la I+D nacional para que incentiven estas tecnologías abiertas a fin de garantizar una cierta independencia tecnológica. Estando al día de las distintas iniciativas, podremos contribuir a las mismas y participar de sus convocatorias.

- Facilitar la organización sostenida de seminarios científico-tecnológicos para la actualización de nuestro entorno y una visión prospectiva sobre la proyección de futuro de estas tecnologías.

Articulación con la Universidad de Zaragoza

El Centro de investigación CereCoN ha establecido formalmente un Convenio de Colaboración mutua con los Grupos de Investigación de la Universidad de Zaragoza. En este caso, con el Dr. Víctor Viñals Yúfera, del Grupo de Arquitectura de Computadores de la Universidad de Zaragoza, especializados en “Jerarquía de memoria, gestión de tareas, y optimización de aplicaciones”, con 14 profesores UZ, 6 doctorandos y colaboradores en UPC, UPV/EHU y UVA. El Grupo está reconocido por el Gobierno de Aragón desde 2003. Áreas: memorias cache, eficiencia energética y fiabilidad, tiempo real, aceleradores heterogéneos y aceleración de aplicaciones.

Además, el Dr. Víctor Viñals y el Dr. Jesús Alastruey son docentes del Doctorado en Ingeniería, mención Computación, de la UTN Mendoza, en el Curso Computación de Altas Prestaciones.

Estos docentes son parte del personal involucrado en el Proyecto de Investigación de referencia.

Referencias

- [1] Instruction sets should be free: The case for RISC-V.
<https://people.eecs.berkeley.edu/~krste/papers/EECS-2014-146.pdf>
- [2] Memoria-Red-RISCV
<http://www.cnm.es/~icas/RISCV/Memoria-Red-RISCV.pdf>

- [3] MIPS Goes Open Source
https://www.eetimes.com/document.asp?doc_id=1334087
- [4] RISC-V: More Than a Core
<https://semiengineering.com/risc-v-more-than-a-core/>
- [5] RISC-V on the Verge of Broad Adoption
https://www.eetimes.com/document.asp?doc_id=1334311
- [6] Open Source Hardware Benefits Procurement Practices
https://www.ebnonline.com/author.asp?section_id=3219&doc_id=%20283851&
- [7] Can Arm Survive RISC-V Challenge?
https://www.eetimes.com/author.asp?section_id=36&doc_id=1334306
- [8] European Processor Initiative (EPI)
<https://ec.europa.eu/digital-single-market/en/news/european-processor-initiativeconsortiumdevelop-europes-microprocessors-future-supercomputers>
- [9] Patterson, D. A., & Hennessy, J. L. (2017). Computer Organization and Design. RISC-V Ed: The Hardware/Software Interface. Morgan Kaufmann.
http://home.ustc.edu.cn/~louwenqi/reference_books_tools/Computer%20Organization%20and%20Design%20RISCV%20edition.pdf
- [10] <https://riscv.org/2014/10/launching-the-open-source-rocket-chip-generator-2/>
- [11] <https://www.pulp-platform.org/>
- [12] <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-157.html>
- [13] <https://www.esperanto.ai/wp-content/uploads/2018/12/Esperanto-Maxes-Out-RISC-V.pdf>

Entorno de Contenedores de Emuladores que contienen Sistemas Embebidos

Waldo Valiente, Esteban Carnuccio, Mariano Volker, Graciela De Luca, Raúl Villca, Matías Adagio

*Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza*

Dirección: Florencio Varela 1703 – CP 1754 – {wvaliente, ecarnuccio, mvolker, gdeluca}@unlam.edu.ar, {raul.villcasd, mati.adagio}@gmail.com

RESUMEN

Aunque la población no se percate, cada vez más, los sistemas embebidos forman parte de la vida cotidiana de las personas. Como consecuencia de dicho auge de la tecnología, surgieron en el mercado diferentes sistemas embebidos conformados por placas de desarrollo, que inicialmente se utilizan para prototipado. Estas son de fácil aprendizaje, lo que permitió que sean muy utilizadas en los últimos años. Por ese motivo para un usuario sin experiencia, resulta un desafío la elección de la placa correcta para sus proyectos. Esto se debe a que muchas veces se puede equivocar en dicha selección, ya que no cumpliría con sus expectativas. Por ende, en esta investigación se pretende minimizar esos riesgos, a través de la utilización de emuladores que permitan imitar el comportamiento de estos sistemas embebidos. Aprovechando estas características se procura generar un entorno de integración, empaquetado y automatizado mediante contenedores. Los cuales permitirán al usuario probar determinados programas rápidamente, en el hardware emulado del sistema embebido, sin necesidad de incurrir en costos. De forma tal de permitirle determinar si la placa elegida puede

llegar a cumplir sus expectativas, sin necesitar adquirir el hardware físico.

Palabras clave: Sistemas Embebidos, Emuladores, Qemu, Contenedores, Docker.

CONTEXTO

Nuestra Línea de Investigación es parte del proyecto “*Entorno de integración continua para validación de sistemas embebidos de tiempo real*”, dependiente de la Unidad Académica del *Departamento de Ingeniería e Investigaciones Tecnológicas*, perteneciente al programa de Investigaciones CyTMA2 de la Universidad Nacional de La Matanza, el cual es formado por docentes e investigadores de la carrera de ingeniería en informática. Este proyecto es continuación de los trabajos que viene realizando el grupo de investigación en sistemas operativos, computación de alto rendimiento y en el área de Internet de las cosas.

1. INTRODUCCIÓN

En los últimos años los sistemas embebidos han tenido un crecimiento exponencial, debido a que son utilizados en la vida cotidiana de las personas con mayor frecuencia. Este gran auge se debe principalmente a sus características que son su miniaturización, su menor consumo

energético, su gran capacidad de procesamiento y por la interconexión entre ellos. Gracias a su evolución, han surgido nuevos conceptos tecnológicos, los cuales hace décadas atrás eran impensados. Tales como dispositivos vestibles, domótica, telefonía móvil, internet de las cosas, ciudades inteligentes, industria 4.0, telemedicina, entre otros. A su vez algunos de ellos se pueden combinar con otras nuevas herramientas de procesamientos de datos, tales como Big Data o Machine Learning. Esto se debe a que existen sistemas embebidos que recompilan gran cantidad de información, para obtener conocimiento de su entorno y así aprender de los datos en forma automática.

Durante la evolución de los sistemas embebidos, hubo un punto importante en su línea de tiempo. Massimo Banzi en el año 2005 inicia, junto a sus alumnos, el proyecto Arduino en el instituto IVRAE en Italia. Este tenía como premisa permitirles aprender a trabajar, en muy poco tiempo y de forma económica, con microcontroladores a sus estudiantes. El proyecto se fundamentó en el diseño, construcción y prototipado de hardware libre. Además de ser multiplataforma y de fácil aprendizaje a través del lenguaje Wiring. [1] Estas características permitieron que en poco tiempo las placas Arduino tuvieran una gran aceptación, de forma tal que actualmente existe una gran comunidad de usuarios que aportan sus conocimientos y constantemente van optimizando el proyecto. [2]

A partir del surgimiento de las placas de prototipado Arduino, emergieron otros sistemas embebidos que tomaron como base las premisas del proyecto de Massimo Banzi, ya que permiten el rápido aprendizaje de su uso e intentan ser de costo accesible para los usuarios. Actualmente existen una gran cantidad de familias de placas de desarrollo en el mercado

que pretenden seguir dicho principio. De esta manera, junto con Arduino, las placas más importantes y buscadas por los usuarios para sus proyectos son las familias de Raspberry Pi, las familias de placas que poseen los microcontroladores STM32, las ESP8266 y su sucesor, que poseen el microprocesador ESP32. Siendo Arduino la más buscada entre todas las demás. Esta preponderancia se puede visualizar en el gráfico siguiente [3].

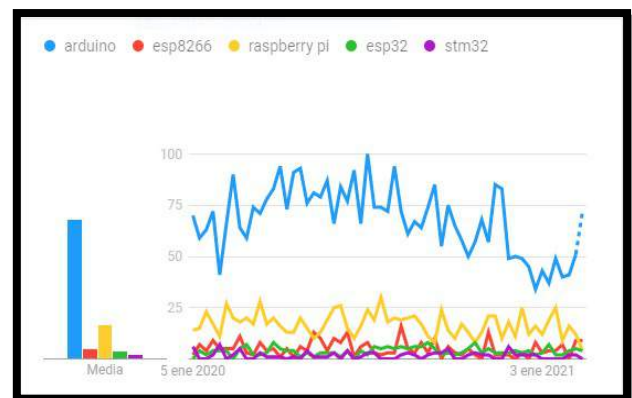


Fig. 1 interés a lo largo del tiempo de los usuarios de distintas placas de desarrollo en Argentina

En Argentina las placas de la familia Arduino es la más buscada y consultada en Internet. Mientras que el resto se mantienen muy por debajo de estas, en cantidad de búsquedas. Diferente es la situación en los países de América del Norte y Europa, en donde la cantidad de búsquedas están lideradas no solo por Arduino, sino también por la Raspberry Pi. Mientras que el resto se mantienen muy por debajo de ellas. Esta tendencia se debe a diversos factores. Por un lado, los STM32 son más potentes que las Arduino, pero a veces no presentan bibliotecas drivers desarrolladas para determinados dispositivos. Mientras que, para Arduino, los drivers son compatibles con la gran mayoría de los periféricos. A su vez los microcontroladores de Arduino, pueden continuar funcionando hasta que las baterías se agoten por completo. En cambio, los STM32 funcionarán hasta alcanzar el umbral de carga

de energía definido por el fabricante [4]. Por su parte las Raspberry Pi, debido a su gran capacidad de procesamiento, posibilidad de ejecución de distintos Sistemas Operativos de escritorio y de conexión de periféricos y conectividad al exterior, se utilizan principalmente como Servidores, Gateway o central de procesamiento de datos e imágenes. No obstante, debido a su alto costo no suelen ser utilizadas para la lectura de sensores y la activación de actuadores, ya que se emplean como intermediarios. Por otro lado, comparando las placas Arduino con las ESP8266, desde el punto de vista de la comunicación, se puede apreciar que una de las características que presentan las placas tradicionales de Arduino, es su falta de conectividad al exterior. Para realizar dicha comunicación se deben utilizar módulos adicionales de comunicación de Wifi, Bluetooth, entre otros. Mientras que el microcontrolador ESP8266, trae incorporado un módulo de Wifi. Por su bajo precio lo hace ideal para proyectos de Internet de las Cosas. No obstante, no tiene una comunidad tan grande de usuarios, como los de las plataformas de Arduino. Además, presenta algunas limitaciones en cuanto a compatibilidad y uso de periféricos. Para resolver dichas falencias en 2016 Espressif, lanzó el microprocesador de doble núcleo ESP32 que amplía las capacidades de cómputo de su predecesor. Al mismo tiempo presenta un módulo WIFI y Bluetooth incorporado. Lo que permite un abanico de utilidades.

En las placas indicadas anteriormente, a excepción de la Raspberry Pi, la forma tradicional de programarlas es con el mecanismo de Bare Metal. Esto quiere decir que funcionan sin Sistema Operativo. No obstante, si el usuario lo desea, se puede instalar un Sistema Operativo, incluso de Tiempo Real, como por ejemplo FreeRTOS.

Estas placas fueron especialmente diseñadas para realizar prototipos de proyectos, pero en algunas situaciones también se los utilizan como producto final. Todas ellas siguen la premisa del proyecto Arduino, la cual consiste en que su uso sea de rápido aprendizaje, de forma tal que cualquier persona con poco conocimiento en electrónica pueda aprender a usarlas [5] [6] [7]. Muchas veces, al plantear un proyecto, un usuario común no puede saber que placa de desarrollo es adecuada para su trabajo. Para saber esto, el usuario necesita contar con el hardware físico para hacer las pruebas necesarias de funcionamiento, lo que incurre en un costo de inversión para adquirir la placa elegida. Muchas veces ocurre que el hardware seleccionado no es el adecuado para el proyecto, lo que deriva en un cambio de plataforma, conllevando así a un costo extra para la adquisición del nuevo producto, que se pueda conseguir en el mercado local. Por ese motivo, una posible alternativa para minimizar los riesgos de este impedimento, en este trabajo se investigan los emuladores para dichas placas de desarrollo.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Los emuladores permiten a los usuarios ejecutar en una computadora de escritorio programas de un sistema embebido, sin necesidad de contar con él físicamente. Esto se ejecuta dentro de un ambiente que imita su comportamiento [8]. A través de su utilización, una persona puede probar el sistema embebido en su computadora, antes de adquirir el hardware físico, y así verificar que le es de utilidad. Por consiguiente, teniendo en cuenta dichas características, la presente investigación se centra en construir un entorno emulado que ejecuta distintos sistemas embebidos. Todo esto funcionando con el esquema de contenedores de Docker. De manera tal, que permita realizar las pruebas de

rendimiento y las verificaciones de distintas optimizaciones de su ejecución, antes de llegar a adquirir el hardware del dispositivo.

Emulador que se utiliza en la investigación:

Existen en el mercado diferentes emuladores que imitan el comportamiento de distintos sistemas embebidos. El más utilizado es Proteus. Si bien es considerado un simulador de circuitos, también permite emular distintas placas de desarrollo, con sus actuadores y sensores de manera óptima. Este software puede emular placas de la familia Arduino, STM32, ESP8266 y Raspberry Pi. El punto negativo de este software es que su uso debe ser a través de licenciamiento pago. Por otro lado, se encuentra el simulador de Arduino Thinkercad, que es muy utilizado en ambientes educativos. La desventaja que presenta esta herramienta es que se limita únicamente a simular una placa de desarrollo Arduino UNO. Adicionalmente existe el emulador Qemu, el cual va a ser el utilizado en esta investigación. Este emulador es muy utilizado en el mercado. Entre sus principales ventajas [9], se destaca que permite trabajar con una gran cantidad de modelos de dispositivos. Además, permite detectar problemas lógicos en etapas tempranas del desarrollo, ya que brinda la depuración del dispositivo. Al mismo tiempo, presenta licenciamiento GPL. Así como también permite generar script de automatización de programas al poder ser ejecutadas desde línea de comandos. Esta característica resulta muy útil para esta investigación. De esta manera se pueden emular distintas familias de dispositivos Arduino, STM32, ESP32, ESP8266 y Raspberry Pi.

Emulador dentro de un contenedor:

Un contenedor de software es un ambiente aislado que permite compilar y ejecutar paquetes de software sin utilizar virtualización del hardware. Estos se diferencian de una

máquina virtual, que realiza la completa virtualización del dispositivo físico y necesita ejecutarse sobre un Sistema Operativo que es redundante. Los contenedores se ejecutan sobre el S.O anfitrión y encima del motor de contenedores, que los emplean como base para poder funcionar. Estos contenedores a su vez pueden tener o no instalados un Sistema Operativo propio. Esta es la gran diferencia con las máquinas virtuales [10] [11].

En esta investigación se utiliza el sistema Docker, como motor de contenedores. De esta forma se está desarrollando un entorno de integración empaquetado y automatizado. Este motor va a generar contenedores de software, donde se ejecutará el emulador Qemu, que estará configurado para emular el hardware físico de las placas de desarrollo seleccionadas. El diseño de esta implementación se puede visualizar en la siguiente figura.

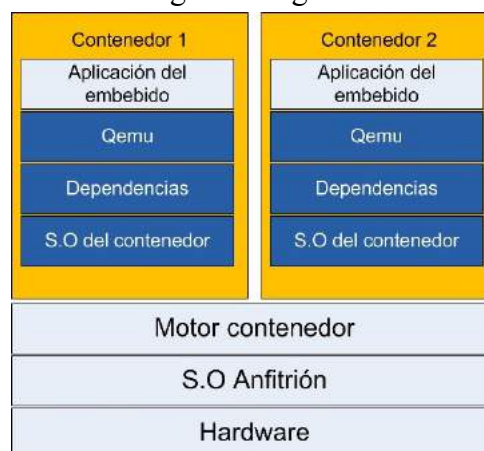


Fig. 2 Contenedor con el emulador de Sistemas embebidos

Como se puede observar en la figura, en la base de la pila se encuentra el hardware de la computadora. Sobre este funciona el Sistema Operativo Anfitrión y por encima del mismo el motor Docker. Dentro del contenedor se apilan las aplicaciones. En la parte inferior se encuentra instalado el Kernel mínimo del S.O Ubuntu. Luego se instalan las dependencias necesarias que necesita el emulador. Esto se debe a que Qemu necesita sus componentes para

poder ejecutarse. Posteriormente, en la capa siguiente se instala Qemu, ya preparado para que emule determinadas placas de desarrollo. Finalmente, en la capa superior se instala las aplicaciones del embebido, que el usuario desea ejecutar sobre la placa emulada en Qemu. De esta forma se pretende generar contenedores con imágenes de los sistemas embebidos, que permitan realizar pruebas en un entorno estandarizado, de manera tal que las aplicaciones del embebido puedan funcionar tanto en un sistema real como virtual.

3. RESULTADOS OBTENIDOS/ESPERADOS

En esta investigación se espera generar ambientes de trabajos automatizados, utilizando contenedores Docker que posean instalado y configurado el emulador Qemu, para poder utilizar distintos modelos de placas de desarrollo STM32, ESP32 y Arduino en forma virtual. De esta manera permitirá generar entornos de trabajos para automatizar pruebas de rendimiento y procesamiento simulado, de forma que los usuarios usen sus aplicaciones en el hardware físico adecuado de acuerdo con sus necesidades, incluso antes de adquirirlos.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación dentro del departamento de Ingeniería e Investigaciones Tecnológicas forma parte del trabajo que uno de los investigadores se encuentra realizando para su maestría. Completan el grupo de investigación dos docentes de categoría V y dos ingenieros en formación de investigador.

5. BIBLIOGRAFÍA

- [1] J. Novillo-Vicuña, D. H. Rojas, B. M. Olivo, J. M. Ríos y O. C. Villavicencio, *Arduino y el Internet de las cosas*, Alicante: Editorial Area de Innovación y Desarrollo, 2018.
- [2] G. L. Nicolas Goilav, *Arduino: Aprender a desarrollar para crear objetos inteligentes*, Barcelona, España: Ediciones ENI, 2016.
- [3] G. Trends. [En línea]. Available: <https://trends.google.com/trends/explore?date=2020-01-01%202021-02-18&geo=AR&q=arduino,esp8266,raspberry%20pi,esp32,stm32>.
- [4] E. Carnuccio, G. De Luca, W. Valiente, M. Volker, R. Villca y M. Adagio, «Análisis de rendimiento y consumo para sistema embebido con requisitos de tiempo explícitos,» de Libro de Actas del XXVI Congreso Argentino de Ciencias de la Computación, San Justo, Universidad de La Matanza, 2020.
- [5] J. Teel, *How to Turn Your Arduino Project into a Sellable Product*, Predictable Designs, 2016.
- [6] M. Montero. [En línea]. Available: <https://tecnoticias.net/2019/07/20/porque-la-raspberry-pi-no-es-una-buena-opcion-para-productos-comerciales/#8230>.
- [7] D. Aranda, *Electronica: Plataformas Arduino y Raspberri Pi*, Buenos Aires: RedUsers, 2014.
- [8] R. Milo-Sanchez, A. Fajardo-Moya y W. Rodriguez, «QEMU, una alternativa

libre para la emulación de arquitecturas de [hardware,» de V Taller de Software Libre, 2014.

- [9] A. Kotovsky, How to Develop Embedded Software Using The QEMU Machine Emulator, Apriorit Inc, 2019.
- [10] J. M. Ortega Candel, DOCKER. Seguridad y monitorización en contenedores e imágenes, RC-Libros, 2019.
- [11] Microsoft. [En línea]. Available: <https://docs.microsoft.com/es-es/learn/modules/intro-to-docker-containers/3-how-docker-images-work>.

Estudio y Simulación de Redes Definidas por Software y Automatización de Red

Raúl Blanchet, Santiago Pérez, Higinio Facchini
CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería)
Facultad Regional Mendoza, Universidad Tecnológica Nacional
Rodríguez 273, Mendoza, Argentina - 0261-5244576
(santiagocp,higiniofac)@frm.utn.edu.ar

Resumen

En los últimos años las tecnologías de Información y Comunicación (TIC) han ofrecido nuevos desafíos para el futuro de las redes. El crecimiento exponencial de tráfico que circula por la red, junto con la demanda cambiante de ancho de banda, tipos de tráfico, y la necesidad de rápido despliegue de nuevos servicios, han puesto en foco los mecanismos tradicionales en los que las configuraciones de dispositivos se realizan manualmente y de forma estática. En consecuencia, se plantea la necesidad de cambiar los mecanismos de diseño y administración para dar un mayor dinamismo e “inteligencia” a la red. Las Redes Definidas por Software (SDN, Software Define Network) y la Automatización de Red surgen como nuevas tecnologías de red revolucionarias, que promete dar soluciones concretas a estas demandas.

En este trabajo se trata de relacionar las necesidades de las redes tradicionales con lo que aporta SDN y la Automatización de Red, construyendo escenarios de prueba simulados para el análisis de desempeño y de nuevas capacidades de gestión.

El personal principal relacionado a esta línea de investigación es docente de grado y posgrado de la Universidad Tecnológica Nacional, Regional Mendoza, y tesis

de posgrado de la Especialización en Redes de Datos de la misma institución.

Palabras clave: SDN, Automatización de Red, Networking,

Introducción

Las Redes Definidas por Software (SDN), y la Automatización de Red, son tecnologías de red en crecimiento, que han ido ganando terreno al dar soluciones concretas a problemas de las redes tradicionales, dados por el actual crecimiento de tráfico, las demandas cambiantes y la necesidad de rápido despliegue de nuevos servicios, entre otros. Se basan en la separación entre el plano de control y el plano de datos en la infraestructura de la red, lo que posibilita un control centralizado y programabilidad en la gestión, haciendo que las redes sean más eficientes, flexibles y escalables. Las SDN y la Automatización de Red permiten centralizar la inteligencia de la red en un conjunto de controladores, y estos, además, se pueden configurar con aplicaciones escritas en lenguajes de alto nivel.

Estado del Arte

Las redes de datos, como actualmente se conocen, surgieron en los años 1970 como resultado de los requerimientos de

las primeras redes militares y de la aparición de Internet. Las mismas fueron concebidas con un funcionamiento basado en la conmutación de paquetes, donde cada componente funciona de manera inteligente, con capacidad para la toma de decisiones autónomas, que aseguren la transmisión de información de extremo a extremo. En este proceso, los componentes involucrados en la comunicación, no poseían una visión integral de la red. Los componentes físicos requeridos para el funcionamiento de las redes nacieron como estructuras monolíticas con hardware, software e interfaces propietarias. Entre otras, dicha característica generó que estos dispositivos evolucionen de manera lenta con respecto a otras tecnologías, como por ejemplo servidores y soluciones de almacenamiento. La consecuencia directa de esta situación se ve reflejada en la incapacidad de las redes para soportar los requerimientos actuales de “time to market” en la provisión de nuevos servicios, dificultad en su gestión, complejidad y elevados costos de operación, entre otros.

Las tecnologías móviles, la continua interacción de los usuarios con contenido en línea, las nuevas tecnologías de virtualización de servidores y el advenimiento de servicios en la nube, obligan a realizar una revisión de las premisas de diseño y funcionamiento de las actuales redes.

Hoy en día, las redes de datos están compuestas por switches y routers, que posibilitan las comunicaciones entre clientes y servidores físicos o virtuales, conformando redes complejas y difíciles de administrar. Para alcanzar tales funcionalidades, los operadores requieren configurar cada dispositivo en forma individual, utilizando comandos específicos que en la mayoría de los casos son propietarios del fabricante. La

característica distribuida de las redes actuales se basa en un diseño descentralizado en el cual la lógica de control y la función de distribución de paquetes (ruteo/forwarding) está embebido en cada uno de los componentes de la red. Es así que cada router o switch soporta una serie de protocolos distribuidos que facilitan la toma de decisión en el direccionamiento de paquetes a lo largo de la red.

SDN es un paradigma en desarrollo que permite el control centralizado y programabilidad en la gestión de las redes mediante la separación del plano de control (software) y del plano de datos (hardware). El objetivo principal es obtener mayor rendimiento, flexibilidad y escalabilidad en la implantación de servicios de red, a la vez que facilita la labor de gestión de dichos servicios por parte del administrador. Se basa en la interacción entre tres capas: capa de aplicación, donde residen las aplicaciones que determinan el comportamiento deseado en la red; el plano de datos, que contiene los switches, físicos o virtuales; y la capa de control, que sirve de nexo entre las dos anteriores. En esta capa se encuentran los controladores, que proporcionan una abstracción de las capas inferiores a los desarrolladores de aplicaciones. SDN se vincula con la Automatización de Red, que es una metodología en la cual el software configura, aprovisiona, administra y prueba automáticamente los dispositivos de red. Lo utilizan las empresas y los proveedores de servicios para mejorar la eficiencia y reducir los errores humanos y los gastos operativos.

Los equipos que posibilitan el encaminamiento de paquetes a través de las redes en la actualidad (switches y routers) están diseñados con una arquitectura del tipo monolítica, de acuerdo al funcionamiento tradicional de

las redes de datos, es decir, incorporando en un mismo dispositivo físico el plano de datos, de control y gestión, para así posibilitar el ruteo individual de los paquetes entrantes. En el caso de un dispositivo de capa 2 del modelo OSI (switch), el plano de datos o forwarding, está constituido por los puertos físicos a través de los cuales se reciben y transmiten las tramas de datos. Su principal función es la de encaminar éstas hacia los puertos de salida, utilizando para ello la información contenida en la tabla de forwarding embebida. Si la información de cabecera de una trama entrante, es encontrada en dicha tabla, esta puede ser susceptible de modificaciones para luego ser transmitida a través del puerto correspondiente, sin intervención de los otros planos. Esta situación no sucede en todos los casos. El ejemplo más claro, es cuando la información de la cabecera no se encuentra aún mapeada en la tabla. Es en este caso, es necesaria la intervención del plano de control, cuya principal responsabilidad es la de mantener actualizada la información de forwarding, de manera tal de que el plano de datos pueda retransmitir las tramas sin su intervención.

Una de las principales organizaciones que contribuyen al desarrollo de SDN es Open Network Foundation (ONF) que es una organización compuesta por los mayores referentes del mercado de las telecomunicaciones y cuyo objetivo es la promoción, adopción y estandarización de SDN.

Protocolo OpenFlow

Es el protocolo más utilizado para la comunicación entre la capa de datos y el controlador. Surgió en 2008 con fines de investigación, pero pronto fue incorporado por empresas ya que permite

gestionar la red sin depender de los fabricantes. Los conmutadores que funcionan con OpenFlow se denominan switches OpenFlow. Los switches SDN tienen al igual que los tradicionales, la funcionalidad de forwarding que es la que permite decidir qué hacer con cada flujo que ingresa por sus puertos. Sin embargo, a diferencia de los tradicionales que funcionan a través de MAC learning (proceso mediante el cual construyen su tabla de forwarding), los switches de SDN funcionan a partir de una tabla de flujo. Dicha tabla de flujo es gestionada a través del controlador. Un flujo es un conjunto de paquetes que comparten una serie de características. El protocolo OpenFlow es el encargado de transmitir al switch las entradas que debe añadir, modificar o quitar en su tabla de flujos.

Metodología

El presente trabajo abordará la temática de las Redes Definidas por Software (SDN) y la Automatización de Red en contraposición a las redes tradicionales. Comenzará con una introducción conceptual sobre SDN, y la Automatización de Red, sus fundamentos tecnológicos y su aporte a la flexibilización e inteligencia de las redes. En segunda instancia se creará un ambiente de simulación que permita representar distintos escenarios, con el fin de obtener medidas de desempeño en redes tradicionales y redes SDN-Automatización de Red en cuanto a: (i) funcionamiento normal, (ii) caída de enlaces, (iii) distintas demandas de calidad de servicio (QoS), y (iv) nuevas capacidades de red. Por último, se analizarán los resultados obtenidos y pondrá en contraste el desempeño y limitaciones de ambas tecnologías, con el

fin de obtener conclusiones y posibles trabajos futuros.

Para el desarrollo del proyecto se ha previsto las siguientes tareas:

Tarea 1: Recopilar de información y estudio de los protocolos y topologías disponibles.

Tarea 2: Compilar trabajos de investigación sobre la temática, determinando escenarios experimentales, herramientas de simulación, tráfico utilizados, métricas medidas, resultados y conclusiones.

Tarea 3: Establecer mecanismos comparativos de los trabajos de investigación compilados, usando cuadros, índices, ponderaciones, etc.

Tarea 4: Definir escenarios de experimentación que faciliten el contraste entre ambos tipos de tecnologías.

Tarea 5: Construir tablas comparativas de las métricas para cada escenario. Se construirán tablas y gráficas comparativas de las prestaciones o métricas para cada caso, según las combinaciones que permitan los diversos escenarios.

Tarea 8: Documentar, publicar y difundir resultados.

Tarea 9: Redactar Informe Final.

Estado de Avance

Debido a los requerimientos de hardware para el ambiente de simulación, se ha montado todo un laboratorio sobre un escenario completamente virtualizado.

Los servidores de virtualización utilizados con 6.0, con vCenter Server Appliance 6.5. Sobre VMware ESXi se creó una Máquina Virtual (VM), para alojar a GNS3 VM versión 2.2.11, y otra VM con un Windows 10 que tiene instalado dentro del propio Sistema Operativo, el software GNS3 para Windows, versión 2.2.11. Esta forma de instalación es la recomendada por el desarrollador para un óptimo desempeño,

y permite utilizar GNS3 para Windows como entorno gráfico y cliente, mientras que GNS3VM cumple la función de servidor. Para esto, cuando se crean topologías de red y se van agregando los distintos elementos, se puede indicar que los mismos se ejecuten sobre GNS3VM.

GNS3 (Graphical Network Simulator 3) es un software gratuito y de código abierto que se utiliza para emular, configurar, probar y solucionar problemas de redes virtuales y reales. Permite ejecutar desde una pequeña topología que consta de unos pocos dispositivos en una computadora portátil, hasta aquellas con múltiples dispositivos alojados en distintos servidores o incluso en la nube.

Entre otras herramientas, se incluye el uso de Mininet. Mininet es un proyecto de código abierto liderado por miembros de la ONF que surgió en las universidades de Berkeley y Stanford. Permite emular redes en un único kernel de Linux. Utiliza virtualización ligera para hacer que un solo sistema parezca una red completa, ejecutando el mismo kernel, sistema y código de usuario. De esta forma los usuarios pueden experimentar con redes que incluyen hosts finales, switches y enlaces virtualizados sobre un único sistema operativo. Es utilizado en entornos de prueba y prototipado ya que su comportamiento es similar al de los respectivos equipos hardware, con las precauciones debidas en el apartado de prestaciones

Estos avances parciales han sido realizados por el equipo de trabajo integrado por docentes investigadores, y becarios graduados y alumnos que son miembros del Centro UTN CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería) de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional, dentro del Área de Análisis de Tráfico y Seguridad en Redes de Datos. Entre ellos

están Docentes de grado en la Redes de Datos de la Carrera de Ingeniería en Electrónica, y de la Especialización en Redes de Datos.

Objetivos

Objetivo General:

El presente trabajo abordará el contraste cualitativo y cuantitativo de las Redes Definidas por Software (SDN) y la Automatización de Red en contraposición a las redes tradicionales desde la simulación experimental.

Objetivos Específicos:

Para el presente proyecto se buscan los siguientes objetivos:

- Determinar el impacto de avance de estas tecnologías, y la diversidad de modelos arquitectónicos disponibles.
- Determinar las características a detalle de los recursos de simulación, y de los escenarios de experimentación.
- Determinar las configuraciones óptimas SDN y de red tradicionales, y sus ámbitos de aplicación, para cada uno de los escenarios de experimentación.
- Definir las necesidades actuales locales a fin de iniciar desarrollos y experiencia, para alcanzar una masa crítica de conocimientos y poder interactuar, con otros investigadores y profesionales.

En general se definirán métricas a obtener para cuantificar los distintos parámetros de trabajo, en cuanto a rendimiento, velocidad, consumo de recursos, uso de memoria interna, etc para poder realizar tablas, cuadros y gráficos comparativos para cada escenario planteado.

Formación de Recursos Humanos

Mediante el presente proyecto se busca la formación de Recursos Humanos con el fin de:

- Aglutinar RRHH calificados en torno a estas nuevas tecnologías y crear las sinergias para facilitar su participación en proyectos e iniciativas nacionales (2021-2022) e internacionales (2023-2024).
- Incidir en los gestores de la I+D nacional para que incentiven estas tecnologías a fin de garantizar una cierta independencia tecnológica. Estando al día de las distintas iniciativas, podremos contribuir a las mismas y participar de sus convocatorias.
- Facilitar la organización sostenida de seminarios científico-tecnológicos para la actualización de nuestro entorno y una visión prospectiva sobre la proyección de futuro de estas tecnologías.

Referencias

- [1] Stallings, W. (2016). Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud. Pearson Education.
- [2] Chuck Black, Paul Goransson (2017). Software Defined Networks: A Comprehensive Approach, Second Edition. Elsevier Inc.
- [3] Neumann, J. (2015). The Book of GNS3 Build Virtual Network Labs, 1-6.
- [4] Pujolle, G., (2015). Software Networks - Virtualization, SDN, 5G and Security, 25-26.
- [5] Techopedia: Educating IT Professionals To Make Smarter <https://www.techopedia.com/> (Accessed July 16th, 2020)
- [6] Webopedia: Online Tech Dictionary for Students, Educators and IT Professionals,

- <https://www.webopedia.com/>, (Accessed July 16th, 2020)
- [7] MIT Technology Review, <https://www.technologyreview.com/>, (Accessed July 16th, 2020)
- [8] Ericsson, <https://www.ericsson.com/en>, (Accessed July 16th, 2020)
- [9] Bodafone, <https://www.vmware.com/content/microsites/possible/stories-uk/vodafone.html>, (Accessed July 16th, 2020)
- [10] Blue Planet Technology, <https://www.blueplanet.com/technology/>, (Accessed July 16th, 2020)
- [11] China Mobile, <https://www.chinamobileltd.com/en/global/home.php>, (Accessed July 16th, 2020)
- [12] Catalyst Digital Showcase, <https://inevent.com/en/TMForum-1588605874/catalyst-digital-showcase/hotsite.php>, (Accessed July 16th, 2020)
- [13] China Telecom Corporation Limited, <https://www.chinatelecom-h.com/en/global/home.php>, (Accessed July 16th, 2020)
- [14] Global Architecture Forum, <https://www.tmforum.org/global-architecture-forum/>, (Accessed July 16th, 2020)
- [15] Orange, <https://www.orange-business.com/en/solutions/connectivity>, (Accessed July 16th, 2020)
- [16] OpenAI, <https://openai.com/>, (Accessed July 16th, 2020).
- [17] Chafloque Mejia, J.: Propuesta de diseño de una red LAN bajo la arquitectura SDN para la Red Telemática de la Universidad Nacional Mayor de San Marcos, 2018
<https://cybertesis.unmsm.edu.pe/handle/20.500.12672/10017>
- [18] Ibañez Garcia, F.: Estudio de las tecnologías SDN y NFV, 2016
<http://openaccess.uoc.edu/webapps/o2/bitstream/10609/45441/6/fibanezgarTFG0116memoria.pdf>
- [19] ONF: OpenFlow Switch Specification, 2012,
<https://opennetworking.org/wp-content/uploads/2014/10/openflow-specv1.3.0.pdf>
- [20] Osaba M: Virtualización en redes definidas por software, 2016
https://ri.itba.edu.ar/bitstream/handle/123456789/785/TELCOM.%20Osaba%20Virtualizaci%C3%B3n_de_Redex_Def_por_SW_.pdf?squence=1&isAllowed=y

Protocolo TCP: El RTT como un factor de evaluación del rendimiento

Carlos A. Talay, Diego R. Rodríguez Herlein, Marycarmen Díaz Labrador,

Claudia N. González

Campus Universitario – Oficina 18/ Dpto. Ciencias Exactas e Informática
UARG/UNPA

{ctalay, dherlein}@uarg.unpa.edu.ar, marydl920911@gmail.com, cgonzalez@uarg.unpa.edu.ar

Luis A. Marrone

L.I.N.T.I. – Universidad Nacional de La Plata
Calle 50 y 120 – 2 do. Piso – Edificio Bosque Oeste
lmarrone@linti.unlp.edu.ar

Resumen

La latencia es una de las principales variables a considerar en una red de datos. Su fluctuación afecta notablemente la transmisión de información, impactándola negativamente. La razón por la cual esto sucede puede ser muy variada: congestión de datos, degradación del medio de transmisión, el uso de redes heterogéneas que poseen distintas características de rendimiento, etc. Teniendo en cuenta este hecho, es interesante analizar que característica de la red puede alertarnos al respecto. En este artículo intentamos mostrar un ejemplo de cómo podemos utilizar el RTT (Round Trip Time), para ayudarnos a determinar la existencia de factores que degradan las propiedades de transmisión de una red de datos.

Palabras clave: TCP, Latencia, RTT, Rendimiento

Contexto

Este estudio está enmarcado en el proyecto de investigación 29/A451-1 “Análisis del comportamiento de protocolos de la capa de transporte en redes con enlaces inalámbricos”, y es una

continuación del proyecto 29/A396-1 “Evaluación del protocolo TCP en topologías mixtas cableadas-inalámbricas”, ambos radicados de la UNPA-UARG. El proyecto está compuesto mayoritariamente de docentes de la UNPA-UARG, dirigido por el Sr. Carlos A. Talay, cuenta como Co-director al Sr. Luis A. Marrone perteneciente a la UNLP y lo integran los docentes Claudia N. González, Marycarmen Díaz Labrador, Diego R. Rodríguez Herlein y los alumnos Vanesa Cadín y Franco A. Trinidad, financiándose íntegramente con fondos destinados a proyecto de investigación de la UNPA-UARG.

Introducción

En los proyectos enunciados, se ha analizado el comportamiento de algunas variantes del protocolo TCP (Transmission Control Protocol) [1], con distintas condiciones de tráfico y topología. Con ello intentamos determinar las fortalezas y debilidades de estos protocolos, y así proponer mejoras que optimicen su respuesta según las condiciones de funcionamiento [2].

Es así que, una vez obtenidos los datos resultantes de las simulaciones realizadas, procedemos a analizar los resultados arrojados por el simulador [3]. Para ello recurrimos a los archivos de datos que el simulador nos entrega y en base a estos calculamos las métricas. Estas describen el comportamiento del tráfico en esa topología, pudiendo citar entre las más utilizadas: *Throughput*, *Goodput*, *Delay*, *RTO (Retransmission Time Out)*, *CWND (Congestion Window)*, *PLR (Packet Lost Ratio)* y *RTT (Round-trip Time ó Round-Trip delay Time)*. En particular esta última es utilizada para verificar, en forma instantánea, el estado del enlace entre dos nodos. Hay que tener en cuenta que en la actualidad dominan la escena las redes heterogéneas [4] y, si bien el RTT nos da una idea del estado del enlace extremo a extremo, no podemos conocer las condiciones particulares de los estadios intermedios. A pesar de ello, el RTT nos brinda una importante información del tiempo que emplea un paquete de datos, desde que sale del nodo origen, hasta la recepción de su correspondiente ack (acknowledge), que confirma su llegada a destino.

Uno de los mecanismos que más influyen en el rendimiento de TCP es el de control de congestión [5]. Los algoritmos de control de congestión [6] permiten regular la tasa de paquetes enviados, y de esta manera, evitar el colapso de una comunicación por congestión. En este sentido el RTT [7] [8] nos proporciona información respecto al estado del camino del flujo de datos, dado que, por ejemplo, comparando el valor

obtenido respecto de un valor medio, podemos inferir el estado de sobrecarga de toda la red.

El siguiente esquema (fig.1), nos permite referenciar de una manera simple la situación en que se encuentra el tráfico de una red en un momento determinado.

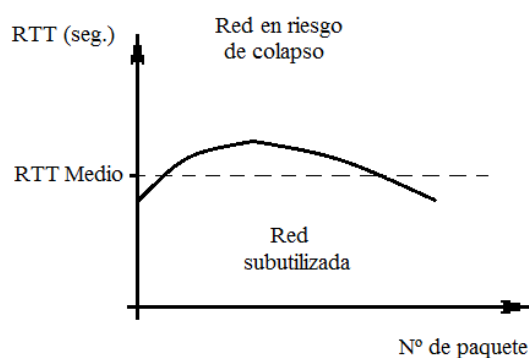


Figura 1. RTT vs. Número de paquete

En este esquema, podemos apreciar como el RTT nos da una idea de cómo se ve afectado el tráfico en la red. Veamos un caso de estudio donde el RTT nos permite apreciar la degradación que se produce en un flujo de datos cuando la línea sufre distintos retardos en el tráfico.

Caso de estudio

Como mencionábamos proponemos un caso de estudio en donde se tienen dos flujos de datos, originados en distintos nodos, debido al diseño ya referenciado. También tenemos un nodo intermedio que podría atribuirse a un access point de una red inalámbrica típica (WiFi). El modelo implementado en el simulador ns-3 (Network Simulator 3) representa una topología híbrida de 5 nodos (2 nodos cableados, 1 estación base y 2 nodos

inalámbricos), como se observa en la figura (fig.2)

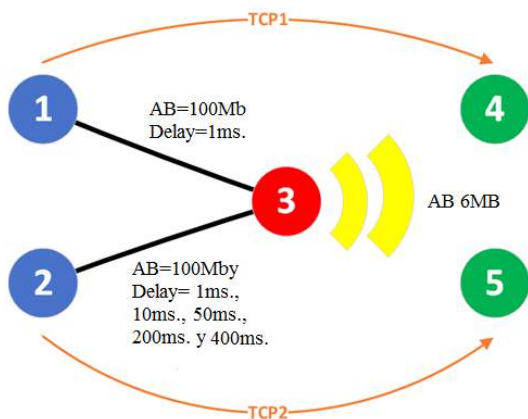


Figura 2. Modelo empleado en el caso de estudio

Los nodos 1 y 2 están conectados a la estación base (nodo 3), mediante enlaces cableados full dúplex de 100Mb/s. El retardo de propagación indicado para cada caso se recrea en los gráficos subsiguientes y política de cola DropTail. Para el enlace inalámbrico tenemos un modo de propagación TwoRayGround, capa física WirelessPhy, MAC 802.11, antena OmniAntena y los dos nodos (4 y 5), que no poseen movimiento.

Para la realización de la prueba se realizaron 5 (cinco) ensayos con dos tráficos de tipo FTP, que envían 5.000 paquetes con un tamaño de 1.000 bytes cada uno. Los protocolos utilizados en ambos casos es New Reno. Los flujos se establecieron del nodo 1 al 4 y simultáneamente se estableció un flujo del nodo 2 al 5. Es decir, ambos flujos comparten su tránsito por el nodo 3.

Línea de investigación y desarrollo

Para el primer ensayo, retardo de línea (delay), asignado para los 2 nodos

cableados, es de 1 ms. Las condiciones de funcionamiento del segmento inalámbrico de la red (WiFi) es común a los nodos 4 y 5.

Bajo estas condiciones presentamos los resultados obtenidos para RTT, calculados para los nodos 1 y 2.

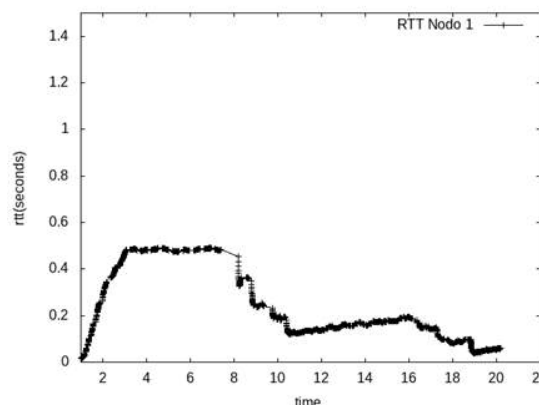


Figura 3. RTT vs. Tiempo. Nodo 1 y retardo=1ms. NewReno-NewReno.

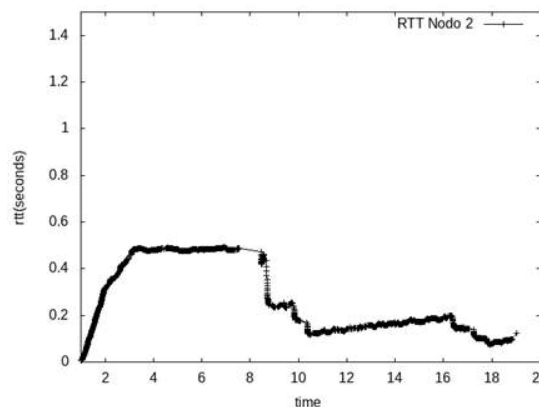


Figura 4. RTT vs. Tiempo. Nodo 2 y retardo=1ms. NewReno-NewReno.

Luego de obtener estos valores, que se presentan como estado base de comparación, procedemos a incrementar los retardos sobre la línea que une los nodos 2 y 4 y generamos 4 nuevos ensayos, en donde la línea que une los nodos 1 y 3 permanece en 1ms. (fig. 5, 7, 9 y 11), medido en el nodo 1 y por otro lado la línea que une los nodos 2 y 3, que

en estos ensayos se definen con retardos de 10ms. (fig. 6), 50ms. (fig. 8), 200ms. (fig. 10) y 400ms. (fig. 12), medidos en el nodo 2.

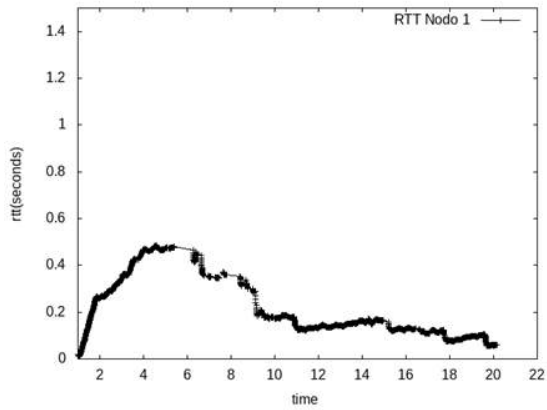


Figura 5. RTT vs. Tiempo. Nodo 1 y retardo=1ms. NewReno-NewReno.

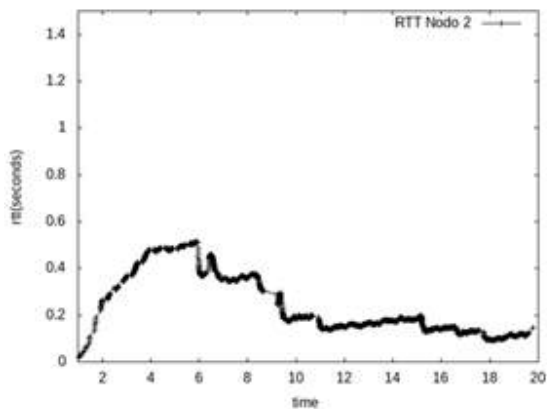


Figura 6. RTT vs. Tiempo. Nodo 2 y retardo=10ms. NewReno-NewReno.

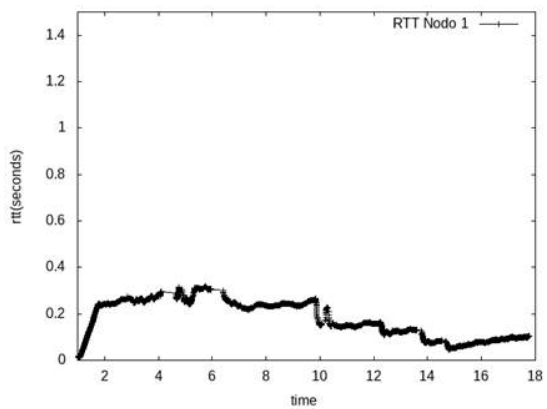


Figura 7. RTT vs. Tiempo. Nodo 1 y retardo=1ms. NewReno-NewReno.

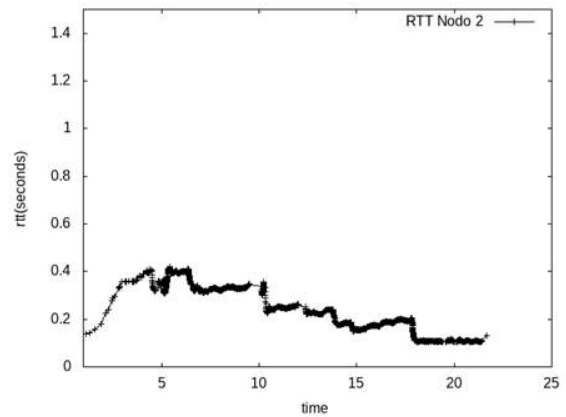


Figura 8. RTT vs. Tiempo. Nodo 2 y retardo=50ms. NewReno-NewReno.

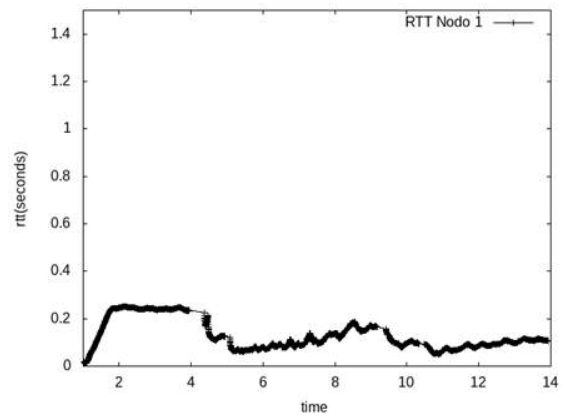


Figura 9. RTT vs. Tiempo. Nodo 1 y retardo=1ms. NewReno-NewReno.

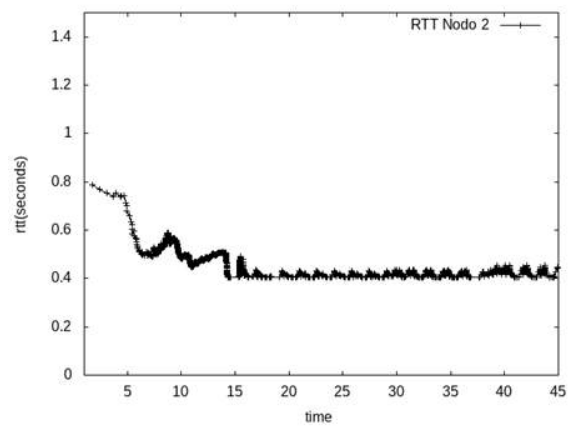


Figura 10. RTT vs. Tiempo. Nodo 2 y retardo=200ms. NewReno-NewReno.

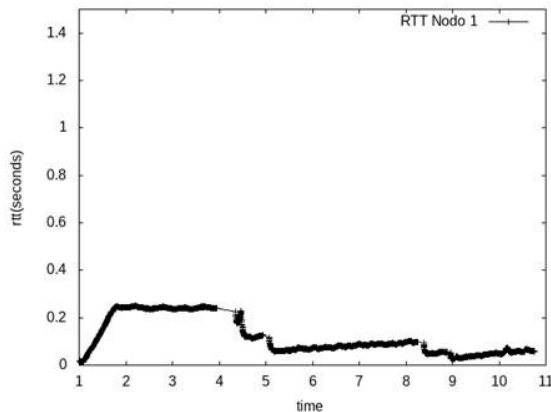


Figura 11. RTT vs. Tiempo. Nodo 1 y retardo=1ms. NewReno-NewReno.

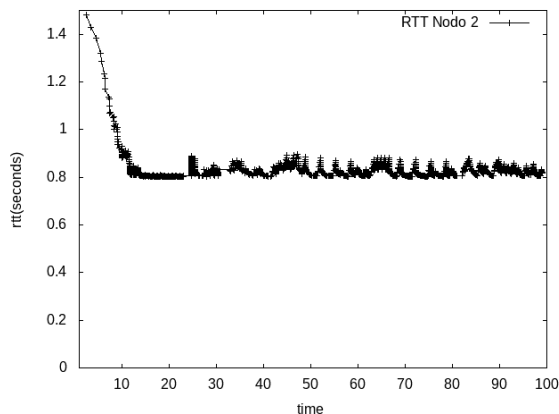


Figura 12. RTT vs. Tiempo. Nodo 2 y retardo=400ms. NewReno-NewReno.

Resultados obtenidos y futuras líneas de investigación

Como se observa la variación del RTT es un indicativo de la degradación del tráfico. Cuando el RTT aumenta, esto nos indica problemas en el tránsito de los datos. Si bien en este caso se forzó la degradación introduciendo retardos en la línea, los problemas de degradación puede tener orígenes más complejos y su determinación es motivo de un análisis que puede incluir más factores.

En el futuro se tiene previsto ampliar la cantidad de nodos incluidos en la topología y utilizar más variantes del

protocolo TCP. También nos encontramos en el proceso de migración del simulador ns-2 a ns-3, básicamente por la discontinuidad en el soporte de ns-2 y la mayor versatilidad de ns-3 para construir escenarios y simularlos.

Formación de Recursos Humanos

Este proyecto se ha conformado con investigadores consolidados en el área de protocolos de comunicaciones. Así mismo se ha incorporado a un nuevo docente-investigador que se formará en este área de investigación, reforzando las características heterogéneas del grupo de profesionales vinculados al grupo.

Referencias

- [1] Saleem-ullah Lar, Xiaofeng Liao, "An initiative for a classified bibliography on TCP/IP congestion control", *Journal of Network and Computer Applications* 36, 126–133, 2013.
- [2] Teja F. R., Vidal, L., Alves, L., "TCP sobre enlaces wireless – Problemas y algunas posibles soluciones existentes", Curso de posgrado y actualización, Instituto de Ingeniería Eléctrica, Facultad de la República, marzo 2004.
- [3] ns-2/ns-3. <https://www.nsnam.org/support/faq/ns2-ns3/> (accessed Sep. 10, 2020).
- [4] HALA ELAARAG, "Mejora del rendimiento de TCP sobre redes móviles" Stetson University - ACM Computing Surveys, Vol. 34, No. 3, September 2002.
- [5] M. Handley, J. Padhye and S. Floyd, "TCP Congestion Window Validation", RFC 2861, June 2000.
- [6] Diego R. Rodríguez Herlein, Carlos A. Talay, Claudia N. González, Franco A. Trinidad, María L. Almada and Luis A. Marrone, "Un Análisis de contienda de mecanismos de control de congestión en un escenario que incluye acceso inalámbrico". 6º Encuentro de investigadores UNPA-UARG. ISBN: 978-987-3714-88-7, pág. 705-9. 2020.
- [7] M. Daniel Docmac, Eduardo I. Silva, and Agustín J. González. "Evaluación de algoritmos para la estimación del Round-Trip Delay Time y una propuesta de mejora" *Revista Iberoamericana de Automática e Informática Industrial RIAI*, vol. 10, no. 1, 2012. doi:10.1016/j.riai.2012.11.006
- [8] Lei, Y., Zhu, R., Wang, W., 2006. "A survey on TCP protocol and RTT estimation. In: *Intelligent Control and Automation*", 2006. WCICA 2006. The Sixth World Congress on. Vol. 1. pp. 4410–4414. DOI: 10.1109/WCICA.2006.1713211

Redes de acceso de 5ta generación

Peliza Carlos, Dufour Fernando, Serra Ariel, Micieli Gustavo, Machaca Darío

Departamento de Informática e Investigaciones Tecnológicas

Universidad Nacional de La Matanza

Florencio Varela 1908, San Justo

pelizac@gmail.com; fdufourf@hotmail.com ; aserra@unlam.edu.ar ;
gmicieli@gmail.com

RESUMEN

Las expectativas de crecimiento en las redes móviles siguen siendo enormes. El mundo de las comunicaciones avanza hacia la próxima generación de redes inalámbricas 5G, en base a la necesidad de los usuarios de superar las prestaciones que tienen las tecnologías 4G y sus predecesoras.

Para la implementación completa de esta nueva generación de redes móviles, existe el problema de integrar diversas tecnologías que deben interactuar en forma adecuada. Como ejemplo de lo citado podemos mencionar <https://www.iprofesional.com/tecnologia/287385-celular-Falla-de-seguridad-en-4G-y-5G-permitiria-a-hackers-interceptar-llamadas>.

La responsabilidad de integración recae, como siempre ocurre en el mundo de las comunicaciones, sobre las nuevas tecnologías y arquitecturas que se incorporarán a los servicios existentes.

La metodología a utilizar será la investigación documental, entrevistas a especialistas de distintos fabricantes de equipamiento para la generación 5G y operadoras de servicios (Carriers), pruebas y análisis de normativa generada por entidades de estandarización.

Se buscarán las incompatibilidades y a partir de un debate interno se sacarán las conclusiones respecto de las mejores prácticas que se contrastarán con

pruebas hechas por los Carriers y/o pruebas piloto.

Se generará un informe final en el que entre otros aspectos se detallarán las mejores prácticas para la implementación de estas tecnologías en las redes y servicios actuales. Se realizará también una prospectiva de esta tecnología para los próximos 3 años.

El impacto que producirá la investigación será en el entorno de los Carriers de comunicaciones que tendrán una guía para determinar qué tipos de implementaciones hacer para mejorar sus servicios actuales de comunicaciones. En el ámbito académico, el impacto de la investigación estará dado por la incorporación de un tema de punta en las cátedras de redes y comunicaciones del Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLAM.

Palabras Clave: Modulación, Frecuencias, Espectro Radioeléctrico, Prospectiva, Dispositivos

CONTEXTO

La secretaría de investigaciones tecnológicas, en conjunto con del Departamento de Ingeniería e Investigaciones Tecnológicas, son quienes dictan las políticas de investigación y las entienden como la base de producción y distribución de conocimiento. Particularmente este grupo de investigación ha venido desarrollando una serie de

investigaciones orientadas a las redes de comunicaciones o a las telecomunicaciones en general, como antecedentes se pueden citar: C164 (Carrier- Ethernet), C189 (LTE), C210 (Virtualización de funciones de red) y trabajos en diferentes congresos como “Carrier Aggregation”, “Narrow Band IoT”, “VoLTE”.

Temáticas que luego son introducidas en cátedras de comunicaciones de la Universidad Nacional de La Matanza, en carreras de Informática como de Electrónica. Adicionalmente, se brindan charlas de divulgación para la comunidad educativa de acuerdo a las necesidades del centro de estudiantes de dicha institución

1. INTRODUCCION

Las redes de acceso de 5ta generación son la evolución de las redes de datos móviles y brindarán un mayor ancho de banda de navegación, una disminución de la latencia y una mejora notable en la calidad de las comunicaciones entre otros beneficios.

El despliegue de estas redes por parte de las diferentes operadoras de telefonía móvil en Argentina se encuentra en un estado de desarrollo primario, existiendo configuraciones y parametrizaciones aún no depuradas, denominadas PoC (pruebas de concepto) o demostraciones comerciales.

Al ser una tecnología en auge, en el mercado no hay gran cantidad de profesionales especializados ni abundancia de documentación (más allá del estándar 3GPP). Esto hace que el presente protocolo de investigación cobre suma importancia para el mercado de las telecomunicaciones en Argentina.

La problemática a investigar incluye, entre otras, las percepciones de los principales especialistas en estas tecnologías, la determinación de las

mejores prácticas para la implementación de redes de acceso de 5ta generación, y lo que ocurrirá con esta tecnología de acceso durante los próximos 3 años.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Los estudios permitirán vislumbrar distintas soluciones de implementación que llevarán asociadas cada una de ellas, ventajas y limitaciones, que determinarán la capacidad de los servicios que se presten.

El objeto de investigación es determinar dentro del universo de soluciones posibles de arquitecturas de redes, cuáles son las mejores prácticas para su implementación que impliquen la mejor utilización del espectro radioeléctrico, de manera de poder obtener en los terminales la mejor experiencia en el uso de los servicios.

El campo específico de la investigación consiste en trabajar sobre los aspectos de interoperabilidad y capacidades de los elementos de red de distintos proveedores determinando su compatibilidad de funcionamiento frente a los estándares, concentrándonos en el comportamiento del dispositivo y la red, frente a los servicios de este nuevo escenario tecnológico.

El modelo teórico conceptual se estructura a partir del conocimiento de incompatibilidades históricas que presentan las tecnologías en sus estados iniciales de desarrollo con las necesidades de buen funcionamiento de los servicios.

Esto requiere un análisis de dichas incompatibilidades de acuerdo con el estado del arte, con el objeto que los proveedores de tecnologías trabajen sobre las mismas y que los Carriers de comunicaciones tengan claro cuáles son

las limitaciones a la hora de implementar sus servicios.

La validez de los resultados se obtendrá a través de la verificación de la compatibilidad de las tecnologías para los servicios convergentes y la búsqueda de pruebas exitosas que convaliden nuestras conclusiones, para lo cual, una vez obtenido el resultado se lo contrastará con pruebas de laboratorio o pruebas piloto en operadoras de telefonía móvil que operan en el mercado.

3. OBJETIVOS

Objetivo principal

Esclarecer el grado de madurez de la tecnología, las características de las redes desplegadas y a desplegar en Argentina, verificando la compatibilidad e interoperabilidad con las redes actualmente desplegadas de 2G, 3G y 4G generando una recomendación de mejores prácticas para su implementación y despliegue de acuerdo con el estado del arte de esta tecnología, complementándolo con una prospectiva de evolución a 3 años.

Lograr reconocer las implementaciones más convenientes (mejores prácticas) que podrán ser utilizadas por los operadores para reducir tiempos de implantación

4. RESULTADOS OBTENIDOS/ ESPERADOS

Reconocer los proveedores de tecnología que existen en el mercado, así como las distintas Operadoras que brindan servicios sobre esta red.

Detectar los problemas de interoperabilidad entre proveedores.

Determinar los 5 principales dispositivos con capacidad de operar sobre la nueva tecnología.

Determinar las limitaciones y potencialidades que tendrán las nuevas redes.

Formar recursos humanos en esta novedosa tecnología.

Obtener información sobre cuáles serán los servicios que officiarán de apalancadores del negocio de las telecomunicaciones de la región.

5. FORMACION DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente del Departamento de Ingeniería de la Universidad de La Matanza, con miembros de las carreras de Electrónica e Informática, todos con gran experiencia en el área de telecomunicaciones.

Pero en esta investigación se ha incorporado un alumno, como parte de un plan para mejorar sus habilidades blandas o soft skills, a quien se le asignó la tarea supervisada de redacción científica de textos.

El alumno ha mejorado su capacidad de lectura y escritura en formato académico con especial énfasis en la generación de síntesis y análisis documentados, permitiéndonos acelerar la producción escrita de artículos científicos.

6. BIBLIOGRAFIA

- Ahmadi, S. (2019). Architecture, Technology, Implementation, and Operation of 3GPP New Radio Standards. Academic Press.
- Chen, A. Z. (2018). 5G Physical Layer Principles, Models and Technology Components. Academic Press.
- Dahlman, E., Parkvall, S., & Skold, J. (2018). 5G NR: The Next Generation Wireless Access Technology. Academic Press.

- Gosling, W. (1999). *Radio Spectrum Conservation*. Newnes 1999.
- Rodriguez, J. (2015). *Fundamentals of 5G Mobile Networks*. John Wiley & Sons, Ltd.
- Xiang, W., & Zheng, K. (2017). *5G Mobile Communications*. Switzerland: Springer International Publishing.

Salón de Clases Inteligente Eficiente Energéticamente

Diego A. Godoy^{a,1}, Hernán Bareiro^{b,1}, Fabian Favret^{c,1}, Guillermo Colloti^{d,1}, Juan Pablo Blariza^{e,1},
E. Marcelo Albornoz^{f,2}

¹Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.)
Departamento de Ingeniería y Ciencias de la Producción-Universidad Gastón Dachary
Av. López y Planes 6519- Posadas, Misiones, Argentina. Teléfono: +54-376-4438677

1

²Instituto de Investigación en Señales e Inteligencia Computacional, Facultad de Ingeniería y Cs.
Hídricas, Universidad Nacional del Litoral - CONICET

^a diegodoy@citic.ugd.edu.ar, ^b hbareiro@citic.ugd.edu.ar, ^c efabianfavret@citic.ugd.edu.ar,
^d gcolloti@citic.ugd.edu.ar, ^e jblariza@citic.ugd.edu.ar, ^f emalbornoz@sinc.unl.edu.ar

Resumen

En este trabajo se presenta un proyecto de investigación denominado “Tecnologías para Desarrollos Sostenibles de Ciudades Inteligentes”. Particularmente en este artículo se presentan los avances realizados en relación a “La construcción de un salón de clases eficiente energéticamente”. Para ello se trabaja en el desarrollo de un prototipo que permite reducir el uso de la energía eléctrica en función de la ocupación del aula y las necesidades del docente.

Palabras claves: Internet de las Cosas, Frameworks, Eficiencia Energética.

Contexto

Este trabajo se enmarca en el proyecto de investigación denominado “Tecnologías para Desarrollos Sostenibles de Ciudades Inteligentes”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) con el número Código IP A10002/19 y radicado en el Centro de Investigación

en Tecnologías de la Información y Comunicaciones de dicha Universidad.

El mismo fue incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°10 mediante la Resolución Rectoral (R.R.) 44/A/2019 y es una continuidad de los Proyectos Simulación en las TICs: Diseño de Simuladores de Procesos de Desarrollo de Software Ágiles y Redes de Sensores Inalámbricos para la Industria y la Academia. R.R. UGD N° 07/A/17 y Simulación como herramienta para la mejora de los procesos de software desarrollados con metodologías ágiles utilizando dinámica de sistemas, R.R. UGD N° 18/A/14 y R.R. UGD N° 24/A/15.

Introducción

Una institución educativa moderna, cuenta con una gran cantidad de aulas, equipadas con numerosos dispositivos de iluminación y refrigeración (aires acondicionados). “Diversos organismos, comprometidos con el uso eficiente de la energía y la conservación de nuestro medio ambiente, han reportado que los edificios son responsables por el

consumo del 40% o más de toda la energía primaria producida a nivel mundial...” [1]. De acuerdo a esto, es razonable afirmar que la energía consumida en un día de clases en horario pico es muy elevada.

En el mismo sentido, el control de temperatura (mediante el encendido, apagado y regulación en aires acondicionados) e iluminación (apagado y encendido de luces) en las aulas, se efectúe de forma manual. En instituciones como la nuestra, estas tareas están asignadas a una persona, que debe realizar las configuraciones apropiadas en cada aula.

De la misma manera, se observa que cada profesor suele tener su propio requerimiento de iluminación en aula. Por ejemplo, hay profesores que requieren la máxima iluminación posible (todas las luces del aula encendidas), mientras que otros no utilizan las luces próximas al pizarrón ya que utilizan proyectores con presentaciones y/o filminas. Respecto de la climatización, la configuración de la temperatura claramente dependerá de la cantidad de personas, de los equipos en la sala y de las condiciones climáticas del día.

Entonces, es necesario encontrar un método para hacer eficiente el consumo energético, reducir los gastos de la institución y para atenuar el impacto ambiental. Es por ello que haciendo uso de los avances tecnológicos disponibles, se propone diseñar una solución al problema mediante la utilización de un Framework de IoT presentado en [2], una aplicación Web y hardware específico [3]. Esto permitirá a cada profesor configurar de forma independiente su perfil de iluminación y temperatura para el aula que va a utilizar, el cual será aplicado de forma automática mientras se encuentre dando clases. Como trabajos relacionados se pueden ver mencionarlos siguientes.

En [1] se establece una línea referida a costumbres y políticas de buen uso de la energía que no solamente promuevan el desarrollo, implementación y adaptación de software y hardware; sino que estas permitan un ahorro de dinero en la Universidad Nacional de Misiones. En [4] se describe el proceso de investigación para el desarrollo de un sistema de IoT, para promover un servicio de iluminación inteligente en un ambiente académico. El sistema orquesta una serie de sensores, sistemas de monitoreo y acciones controladas, basadas en el principio de hacer disponibles las funciones del sistema y el registro de consumo en tiempo real por medio de servicios web.

Asimismo, en [5] se propone el diseño y la implementación de un sistema automatizado inteligente para la conservación de la energía eléctrica. El mismo permite que los dispositivos eléctricos y los interruptores puedan ser controlados y monitoreados remotamente sin ninguna intervención humana.

En trabajo realizado en [6] se utiliza la tecnología basada en IoT para lograr la automatización de las aulas y propone un enfoque para controlar y manejar equipos eléctricos como ventiladores y luces basado en la presencia de personas.

Línea de Investigación

Para esta línea de investigación el objetivo principal es: determinar el Framework de IoT más adecuado en cuanto a métricas de software para diseñar una solución que contribuya a la eficiencia energética en organizaciones.

Como objetivos específicos se propusieron los siguientes: 1) Analizar bibliografía y trabajos existentes sobre IoT aplicada a la eficiencia energética. 2) Definir y aplicar un proceso de

selección de dos frameworks de IoT basado en las facilidades de implementación, seleccionados de entre cinco de los más usados. 3) Diseñar un prototipo de aplicación que contribuya a la eficiencia energética en ambientes organizacionales a implementarse en los dos frameworks seleccionados. 4) Elaborar diversos escenarios que podrían ocurrir y realizar pruebas en éstos de implementación de los prototipos con datos estáticos previamente elaborados. 5) Determinar el framework de IoT que mejor se adapte al escenario propuesto, considerando las métricas de software de los prototipos desarrollados.

Resultados

La solución tecnológica al problema tiene su núcleo en el framework de IoT, que cumple la función de efectuar la comunicación entre la aplicación web y el hardware (Figura 1) brindando un método para almacenar la información permitiendo la disponibilidad para la lectura y que pueda ser consumida en el momento requerido. Por su parte, la Aplicación Web define todo lo referido al comportamiento a seguir y luego, el hardware únicamente ejecuta las acciones ordenadas desde la Aplicación Web. Un ejemplo de su funcionamiento sería: la aplicación web lee la información del sensor de temperatura que envía el hardware, y dice en qué momento prender o apagar un aire acondicionado. Podemos establecer un esquema del sistema solución como se muestra en la Figura 1.

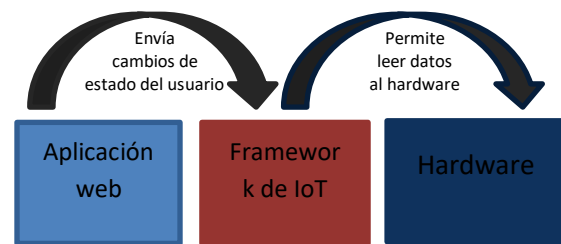


Figura 1. Esquema del prototipo

Cada flecha del gráfico anterior, representa un Request HTTP que realizan la aplicación web y el hardware respectivamente. En base a esto, tenemos cuatro situaciones:

1. por ejemplo, con el sensor La aplicación web envía datos de cambios de estados al framework de IoT. Estos cambios de estados se refieren a cambios en los perfiles según las preferencias de cada profesor. Por ejemplo, si el aire acondicionado en el aula 1, debe encenderse en determinado momento, el framework de IoT almacena estos datos, y los deja disponibles para ser consumidos. Cabe aclarar que el framework de IoT en este caso no maneja la lógica de cuándo debe encenderse un aire o no, simplemente recibe el dato, por ejemplo, "Aire1: 1", ya generado por la aplicación web, y lo deja disponible para ser consumido por quien lo requiera.
2. Los datos disponibles son consumidos por el hardware, que lee los datos mediante un Request HTTP y se limita a ejecutar la acción. Por ejemplo, si leyera que las luces del aula 1 deben estar encendidas y estas se encuentran apagadas, las encenderá.
3. Existen ciertos momentos donde el hardware precisa alimentar de información al sistema, de temperatura. En este caso, tomará la temperatura del ambiente y la enviará hacia el framework de IoT, para que ésta se encuentre disponible para quien la requiera.

4. Por último, existirán casos donde la Aplicación web requiere retroalimentarse de información del sistema, como ser cuando el hardware informa la temperatura actual de determinada aula. De esta manera, la Aplicación Web sabrá qué información producir y enviar al framework de IoT nuevamente.

A continuación se menciona la tecnología utilizada en este trabajo.

El hardware utilizado es una plataforma IoT de código abierto, el NodeMCU ESP8266 (Figura 2). Incluye el firmware que se ejecuta en el SoC (Sistema en chip) WiFi ESP8266 de Espressif Systems y el hardware que se basa en el módulo ESP-12 [7].

El ESP8266 es un chip de bajo costo WiFi con una pila TCP/IP completa y un microcontrolador. Puede ser programado utilizando el lenguaje de scripts Lua o con el IDE de Arduino en lenguaje C [8].

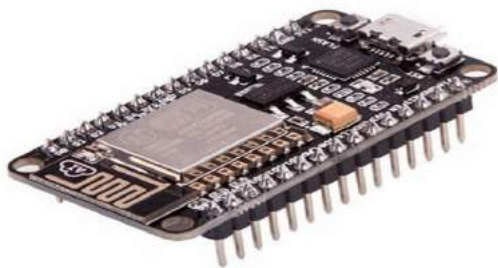


Figura 2. NodeMCU ESP8266.

Para la realización de la aplicación web se utilizaron distintos framework: Laravel, VueJs y Postman. Laravel es un framework PHP de código abierto para desarrollar aplicaciones y servicios web mediante la arquitectura de capas. VueJs es un framework progresivo de JavaScript para crear interfaces de usuario. Es una alternativa a frameworks como Angular o React [9]. Postman es una herramienta que

permite realizar peticiones HTTP hacia cualquier API REST, ya sea de terceros o propias para ir testeando el funcionamiento de la API por medio de su interfaz gráfica. Finalmente, el Ubidots [10] es una plataforma que permite construir, desarrollar, probar, aprender y explorar el futuro de las aplicaciones y soluciones conectadas a Internet. Entre algunas características existentes, se incluyen:

- Conectar hardware a la nube de Ubidots fácilmente con más de 200 bibliotecas, SDK y tutoriales comprobados por el usuario para guiar su integración a través de HTTP, MQTT, TCP, UDP o protocolos personalizados.
- Crear una API propia y analizar los datos en una función GET o POST HTTP de Node.js en la nube que extiende su solución de conexión más allá de la arquitectura y las capacidades de Ubidots.
- Analizar datos y agregar protocolos de terceros o nubes de dispositivos como Sigfox, Particle Cloud, LoRaWan y otras.
- Ampliar el monitoreo y el análisis de datos de las aplicaciones con integraciones de API como Weather Underground, Watson de IBM, Google Locations, Zapier, etc.

El stack de servicios se puede ver en la Figura 3.

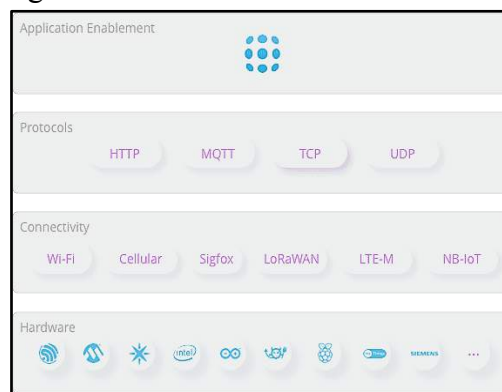


Figura 3. Stack de servicios de Ubidots

Formación de Recursos

Humanos

El equipo de trabajo se encuentra formado por cuatro investigadores, un Doctor en Tecnologías de la Información y Comunicación, un Doctor en Ingeniería; dos Doctorandos en Informática y ocho estudiantes en período de realización de trabajos finales de grado de Ingeniería en Informática de la UGD. Actualmente, el número de tesinas de grado aprobadas en el contexto de este proyecto, es de tres, y otras dos en proceso de desarrollo. En la UGD se lanzando recientemente la carrera de posgrado de Especialización en Gestión de Tecnologías de la Información y Comunicación y contara en su comité académico con un doctor de Universidad Nacional del Litoral.

Bibliografía

- [1] E.O Sosa, D.A Godoy, J. Benítez, and M.E. Sosa, "Eficiencia Energética y Ambientes Inteligentes. Investigación y Desarrollo Experimental en la UNaM," Posadas, Misiones, Argentina, 2015.
- [2] Diego Alberto Godoy, Hernán Bareiro, Fabian Fabret, and et.al., "Propuesta de métricas para comparación de Frameworks IoT," in *Workshop de Investigadores en Ciencias de la Computación.*, El Calafate, 2020.
- [3] Diego Alberto Godoy, Hernán Bareiro, Fabian Fabret, and et.al., "Análisis de componente de hardware para la utilización con Frameworks de IoT," in *Workshop de Investigadores en Ciencias de la Computación.*, El Calafate, 2020.
- [4] C.-L.Cárdenas-García,J.-A. Caicedo-Muñoz, andM.-A.Mendoza-Moreno C.-A.González-Amarillo, "Smart Lumini: A Smart Lighting System for Academic Environments Using IOT-Based Open-Source Hardware," *Revista Facultad de Ingeniería*, vol. 29, no. 54, 2020.
- [5] Anisha Gupta, Punit Gupta, and Jasmeet Chhabra, "IoT based power efficient system design using automation for classrooms," in *Third International Conference on Image Information Processing (ICIIP)*, Wagnaghat, India, 2015.
- [6] "An Approach Towards Building an IoT Based Smart Classroom," in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, India, R. Ani ; S. Krishna ; H. Akhil ; U Arun.
- [7] NodeMCU. (2020, Mar.) NodeMCU. [Online]. https://www.nodemcu.com/index_en.html
- [8] Germán Martín. (2019, May) NodeMCU con el IDE de Arduino. [Online]. <https://www.programarfacil.com/esp8266/como-programar-nodemcu-ide-arduino/>
- [9] Vuejs. (2020, May) Vuejs. [Online]. <https://vuejs.org/v2/guide/>
- [10] UBIDOTS. (2020, Aug.) ubidots.com. [Online]. <https://ubidots.com/docs/>

Simulación de enrutamiento BGP con GNS3

Álvaro Ignacio Gamarra, Ernesto Sánchez, Daniel Arias Figueroa.

Universidad Católica de Salta / Facultad de Ingeniería /
Universidad Nacional de Salta / C.I.D.I.A. (Centro de Investigación y Desarrollo en
Informática Aplicada) / Facultad de Ciencias Exactas.
Campus Castañares, 08105558227 / Av. Bolivia 5150, 3874255408
alvaroig@cidia.unsa.edu.ar, esanchez@cidia.unsa.edu.ar, daaf@cidia.unsa.edu.ar

Resumen

El presente trabajo expone una experiencia en la implementación de un Punto de Intercambio de Tráfico de Internet (IXP) de la Cámara Argentina de Internet (CABASE), para la ciudad de Salta, por iniciativa de la Universidad Católica de Salta y Gobierno de la Provincia de Salta. Las tareas se iniciaron en el año 2018 y una vez puesto en funcionamiento tal infraestructura, los esfuerzos se pusieron en invitar a proveedores locales e instituciones a formar parte de este punto de intercambio. Esto significó acercar los conocimientos teóricos y prácticos a fin de alcanzar tal objetivo.

Por todo esto desde el ámbito de la Universidad se propusieron actividades de transferencia mediante charlas y workshops utilizando un entorno de simulación de redes basado en GNS3 sobre el cual se expusieron los casos de uso y se impartieron capacitaciones a los responsables administradores de red los proveedores locales.

Palabras clave: IXP, BGP, GNS3.

Contexto

La línea de investigación se encuentra apoyada por el C.I.D.I.A. (Centro de Investigación y Desarrollo en Informática Aplicada) que depende de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta y por la Facultad de Ingeniería de la Universidad Católica de Salta, por lo tanto, se cuenta con toda la infraestructura disponible para esta investigación. El proyecto cuenta con el financiamiento del Consejo de Investigación de la Universidad Católica de Salta y el CIUNSa – Consejo de Investigación de la Universidad Nacional de Salta.

Introducción

Históricamente la Argentina ha contado con una buena capacidad de Banda Ancha internacional, permitiendo mantener relativamente bajos los precios de tránsito internacional. Sin embargo, la situación al interior del país era un tanto diferente, ya que los ISPs lejanos a las rutas de fibra óptica que conectan las grandes ciudades con Buenos Aires enfrentaban altos costos de tránsito interno, generando costos de acceso mucho más altos en ciudades pequeñas y alejadas que en las grandes ciudades [1].

A partir de la experiencia del IXP de Buenos Aires, CABASE comenzó a desarrollar iniciativas en el interior del país, en mercados más pequeños. Actualmente la Argentina cuenta con 32 IXPs en funcionamiento [2].

En particular en el año 2018, en la ciudad de Salta se crea el punto de intercambio de tráfico de Internet CABASE, por iniciativa de la Universidad Católica de Salta y Gobierno de la Provincia de Salta. y se invita a miembros del ecosistema local de internet a formar parte del mismo. Este primer paso, promovió que proveedores locales, vieran la necesidad de que se les acerquen los conocimientos para poder conectarse y hacer uso de las bondades de este IXP. Por ello desde la universidad se propusieron cursos de extensión y actividades de transferencia a fin de despejar dudas y poner en conocimiento aspectos técnicos para poder formar parte del punto de intercambio.

Puntos de Intercambio de Tráfico – IXPs

Consideramos pertinente acercar una definición sucinta de lo que es un punto de intercambio de tráfico. Según CABASE, *“Son espacios donde los proveedores de acceso a Internet, los organismos de gobierno, las entidades académicas, los grandes proveedores de contenido y otros actores del ecosistema de Internet, se interconectan entre sí para intercambiar el tráfico de sus redes de una manera más eficiente, con el objetivo de acercar el contenido al usuario final. Al estar más cerca, mejora sensiblemente la experiencia del usuario y la velocidad de navegación al reducir el tiempo de viaje del contenido (latencia).”* [3].

Tomando como premisa el mejorar la experiencia del usuario y velocidad de navegación, es que se diligenciaron las acciones para la creación del IXP en Salta.

Ventajas de los IXPs (estabilidad y resiliencia)

- Tráfico local se rutea localmente.
- Menor latencia para las aplicaciones.
- Menores costos.
- Posibilidad de implementar CDNs (Equipos que albergan contenidos de los generadores más importantes).
- El tráfico de una región/país/zona no es visto desde otras regiones/países, (en caso de requerirse).
- Introducción de nuevas tecnologías (IPv6, RPKI, etc.).
- Sentido de "comunidad". Compartir problemas, estrategias, acciones en común. [4]

Relevancia de los IXPs

- Mayor estabilidad y resiliencia al tráfico de Internet
 - Mejora la conectividad, redundancia, conexiones estables.
 - Facilidad para instalar servidores raíz de DNS (root servers).
- Seguridad
 - Tráfico más controlado, localizado, menos sujeto a inspección.
 - Posibilidad de intercambio de información de incidentes.
 - Acciones en común a fin de mitigar ataques.
- Entrenamientos y capacitaciones: grupo focalizado.

- Facilidad de introducción de pilotos de RPKI e IPv6.
- Posibilidad de investigación y estadísticas de tráfico.
- Crecimiento y desarrollo de Internet en la región. [4]

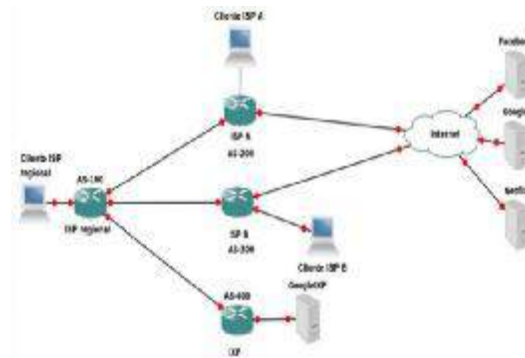
Entorno de prueba utilizado.

Al momento de plantearnos el escenario de pruebas a utilizar, nos encontramos con la complejidad que implica el despliegue de una topología donde intervengan ISPs tier 1 y tier 2, que implementen sesiones BGP, y el manejo de más de 800000 rutas que componen la tabla de enrutamiento de Internet. [5]

Toda la complejidad mencionada anteriormente se logró abstraer mediante el uso de la herramienta de software para simulación de redes GNS3. El escenario de pruebas se simuló sobre plataforma Windows 10 con 8gb de memoria RAM, y microprocesador Intel I7 de 6ta generación. La primera observación que pudimos realizar fue que para un uso eficiente de recurso de procesador se hizo necesario iniciar uno a uno los dispositivos simulados y no todos juntos, ya que el consumo de procesador es demasiado alto en el segundo caso.

Para la simulación de los routers de borde se utilizó la versión Mikrotik Cloud Hosted Router, la cual es de libre distribución, con la limitación de poder traficar hasta 1Mb por interfaz, lo cual para una simulación no implica ningún inconveniente. La simulación de servidores de contenidos se realizó sobre la versión de Linux denominada Tiny, que tiene como principal característica el consumo mínimo de recursos, tanto de memoria, procesador, como de espacio en disco.

A continuación, se describe la funcionalidad de cada uno de los dispositivos que forman parte del escenario de pruebas:



Dispositivo	Descripción
AS-XXX	Nº Sistema Autónomo
IXP	Punto de Intercambio
ISP X	Proveedor de Internet tier 2
ISP regional	Proveedor de Internet local
Facebook	Servidor de contenidos
Google	Servidor de contenidos
Netflix	Servidor de contenidos
Google IXP	Servidor de contenidos conectado a IXP
Nube Internet	Proveedor de Internet tier 1

El escenario de pruebas se completa con la configuración del protocolo BGP en los routers de borde necesario para: [6]

- Establecimiento de sesiones.
- Intercambio de rutas.
- Creación de filtros de seguridad
- Aplicación de atributos BGP para influenciar sobre el tráfico.

Líneas de Investigación, y Desarrollo

Los principales ejes temáticos que se están investigando son los siguientes:

- Puntos de Intercambio de tráfico.
- Border Gateway Protocol.
- Internet.
- GNS3.
- Simulación.

Resultados Obtenidos/Esperados

En el escenario propuesto se expone el caso de mejora en la experiencia de navegación para el Cliente conectado al ISP regional, en situaciones donde solicita contenidos del servidor Google IXP, ya que este desplegó tales contenidos en la infraestructura del IXP. Así mismo, se pudo probar y exponer todas las configuraciones antes mencionadas, donde se establecen sesiones BGP entre routers, se intercambian rutas y mediante filtros de seguridad, por ejemplo, el Cliente conectado al ISP A, no podría acceder a los contenidos desplegados en el servidor Google IXP.

Cabe destacar que todas las pruebas de ruteo y configuraciones, se realizaron sobre IPv4, porque al momento de la presentación alrededor de un 30% del contenido de Internet no es alcanzable con IPv6, en consecuencia, es una necesidad para los proveedores de internet seguir implementando IPv4 para poder alcanzar estos contenidos. Sin embargo, se realizaron pruebas sobre IPv6 y el caso expuesto se puede trasladar a tal alternativa sin ningún inconveniente.

Formación de Recursos Humanos

El grupo de investigación conformado se caracteriza por una constitución heterogénea de profesionales vinculados a

la informática. El director es Doctor en Ciencias Informáticas por la UNLP y el Codirector Master en Ingeniería de Software. Dentro de los investigadores se cuenta con un Magister en Redes de Datos, una Magister en Ingeniería de Software, un Ingeniero en Sistemas y un Técnico Universitario en Programación. También integran el grupo dos alumnos avanzados de la carrera Licenciatura en Análisis de Sistemas de la UNSa. En el transcurso del proyecto se tiene como objetivo consolidar la formación en investigación de los integrantes de menos antecedentes y también está contemplado que uno de los integrantes complete el cursado de la Maestría en Redes de Datos y la certificación Mikrotik.

Bibliografía

[1] Paper “Modelos e impactos de los puntos de intercambio de tráfico (IXPs) en América Latina y Caribe”. Alejandro Prince, Lucas Jolías. 44 JAIIO - STS 2015 - ISSN: 2451-7631.

[2] Poster “Red Nacional de IXPs 2020” Cámara Argentina de Internet. <https://www.cabase.org.ar/wp-content/uploads/2020/09/PosterCabase2020interactivoOK.pdf>

[3] Cámara Argentina de Internet. Puntos de intercambio de tráfico. <https://www.cabase.org.ar/servicios/#ixps>

[4] Puntos de Intercambio de tráfico: IXPs Importancia y Beneficios. Guillermo Cicileo. Estrategia de Seguridad y Estabilidad LACNIC.

[5] Tesis de Magister “Diseño e implementación de una solución de administración de tráfico de red basada en DNS y chequeos de disponibilidad”. Nicolas del Rio. <http://sedici.unlp.edu.ar/handle/10915/51098>.

[6] Simulación de redes con mikrotik y gns3 ejemplos de BGP. Ing. Álvaro Gamarra. Mikrotik MUM 2018. https://mum.mikrotik.com/2018/AR/agenda#431_6297.

Simulación en la enseñanza de redes de computadoras en el nivel universitario en tiempos de pandemia

Daniel Arias Figueroa, Ernesto Sánchez, Loraine Gimson, Álvaro Gamarra, Gustavo Gil, Nelson Rodríguez, Rodolfo Baspineiro

Facultad de Ingeniería de la Universidad Católica de Salta
C.I.D.I.A. (Centro de Investigación y Desarrollo de Informática Aplicada)
Facultad de Ciencias Exactas – Universidad Nacional de Salta
Facultad de Ciencias Exactas Físicas y Naturales – Universidad Nacional de San Juan
daaf@cidia.unsa.edu.ar, esanchez@cidia.unsa.edu.ar

Resumen

Nuestra línea de investigación, tiene como principal objetivo evaluar la influencia de la utilización de software de simulación en la enseñanza de contenidos de redes de computadoras en carreras de grado y cursos de postgrado a fin de contrastar con la enseñanza en laboratorios con equipo real Mikrotik. Para ello, nos planteamos un estudio aplicado, longitudinal y experimental que considera aspectos cuantitativos y cualitativos, que permitan determinar si los estudiantes que realizan prácticas en un entorno de simulación aprenden significativamente más que aquellos estudiantes que se capacitan en laboratorios con equipos reales de red.

Por otro lado, la educación superior en tiempos de pandemia tuvo que transformarse de manera brusca a un esquema virtual de emergencia para dar respuesta a las necesidades. En ese contexto, los objetivos del proyecto de investigación, al menos durante el período 2020, tuvieron que adaptarse a dichos cambios ya que no disponíamos de la posibilidad de realizar

prácticas en laboratorios presenciales de redes durante la cursada.

Palabras clave: protocolo TCP-IP, emulación, simulación, enseñanza, redes, Mikrotik, GNS3, Packet Tracer.

Contexto

La línea de investigación se encuentra apoyada por el C.I.D.I.A. (Centro de Investigación y Desarrollo de Informática Aplicada) que depende de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta y por la Facultad de Ingeniería de la Universidad Católica de Salta, por lo tanto se cuenta con toda la infraestructura disponible para esta investigación. El proyecto cuenta con el financiamiento del CIUNSa – Consejo de Investigación de la Universidad Nacional de Salta y el Consejo de Investigación de la Universidad Católica de Salta.

Introducción

En otras circunstancias podríamos haber encarado la escritura de este trabajo

abordando las áreas de competencia de nuestra tarea de investigación científica, pero hoy nos encontramos navegando en áreas de la educación a distancia.

Reconociendo el carácter excepcional del momento que atravesamos, este artículo propone una reflexión del complejo proceso de enseñanza y aprendizaje y de las prácticas pedagógicas que hemos adoptado, y de una experiencia de cambios o modificaciones para alcanzar los objetivos académicos de manera efectiva.

Los procesos de tutorización evolucionaron desde el rol tradicional del docente como eje de la comunicación y la información hacia planteamientos que fomentan el uso de las metodologías activas y el aprendizaje autónomo del alumno.

Uno de los objetivos claves en la enseñanza de las redes de computadoras en las carreras de sistemas es transmitir conceptos básicos y fundamentos a los estudiantes. Sin embargo, desde hace un tiempo, se viene investigando cómo facilitar la relación entre la realidad y las teorías y modelos, es decir, entre lo concreto y lo abstracto. Así, las computadoras personales, con la variedad de software que se ha desarrollado, tienen en sí mismas un gran potencial para mejorar el proceso de enseñanza y aprendizaje haciendo que:

- sea más interesante el aprendizaje,
- sea un aprendizaje activo y no pasivo en las aulas,
- sea al ritmo del estudiante en forma personalizada,
- los estudiantes estén más motivados y la educación sea permanente.

Líneas de Investigación, Desarrollo e Innovación

En los años previos a la pandemia, durante las cursadas de la asignatura Redes de Computadoras I, la tutorización siempre se realizó de manera presencial o semipresencial. La idea de la tutorización virtual surge en el año 2020 fruto de la coyuntura, y la imposibilidad del habitual cursado presencial.

Teniendo en cuenta, tanto la dedicación que requiere el seguimiento de las diferentes actividades durante la cursada, así como el hecho de que algunas actividades como la simulación puedan desarrollarse de manera autónoma por parte del alumno, hemos llegado a la conclusión de que tal vez el uso de las TICs aplicadas al aprendizaje y al conocimiento (TAC) permitiría llevar a cabo una mejor gestión del proceso de la que se realiza cuando todo el trabajo se desarrolla de manera presencial o semipresencial. Del mismo modo, hemos considerado que, a través de este método, se facilitarían estrategias que verdaderamente empoderarán al estudiante en el desarrollo y la toma de decisiones sobre las diferentes actividades.

Tomando esta decisión como punto de partida, el reto consistió en desarrollar una propuesta de tutorización que lograra aunar ambos intereses, La opción de plantear un proceso de tutorización virtual en una asignatura a priori presencial suponía una experiencia de innovación pensada como una primera aproximación hacia un método que permitiera al cuerpo docente implicarse en la tutorización del trabajo de los estudiantes, pero de una manera eficiente y sostenible.

Así pues, para la propuesta de un proceso de tutorización virtual, el método se basó en la puesta en práctica y consolidación de un

sistema de interacción virtual a través de la web 2.0, basado en sistemas de comunicación formal e informal mediante el uso de cuatro canales de contacto:

1- Comunicación formal:

- Moodle® como plataforma de intercambio de recursos (lecciones, cuestionarios, wikis, etc.)
- Foros de Moodle® como instrumento de vinculación entre docente, estudiantes para debates y evacuación de dudas
- Clases interactivas en línea con la plataforma Zoom® para presentación de los temas más importantes del contenido.
- Clases interactivas en línea con el módulo de Moodle BigBlueButton como una alternativa a la disponibilidad de la plataforma Zoom.

2- Comunicación informal:

- Servicios de mensajería instantánea con WhatsApp®. Grupo específico con alumnos y docentes para coordinación de las diferentes actividades.
- Chat propio de la plataforma Moodle® para consultas.

Durante el curso 2019, se había optado por no utilizar herramientas por fuera de la plataforma Moodle®, pero en la cursada 2020, vimos la necesidad de gestionar el grupo de estudiantes a través de WhatsApp®, dado que el volumen de alumnos fue más alto en este curso. Como puede verse en la figura 1, en la plataforma Moodle® se publicó el programa analítico vigente, la bibliografía básica a seguir, un cronograma de tiempo dedicado a cada tema, y cuestiones generales sobre la manera de estudiar con el material disponible en la plataforma. Asimismo, la

función básica de la plataforma Moodle® fue la de repositorio de todo el material necesario para el desarrollo progresivo del trabajo propuesto.



Figura1: vista del curso dentro de Moodle ®

Para cada unidad del programa, el alumno disponía del siguiente material de estudio:

- Una presentación en PowerPoint, la misma constituyó un guía de lectura para el capítulo correspondiente del libro de base propuesto. Esta guía contiene los temas que abarca la asignatura ya que no todos los temas del libro son analizados como parte de la asignatura.
- Una Guía de Trabajos Prácticos en formato pdf para su desarrollo teórico, como un conjunto de preguntas y ejercicios propuestos a desarrollar.
- El Libro de la Cátedra “Redes de Computadoras I con Packet Tracer”, que contiene no solo una guía de ejercicios para resolución con el simulador de redes Packet Tracer de Cisco, sino que también propone una solución posible, material autocontenido pensado para la autoaprendizaje por parte del alumno. El objetivo de este material es llevar al plano de lo concreto, lo abstracto que puede significar el estudio de redes de datos. Si bien el simulador es limitado en algunos aspectos, consideramos que es suficiente

para asimilar los conceptos y fundamentos sobre redes de computadoras.

Durante todo el cursado de la asignatura, se dieron instancias de evaluación continua y real de seguimiento del proceso de enseñanza-aprendizaje, con los aciertos y errores, fortalezas y debilidades, por cada tema abordado. Este proceso se desarrolló teniendo en cuenta el conocimiento, las habilidades, las actitudes y el desempeño del alumno. Se resaltan los aspectos positivos de su desempeño en general, donde más del 78% de los estudiantes regularizaron la asignatura.

Al finalizar el primer cuatrimestre del año 2020, se realizó una encuesta anónima y de carácter no obligatorio a los alumnos que cursaron la asignatura a fin de sondear las principales acciones, plataformas y estrategias pedagógicas aplicadas en la asignatura, relevar resultados sobre la disponibilidad de recursos tecnológicos, percepción sobre la metodología empleada y cuestiones más vinculadas al acompañamiento por parte de los docentes en el marco de esta pandemia.

Las evaluaciones consistieron en cuestionarios cortos, con tiempos acotados de resolución y estuvieron compuestas por diferentes tipos de preguntas: verdadero o falso, opción múltiple, abiertas, etc. De esta forma, se buscó construir conocimiento de una forma creativa y dinámica, realizando una evaluación continua, completa, adecuada y elaborada para que el proceso enseñanza logre consolidar el aprendizaje de cada estudiante.

Resultados

A partir del trabajo con los alumnos durante el primer cuatrimestre, al finalizar las

actividades se realizó una encuesta dividida en cuatro partes. En la primera parte de la encuesta se indagó respecto del material disponible para la cursada, e-book con práctica para la simulación, guías prácticas para resolución en papel, guía de lectura de la bibliografía de base, etc. En la segunda parte de la encuesta se preguntó sobre la comunicación formal e informal utilizada, plataformas Moodle, Zoom, BigBlueButton®, grupo de WhatsApp® y chat de Moodle®. En la tercera parte, sobre la tutorización y evaluación, estructura de trabajo propuesta, comunicación con el equipo docente, etc. En la cuarta parte se indagó sobre la herramienta de software de simulación utilizada, su facilidad para configurar una topología, configurar dispositivos, verificar la funcionalidad, realizar el seguimiento de los eventos y otros.

Dentro de la encuesta se pedía contestar en una escala de 1 al 5 donde los valores equivalían a los siguientes (ver tabla 1):

5	En gran medida
4	En buena medida
3	Lo suficiente
2	Poco
1	Prácticamente nada

Tabla 1: Escala de la encuesta

Análisis estadístico de los datos

Para el conjunto de datos analizado, la media aritmética supera, en todos los casos, el valor medio de la escala (3.00), lo que confirma la validez del contenido de todos los ítems incluidos en la encuesta. El coeficiente Alfa de Cronbach calculado es de 0,82 que supera el valor de 0,7, por lo que se puede afirmar el grado de fiabilidad del cuestionario y por lo tanto una mayor polarización de

respuestas confiables por parte de los estudiantes. Sin entrar en detalles estadístico de los datos, para cada una de las secciones de la encuesta realizada, analizando los valores medios (ver figuras 2 y 3) podemos mencionar que:

- 1- Sobre el material disponible para la cursada, en promedio consideran que fue adecuado en buena medida.
- 2- Sobre los Canales para la comunicación formal e informal utilizada en promedio consideran que fue adecuado en buena medida salvo para el chat de Moodle® donde un más de un 40% no considera que sea muy relevante (ver Figura 3).
- 3- Sobre la tutorización y evaluación, estructura de trabajo propuesta, comunicación con el equipo docente, en promedio consideran que fue adecuado en buena medida
- 4- Sobre la herramienta de software de simulación utilizada, su facilidad para configurar una topología, configurar dispositivos, verificar la funcionalidad, realizar el seguimiento de los eventos y otros también se obtuvo una media aritmética que manifiesta que fue adecuado en buena medida.

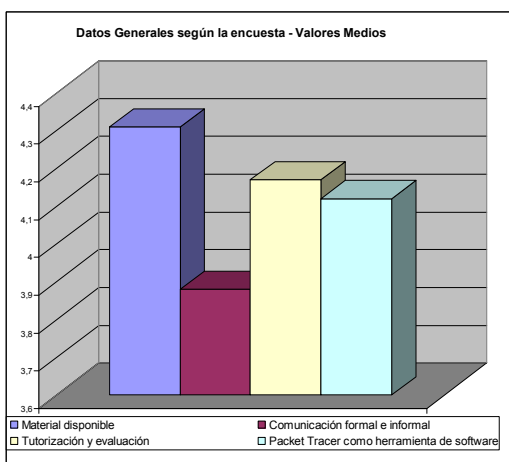


Figura2: Valores medios sobre datos de la encuesta.

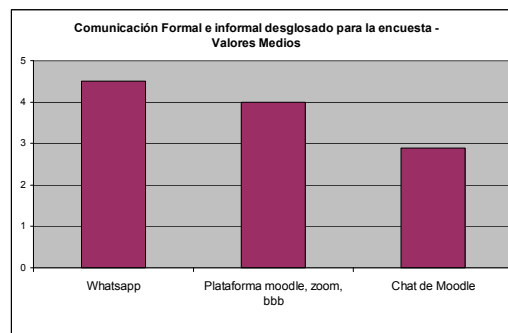


Figura3: Valores medios sobre comunicación formal e informal de la encuesta.

A continuación se enumeran textualmente algunas opiniones de los estudiantes sobre la experiencia realizada:

- *La virtualidad en mi caso conlleva mucha más disciplina a la hora de estudiar, ya que al no tener unas horas específicas para realizar las prácticas como en lo presencial, me retraso un poco.*
- *A pesar de las condiciones en la que nos encontramos, la vocación y las ganas de enseñar estuvieron siempre presentes.*
- *El uso de las herramientas proporcionadas por la cátedra fue muy bueno, principalmente debido al soporte virtual que tuvimos en cada instante del cursado.*
- *Lo que resulto muy exigente en está cursada y me gustó, fue la entrega de los trabajos prácticos. Si bien antes se discutían las respuestas en clase práctica, las notas que tomaba no estaban del todo estructuradas, por lo que al tener que entregarlos, me esforzaba mas en mis respuestas para que sean lo más claras posibles.*

El proceso vivenciado durante esta particular cursada, nos permitió descubrir nuevas alternativas metodológicas. Vimos que es necesario mejorar la interactividad

entre docentes y alumnos, tanto en herramientas tecnológicas como en el manejo de los recursos.

Si bien consideramos de mucha utilidad el uso de plataformas como Zoom y BigBlueButton, observamos, en nuestro contexto en particular, que no son adecuadas para que el alumno se exprese y realice consultas interrumpiendo la exposición y propiciándose de ese modo una diálogo entre los asistentes, como si ocurre en una instancia presencial. Esto no lleva a reflexionar y replantearnos si no es conveniente subir videos pregrabados en la plataforma Youtube® por ejemplo, con las clases teóricas y dejar las plataformas interactivas para instancias de consultas y esclarecimiento de dudas una vez que el estudiante ya realizó como actividad el seguimiento del video. De esta manera, el recurso escaso en nuestra universidad (plataforma Zoom®) se utilizaría de manera mucho más concreta y por menos espacio de tiempo, incluso con la versión de prueba de 40 minutos máximos se podría resolver.

Respecto de la utilización de la herramienta de simulación, cabe decir que en la cursada 2019 habíamos migrado a GNS3 y Mikrotik®, donde el alumno realizaba prácticas con simulación y en laboratorios con equipos reales Mikrotik®. En la cursada 2020 tuvimos que volver a utilizar Packet Tracer® debido a la sencillez de instalación y los escasos recursos necesarios para su funcionamiento. De todas maneras, advertimos que la herramienta fue bien valorada por los alumnos y estamos convencidos que acerca el estudio de las redes de lo abstracto a lo concreto para comprender conceptos y fundamentos básicos.

Formación de Recursos Humanos

El grupo de investigación conformado se caracteriza por una constitución heterogénea de profesionales vinculados a la informática. El Director es Doctor en Ciencias Informáticas por la UNLP y el Codirector Master en Ingeniería de Software. Dentro de los investigadores se cuenta con un Magister en Redes de Datos, una Magister en Ingeniería de Software, un Ingeniero en Sistemas y un Técnico Universitario en Programación. También integran el grupo dos alumnos avanzados de la carrera Licenciatura en Análisis de Sistemas de la UNSa. En el transcurso del proyecto se tiene como objetivo consolidar la formación en investigación de los integrantes de menos antecedentes y también está contemplado que uno de los integrantes complete el cursado de la Maestría en Redes de Datos y la certificación Mikrotik®.

Referencias

- SAKAR, N. I. (2006) Teaching TCP/IP Networking Using Practical Laboratory Exercises, *International Journal of Information and Communication Technology Education*, Vol. 2, No. 4, pp. 39-50.
- GOLDSTEIN, G., M LEISTEN, S, STARK, K., & TICKLE, A. (2005) Using a Network Simulation Tool to Engage Students in Active Learning Enhances Their Understanding of Complex Data Communications Concepts, *Proceedings of the 7th Australasian conference on Computing Education*, pp. 223-228.
- JAVIDI, G. & SHEYBANI, E. (2008) Content-Based Computer Simulation of a Networking Course: An Assessment, *Journal of Computers*, Vol. 3, No. 3, pp. 64-72.
- CAMERON, B. (2003): Effectiveness of simulation in a hybrid online networking

course. *Quarterly Review of Distance Education*, 4(1), 51.

GATTO, D. (1993): The use of interactive computer simulations in training. *Australian Journal of Educational Technology*, 9(2), 144-156.

YAVERBAUM, G., & NADARAJAN, U. (1996): Learning basic concepts of telecommunications: an experiment in multimedia and learning. *Computers & Education*, 26(4), 215-224.

ZHU, S. Y. (2011). Teaching Computer Networks through Network Simulation Programs. Faculty of Business, Computing and Law – School of Computing. University of Derby. *Learning Teaching & Assessment Conference*.

KUROSE, J.F. & ROSS, K.W. (2015). Computer Networking: A Top-Down Approach. 6th Edition. *Pearson Education*. ISBN: 9780132856201.

AVILA BLAS, Orlando José (2003). Probabilidad y estadística inferencial: teoría y aplicaciones. ISBN: 978-987-9381-23-6. *Editorial: Univ. de Salta*.

GLIEM, J & GLIEM, R. (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. Conference in Adult, Continuing, and Community Education. Midwest Research to Practice.

CÁMARA, ALZUGARAY (2011). Trabajos Prácticos, Métodos de Simulación y aprendizaje significativo. La Tecnología Educativa al servicio de la Educación Tecnológica.

ARIAS FIGUEROA, D. SÁNCHEZ, E. Colaboración de BASPINEIRO (2018). Redes de Computadoras II con Packet Tracer. Editorial de la Universidad Nacional de Salta – EUNSa – ISBN 978-987-633-527-0; 1a ed. - Salta - E-Book - CDD 004.678.

ARIAS FIGUEROA, D. (2015). Redes de Computadoras I con Packet Tracer. Editorial de la Universidad Nacional de Salta – EUNSa – ISBN 978-987-633-132-6; 1a ed. - Salta - E-Book - CDD 004.68.

TECNOLOGÍA DE DISPOSITIVOS DE LÓGICA RECONFIGURABLE APLICADA EN LA IMPLEMENTACIÓN SEGURA DE SISTEMAS DE IOT

Oswaldo L. Marianetti ⁽¹⁾⁽²⁾, Pablo D. Godoy ⁽¹⁾, Ernesto E. Chediak ⁽¹⁾, Daniel S. Fontana ⁽¹⁾, Carlos García Garino ⁽¹⁾

⁽¹⁾Universidad Nacional de Cuyo. Facultad de Ingeniería. ⁽²⁾Universidad de Mendoza. Facultad de Ingeniería.

olmarianetti@gmail.com, pablodgodoy@gmail.com, ernestochediack@gmail.com, danielsantiagofontana@gmail.com, cgarcia@itu.uncu.edu.ar

RESUMEN

Internet de las Cosas (IoT) presenta un escenario en el cual miles de millones de dispositivos se encuentran interconectados y distribuidos por casi cualquier lugar, desde el cuerpo de un ser humano hasta las áreas más remotas del planeta. Gracias al abaratamiento de los costos de los microprocesadores, al aumento de las capacidades de cómputo y a las nuevas tecnologías inalámbricas, se prevé en 2021 más de 50.000 millones de dispositivos conectados. Los ataques informáticos en general, pueden robar o modificar datos importantes, hacer caer servicios críticos online o conseguir dinero de forma ilícita. En cambio, en un contexto de IoT, además de todas estas acciones, existen posibilidades de hacer daño físico a personas a distancia y/o manipular infraestructuras críticas.

Este proyecto propone demostrar que, a partir del paradigma de Computación en la Niebla, los dispositivos lógicos programables como las FPGA (Field Programmable Logic Array), con capacidades de reconfiguración y gran potencia computacional, más la posibilidad de adaptar el procesamiento al tipo de información que presente en estas aplicaciones (eventos, imágenes, etc.), se pueden considerar una alternativa de desarrollo frente a las problemáticas que presenta la implementación segura de sistemas de IoT.

Palabras clave: IOT, seguridad, FPGA, reconfigurable.

CONTEXTO

Desde la Secretaría de Investigación, Internacionales y Posgrado se desarrollan convocatorias a proyectos bienales que tienen por finalidad la promoción de la investigación científica y tecnológica.

Este proyecto está inserto en la convocatoria Proyectos de investigación SIIP 2019, http://www.uncuyo.edu.ar/ciencia_tecnica_y_posgrado/proyectossiip2019.

1. INTRODUCCIÓN

Internet de las Cosas (IoT) o la evolución a IOE (M2M, P2M Y M2P) presenta un escenario en el cual miles de millones de dispositivos se encuentran interconectados y distribuidos por casi cualquier lugar. Podemos imaginar un mundo donde todos los objetos tienen identidad propia, desde una cafetera, un auto que se prepara para arrancar o un servicio médico capaz de monitorizar la salud de un individuo salud en forma remota. Internet de las cosas facilitará el despliegue de ciudades inteligentes, Industria 4.0, vigilancia inteligente y grandes cambios en la gestión de infraestructuras, agricultura, medicina y casi cualquier otro sector económico.

En el marco de lo que se denomina la nueva Revolución Industrial o Industria 4.0, se hace referencia al concepto de “smart factory”. En

el concepto de la fábrica inteligente, una planta industrial completamente conectada, la cual podría generar diariamente varios centenares de gigabytes, volúmenes enormes de información que dificultarían su procesamiento en la nube o en de forma centralizada. Estos requerimientos permiten replantear la infraestructura actual de la nube tal como la conocemos.

Una alternativa que propone posibles soluciones a la problemática que acompaña la implantación total de la IoT, se conoce como computación en la niebla (fog computing). [1] Fog computing es el nombre de una tecnología cloud por la cual los datos que generan los dispositivos no se cargan directamente en la nube, sino que se preparan primero en centros de datos descentralizados más pequeños. El concepto engloba a una red que se extiende desde sus propios límites, que es donde los terminales o sensores generan los datos, hasta el destino central de los datos en la nube pública o en un centro de datos o nube privada.

El objetivo de la computación en la niebla es acortar las vías de comunicación entre la nube y los dispositivos y reducir el caudal de datos en redes externas. Los nodos cumplen el rol de capa intermedia en la red en la que se decide qué datos se procesan localmente y cuáles se envían a la nube o a un centro de datos para ser analizados o procesados. Las tres capas (layer) de una infraestructura de computación en la niebla son:

- Edge layer: la capa del borde comprende a todos los dispositivos inteligentes (dispositivos en el borde de la red) de una arquitectura de IoT. Los datos que se generan en esta capa se procesan en el mismo terminal o se envían a un servidor (nodo fog) en la capa niebla.
- Fog layer: la capa en la niebla está compuesta por una serie de servidores de gran rendimiento que reciben los datos de la primera capa, los preparan y los envían a la nube si es necesario.
- Cloud layer: la capa en la nube constituye el punto final de una arquitectura de fog computing.

En la fog computing los recursos para el almacenamiento y la preparación de los datos abandonan la nube pública o el centro de datos y se distribuyen en una capa intermedia en la red por medio de nodos fog o unidades de preprocesamiento.[2]

El OpenFog Consortium, que reúne a algunas de las empresas e instituciones académicas más innovadoras del sector con el objetivo de desarrollar un marco común de desarrollo de la tecnología, está trabajando en una arquitectura de referencia para sistemas de fog computing.

La computación en la niebla se diferencia de la tecnología cloud, sobre todo, por el lugar donde se accede a los recursos y se procesan los datos. La computación en la nube suele basarse en centros de datos centralizados. Aquí, los servidores en el backend son los que suministran recursos como la potencia de procesamiento y la memoria, que utilizan los clientes a través de la red. La comunicación tiene lugar entre dos o más terminales siempre mediante un servidor en un segundo plano.

Con conceptos como el de la fábrica inteligente esta arquitectura tiene limitaciones, puesto que en ella hay un gran número de dispositivos que están intercambiando datos constantemente. Apoyándose en el procesamiento de los datos cerca de la fuente, la computación en la niebla logra reducir el tráfico de datos.

Pero la computación en la nube no solo se ve saturada por el tráfico de datos que generan las grandes infraestructuras de IoT, sino también por la latencia, porque el procesamiento centralizado de los datos implica depender de las rutas de transferencia y esto ocasiona siempre cierto desfase. Los dispositivos y los sensores han de estar siempre en contacto con el servidor en el centro de datos para poder comunicarse y esperar tanto el tratamiento externo de la petición como la propia respuesta, de modo que este tiempo de latencia se convierte en un problema en procesos de fabricación apoyados en IoT que necesitan el procesamiento inmediato de la información

para que las máquinas puedan reaccionar inmediatamente a cualquier incidente.

La fábrica inteligente no es el único campo de aplicación en que la computación en la niebla puede aportar sus ventajas. Otros proyectos como los autos semiautónomos o completamente autónomos, o la ciudad inteligente, integrada por redes inteligentes de suministro, también necesitan que los datos se analicen en el momento. Un coche inteligente, por ejemplo, está constantemente recogiendo datos sobre el entorno, las condiciones de conducción y la situación del tráfico que se han de evaluar sin latencia para que pueda reaccionar ante cualquier imprevisto de la forma correcta. La computación en la niebla permite procesar los datos del vehículo tanto en el automóvil mismo como en el proveedor del servicio.

IoT augura un futuro muy prometedor e interesante, pero respecto a lo relacionado a la seguridad, la evolución no presenta el mismo nivel de progreso. [3]

Existen varias organizaciones y gobiernos que están promoviendo y obligando a que la seguridad esté presente desde el diseño de los dispositivos, forzando a los fabricantes a que tengan en cuenta todos los aspectos. (Congreso de los EEUU, Department of Homeland Security y National Science Foundation)

Los principales problemas de seguridad del Internet de las Cosas son:

1. La heterogeneidad de tecnologías. Son necesarios conversiones de protocolos y hacer compatibles los mecanismos de seguridad implementados por distintos fabricantes.
2. Los dispositivos IoT no cuentan en la actualidad con la capacidad de computación que requieren las medidas de seguridad que se adoptan en otras plataformas.
3. Las comunicaciones de toda la tecnología IoT se soporta casi totalmente en el aire, es decir comunicaciones inalámbricas. Esta es la tecnología que más tipos de ataque puede sufrir, cuando precisamente el intercambio de información de los dispositivos IoT son bastante predecibles y su arquitectura y formato no son fáciles de cambiar.

Las soluciones tradicionales donde la seguridad se aplica como una ocurrencia posterior y como un parche contra los ataques conocidos son insuficientes. Se requiere una visión de seguridad desde el diseño, donde las amenazas se abordan de forma proactiva. [4]

La tecnología de lógica reconfigurable es un prestador de soluciones eficientes, escalables y sostenibles. La gran capacidad computacional de las FPGA y la posibilidad de adaptar el procesamiento a distintos tipos de información, y las propiedades que se citan a continuación, también ofrecen una respuesta a las necesidades planteadas:

Gestión de la capacidad y carga dinámica: las FPGAs permiten adaptar en tiempo de ejecución los recursos disponibles para una determinada tarea sin complejas infraestructuras

Seguridad: debido a la naturaleza del hardware reconfigurable, el sistema es más resistente a ataques. Además, pueden añadirse sistemas de encriptación más potentes que se ejecuten en hardware sin afectar al funcionamiento de la aplicación. [5]

Simplificación de la infraestructura software: mejora la mantenibilidad y el costo operativo de la plataforma. Toda la funcionalidad del nodo fog (concentrador o puerta de enlace) IoT puede implementarse en una FPGA.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes del tema de investigación del proyecto son:

- a) Identificar los principales problemas de seguridad en los dispositivos utilizados en plataformas de IoT.
- b) Investigar las características de los dispositivos lógicos programables FPGA que cumplen con los requerimientos de capacidad de cómputo.
- c) Demostrar experimentalmente la vulnerabilidad de dispositivos IOT. (En particular en redes de sensores inalámbricos)
- d) Desarrollar arquitecturas de sistemas programables en un chip (SOPC) con

dispositivos reconfigurables, FPGAs, optimizadas para operar como nodo de una red de sensores inalámbricos y/o como nodo de la capa de borde en aplicaciones de sistemas de IoT y que pueda interactuar con concentradores o nodos de la capa fog, utilizando tecnologías y herramientas de disponibles y accesibles en el contexto local.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos son:

1-Diseño de un prototipo de procesador soft-core para aplicaciones en nodos de WSN. [6] Unos de los problemas a resolver en las WSN es la optimización del consumo de energía disponible en los nodos, de modo de lograr el máximo tiempo de vida útil de la red, optimizar los recursos de procesamiento y adecuar los nodos a la topología dinámica de la red.

Las WSN utilizan en su despliegue nodos con procesadores o microcontroladores de propósito general. Como alternativa a estos, se implementó un prototipo de procesador soft-core. Esto optimiza la arquitectura del procesador para que se pueda adaptar a las necesidades de las aplicaciones de las redes de sensores. Por lo tanto, el procesador puede contar con sólo aquellos componentes de hardware requeridos por la aplicación particular y así reducir el consumo de energía del procesador. Además, el procesador soft-core puede ser integrado en un chip con los otros componentes de hardware necesarios para desarrollar el nodo sensor.

Se desarrolló la descripción mediante código VHDL del prototipo de procesador soft-core. Se concretó la síntesis de este diseño y se programó en un dispositivo FPGA. El funcionamiento de la implementación se verificó con entidades de test con estímulos. Estos correspondían a carga de registros y accesos a memoria, como así también operaciones aritméticas.

El prototipo ofrece en principio una mayor velocidad de procesamiento que un microcontrolador de propósito general con

una implementación más simple (se comparó con 4 microcontroladores comerciales).

Deben considerarse mejoras en la arquitectura, como instrucciones de manejo de bits e incorporar módulos de interface para comunicación con sensores.

2- Vulnerabilidad de los dispositivos utilizados en aplicaciones de IoT. [7]

Se desarrolló un equipo de control de condiciones ambientales. Este sistema se implementó utilizando dispositivos de uso frecuente en hogares y oficinas, que cuentan con sensores, los que podrían ser vulnerables a un uso inadecuado. El producto final tiene por objetivo la medición de los siguientes parámetros ambientales: temperatura, humedad, iluminación, nivel de ruido, y presencia.

El funcionamiento del equipo responde a la consigna de medir los parámetros ambientales y transmitirlos mediante WIFI a una plataforma en la Nube utilizando un protocolo UDP. El firmware embebido en el dispositivo, tiene una rutina principal que realiza una lectura de los sensores conectados por una interface I2C, realizando una actualización de los valores cada 30 segundos. Por otra parte, en el desarrollo del sistema se programó una subrutina embebida en el código que realiza la lectura un puerto analógico, conectado a la salida analógica del micrófono disponible en la placa de sensores. A los usuarios del sistema solo se les da acceso a la parte de la plataforma que presenta los parámetros ambientales. Quienes desarrollaron el equipo pueden tener acceso pleno al sonido de la habitación en donde se encuentra instalado el sensor.

El experimento desarrollado demuestra la vulnerabilidad de los elementos que en la actualidad se utilizan en las aplicaciones de IoT, sino se considera la problemática de la seguridad desde el diseño del sistema.

3-Sistema embebido basado en soft_processors. [8]

Se ha diseñado una arquitectura de un sistema embebido basada en soft_processors. Esta arquitectura presenta la misma funcionalidad de los sistemas comerciales con componentes

de seguridad en su diseño, pero puede incorporar las características de la naturaleza del hardware reconfigurable, para que el sistema sea más resistente a ataques. En la arquitectura basada en FPGA sus unidades funcionales (memorias, puertos, controladores, temporizadores, etc.) son reconfigurables y adaptables a nuevos requerimientos, incluso en forma remota.[9] [10]

Para el desarrollo de la arquitectura del sistema embebido basada en soft_processors, se ha utilizado el entorno de desarrollo Quartus II (versiones 13.1 web edition y Quartus Prime Lite Edition 17.0) La herramienta QSYS de estos entornos para la generación del SOPC (sistema programable en el chip) y el entorno NIOS II software build tool for Eclipse para la programación del soft_processor NIOS II/e.

En esta implementación se debe concluir con el desarrollo del módulo de encriptación del SOPC.

Resultados esperados:

Verificar el funcionamiento del SOPC desarrollado como nodo WSN y como nodo de la capa de borde.

Desarrollar una arquitectura de nodo Fog basada en diseño SOPC sobre tecnología FPGA.

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo actual es: Dos doctores: Dr. García Garino y Dr. Godoy. Un magister: Mgt Marianetti (actual doctorando) y dos ingenieros: Ing. Chediak e Ing. Fontana.

Los conocimientos y experiencias resultantes, debidamente mediados, podrán ser transferidos directamente a los alumnos de la carrera Licenciatura en Ciencias de la Computación, que se dicta en la Facultad de Ingeniería de la Universidad Nacional de Cuyo, ya que todos los integrantes de este proyecto son profesores de diferentes materias de dicha carrera (Arquitectura de

Computadoras, Arquitecturas Distribuidas, Redes), y los temas sobre los que se va a investigar, articulan con los programas de estudios de varias asignaturas de dicha carrera, cuyo cursado comenzó en el año 2017.

En el presente hay una tesis posgrado en curso.

5. BIBLIOGRAFÍA

- [1] Baktyan, A. A., & Zahary, A. T. (2018). A Review on Cloud and Fog Computing Integration for IoT: Platforms Perspective. EAI Endorsed Transactions on Internet of Things.
- [2] Tatiana Delgado Fernández. Un acercamiento a fog computing (Computación en la niebla). <https://sistemas.acis.org.co/index.php/sistemas/articulo/download/121/94>.
- [3] IoT Security White Paper 2018 – Huawei. https://www.huawei.com/.../iot/.../iot_security_white_paper_2018
- [4] Restuccia, F., D’Oro, S., & Melodia, T. (2018). Securing the Internet of Things: New Perspectives and Research Challenges. <https://arxiv.org/abs/1803.05022>.
- [5] A Lattice Semiconductor White Paper IoT Sensor Connectivity and Processing with Ultra-Low Power, Small Form-Factor FPGAs. 2018.
- [6] O. Marianetti, A. Iglesias, L. Arce. Diseño de un prototipo de procesador soft-core para aplicaciones en nodos de WSN. <https://doi.org/10.18682/cyt.v1i17>. Online ISSN 2344-9217 | Print ISSN 1850-0870. Universidad de Palermo. Facultad de Ingeniería
- [7] O. Marianetti, Pablo D. Godoy, E. Chediak, Daniel S. Fontana. Vulnerabilidad de los dispositivos utilizados en aplicaciones de IoT. CASE 2019. Libro de Trabajos. P. 135. ISBN 978-987-46297-6-0)
- [8] O. Marianetti, Pablo D. Godoy, E. Chediak, Daniel S. Fontana. La tecnología de lógica reconfigurable como alternativa en la solución a los problemas de seguridad en Internet de las Cosas. XXVI Jornadas de investigación: “Avances y desafíos de la ciencia en pandemia”. 2020. UNCUIYO.
- [9] Saar Drimer, Markus G. Kuhn. A Protocol for Secure Remote Updates of FPGA

Configurations.. Computer Laboratory, University of Cambridge. 2018.

[10] Lei Zhou, Qingxiang Liu, Bangji Wang, Peixin Yang, Xiangqiang Li and Jianqiong Zhang. Remote System Update for System on Programmable Chip Based on Controller Area Network. School of Physical Science and Technology, Southwest Jiaotong University, 2017.

Tecnologías de IoT y aprendizaje automático para la solución de problemas en el medio productivo y el cuidado del medioambiente

Jorge Osio^{1,2}, Juan Salvatore¹, Mauro Salina¹, Diego Montezanti^{1,3}, Nicole Denon¹, Santiago Doti¹, Lucas Olivera, Matías Busum Fradera¹, Daniel Alonso¹, Marcelo Cappelletti^{1,2}, Diego Encinas^{1,3}, Martín Morales^{1,4}

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 LIDI - Fac. de Informática -UNLP.

⁴ Centro UTN CODAPLI-FRLP, La Plata (1900), Argentina.

{josio, mcappelletti, jsalvatore, dalonso, dencinas, martin.morales}@unaj.edu.ar

Resumen

El presente proyecto se basa en la utilización de internet de las cosas (IoT) como herramienta fundamental para proveer soluciones a problemáticas de interés social, como lo es el cuidado del medioambiente y la innovación en el sector productivo, focalizando la investigación en las técnicas de visión por computadora y aprendizaje automático.

Entre los temas de investigación que se desarrollarán, se incluye el diseño e implementación de técnicas de visión por computadora con el objeto de agregar funcionalidades a un robot móvil, de manera de proveer autonomía en ambientes con obstáculos, con el agregado de control y supervisión remota mediante IoT. En esta línea, también se implementarán técnicas de visión por computadora para la clasificación de residuos reciclables mediante algoritmos de aprendizaje automático. Esta última aplicación se suma a las líneas relacionadas con el cuidado del medioambiente que se desarrollaron en el proyecto anterior.

En esta propuesta se continuará con las líneas del proyecto anterior de procesamiento digital de imágenes con el agregado de técnicas de aprendizaje automático. Teniendo en cuenta que las técnicas de procesamiento de imágenes aplicadas a visión por computadora requieren alto poder de cómputo, se considera necesario determinar la tolerancia a fallos del sistema de

procesamiento utilizado, para asegurar la correcta ejecución de dichos algoritmos.

En relación a la detección de fallos, se propone el perfeccionamiento de la metodología desarrollada de tolerancia a fallos transitorios característicos de las arquitecturas multicore, con el objeto de aplicarlo al sistema de visión por computadora.

Palabras clave: *Aprendizaje automático, procesamiento de imágenes, IoT, Visión por computadora, tolerancia a fallos.*

Contexto

Las líneas de I/D presentadas en este trabajo están incluidas dentro del Programa TICAPPS (TIC en 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

El potencial de los sistemas digitales en IIoT es la integración de diferentes componentes críticos para los procesos industriales que permiten optimizar su rendimiento global (Industria 4.0)[1]. Se pueden obtener grandes beneficios cuando los dispositivos almacenan y analizan sus propios datos de uso, de manera que luego puedan

comunicar esos datos de forma inteligente a otros dispositivos conectados.

La Visión por Computador es la disciplina que estudia cómo procesar, analizar e interpretar imágenes de forma automática. Estas técnicas tienen aplicaciones en muchos ámbitos, como la seguridad, la medicina, la inspección automática, o la navegación automática.

Para explotar las ventajas de las nuevas tecnologías, queda de manifiesto que internet de las cosas es necesario para el almacenamiento y acceso a grandes cantidades de datos en la nube, por otro lado el aprendizaje automático permite el entrenamiento de máquinas para la realización de tareas y la visión por computadora permite la manipulación de objetos para automatizar actividades y procesos.

Dentro de la Visión por Computador, la detección de objetos es uno de los temas más candentes. El problema parece sencillo, pues dada una imagen se desea diseñar un sistema capaz de encontrar en ella un objeto determinado. Para un humano esta tarea parece algo obvio pero para una máquina no lo es en absoluto. Para entender el problema hay que pensar en cómo queda codificada una imagen digital. En general, para una máquina las imágenes son enormes cajas tridimensionales llenas de números. Cada píxel (o punto) de la imagen queda representado con tres valores, que modifican su color como una combinación de la cantidad de rojo, verde y azul. Así pues, cuando una máquina busca un objeto dentro de una imagen lo que realmente hace es buscar patrones que se correspondan con el objeto en particular.

Si dotamos a estas herramientas inteligentes con tecnologías como visión 3D, imágenes multispectrales y lo combinamos con aprendizaje automático, (una rama de la Inteligencia Artificial), para un análisis de datos más profundo y una predicción de modelos más precisa, conseguiremos una inteligencia artificial que dote de mayor flexibilidad a los equipos industriales para la toma de decisiones. De esta manera, se desarrolla la capacidad de configurar un robot

a través del aprendizaje en lugar de usar programación prediseñada.

El concepto de internet de las cosas ha adquirido gran relevancia en los últimos años, debido a la posibilidad que ofrece de interconectar objetos entre sí y la conectividad a internet que provee a las redes de objetos [2]. Aunque los sistemas IoT son más que conocidos, su implementación y desarrollo no ha sido completamente explotado en determinados sectores productivos. Para este tipo de ambientes se dispone de un conjunto de protocolos de comunicación e interfaces como GSM/GPRS, Wifi, bluetooth y zigbee que se pueden implementar mediante herramientas de sistemas embebidos como las desarrolladas en [3-5]. Las investigaciones realizadas mediante IoT se podrían aplicar en el monitoreo de procesos, almacenamiento masivo de datos de sensores y dispositivos de cómputo (Big Data).

Respecto al aprendizaje automático [6], una de las áreas en donde se avanzó notablemente es en la de detección de objetos y clasificación de imágenes. Esto se debe en su mayor parte al desarrollo de nuevas técnicas de Machine Learning (Aprendizaje Automático) como el Deep Learning o Aprendizaje Profundo [7], además de las innovaciones en el manejo de Big Data (datos a gran escala) y el aumento en la capacidad de cómputo mediante el uso de diferentes tecnologías como cloud computing (computación en la nube) o el uso de GPU (unidad de procesamiento gráfico) para el análisis de información. Este avance puede verse en distintas áreas como: medicina, seguridad, turismo, finanzas, robótica, entre otras. Algunos ejemplos de dichos avances en el área se aplican en: control de vehículos autónomos, detección de rostros, detección de matrículas, diagnóstico de enfermedades, realidad aumentada, etc.

Machine Learning es un subcampo de la IA en el que se utilizan diferentes algoritmos para recolectar datos, y con estos realizar un aprendizaje para luego hacer una predicción o sugerencia sobre algo [8]. De esta manera se permitirá resolver problemas de forma intuitiva y automatizada, sin que el

mecanismo de elección se encuentre previamente programado. En la práctica esto se traduce en una función matemática en la que se parte de una entrada y se obtiene una salida, por lo que el desafío reside en construir un modelado automático de esta función matemática.

Deep Learning es un subcampo de Machine Learning, pero existen técnicas de Machine Learning que no utilizan Deep Learning. Este último es utilizado para realizar procesos de Machine Learning empleando redes neuronales artificiales compuestas por varios niveles jerárquicos [7]. En el nivel inicial la red aprende patrones simples, y esta información se envía al siguiente nivel de la jerarquía. Este segundo nivel toma la información obtenida en el primero y la combina con nuevos patrones aprendidos en este, generando información un poco más compleja, la cual es pasada a un tercer nivel, y así sucesivamente. Las técnicas de machine learning y deep learning proveen gran soporte para el diseño de aplicaciones de visión por computadora, es por eso que son parte fundamental de la propuesta.

Por último, teniendo en cuenta que la propuesta será ejecutada en sistemas de cómputo de altas prestaciones y la complejidad incorporada al implementar sistemas multi-cores, incrementa la vulnerabilidad a los fallos transitorios, estos fallos podrían corromper los resultados de las aplicaciones [9]. El alto costo (en términos temporales y de utilización de recursos) que implica volver a lanzar la ejecución de una aplicación desde el comienzo, en caso de que un fallo transitorio produzca la finalización de la aplicación con resultados incorrectos, justifica la necesidad de desarrollar estrategias específicas para mejorar la confiabilidad y robustez en sistemas de múltiples procesadores [10-12]. En cuanto al estado actual de la tolerancia a fallos, la recuperación es correcta pero existe un grado de ineficiencia debido a que no está determinado si un fallo en la reejecución es un fallo nuevo o una repetición del fallo previamente detectado.

Aplicaciones

Entre las aplicaciones que se están desarrollando mediante IoT, procesamiento de imágenes y machine learning se tiene; la detección automática de residuos reciclables, La detección de malezas, técnicas de eficiencia energética, calidad del aire y visión artificial aplicada a robótica.

El reciclado de desechos es considerado de suma importancia para el cuidado del medioambiente, debido a que, supone la reutilización de elementos u objetos ya utilizados, los que de otro modo serían desechados contribuyendo al incremento de la basura y al daño ambiental permanente. En nuestro país, aunque se han tomado medidas para fomentar el reciclado, solo un 24% de la población se esfuerza por separar los residuos para minimizar su generación y la contaminación. Gran parte del problema radica en el esfuerzo que requiere clasificar y separar los residuos inorgánicos, es por eso que en esta línea de investigación se propone desarrollar un sistema basado en redes neuronales artificiales que permita detectar y clasificar los objetos reciclables más comunes como papel, cartón, botellas, latas, etc, y los materiales con los que están hechos (plástico, vidrio, metal, papel).

Respecto a la visión por computadora ([13] y [14]), 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 [15].

Adicionalmente, se ha desarrollado un sistema de control para el sensado de gases y el control automático de iluminación y ventilación, de tal forma de contribuir con la eficiencia energética y calidad del aire y calidad del aire en ambientes cerrados [16].

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, tolerancia a fallo, procesadores embebidos, IoT, aprendizaje automático, y robótica.

En cuanto a la robótica ([15] y [17]), esta línea está en pleno desarrollo y se buscan aplicaciones innovadoras como la detección de objetos para reciclado, la capacidad de detección de señales y obstáculos para el funcionamiento autónomo mediante procesamiento de imágenes ([18] y [19]).

Las líneas de I/D que se presentan en este trabajo están basadas en el estudio y desarrollo de herramientas alternativas para el procesamiento de imágenes a partir de algoritmo de procesamiento, aprendizaje automático y visión por computador.

Temas de Estudio e Investigación

Los datos recopilados de la robótica industrial pueden suponer grandes cantidades de información dependiendo de los sensores de los que disponga y de la integración con otros sistemas. Este Big Data Industrial nos permitirá conseguir una mayor precisión en los sistemas de Machine Learning y, cuanto mayores sean los datos útiles con los que alimentemos el modelo, mejor será su aprendizaje y, por ende, sus predicciones.

Los aportes originales no solo tienen que ver con el manejo de la información, sino también con las aplicaciones innovadoras que se le puede dar al aprendizaje automático en visión por computadora y robótica, donde el ejemplo más claro se muestra en la posibilidad de disponer de un sistema de reciclaje automático. También se puede innovar respecto a la visión por computador para conseguir la realización de tareas específicas mediante dispositivos robóticos en procesos particulares de la industria local.

Respecto a la tolerancia a fallos, partiendo del sistema desarrollado previamente, se buscará mejorar la eficiencia en la detección de fallos.

3. Resultados y Objetivos

Resultados alcanzados:

La detección y clasificación de objetos reciclables, son líneas de estudio comenzadas recientemente, aunque, ya se cuenta con la implementación de los primeros modelos basados en la arquitectura de redes neuronales convolucionales (CNN). Estos modelos se desarrollaron en un entorno virtualizado obteniendo como resultado tasas de acierto cercanas al 75%.

Por otro lado, se iniciaron pruebas desarrollando modelos que hacen uso de las técnicas de aprendizaje por transferencia (Transfer Learning), en los que los resultados son prometedores, con aciertos cercanos al 84%. También se está trabajando en la mejora continua de los set de datos utilizados para el entrenamiento de los modelos de CNN desarrollados.

En el trabajo reciente, se llevó a cabo el desarrollo de una aplicación (versión Beta), implementada en un miniordenador Raspberry Pi 3 Model B+, en donde se hace uso del módulo de la cámara (pi camera) de la Raspberry para tomar fotos en tiempo real y realizar la clasificación de dicha imagen determinando que tipo de objeto reciclable se encuentra en ella [13]. Dentro de los resultados obtenidos se está evaluando no solo el porcentaje de acierto, sino también, los tiempos de predicción. Si bien se detectó que la carga inicial del modelo de red neuronal tiene una latencia de entre 30 y 40 segundos, luego, la captura de la imagen y posterior clasificación arrojó tiempos aproximados a los 10 segundos.

En cuanto a las aplicaciones en robótica, actualmente se está trabajando con un robot móvil con cámara y un brazo robótico, se espera obtener resultados a lo largo del año.

Por último, se debe destacar que en cuanto a la tolerancia a fallos, se concluyó con los estudios de doctorado del Profesor Montezanti, obteniendo muy buenos resultados en sistemas de cómputo de altas prestaciones. De cualquier manera, se

continúa con la línea en el nuevo proyecto de investigación.

Objetivos esperados:

- Promover la generación de conocimiento teórico y aplicado, como el desarrollo de instrumentos que puedan aplicarse en las materias de posgrado y la formación de recursos humanos en investigación.

- Fortalecer la actividad de investigación y vinculación en el área de las TIC que contribuyan a remediar las problemáticas existentes dentro del territorio de influencia de la UNAJ en las áreas de medio ambiente y medio socio productivo.

- El objetivo principal de esta línea de investigación es desarrollar nuevas técnicas y obtener resultados favorables respecto a la detección de objetos y características del ambiente aplicando técnicas de visión por computadora y aprendizaje automático.

- Realizar el diseño y desarrollo de una metodología que permita tolerar fallos transitorios que se producen en las arquitecturas multicore (sistemas de múltiples procesadores), y que afectan especialmente la ejecución de aplicaciones paralelas de cómputo intensivo.

Se espera contribuir con el cuidado del medio ambiente, a través de la detección, reconocimiento y clasificación de objetos reciclables. Además, se pretende contribuir en el desarrollo de tecnología para el sector agroindustrial, mediante diferentes aplicaciones de procesamiento de imágenes y aprendizaje automático.

4. 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 dos Doctores en Ingeniería, un Magister, un Especialista, un Doctorando y un integrante realizando Maestría en temas relacionados.

Además hay dos estudiantes avanzados realizando las PPS de final de carrera en la temática. Actualmente, se encuentran en curso dos Tesis de Doctorado y una beca de Estímulo a las Vocaciones Científicas del Consejo Interuniversitario Nacional (Becas EVC – CIN), relacionadas directamente con las líneas de I/D presentadas.

5. Bibliografía

[1] Luis joyanes Aguilar, “Industria 4.0, la cuarta revolución industrial, Alfaomega, 2016. ISBN: 978-607-622-942-2

[2] Leila Fatmasari Rahman, ”Choosing your IoT Programming Framework: Architectural Aspects”, 2016 IEEE 4th International Conference on Future Internet of Things and Cloud

[3] J. Osio, M. Cappelletti , G. Suárez, L. Navarro, F. Ayala, J. Salvatore , D. Alonso ,D. Encinas, M. Morales, “Diseño de aplicaciones de IoT para la solución de problemas en el medio socio productivo”, UNSJ, San Juan, WICC 2019.

[4] Osio J., C. Acquarone, E. Hromek, J. Salvatore, “Plataforma de desarrollo para IoT”, IV conaiisi, 2017

[5] J. Osio, J. Salvatore , D. Alonso , V. Guarepi , M. Cappelletti , M. Joselevich , M. Morales, “Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo”, UNNE, Ciudad de Corrientes, WICC 2018.

[6] J. Hurwitz, D. Kirsch, “Machine Learning for Dummies”, Ed. Wiley, 2018.

- [7] Nikhil Buduma, “Fundamentals of Deep Learning”, Editorial O’reilly. 2017
- [8] Aurelien Gerón, “Hands-On Machine Learning with scikit Learn & TensorFlow”, Editorial O’reilly, 2017
- [9] D. Montezanti, A. De Giusti, M. Naiouf, J. Villamayor, D. Rexachs, E. Luque, “AMethodology for Soft Errors Detection and Automatic Recovery”, in Proceedings of the 15th International Conference on High Performance Computing & Simulation (HPCS). ISBN: 978-1-5386-3250-5/17. IEEE, 2017, pp. 434
- [10] J. Osio, J. Salvatore, E. Kunysz, D. Montezanti, D. Alonso, M. Morales, “Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos”, ITBA, CABA, WICC 2017
- [11] J. Osio, D Montezanti, E. Kunysz, Morales M., “Análisis de eficiencia y tolerancia a fallo en Arquitecturas Multiprocesador para aplicaciones de procesamiento de datos ”, UNNE, Corrientes, WICC 2018.
- [12] J. Osio, D Montezanti, E. Kunysz, Morales M., “Determinación de la eficiencia y Estrategias de Tolerancia a Fallos en Arquitecturas Multiprocesador para aplicaciones de procesamiento de datos”, UNSJ, San Juan, WICC 2019.
- [13] Ashwin Pajankar , Raspberry Pi Computer Vision Programming Second Edition, (2015).
- [14] Francisco Moreno y Esmitt Ramírez, “Algoritmos de Visión por Computador para un SBC”, (2017).
- [15] Mohammad Jamshidi y Patrick J. Eicker, Robotics and remote system for hazardous environments, Prentice Hall, New Jersey, Estados Unidos (1993)
- [16] Lo E., Fain F., Osio J., Cappelletti M., Aróztegui W., “Machine learning aplicada a variables ambientales”, VII CONAIIISI, 2020.
- [17] Jorge Angeles, “Fundamentals of Robotic Mechanical Systems”, (2014).
- [18] Gary Bradski and Adrian Kaehler, “Learning OpenCV”, (2008).
- [19] Wilhelm Burger & Mark J. Burge, “Digital Image Processing”, Second Edition (2008).

UNA ARQUITECTURA DE MICROSERVICIOS PARA DAR SOPORTE A LA CREACIÓN Y EJECUCIÓN DE ACTIVIDADES DE RECOLECCIÓN DE DATOS CON INTERVENCIÓN HUMANA

Jose Arcidiacono¹, Patricia Bazán², Alejandra B. Lliteras^{3,4}

¹ UNLP, Facultad de Informática, ² UNLP, Facultad de Informática, LINTI, ³ UNLP, Facultad de Informática, LIFIA, ⁴ CICPBA

j2arcidiacono@outlook.com, pbaz@info.unlp.edu.ar, alejandra.lliteras@lifia.info.unlp.edu.ar

RESUMEN

DEHIA es una plataforma para la creación y ejecución de actividades de recolección de datos que requieren de la intervención humana para realizar este objetivo.

La plataforma está definida sobre una arquitectura distribuida, basada en microservicios y cuenta con un prototipo funcional probado con un conjunto de usuarios [1].

Este trabajo presenta una versión revisada de la arquitectura respecto de lo presentado en [2] donde se describen cada uno de sus componentes y las capas en que fueron estructurados.

Palabras clave: Arquitecturas distribuidas, Microservicios, Recolección de datos, Intervención Humana

CONTEXTO

En el año 2020 se presentó una tesina de grado [1] proponiendo el diseño de una arquitectura para una plataforma de creación y ejecución de actividades de recolección de datos con intervención humana, junto con un prototipo implementando parte de la funcionalidad propuesta.

La motivación inicial de la tesina respondía a una iniciativa de un Proyecto de Extensión de la Facultad de Informática, “Recicla tu Compu-Recicla tu Mundo”¹ [3], que buscaba la creación de herramientas basadas en

tecnología para acercar a las escuelas las problemáticas de los RAEE (Residuos de Aparatos Eléctricos y Electrónicos).

Para encarar ese trabajo se tomó el enfoque de dos proyectos de investigación ([4], [5]) que proponen distintas estrategias para llevar a cabo actividades con intervención humana mediadas por tecnología, aplicando conceptos de Ingeniería de Software e Interacción Humano-Computador (HCI).

Existe otro trabajo de investigación que presenta una herramienta de autor de actividades educativas mediadas por tecnología móvil, basada en este mismo enfoque [6]. La tesina buscaba presentar un desarrollo nuevo que amplíe la funcionalidad de la herramienta propuesta en [6], tomando nuevos elementos del enfoque [4] y proveyendo mayor robustez así como la posibilidad de abarcar otros campos como la recolección de datos para lo que se conoce como *Ciencia Ciudadana* [7], [8].

Se presentaron los avances de la tesina en WICC 2020 [2] así como en JAIIO 2020 ([9], [10]).

1. INTRODUCCIÓN

Los sistemas distribuidos pueden concebirse como aquellos cuya funcionalidad se encuentra fraccionada en componentes que al trabajar sincronizada y coordinadamente otorgan la visión de un sistema único, siendo la distribución transparente para quien hace uso del sistema. [11]. También se puede definir un sistema distribuido como aquel en

¹ Aprobado y subsidiado en la convocatoria de Proyectos de Extensión Universitaria 2018, de la UNLP. El mismo está destinado a promover la educación ambiental sobre los RAEE en escuelas de la región en el período 2018 al 2019.

el cual componentes de software o hardware ubicados en computadoras en red se comunican y coordinan sus acciones sólo por pasaje de mensajes [12].

Dentro de las arquitecturas distribuidas, se encuentran las arquitecturas de microservicios (MS), las cuales son aquellas en las que una única aplicación está compuesta por un conjunto de pequeños *servicios*, cada cual ejecutándose en su propio proceso y comunicándose por mecanismos ligeros, a menudo una API HTTP [13].

Los microservicios son una alternativa a los sistemas monolíticos, que consisten en un único proceso o unidad de código que resuelve tanto la interfaz, como la lógica de negocio y el acceso a los datos. En el caso de los microservicios, el sistema se descompone en componentes débilmente acoplados y comunicados por pasaje de mensajes.

Entre las ventajas de los microservicios a nivel mantenimiento, se tiene que la funcionalidad relacionada está en el mismo servicio, y un cambio en esa funcionalidad no afecta a los demás servicios [14].

Otra característica de los microservicios es que permiten usar una o varias tecnologías diferentes en cada servicio, lo que permite adaptarse a las necesidades de cada funcionalidad. Sin embargo, puede ser beneficioso mantener una diversidad limitada de tecnologías cuando se piensa que el equipo de trabajo debe poder lidiar con todas ellas.

Se diseñó e implementó una plataforma con una arquitectura basada en microservicios. En este trabajo se completa la propuesta que se hizo en [2].

Se tiene un modelo en cuatro capas, las cuales pueden observarse en la Figura 1 y se describirán a continuación:

- La primera capa, la de *clientes*, contiene los componentes que permiten a los usuarios la interacción con el resto de la plataforma. En esta

capa se cuenta con dos tipos de clientes. Desde el cliente Web se definen las actividades (funcionalidad a la cual pueden acceder los usuarios denominados “autores”) y también pueden resolverse de manera web (funcionalidad a la cual pueden acceder los “clientes web”). El cliente Móvil, es decir una aplicación para celular, también permite la ejecución de actividades. Este cliente es accedido por los “usuarios móviles”.

- La capa de compuertas contiene los componentes tecnológicos que permiten a los clientes comunicarse de manera transparente con los servicios (que serán descritos más adelante). Se tiene una compuerta por cada tipo de cliente de acuerdo a las necesidades de comunicación de cada uno.

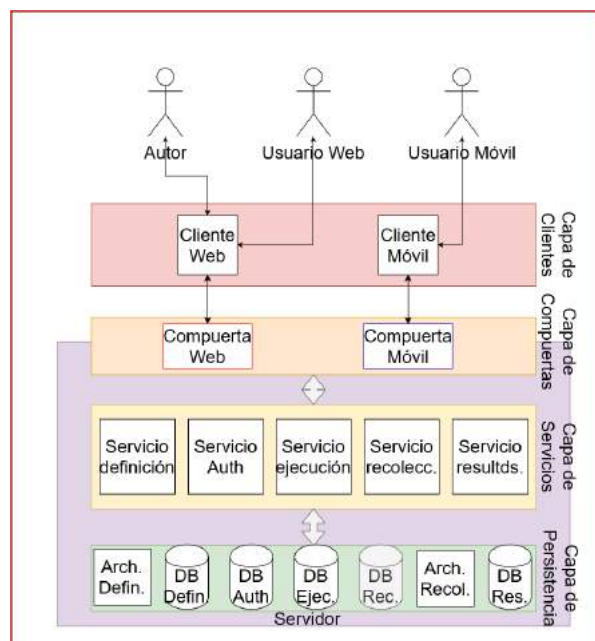


Figura 1. Arquitectura por capas de la plataforma

- La capa de servicios contiene los servicios, en este caso *microservicios* que brindan la funcionalidad principal de la plataforma: el servicio de

definiciones gestiona la gestión, listado y descarga de actividades; el servicio Auth (autenticación-autorización) gestiona los tokens de acceso para autenticar a los usuarios y autorizarlos a realizar determinadas tareas; el servicio de ejecución que mantiene el estado de ejecución de las actividades en Cliente Web; el servicio de recolección que recibe las respuestas producto de la resolución de las actividades; y finalmente, el servicio de resultados, que expone los datos de forma de poder ser consumidos para generar resúmenes, estadísticas y tablas.

- La última capa, la de persistencia, contiene los mecanismos para hacer permanentes los datos de los distintos servicios. En una arquitectura de microservicios, los datos son privados a los servicios, de modo que cada servicio tiene su propia base de datos, y dos de ellos además tienen sistemas de archivos. Además, son todas bases de datos SQL a excepción del servicio de recolección, que necesita una base de datos NoSQL.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La tesina de grado que sustenta este trabajo dio origen a la definición de varias líneas de investigación y desarrollo y que abordan diferentes áreas, como 1- Ingeniería de Software, 2- Interacción Humano-Computador, 3 - Usabilidad y Accesibilidad tanto web como móvil, 4- Arquitecturas Distribuidas, 5- Workflows, 6 - Herramientas de autor, 7 - Testing y 8- Recolección de datos.

De todas las áreas mencionadas anteriormente, este trabajo aborda la de arquitecturas distribuidas y define una línea

de investigación que apunta a mejorar aspectos de los microservicios.

Una arquitectura distribuida con microservicios trae beneficios pero también conlleva nuevos desafíos. Un ejemplo es la distribución por componentes en contenedores independientes que aportan a la interoperabilidad pero que aumenta la complejidad operacional y requiere automatización y monitoreo.

En este sentido, la mejora en torno a los microservicios que aborda la línea de trabajo apunta a considerar requerimientos no funcionales (escalabilidad, seguridad) a partir de un mejor aprovechamiento de las capacidades que ofrecen los microservicios, sin perder de vista el mantenimiento, despliegue y monitoreo de los mismos.

El estilo arquitectónico de los microservicios aborda el desafío de las arquitecturas monolíticas tradicionales donde el software se construye en forma de bloque complejo e indivisible. Para aprovechar los beneficios del estilo arquitectónico de microservicios, es necesario utilizar tecnologías alineadas con las características de los microservicios [15]. La contenerización (*containerization*) se ha convertido en un formato de implementación popular para microservicios, y Docker [16] es el contenedor líder siendo una plataforma que empaqueta todo lo necesario para implementar y ejecutar microservicios.

Por otra parte, para permitir orquestar y gestionar la implementación de contenedores, se requieren plataformas específicas, siendo Kubernetes, un caso de ellas. Kubernetes [17] es la plataforma líder para la automatización implementación, escalado y administración para contenedores aplicaciones y facilita la tarea de los desarrolladores, manejando la complejidad de implementar la resiliencia de sus aplicaciones y les permite concentrarse en la lógica de negocio de las aplicaciones [18].

3. RESULTADOS ESPERADOS / OBTENIDOS

Se propuso una arquitectura distribuida basada en microservicios que permitirá la gestión y ejecución de actividades recolección de datos con intervención humana. Se plantearon los lineamientos para una plataforma que cumpliera con esta funcionalidad y fuera accesible para usuarios sin conocimientos de programación. La plataforma debía favorecer la reutilización de conocimiento (actividades, tareas, planificaciones) propio y de otros usuarios.

Se pretendía desarrollar un prototipo funcional de esta plataforma aplicándolo en un caso de estudio que permitiera hacer pruebas con usuarios.

Finalmente se pudo implementar el prototipo basado en la arquitectura y los lineamientos propuestos, y se tomó como caso de estudio una necesidad del Programa E-Basura de la Universidad Nacional de La Plata². Se instanció el prototipo para representar con actividades uno de los formularios³ de recepción de donaciones del Programa. Se realizaron pruebas de usabilidad con usuarios experimentados y no experimentados en la configuración de herramientas. Las pruebas indicaron que la plataforma podría ser usada por personas sin conocimientos de programación, llegando a niveles de usabilidad aceptables [1].

4. FORMACIÓN DE RECURSOS HUMANOS

El trabajo de investigación realizado por la primera autora de este artículo ha permitido consolidar varias de las líneas de investigación del actual proyecto de investigación del LINTI “De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje

tecnológico y ético en nuestra región”⁴ que entre sus amplios objetivos específicos propone promover buenas prácticas de seguridad en las distintas plataformas de despliegue de aplicaciones garantizando la disponibilidad, integridad y confidencialidad de la información.

En particular la temática de la tesina será continuada por su autora desde el enfoque de la arquitectura a través de una propuesta de doctorado a presentarse en 2021.

5. BIBLIOGRAFÍA

[1] Arcidiacono, J (2020). DEHIA: una plataforma liviana para definir y ejecutar actividades con intervención humana basadas en workflows. Tesina de grado. Facultad de Informática, UNLP

[2] Arcidiacono, J., Lliteras, A. B., & Bazán, P. A. (2020). Plataforma para la definición y ejecución de actividades orientadas a la recolección y análisis de datos, con intervención humana. In XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz).

[3] Proyecto Recicla tu Compu-Recicla tu mundo. Proyecto de extensión de la UNLP, Facultad de Informática, LINTI. <http://bit.ly/3eoYw0e>

[4] Lliteras, A. B. (2015). Un enfoque de modelado de actividades educativas posicionadas que contemplan elementos concretos. Tesis de maestría. Facultad de Informática, UNLP.

[5] Lliteras, A. B., Grigera, J., dal Bianco, P. A., Corporaal, F. M., & Gordillo, S. E. (2018). Challenges in the design of a customized location-based mobile learning

²<https://e-basura.unlp.edu.ar/>

³<https://encuestas.linti.unlp.edu.ar/index.php/369463/lang-es>

⁴ Programa de Incentivos a la Investigación del Ministerio de Educación de la Nación

application. In 2018 XIII Latin American Conference on Learning Technologies (LACLO) (pp. 315-321). IEEE.

[6] Dal Bianco, P. A., Mozzon Corporaal, F., Lliteras, A. B., Grigera, J., & Gordillo, S. E. (2019). MoLE: A web authoring tool for building mobile learning experiences. In XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019, Universidad Nacional de Río Cuarto).

[7] Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977-984.

[8] Strasser, B. J., Baudry, J., Mahr, D., Sanchez, G., & Tancoigne, E. (2019). "Citizen Science"? Rethinking Science and Public Participation. *Science & Technology Studies*, 52-76.

[9] Arcidiacono J, Bazán P y Lliteras AB (2020). Arquitectura de microservicios distribuidos para una plataforma que orquesta actividades orientadas a la recolección y análisis de datos, con intervención humana en JAIIO 2020.

[10] Arcidiacono J, Lliteras AB. y Bazán P (2020). DEHIA, una plataforma para la generación y ejecución de actividades de recolección de datos con intervención humana aplicada en el Programa E-Basura en JAIIO 2020.

[11] Bazán P. et al. (2017). *Arquitecturas, Servicios y Procesos Distribuidos. Una Visión desde la construcción del software*. Libro de Cátedra. Editado por EDULP (Editorial de la

UNLP). ISBN 978-950-34-1520-7
<https://bit.ly/3chD0rh>

[12] Colouris, G et al. (2000) *Distributed Systems Concepts and Design 3e*. Addison-Wesley. ISBN- 13: 978-0132143011

[13] Fowler, Martin, and James Lewis. "Microservices a definition of this new architectural term." URL: <http://martinfowler.com/articles/microservices.html> (2014).

[14] Newman, Sam. *Building microservices: designing fine-grained systems*. O'Reilly Media, Inc., 2015.

[15] N. Dragoni et al., "Microservices: Yesterday, Today, and Tomorrow," in *Present and Ulterior Software Engineering*, M. Mazzara and B. Meyer, Eds. Cham: Springer International Publishing, 2017, pp. 195–216.

[16] "Docker" [Online]. Disponible: <https://www.docker.com/>.

[17] "Kubernetes," Kubernetes. [Online]. Disponible: <https://kubernetes.io/>.

[18] Vayghan, L. A., Saied, M. A., Toeroe, M., & Khendek, F. (2019, July). Microservice based architecture: Towards high-availability for stateful applications with Kubernetes. In 2019 IEEE 19th International Conference on Software Quality, Reliability and Security (QRS) (pp. 176-185). IEEE.

Agentes y Sistemas Inteligentes

Análisis de textos con estructura

Marina Cardenas, Julio Castillo

Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional

{ing.marinacardenas, jotacastillo}@gmail.com

RESUMEN

Este artículo describe un proyecto de investigación relacionado con el análisis de textos que presentan cierta estructura, en particular se centra en el procesamiento de archivos de código fuente en un lenguaje de programación determinado. El proyecto aborda el problema de la detección de reutilización de código y presenta otras aplicaciones derivadas, como la detección de plagio en el código utilizado.

Se utilizan técnicas adaptadas del área lingüística computacional, y de una subtarea que se denomina implicación textual que es definido originalmente para textos sin estructura subyacente.

El proyecto se encuadra dentro de una línea de investigación en aprendizaje automático por computadora.

Palabras clave: análisis de textos, corpus, extracción de información.

CONTEXTO

En este artículo se presenta el proyecto denominado *Modelado para el procesamiento de textos estructurados*, el cual se trata de un proyecto acreditado por la Secretaria de Ciencia y Tecnología de la UTN con código: UTN4518.

El mismo aborda temáticas específicas del área de lingüística computacional, inteligencia artificial y compiladores. Su desarrollo se lleva a cabo en el Laboratorio

de Investigación de Software LIS¹ del Departamento de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba.

Este proyecto se encuentra dentro del grupo formal UTN denominado GA2LA: Grupo de Aprendizaje Automático, Lenguajes y Autómatas y del LIS¹.

El grupo GA2LA [1] reúne a diversos proyectos de investigación relacionados con redes neuronales, gramáticas y lenguajes de programación, como así también de calidad y trazabilidad en el desarrollo de aplicaciones informáticas. A su vez, trabaja desde un abordaje práctico en colaboración con profesionales de ciencias sociales, a los efectos de poder resolver problemáticas de salud comunitaria desde un enfoque de aprendizaje de computadoras y de la construcción de modelos teóricos computacionales.

Este grupo está compuesto por los integrantes de los diferentes proyectos que nuclea, y que incluye a docentes investigadores, doctores, pasantes y becarios.

1. INTRODUCCIÓN

El proyecto *Modelado para el procesamiento de textos estructurados* tiene como objetivo construir modelos y aplicaciones software que permitan determinar las similitudes a nivel de archivos de códigos fuentes, utilizando para ello diversas técnicas de procesamiento del lenguaje natural y de minería de textos en

¹ www.investigacion.frc.utn.edu.ar/mslabs/

general [2]. Debido a ciertas similitudes de los lenguajes de programación con los lenguajes naturales, y al ser objeto de estudio dentro de la clasificación general de lenguajes y gramáticas atribuidas a Noam Chomsky, es tomado como punto de partida para integrar conocimientos de áreas de compiladores con otros de lingüística computacional [3] [4].

La utilidad de la detección de similitudes en códigos fuentes es basta [5], y entre algunas aplicaciones podemos mencionar: la evolución de un archivo de código fuente, y de un proyecto de desarrollo de software en general, la detección de código reutilizado en un mismo proyecto (útiles al momento de la "refactorización" del código y para el seguimiento de defectos), la detección de prácticas de plagio, entre otras [6] [7].

En el proyecto abordamos la problemática de detección de similitudes de código fuente con fines de reutilización, que de ser hecha manualmente requeriría de altos costos debido a lo laborioso de dicha actividad, en particular con proyectos de escala mediana - grande.

En la bibliografía pueden encontrarse tres tipos común de abordajes a esta tarea, que las describimos brevemente a continuación:

- Aproximaciones basadas en atributos, donde las métricas se calculan a partir del código fuente y se utilizan para la comparación de los distintos archivos. Por ejemplo, se puede utilizar el tamaño del código fuente (número de caracteres, palabras y líneas) como atributo comparable de tamaño [6], o también el número de variables, el número de funciones, el número de clases, entre otros atributos.
- Aproximaciones basadas en Tokens, en las cuales se convierte el código fuente en una secuencia de "fragmentos", para una evaluación posterior y luego la selección de estas secuencias de tokens según ciertas métricas [7] [8] [9].
- Aproximaciones basadas en la estructura, en este caso, el código fuente es

convertido a una representación intermedia interna (IR), que luego es la que se utiliza para la comparación [10] [11].

Dentro de este contexto, en este proyecto se utiliza una combinación de los dos primeros enfoques, que permitan la detección de similitudes de código fuente en los lenguajes de programación Java y Python en base a un corpus elaborado específicamente para tal fin.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación en las que se enmarca el proyecto de modelado para el procesamiento de textos estructurados es el área de inteligencia artificial, más concretamente una sub-especialidad que se denomina lingüística computacional.

En particular, nos centramos en enfoques basados en el aprendizaje automático. Los desarrollos en esta línea de investigación están constituidos, por un lado, por las herramientas desarrolladas para facilitar el análisis y procesamiento de archivos de texto, en este caso código fuente, y por otro, los sistemas de reconocimiento de similitudes entre dos archivos de código fuente.

La innovación del proyecto radica en los nuevos métodos propuestos para el análisis y procesamiento de textos, así como a los algoritmos creados para abordar los problemas antes mencionados. Los algoritmos diseñados aprovechan las diferentes características que se pueden aprender de los textos y que se recopilan y crean a partir de las herramientas de procesamiento de textos.

En cuanto a las posibles aplicaciones de los resultados, los mismos pueden ser de utilidad en tareas de extracción de información [12] [13], evaluación de las traducciones automáticas, evaluación de la calidad de las traducciones [14], reconocimiento de paráfrasis [15] [16] e implicación de textos [17][18][19][20].

3. RESULTADOS OBTENIDOS/ESPERADOS

En esta sección mencionaremos los resultados obtenidos hasta el momento, en el proyecto que ha concluido su primera etapa, pero que continúa en desarrollo.

Se ha desarrollado una aplicación web que permite realizar la comparación entre archivos de códigos fuentes (en múltiples lenguajes de programación), brindando una métrica de similitud entre dos archivos de código fuente en un lenguaje específico.

La aplicación web brinda la posibilidad de la ejecución "en lotes" de un conjunto grande de archivos. Esta actividad es muy útil cuando se necesita comparar un archivo contra un conjunto a los efectos de identificar una copia (plagio) o bien archivos que representen reutilización de código fuente.

Los archivos de códigos fuentes permitidos en la comparación son por el momento:

- .java, del lenguaje JAVA
- .cpp, del lenguaje de Programación C++
- .c, del lenguaje de Programación C.

Adicionalmente esta aplicación genera reportes con las salidas de las ejecuciones realizadas, que permitirán al decisor identificar rápidamente un subconjunto de archivos que presenten gran similitud entre sí, dentro de un conjunto grande de archivos.

Se está trabajando en el tiempo de comparación, el cual se torna demasiado costoso cuando se quieren comparar decenas de archivos, dependiendo de su tamaño en kilobytes y de las características con las que se lo desee comparar.

Para lograr las clasificaciones de los archivos se trabaja principalmente (por el momento) a nivel léxico con una exploración sintáctica superficial.

Otro de los resultados obtenidos hasta el momento consiste en un corpus de pares de archivos de código fuente. Se trata de un corpus de texto estructurado, realizado

manualmente y requirió un considerable tiempo de clasificación, de dos etiquetadores por el lapso de dos años.

Este corpus fue generado con el Programa Asistente de Creación de Corpus v3 [21] que es producto de otro proyecto de investigación de procesamiento de textos [22].

Si bien inicialmente se pensaba en la construcción de un Corpus de 1000 pares, se pudo construir un corpus con 300 pares de códigos fuentes en lenguaje Java en los que se los ha etiquetado en base a las características presentes en los mismos.

Este corpus sirve de material de entrenamiento para el sistema. Actualmente se encuentra en un trabajo en progreso y se está entrenando el sistema con este conjunto de entrenamiento.

Como trabajos futuros mencionamos la necesidad de modelar el fenómeno de similitud textual en base a diferentes niveles de abstracción, que involucre características a nivel léxico, sintáctico y semántico, cuidando al mismo tiempo de no incrementar su complejidad computacional para que pueda ser utilizado en la práctica y no como un mero desarrollo teórico.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto físicamente se lleva adelante en Laboratorio de Investigación de Software LIS² del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba, en el cual realizan su labor diaria docentes-investigadores, becarios y pasantes de diferentes proyectos de investigación.

En cuanto a las personas que constituyen este proyecto mencionamos:

- Un doctor en ciencias de la computación formado en el área de lingüística computacional y que tiene experiencia en el área de detección y generación de paráfrasis, y en la

² www.investigacion.frc.utn.edu.ar/mslabs/

detección de implicaciones textuales. Entre las actividades que realiza se encuentra la formación de becarios, pasantes y la dirección de investigadores.

- Un doctorando en ingeniería con mención en sistemas de información cuyo tema de tesis se relaciona con el proyecto actual. A su vez, realiza la dirección de becarios, pasantes, y otros miembros del equipo.
- Participan del proyecto alumnos que necesitan realizar su práctica supervisada que es uno de los requisitos para la obtención del grado de Ingeniero. Su asignación depende de las actividades del proyecto. Siguen un plan de trabajo y colaboran con los demás integrantes del proyecto.
- Participa un becario de posgrado como parte de un programa de formación inicial de investigadores en proyectos homologados en UTN.
- Participan anualmente, uno o dos becarios alumnos a los que se les asigna tareas específicas del proyecto, se les enseña a trabajar en el contexto de un proyecto de investigación, y se les enseña la metodología de investigación científica.

5. BIBLIOGRAFÍA

- [1] Juan C Vázquez, Julio J Castillo, Leticia Constable, Marina E Cardenas. (2018). XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).
- [2] Cardenas, Marina E., Castillo, Julio J. Procesamiento de textos estructurados. (2018). XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018).
- [3] Frantzeskou, G., MacDonell, S., Stamatatos, E., Gritzalis S. (2008). Examining the significance of high-level programming features in source code author classification. *The Journal of Systems and Software*, 81(3):447–460.
- [4] M. Craven y J. Shavlik. (1997). Using Neural Networks for Data Mining. *Future Generation Computer Systems*, 13, págs. 211-229.
- [5] Smith, R. y Horwitz, S. (2009). Detecting and Measuring Similarity in Code Clones. *International Workshop on Software Clones (IWSC'09)*, pp. 28-34.
- [6] Wise M. (1992). Detection of similarities in student programs: YAP'ing may be preferable to plaguing. In *ACM SIGCSE Bulletin*, volume 24, pp. 268–271.
- [7] Wise M. (1993) Running Karp-Rabin matching and greedy string tiling. *Basser Dept. of Computer Science, University of Sydney, Sydney*.
- [8] Schleimer, S., Wilkerson, D., y Aiken, A. (2003). Winnowing: Local Algorithms for Document Fingerprinting. En: *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data*, pp. 76-85.
- [9] Li, X., y Zhong, X. (2010). The source code plagiarism detection using AST. In *International Symposium IPTC*, pp. 406–408.
- [10] Baxter I., Yahin, A., Moura, L., Sant'Anna, M., y Bier, L. (1998). Clone detection using abstract syntax trees. En *Proceedings de IEEE ICSM 1998*, pp. 368–377.
- [11] Jadon, S. (2016). Code clones detection using machine learning technique: Support vector machine. *2016 International Conference on Computing, Communication and Automation (ICCCA)*. IEEE. Noida, India.
- [12] Feldman, R. y Hirsh, H.. (1996). Exploiting Background Information in Knowledge Discovery from Text. *Journal of Intelligent Information Systems*.
- [13] Lewis, D. (1995). Evaluating and optimizing autonomous text classification systems. In *Proceedings of SIGIR-95, 18th ACM International Conference on Research*

and Development in Information Retrieval. Seattle, US, págs. 246-254.

[14] Julio Castillo and Paula Estrella. (2012). Semantic textual similarity for MT evaluation. In Proceedings of the Seventh Workshop on Statistical Machine Translation, pages 52–58, Montréal, Canada.

[15] Castillo J., Cardenas M. (2010). Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. Iberamia, LNCS, vol. 6433, pp. 366-375, 2010.

[16] Castillo J. (2010). An approach to Recognizing Textual Entailment and TE Search Task using SVM. Procesamiento del Lenguaje Natural 44, 139-145, 2010. 4.

[17] Castillo J. (2010). Using Machine Translation Systems to Expand a Corpus in Textual Entailment. Proceedings of the Iccetal 2010, LNCS, vol. 6233, pp.97-102.

[18] Julio J. Castillo. (2010). Recognizing textual entailment: experiments with machine learning algorithms and RTE corpora. Special issue: Natural Language Processing and its Applications, Research in Computing Science, 46:155–164.

[19] Castillo, J. J. (2010). Textual entailment search task: An initial approach based on coreference resolution. Intelligent Computing and Cognitive Informatics, International Conference on: 388-391.

[20] Castillo, J. J. (2010). Sagan in TAC2010: A machine learning approach to RTE within a corpus. In Proceedings of the Text Analysis Conference (TAC'10).

[21] Julio J Castillo, Marina E Cardenas, Martin Navarro, Nicolás A Hernández, Melisa Velazco. (2018). Sistemas de análisis textual en formato no estructurado. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).

[22] Julio J Castillo, Marina E Cardenas, Adrián Curti, Osvaldo Casco, Martin Navarro, Nicolás A Hernández, Melisa

Velazco. (2017), “Desarrollo de sistemas de análisis de texto,” in XIX Workshop de Investigadores en Ciencias de la Computación, 2017, pp. 58–62.

Aplicación de técnicas avanzadas de aprendizaje automático para identificar emociones en textos

Alejandra Cardoso¹, Lorena Talame¹, Matias Amor¹, Agustina Monge¹

Grupo de Análisis de Datos /Facultad de Ingeniería e IESIING

¹ Universidad Católica de Salta

Campo Castaños s/n, 4400 Salta, (0387) 426 8536.

{acardoso, mltalame, mnamor}@ucasal.edu.ar, agum_96@hotmail.com

RESUMEN

En este proyecto se pretende aprender a clasificar textos cortos (opiniones) generados en la red social Twitter según el sentimiento que expresan, aplicando técnicas avanzadas de aprendizaje automático como redes neuronales. El proyecto se encuentra en desarrollo. En la primera etapa se exploró en la clasificación de textos usando una red neuronal LSTM y en la etapa actual, se está analizando algunas formas de representar los textos para crear un corpus de palabras embebidas que serán utilizadas en otros experimentos.

Palabras clave: emociones, aprendizaje automático, redes neuronales, Twitter.

CONTEXTO

Este proyecto continúa la línea de investigación que el Grupo de Análisis de Datos de la Facultad de Ingeniería de la Universidad Católica de Salta viene desarrollando en minería de textos. En los proyectos anteriores se exploraron diversos problemas relacionados a minería de textos, entre ellos, clasificación de documentos, reconocimiento de entidades con nombres, generación automática de resúmenes y búsquedas de respuestas.

En el último proyecto se exploró el análisis de sentimientos para la identificación de cuatro emociones en textos extraídos de la red social Twitter con dos enfoques: aplicando aprendizaje automático supervisado y semi supervisado [1] y a partir de lexicones de dimensiones emocionales [2]. En el proyecto actual, se continúa el aprendizaje de detección de emociones en textos, a una colección de mensajes de esa red social.

1. INTRODUCCION

La gran cantidad de información contenida en los medios sociales ha propiciado que la comunidad científica dedique grandes esfuerzos a analizar, estructurar y procesar esta información. Estos medios se suelen utilizar para expresar opiniones y sentimientos diversos sobre diferentes aspectos de la sociedad, productos, servicios, política, celebridades, etc. Empresas, organizaciones y gobiernos han mostrado su interés en conocer las opiniones que los usuarios tienen sobre sus actividades. El marketing digital se basa, principalmente, en las opiniones vertidas en las redes sociales. La detección de polaridad en la opinión textual, es una tarea ampliamente investigada, sobre todo para el idioma inglés. La identificación de la emoción expresada en una

opinión es una tarea menos investigada y con amplias posibilidades de investigación.

Al respecto, los enfoques frecuentemente utilizados son: el aprendizaje supervisado, que utiliza grandes cantidades de texto como entrada a los algoritmos, y el uso de diccionario de palabras asociadas a una o más emociones. Estos tipos de aprendizajes fueron abordados por el grupo de investigación en el proyecto anterior.

Los enfoques de aprendizaje profundo (deep learning) han demostrado su capacidad para resolver tareas relacionadas al procesamiento de lenguaje natural y aplicaciones de la inteligencia artificial. Las redes neuronales son efectivas a la hora de realizar tareas de clasificar textos [3]. En [4] aplican un método híbrido con redes neuronales convolucionales y redes neuronales recurrentes a la clasificación de polaridad en tweets. En [5] se utiliza una combinación de redes neuronales artificiales para el reconocimiento de emociones en textos. La técnica word embeddings consiste en representar palabras como vectores de números reales sobre los que es posible realizar operaciones y obtener resultados sorprendentes. En [6] se utiliza esto para incrementar la efectividad en la clasificación de emociones. En [7] se combina word embeddings y un diccionario léxico de emociones.

Previo a la etapa de aprendizaje, es necesario realizar acciones de preprocesado para eliminar aquellas características que pueden producir ruido en las etapas siguientes, por ejemplo:

- Tokenización y lematización
- Eliminación de palabras vacías (stopwords)
- Eliminación de imágenes, links y referencias a usuarios

Generalmente los usuarios de redes sociales suelen usar emojis para destacar lo que se

quiere expresar, como una forma de entonación de la voz o de expresión corporal. En el proyecto anterior, se demostró que mantener emojis y hashtags resulta relevante, por lo cual, deben ser transformados a texto. Estos tipos de elementos, en general, no son considerados en los trabajos citados anteriormente.

Python es uno de los lenguajes de programación más aceptados por la comunidad científica. Es potente y se caracteriza por su simplicidad, distribución open source y la posibilidad de integración con múltiples librerías. Para el procesamiento de textos con Python, existe una variedad de herramientas informáticas. En [8] se analizaron algunas de ellas, demostrando que Freeing y Stanford resultan las más confiables en cuanto a la tokenización y etiquetado gramatical.

2. LINEA DE INVESTIGACION Y DESARROLLO

Este proyecto de investigación propone detectar sentimientos expresados en textos, en particular, opiniones textuales emitidas en una red social. El proyecto se desarrolla en las siguientes etapas:

- Revisión de la literatura relevante al problema de minería de opiniones y sentimientos.
- Evaluación y comparación las técnicas de aprendizaje profundo para la clasificación de textos.
- Evaluación y comparación de otras técnicas de aprendizaje automático para clasificación de opiniones textuales.
- Desarrollo de un prototipo para la clasificación de opiniones.

Se utilizan los tweets capturados para el proyecto anterior y otros obtenidos durante el

último año, totalizando más de 150000 tweets. Muchos de estos tweets fueron descartados por contener solo imágenes, íconos o textos con poca información para el análisis.

3. RESULTADOS OBTENIDOS/ESPERADOS

En la primera parte del proyecto se iniciaron los experimentos con redes neuronales para el aprendizaje de clasificación de textos.

Con el propósito de combatir el desplazamiento del gradiente (vanishing gradient), que ocurre en las Recurrent Neural Network (RNN), surgen las redes LSTM (Long Short-Term Memory) las cuales son un tipo especial de redes recurrentes [9]. La característica principal de las LSTM es que la información puede persistir introduciendo bucles en el diagrama de la red, para decidir cuál será el siguiente [10]. En la Figura 1 se observa la estructura típica de este tipo de red neuronal.

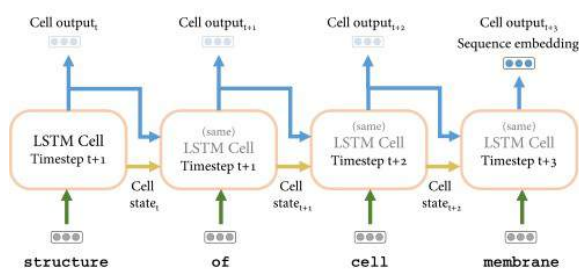


Figura 1. Long Short-Term Memory [9]

La principal diferencia con las redes neuronales tradicionales (RNT), es que estas no tienen persistencia (memoria) de los datos anteriores. Una RNT, no puede usar su razonamiento sobre los eventos anteriores para decidir sobre los posteriores.

Se utilizó el dataset del proyecto anterior [2], un conjunto de tweets en español clasificado por las emociones expresadas, para realizar las pruebas de redes LSTM. Para un primer

acercamiento al proceso, se decidió trabajar con dos categorías, de tal manera de clasificar los tweets por su polaridad. Se agruparon de la siguiente forma:

Positivos = felicidad y sorpresa

Negativo = asco, ira, tristeza y miedo

La Figura 2 muestra la distribución de tweets según las emociones expresadas (positivas o negativas)

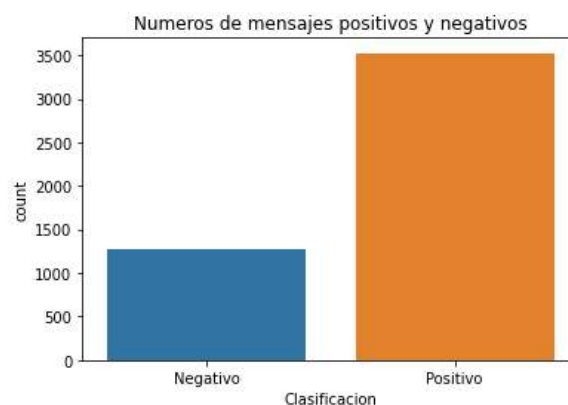


Figura 2. Total de tweets según polaridad

La configuración y parametrización de la RNN fue realizada utilizando la herramienta Colaboratory provista por Google. La Figura 3 contiene parte del código Python utilizado para configurar la RNN

```
#Se define la RNN LSTM
def RNN():
    inputs =
    Input(name='inputs', shape=[max_len])
    layer =
    Embedding(max_words, 120, input_length=max_l
en)(inputs)
    layer = LSTM(64)(layer)
    layer = Dense(1, name='out_layer')
(layer)
    layer = Activation('softmax')(layer)
    model =
    Model(inputs=inputs, outputs=layer)
    return model
```

Figura 3. Configuración de RN

EL proceso continúa con la ejecución de la RNN definida, y la evaluación del modelo construido, tomando como referencia el conjunto de entrenamiento separado en una

etapa anterior. Los primeros resultados obtenidos se muestran en la Figura 4.



Figura 4. Resultados

Actualmente se continúa la experimentación con redes neuronales.

Los próximos objetivos a alcanzar son:

- Crear un corpus de words embeddings a partir de la colección de tweets disponibles
- Experimentar el aprendizaje con el corpus de word embeddings generado y con otros disponibles para investigaciones.
- Seleccionar los algoritmos que mejor clasifiquen las opiniones.

Se espera que esta línea de investigación continúe y amplíe los conocimientos sobre procesamiento de lenguaje natural. Se pretende que este proyecto anime el interés por la investigación y por esta temática a los alumnos de nuestra Facultad.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo está integrado por tres docentes y una egresada de la carrera de Ingeniería en Informática. Se espera incorporar al proyecto alumnos interesados en la temática.

5. BIBLIOGRAFIA

- [1] A. C. Cardoso, M. L. Talamé, M. N. Amor y A. Monge, «Creación de un corpus de opiniones con emociones usando aprendizaje automático,» *Revista Tecnología y Ciencia*, n° 37, pp. 11-23, 03 Abril 2020.

- [2] M. N. Amor, A. Monge, M. L. Talamé y A. C. Cardoso, «Clasificación de sentimientos en opiniones de una red social basada en dimensiones emocionales,» *Revista digital del Departamento de Ingeniería*, vol. 5, n° 1, pp. 1-13, 18 Agosto 2020.
- [3] R. Johnson y T. Zhang, «Effective Use of Word Order for Text Categorization with Convolutional Neural Networks,» *ArXiv preprint*, vol. arXiv:1412.1058, 2014.
- [4] R. Montañés, R. Aznar y R. del Hoyo, «Aplicación de un modelo híbrido de aprendizaje profundo para el Análisis de Sentimiento en Twitter,» *TASS 2018: Workshop on Semantic Analysis at SEPLN*, pp. 51-56, 2018.
- [5] B. Kratzwald, S. Ilić, M. Kraus, S. Feuerriegel y H. Prendinger, «Deep learning for affective computing: text-based emotion recognition in decision support,» *Decision Support Systems*, vol. 115, pp. 24-35, 2018.
- [6] X. Mao, S. Chang, J. Shi, F. Li y R. Shi, «Sentiment-Aware Word Embedding for Emotion Classification,» *Applied Sciences*, vol. 9, n° 7, p. 1334, 2019.
- [7] A. Seyeditabari, N. Tabari, S. Gholizadeh y W. Zadrozny, «Emotional Embeddings: Refining Word Embeddings to Capture Emotional Content of Words,» *ArXiv*, vol. abs/1906/00112, 2019.
- [8] L. Talamé, A. Cardoso y M. Amor, «Comparación de herramientas de procesamiento de textos en español extraídos de una red social para Python,» de *Simposio Argentino de Inteligencia Artificial (ASAI)*, Salta, 2019.
- [9] M. T. Pilehvar y J. Camacho-Collados, *Embeddings in Natural*

Language Processing: Theory and Advances in Vector Representations of Meaning, Morgan & Claypool, 2020.

- [10] J. I. Garzón, «Cómo usar redes neuronales LSTM en la predicción de averías en las máquinas,» 6 Noviembre 2018. [En línea]. Available: <https://blog.gft.com/es/2018/11/06/como-usar-redes-neuronales-lstm-en-la-prediccion-de-averias-en-las-maquinas/>. [Último acceso: Diciembre 2020].

Aprendizaje automático aplicado a la pandemia del virus Covid-19 en Argentina

Carolina Cardoso¹, Lorena Talame¹, Matias Amor¹

Grupo de Análisis de Datos /Facultad de Ingeniería e IESIING

¹ Universidad Católica de Salta

Campo Castaños s/n, 4400 Salta, (0387) 426 8536

{acardoso, mltalame, mnamor}@ucasal.edu.ar

RESUMEN

Este proyecto se desarrolló con el interés de analizar los casos registrados sobre Covid-19 en nuestro país, publicados por fuentes oficiales. Se experimentó con redes neuronales con el fin de predecir casos positivos de la enfermedad, y para encontrar similitudes entre algunos distritos de nuestro país se plantearon relaciones difusas.

Palabras clave: Minería de datos, aprendizaje automático, covid-19

CONTEXTO

Este proyecto continúa la línea de investigación que el Grupo de Análisis de Datos de la Facultad de Ingeniería de la Universidad Católica de Salta viene desarrollando en minería de datos.

1. INTRODUCCION

En diciembre de 2019, Wuhan, China, experimentó un brote de una enfermedad respiratoria causado por un nuevo coronavirus (COVID-19). En marzo de 2020 la OMS (Organización Mundial de la Salud) declaró al mundo en estado de pandemia. A la fecha, fueron confirmados más de 100 millones de personas contagiadas en el mundo.

En Argentina, con el objetivo de evitar la propagación del virus y detección de infectados se tomaron diferentes medidas desde el gobierno nacional [1], entre ellas, el aislamiento social, preventivo y obligatorio.

Si bien, el análisis de datos relacionados a diferentes enfermedades fue explorado ampliamente en la comunidad de investigadores de grandes volúmenes de datos, la exploración de los datos de las personas contagiadas con el nuevo coronavirus nos coloca en una situación especial. No solo por los descubrimientos que se van haciendo día a día por parte de los científicos, sino también por la rápida propagación mundial del virus.

En nuestro país encontramos estudios que toman como base los informes diarios del Ministerio de Salud, es decir, se trata de informes estadísticos que muestran la progresión día a día de la cantidad de casos positivos para Covid-19 y cantidad de decesos [2] [3].

En este proyecto se intentó encontrar similitudes entre algunos distritos argentinos a partir de las cantidades de casos diagnosticados con la enfermedad y con el aprendizaje de relaciones difusas. También se intentó construir una red neuronal para predecir la cantidad de casos positivos de Covid-19 en una provincia argentina.

Existen una variedad de herramientas informáticas para el análisis de datos y desarrollo de modelos para extracción de información. En este trabajo se utilizó el lenguaje de programación Python. Se destaca por ser uno de los más aceptados por la comunidad científica [4] y es uno de los lenguajes más potentes por su simplicidad. Además, es de distribución open source, posibilita la integración con múltiples librerías y por la experiencia de uso de los investigadores de este proyecto. Por otro lado, las herramientas de visualización de datos son complementarias a este proceso, facilitando la lectura y entendimiento de la información detectada [5].

2. LINEA DE INVESTIGACION Y DESARROLLO

El proyecto se desarrolló en las siguientes etapas

- Revisión de la literatura referente a las técnicas de aprendizaje automático aplicadas al problema, en reuniones periódicas del equipo.
- Descarga y almacenamiento de los datos, provenientes de fuentes oficiales (Ministerio de Salud de Nación Argentina, INDEC, etc.)
- Pre-procesamiento de los datos: filtrado, limpieza y selección de datos para entrenamiento y para control del modelo.
- Aplicación de técnicas y algoritmos adecuados para el problema del proyecto
- Validación y evaluación de los modelos obtenidos.

Se utilizó el lenguaje de programación Python y diferentes librerías para machine learning como scikit-learn, scikit-fuzzy y keras.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Desde el mes de Marzo del año 2020, el Ministerio de Salud de la Nación Argentina emite reportes de la situación sanitaria del país, en relación a la situación sanitaria relacionada a Covid-19 [6]. Los casos se registran en el Sistema Nacional de Vigilancia Epidemiológica (SISA) y publicados en la página web de Datos Abiertos del Ministerio [7] en un archivo csv. Cada registro contiene información básica de las personas registradas en SISA, con diagnóstico positivo, negativo (descartado) o fallecido por la enfermedad.

Los datos utilizados en este proyecto se obtuvieron del archivo con fecha de actualización el 31 de octubre de 2020. El archivo consta de 1.048.575 registros y 25 atributos con información básica de las personas registradas: edad, sexo, provincia de residencia, entre otros.

Se realizaron dos tipos de tareas, por un lado, se experimentó con los datos de seis distritos argentinos con el objeto de encontrar similitudes a partir de la definición de relaciones difusas, y por otro, con redes neuronales para la predicción del número de casos positivos de Covid en una provincia.

A través del uso de relaciones difusas se puede definir la similitud entre los elementos de un conjunto como el grado de pertenencia de cada uno de esos elementos a ese conjunto. Se probaron distintos grados de pertenencia para los meses estudiados y se determinó la similitud de seis provincias argentinas.

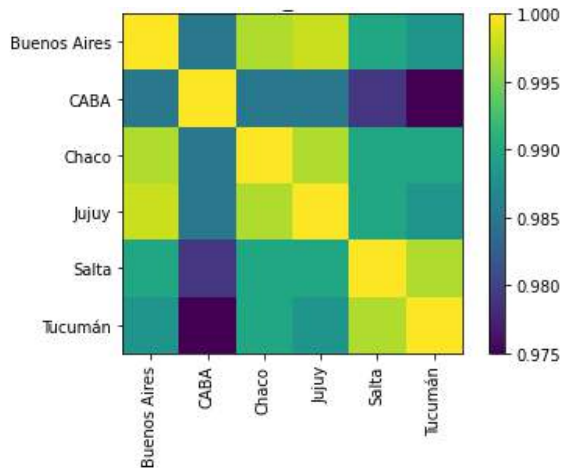


Figura 1. Mapa de calor relaciones de equivalencia mes de Octubre

De esta manera, se detectó por ejemplo, para el mes de octubre 2020, con un grado de pertenencia del 0.990, tres clases de equivalencias donde Buenos Aires y CABA son similares entre sí; Chaco, Jujuy y Tucumán pertenecen a la misma clase y Salta es la única provincia que se distingue del resto (Figura 1).

El modelo de red neuronal para predecir casos de contagios en la provincia de Salta que se probó fue una red secuencial con función de activación hiperbólica. Con esta red y utilizando siete días previos para predecir los casos del octavo se obtuvo un resultado aceptable con el menor error cuadrático medio. tanto en el entrenamiento como en la validación se mantuvo alrededor del 20%. Se pretende probar con otras configuraciones para lograr mejores resultados (Figura 2).

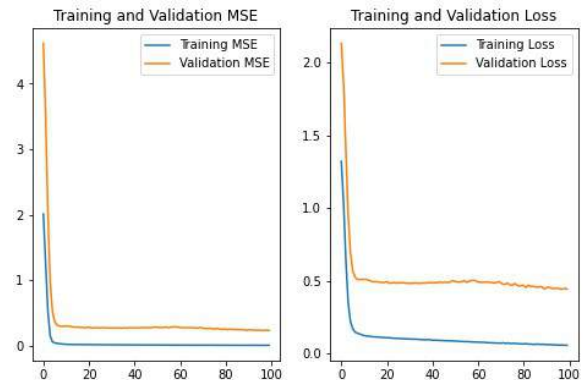


Figura 2. Evaluación

Se espera que esta línea de investigación continúe y amplíe los conocimientos sobre minería de datos. Se pretende que este proyecto anime el interés por la investigación y por esta temática a los alumnos de nuestra Facultad.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo está integrado por tres docentes de la carrera de Ingeniería en Informática. Se espera incorporar al proyecto alumnos interesados en la temática.

5. BIBLIOGRAFÍA

- [1] M. d. S. d. I. N. Argentina, «¿Qué medidas está tomando el gobierno?» 2020. [En línea]. Available: <https://www.argentina.gob.ar/coronavirus/medidas-gobierno>. [Último acceso: Mayo 2020].
- [2] J. Aliaga, «Sobre Covid-19,» 2020. [En línea]. Available: <http://www.jorgealiaga.com.ar/?p=1805>.
- [3] E. Iarussi, «Covid-19 en Argentina,» 2020. [En línea]. Available: <https://observablehq.com/@eiarussi/covid-19-en-argentina/3>.
- [4] G. Piatetsky, «KDnuggets,» Mayo 2019. [En línea]. Available: <https://www.kdnuggets.com/2019/05/poll-top-data-science-machine-learning-platforms.html>.
- [5] J. Hernández Orallo, Introducción a la minería de datos, Madrid: Pearson Prentice Hall, 2004.

-
- [6] M. d. S. d. I. N. Argentina, «Informe diario Covid-19,» 2020. [En línea]. Available: <https://www.argentina.gob.ar/coronavirus/informe-diario>.
- [7] Datos Abiertos, «Ministerio de Salud,» 2020. [En línea]. Available: <http://datos.salud.gob.ar/dataset/covid-19-casos-registrados-en-la-republica-argentina>.
- [8] INDEC, «Instituto Nacional de Estadísticas y Censos - Bases de datos,» 2020. [En línea]. Available: <https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos>.
- [9] R. Arias Montoya, J. J. Santa Chávez y J. D. J. Veloza Mora, «Aplicación del aprendizaje automático con árboles de decisión en el diagnóstico médico,» *Cultura del cuidado*, vol. 10, n° 1, pp. 63-72, 2013.
- [10] I. Witten y E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 3 ed., Morgan Kaufmann, 2011.

APRENDIZAJE PROFUNDO EN APLICACIONES BIOMÉDICAS, AGRONÓMICAS Y AMBIENTALES

Lucas Olivera^{1,2}, Mauro Salina¹, Julissa Atía^{1,2}, Nicole Denon¹, Julián Gómez¹, Jéssica Guzmán¹, Carlos Schenone¹, Jorge Osio^{1,3}, Marcelo Cappelletti^{1,3}, Ramiro Irastorza^{1,4}, Martín Morales^{1,5}

¹ Programa TICAPPS, Univ. Nac. Arturo Jauretche, Florencio Varela (1888), Argentina.

² Comisión de Investigaciones Científicas, Pcia. de Bs. As. (CICPBA), La Plata, (1900), Argentina.

³ Grupo de Control Aplicado, Instituto LEICI (UNLP-CONICET), La Plata (1900), Argentina.

⁴ Instituto de Física de Líquidos y Sistemas Biológicos (CONICET), La Plata (1900), Argentina.

⁵ Centro UTN CODAPLI-FRLP, La Plata (1900), Argentina.

e-mail: lolivera.unaj@gmail.com

Resumen

Las líneas de Investigación y Desarrollo aquí presentadas tienen como objetivo general el de generar conocimiento en cuanto a las nuevas herramientas de las Tecnologías de la Información y la Comunicación para una innovación tecnológica de avanzada aplicadas a la mejora de la calidad de vida. Conjuntamente, se propicia la formación de recursos humanos, tanto de docentes investigadores como de estudiantes, dedicados a entrenarse al máximo nivel en el campo de la Inteligencia Artificial, en particular, en el procesamiento de imágenes mediante técnicas de Aprendizaje Profundo, propendiendo al desarrollo de soluciones innovadoras y de calidad. Específicamente, las líneas de investigación que se describen en este trabajo son: - la evaluación de la salud ósea, lo cual involucra problemas inversos en imágenes de microondas; - la detección e identificación de diversas malezas que pueden afectar a la producción de cultivos de la región; - la detección y clasificación de objetos reciclables con el fin de contribuir con el cuidado del medio ambiente; - la clasificación de diferentes condiciones de cielo como consecuencia de la cobertura de nubes, lo cual será de suma utilidad para la optimización de sistemas que aprovechen la energía solar como recurso energético.

Palabras clave: *Aprendizaje profundo, procesamiento de imágenes, aplicaciones.*

Contexto

Las líneas de I/D presentadas en este trabajo están incluidas dentro del Programa TICAPPS (TIC en 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

El concepto de Inteligencia Artificial (IA) se aplica a cualquier técnica que permita a las computadoras imitar la inteligencia humana a través de expresiones lógicas y esquemas abstractos [1]. Estas técnicas pueden ser utilizadas para modelar, identificar, optimizar, predecir y controlar el comportamiento dinámico de diferentes sistemas reales. Una de las áreas en donde se avanzó notablemente fue en la de detección de objetos y clasificación de imágenes. Esto se debe en su mayor parte al desarrollo de nuevas técnicas tanto de Machine Learning (Aprendizaje Automático) como de Deep Learning (Aprendizaje Profundo), además de las innovaciones en el manejo de Big Data (Datos a gran escala) y el aumento en la capacidad de cómputo mediante el uso de diferentes tecnologías como Cloud Computing (Computación en la Nube) o el uso de GPU (Unidad de Procesamiento Gráfico) para el análisis de información. Este avance puede verse en distintas áreas como: medicina, seguridad, turismo, finanzas, robótica, entre otras. Algunos ejemplos de dichos avances en el área se aplican en: control de vehículos autónomos, detección de rostros, detección de

matrículas, diagnóstico de enfermedades, realidad aumentada, etc.

A continuación, se presenta un breve marco teórico de cada una de las áreas de estudio investigadas.

Evaluación de la salud ósea:

La técnica de tomografía por microondas para obtener imágenes médicas ha despertado un gran interés en la última década, especialmente en aplicaciones de evaluación no invasiva y no ionizante de tejidos [2]. La aplicación que tiene mayor desarrollo en la actualidad es la de evaluación de cáncer de mama. Existen otras aplicaciones tales como la detección de accidentes cerebro vasculares y la evaluación de salud ósea [3]. La idea básica es “iluminar” el tejido con radiación electromagnética en microondas (frecuencias entre 100 MHz y 2 GHz de baja potencia) por medio de antenas emisoras. Estas antenas pueden funcionar también como receptoras, de este modo, el campo eléctrico dispersado es registrado en las mismas y luego mediante algoritmos de reconstrucción tomográfica se construye un mapa de las permitividades y conductividades de los tejidos evaluados resolviendo el denominado problema inverso. En tejidos óseos, esta tecnología se encuentra en reciente surgimiento y será la aplicación en que se enfocará esta línea de I/D. En los últimos años se han desarrollado muchos métodos estocásticos para la resolución del problema inverso, en particular, la utilización de redes neuronales ha tomado gran preponderancia [4].

Detección de malezas en cultivos de la región:

El área de influencia de la UNAJ posee extensas zonas rurales, que cuentan con pequeños, medianos y grandes productores hortícolas, basados fundamentalmente en los cultivos intensivos, donde se busca maximizar la producción en espacios reducidos, utilizando un solo tipo de producto. Los cultivos pueden ser afectados por diferentes amenazas, como por ejemplo las plagas,

inundaciones, incendios, animales silvestres, entre otras. En particular, esta línea de investigación propone el uso de procesamiento de imágenes utilizando redes neuronales para la detección, reconocimiento y clasificación de una amenaza específica, como lo son las diversas malezas que pueden afectar a la producción y a los procesos de cosecha, ya que en algunos casos esto es un impedimento para acceder a los propios cultivos. Existe una gran diversidad de malezas que aparecen dependiendo principalmente de la zona geográfica y de la estación del año. La aparición de estos yuyos viene acompañada con la resistencia a herbicidas y la dificultad para controlarlos. En Argentina, actualmente existen 15 malezas registradas como resistentes a herbicidas (glisofato) según un listado del Servicio Nacional de Sanidad y Calidad Agroalimentaria [5]. Por lo tanto, es importante detectar rápidamente la manifestación de estas hierbas malas, para evitar su propagación y reducir su banco de semillas. Por ejemplo, a finales del año 2019 en la provincia de Córdoba el yuyo colorado ocupaba 18.2 millones de hectáreas, este número es mayor a las hectáreas de cultivo de soja y maíz. Entre las malezas que se proponen identificar se encuentran el yuyo colorado, la rama negra y la roseta. Las dos primeras son comunes en las plantaciones de soja, uno de los cultivos que más se produce en el país, mientras que la tercer maleza es común en huertas y tiene la característica de ser “muy molesta” debido a que sus flores tienen pequeñas púas, lo que hace que en algunas regiones también perjudique al turismo.

Reciclado de objetos:

El reciclado de desechos es considerado de suma importancia para el cuidado del medioambiente en toda sociedad, debido a que, supone la reutilización de elementos u objetos ya utilizados, los que de otro modo serían desechados contribuyendo al aumento de la formación de basura y al daño ambiental permanente. En nuestro país, aunque se han

tomado medidas para fomentar el reciclado, solo un 24% de la población se esfuerza por separar los residuos para minimizar su generación y la contaminación. Gran parte del problema radica en el esfuerzo que requiere clasificar y separar los residuos inorgánicos, es por eso que en esta línea de investigación se propone desarrollar un sistema basado en redes neuronales artificiales que permita detectar y clasificar los objetos reciclables más comunes como papel, cartón, botellas, latas, etc, y los materiales con los que están hechos (plástico, vidrio, metal, papel), para minimizar la cantidad de residuos que se generan diariamente en las zonas urbanas. De esta manera se pretende contribuir no solo al cuidado del medio ambiente, sin la intervención del usuario, sino también a generar conciencia en la población en relación con el reciclado y su importancia para reducir la contaminación.

Estimación de la radiación solar:

La energía solar constituye una fuente limpia y renovable con bajo impacto ambiental, disponibilidad en el sitio de consumo y sostenibilidad, la cual se presenta como uno de los recursos más importantes tendiente a sustituir las fuentes de energía no renovables (carbón, petróleo, gas natural), reducir las emisiones de dióxido de carbono a la atmósfera y proteger nuestro medio ambiente. El conocimiento preciso de la radiación solar en un determinado instante y lugar es un factor clave para la evaluación y dimensionamiento óptimo de sistemas con diferentes principios de operación, como los sistemas fotovoltaicos para la generación de electricidad y los colectores solares para el calentamiento de fluidos. A su vez, la radiación solar es un recurso fundamental en el diseño arquitectónico y en la agricultura.

El problema que se presenta es que no siempre es posible contar con datos experimentales de la radiación solar en los lugares de interés. A su vez, la cobertura de nubes es un factor que atenúa y ocasiona intermitencia en los valores de la energía proveniente del sol. La aparición de grandes

nubes que tapan el disco solar, o bien los cambios repentinos en el tipo de nubosidad (por ejemplo, nube opaca y gruesa en otra nube más tenue y fina, o la división de una nube grande en numerosas nubes de pequeño tamaño), producen cambios repentinos en los valores de la radiación solar. Esta variabilidad provoca que la producción de energía que pueden generar los sistemas que aprovechen esta fuente como recurso energético no sea constante ni totalmente predecible en el tiempo, lo cual es una desventaja con respecto a los combustibles fósiles. Gracias al crecimiento de las capacidades de cálculo y al mejoramiento de los algoritmos implementados, en los últimos años se han comenzado a utilizar técnicas basadas en redes neuronales artificiales, como una buena alternativa para la estimación de la radiación solar con el mínimo error [6-8].

2. Líneas de Investigación y Desarrollo

Las líneas de I/D que se presentan en este trabajo están basadas en el estudio y desarrollo de herramientas alternativas para el procesamiento de imágenes a partir de redes neuronales, que presenten alto rendimiento en la precisión de los resultados, con un bajo costo de cómputo, facilidad de implementación y buena velocidad de convergencia. Las herramientas desarrolladas serán aplicadas para:

- (a) la evaluación de la calidad ósea, a través de la resolución del problema inverso en Tomografía de Microondas.
- (b) la detección e identificación de malezas, las cuales pueden afectar a la producción de los cultivos intensivos de la región.
- (c) la detección y clasificación de objetos reciclables, con el fin de contribuir con el cuidado del medio ambiente.
- (d) la clasificación de diferentes condiciones de cielo como

consecuencia de la cobertura de nubes, lo cual será de suma utilidad para la optimización de sistemas que aprovechen la energía solar como recurso energético.

3. Resultados y Objetivos

Resultados alcanzados:

La detección y clasificación de: malezas en cultivos; objetos reciclables; y diferentes condiciones de cielo, son líneas de estudio comenzadas recientemente.

Mientras que con respecto a la línea de investigación relacionada con la evaluación de la salud ósea, en el grupo se posee experiencia en el estudio de las propiedades dieléctricas de tejido óseo. En [9], estudiamos las propiedades dieléctricas de hueso trabecular humano en estado fisiológico in vitro. Uno de los resultados más llamativos fue la obtención de correlaciones negativas ($R \approx -0,8$) entre la permitividad y la relación volumen óseo / volumen total (BV/TV) para frecuencias entre 700 MHz y 1200 MHz ($p < 0,05$). En el trabajo reciente [10], se aplicaron las técnicas de diferencias finitas (FDTD) para la discretización de las ecuaciones de Maxwell y se simuló imágenes microtomográficas de tejido óseo trabecular en dos dimensiones. Se encontró que las propiedades dieléctricas efectivas podrían ser buenos predictores de la microestructura ósea trabecular. Más recientemente, estudiamos la sensibilidad del problema directo variando las propiedades del tejido óseo y de los tejidos circundantes en una potencial aplicación en el tobillo [11]. También se estudió el problema inverso utilizando algoritmos que se encuadran dentro de la estimación estocástica sin información de fase. Se utilizaron RNA para la evaluación de cilindros homogéneos y no homogéneos en dos dimensiones [12]. Por otro lado, también sin información de fase, se evaluó una potencial aplicación en tobillo (para determinación de salud ósea en el calcáneo) utilizando redes neuronales para resolver el problema inverso [13]. En este último caso, el modelo simplificado estimado

implicaba la estimación del centroide del calcáneo y un radio equivalente (considerándolo un cilindro), actualmente se propone estimar la geometría real.

Objetivos esperados:

Se espera aportar nuevos conocimientos relacionados con la evaluación de la salud ósea de manera no ionizante, a través de la reconstrucción de imágenes de microondas de modelos realistas de tobillo y/o muñeca humanos.

Se espera proporcionar información útil y contribuir a optimizar la producción de los cultivos intensivos, a partir de la detección, reconocimiento y clasificación de diferentes malezas, que pueden ser un impedimento para acceder a los cultivos, afectando a la producción de estos y a los procesos de cosecha.

Se espera contribuir con el cuidado del medio ambiente, a través de la detección, reconocimiento y clasificación de objetos reciclables, lo cual ayudará de manera significativa a la reutilización de elementos u objetos ya utilizados.

Se pretende determinar la atenuación de la radiación solar global, a partir de la clasificación de diferentes condiciones de cielo (despejado, parcial o totalmente nublado) y de variables meteorológicas de sencilla adquisición, lo cual podrá proporcionar información sumamente útil para adaptar la producción de energía a la presencia de nubes, logrando así una optimización de la respuesta de los sistemas que hagan un aprovechamiento de la energía solar.

4. 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, dedicados a entrenarse al máximo nivel en el campo de la Inteligencia Artificial, en

particular, en el procesamiento de imágenes mediante técnicas de Machine Learning como redes neuronales artificiales, propendiendo al desarrollo de soluciones innovadoras y de calidad, lo cual es de una relevancia evidente para el desarrollo tecnológico del territorio de la UNAJ.

El equipo de trabajo de las líneas de I/D que se presentan está conformado por cuatro docentes investigadores formados (tres Doctores y un Magister), cuatro docentes investigadores en formación, y tres estudiantes avanzados, todos autores de este artículo. Actualmente, se encuentran en curso tres Tesis de Doctorado y una beca de Estímulo a las Vocaciones Científicas del Consejo Interuniversitario Nacional (Becas EVC – CIN), relacionadas directamente con las líneas de I/D presentadas.

A su vez, todos los miembros docentes autores de este trabajo, participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

5. Bibliografía

- [1] N. Nilsson. “Principles of Artificial Intelligence”, Springer-Verlag, New York (1980).
- [2] M. Pastorino. “Microwave imaging”. John Wiley & Sons, (2010).
- [3] P.M. Meaney et al. “Clinical microwave tomographic imaging of the calcaneus: A first-in-human case study of two subjects”. *IEEE Transactions on Biomedical Engineering*, Vol. 59 (12), pp.3304-3313, (2012). DOI: 10.1109/TBME.2012.2209202
- [4] L. Li et al.. “DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering”. *IEEE Transactions on Antennas and Propagation*, Vol. 67 (3), pp.1819-1825, (2019). DOI: 10.1109/TAP.2018.2885437
- [5] Listado de malezas resistentes en Argentina <http://www.senasa.gob.ar/casos-confirmados-de-malezas-resistentes-en-argentina>
- [6] Z. Pang, F. Niu, Z. O’Neill. “Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons”. *Renewable Energy*. Vol. 156. 279–289, (2020). DOI: doi.org/10.1016/j.renene.2020.04.042
- [7] V. Çoban, S. Onar. In: Kahraman C. et al., “Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making”. *INFUS 2019. Advances in Intelligent Systems and Computing*, vol 1029. Springer, (2020).
- [8] P. Neelamegam, V. Amirtham. “Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms”. *Journal of Applied Research and Technology* Vol. 14. 206–214, (2016). DOI: doi.org/10.1016/j.jart.2016.05.001
- [9] R.M. Irastorza, E. Blangino, C.M. Carlevaro, F. Vericat. “Modeling of the dielectric properties of trabecular bone samples at microwave frequency”. *Medical & Biological Engineering & Computing*, Vol. 52, pp.439-447, (2014). DOI: doi.org/10.1007/s11517-014-1145-y
- [10] R.M. Irastorza, C.M. Carlevaro, F. Vericat. “Is there any information on micro-structure in microwave tomography of bone tissue?”. *Medical Engineering & Physics*, Vol. 358, pp.1173-1180, (2013). DOI: doi.org/10.1016/j.medengphy.2012.12.014
- [11] J.E. Fajardo, F. Vericat, G. Irastorza, C.M. Carlevaro, R.M. Irastorza. “Sensitivity analysis on imaging the calcaneus using microwaves”. *Biomedical Physics & Engineering Express*, Vol. 5 (4), pp.045039, (2019). DOI: doi.org/10.1088/2057-1976/ab3330
- [12] J.E. Fajardo, J. Galván, F. Vericat, C.M. Carlevaro, R.M. Irastorza, “Phaseless Microwave Imaging of Dielectric Cylinders: an Artificial Neural Networks-Based Approach”, *Progress In Electromagnetics Research*, Vol. 166, pp.95-105, (2019). DOI: [10.2528/PIER19080610](https://doi.org/10.2528/PIER19080610)
- [13] J.E. Fajardo, F.P. Lotto, F. Vericat, C.M. Carlevaro, R.M. Irastorza, “Microwave tomography with phaseless data on the calcaneus by means of artificial neural networks”, *Medical & Biological Engineering & Computing*, Vol. 58 (2), pp.433-442, (2020). DOI: [10.1007/s11517-019-02090-y](https://doi.org/10.1007/s11517-019-02090-y)

Big Data Optimization con algoritmos metaheurísticos utilizando frameworks de computación distribuida

Carolina Salto, Gabriela Minetti, Hugo Alfonso, Carlos Bermúdez, Javier Vargas, Franco Morero

Laboratorio de Investigación en Sistemas Inteligentes (LISI)

Facultad de Ingeniería - Universidad Nacional de La Pampa

Calle 110 Esq. 9 (6360) General Pico - La Pampa - Rep. Argentina

Te. / Fax: (02302) 422780/422372, Int. 6302

e-mail: {saltoc, minettig, alfonsoh, bermudezc}@ing.unlpam.edu.ar

Resumen

La comunidad científica ha encontrado en el uso de los recursos tecnológicos disponibles una aliada para abordar problemas de gran complejidad e identificados como irresolubles. Tales problemas han sido abordados con técnicas exactas o heurísticas para lograr su resolución, o al menos conseguir soluciones de alta calidad, cuando los mismos se clasifican como NP-duros. Inicialmente, los problemas se planteaban en entornos estáticos, pero en los últimos años se les trata de resolver reproduciendo las características dinámicas y de alta dimensionalidad que los alteran. La optimización de estos problemas, conocida como *Big Data Optimization*, se puede realizar diseñando algoritmos metaheurísticos secuenciales y distribuidos (*solvers*) bajo *frameworks* de programación de alto nivel como los que incorporan el paradigma MapReduce para el manejo de *Big Data*. Dichos *solvers*, en principio, serán diseñados y testeados con problemas académicos, con el objetivo de analizar el comportamiento en cuanto a eficiencia y escalabilidad. En consecuencia, nuestro objetivo central es adaptar estos *solvers* para abordar problemas de interés en contextos reales (científico, industrial, entre otros) donde estamos trabajando, y puntualmente en problemas de planificación y de diseño de redes de distribución de agua y de sensores en plantas industriales.

Palabras claves: Big Data, Optimización, Algoritmos metaheurísticos, Solvers

Contexto

Esta línea de investigación surge como una proyección del Proyecto "Técnicas inteligentes avanzadas y sistemas distribuidos aplicados a la resolución de problemas de decisión complejos" que se desarrolló hasta el 2020 en la Facultad de Ingeniería de la UNLPam. Este proyecto se lleva adelante en el Laboratorio de Investigación de Sistemas Inteligentes (LISI) de la Facultad de Ingeniería y es dirigido por la Dra. Salto.

En el LISI, desde su creación en 1998, nos hemos abocado al estudio de algoritmos cada vez más eficientes para la solución de problemas complejos, tanto de optimización como de diseño. En este dominio, el objetivo consiste en obtener algoritmos nuevos que den solución al problema y que necesiten un esfuerzo computacional más pequeño que los algoritmos existentes, así como caracterizar su comportamiento para las clases de problemas que demanda la comunidad científica e industrial en general. Cabe destacar que desde hace varios años, los integrantes de este laboratorio mantienen una importante vinculación con investigadores de las universidades argentinas, Universidad Nacional de San Luis y Universidad Nacional de Río Cuarto, y de la Universidad de Málaga (España), con quienes se realizan trabajos conjuntos.

Introducción

La optimización de los recursos es una exigencia que deben hacer frente la mayoría de las organizaciones en la actualidad. Dicha optimización no se

refiere a ahorrar o a suprimir, sino que se define como la mejor forma de realizar una actividad. Por lo tanto, los administradores deben tomar medidas urgentes y con visión a futuro. Estas decisiones se relacionan tanto con la optimización de procesos y recursos como con el pronóstico de las tendencias de mercado. La primera decisión es quizás la más sencilla de interpretar, mientras que la segunda es la más compleja, ya que la creatividad debe anteponerse a la racionalidad e imaginar cuáles serán las tendencias futuras. Hoy en día, los procesos de toma de decisiones son cada vez más complejos y globalizados debido a: (i) las dimensiones de dichos problemas cuando se tratan de abordar en entornos reales (industriales, científicos, entre otros), (ii) el carácter combinatorio de muchos de ellos y (iii) la naturaleza del objetivo que intenta alcanzar. Todo ello teniendo en cuenta criterios vinculados a la eficiencia del sistema, sus costos de explotación y de distribución, además de los tiempos de recepción, de ejecución y de entrega de materiales, servicios y productos. La gran mayoría de los problemas de decisión complejos del mundo real, cuando son modelados como problemas de optimización numérica con restricciones, pertenecen a la clase NP-duros. En las últimas décadas se ha comenzado a considerar los cambios dinámicos que generalmente se presentan tanto en la función objetivo como en las restricciones que se deben considerar en tales problemas, y algunos equipos de investigación han comenzado a abordarlos ([1], [2], [3], [4], [5], [6]). Este tipo de problema es referenciado en la literatura como Problema de Optimización con Restricciones Dinámico (Dynamic Constrained Optimization Problem (DCOP)) ([2], [7], [8], [9]). Se considera DCOP a todo aquel problema de búsqueda en el cual la función objetivo y/o sus restricciones cambian a lo largo del tiempo; cambios que se materializan en el espacio de búsqueda de las soluciones y por ende en cuáles sean las soluciones óptimas a encontrar. Pero su abordaje en entornos dinámicos está siendo recientemente considerado y debe contemplar variantes de tales cambios que requieren distintos tipos de implementación. Esas variantes son: (i) tanto la función objetivo como las restricciones son dinámicas ([10], [11], [12]); (ii) sólo la función objetivo

es dinámica pero las restricciones se mantienen estáticas ([13], [14], [15]) y (iii) la función objetivo es estática y las restricciones son dinámicas ([16], [17], [18]). En los tres casos, el espacio de búsqueda puede contar con áreas de soluciones que no son factibles de acuerdo a las restricciones, por lo tanto tales soluciones pueden requerir reparaciones para trasladarlas a zona factible. Cuando los problemas de optimización poseen un gran número de variables, los cálculos están drásticamente más allá de la capacidad de las computadoras de uso general. Ante estos problemas de optimización a gran escala (*Big Data Optimization*), a menudo se adoptan estrategias tácticas, como la hibridación de diferentes tipos de algoritmos, el uso de búsqueda local, enfoques para la adaptación de parámetros y la relajación de restricciones, todo contribuye a resolver el problema en tiempo polinomial. Por lo tanto, cómo resolver la optimización a gran escala plantea un serio desafío.

La computación distribuida ha cambiado enormemente el panorama de *Big Data*. A diferencia de la computación centralizada tradicional que escala el hardware para aumentar la capacidad de cómputo, la computación distribuida incrementa el hardware con la incorporación de computadoras más pequeñas con la potencia equivalente a una supercomputadora. En consecuencia, para resolver los problemas de *Big Data Optimization*, se descompone el problema original de tal manera que disminuya la complejidad computacional. Uno de los paradigmas de computación distribuida más usados para procesar *Big Data* es MapReduce (MR), propuesto por Google [19]. MR divide grandes volúmenes de datos en porciones más pequeñas (chunks), las cuales son procesadas en paralelo [20]. Hadoop es un *framework open source* y muy popular que implementa el paradigma MR [21], tanto en la industria como en el ámbito académico. Este *framework* provee una plataforma distribuida lista para usar, la cual es fácil de programar, escalable, confiable, tolerante a fallas, además planifica automáticamente las tareas paralelas [22]. La filosofía de Hadoop es enviar el código a los sitios donde están los datos almacenados en el sistema de archivos distribuido y procesar los

datos de forma local y simultánea. Posteriormente ha surgido un *framework open source* derivado de Hadoop, Apache Spark, que está atrayendo cada vez más interesados en la comunidad de *Big Data*. Como contiene resultados intermedios en la memoria en lugar de almacenarlos en el disco (como lo hace Hadoop), los trabajos MR de varias pasadas se pueden llevar a cabo minimizando el número de operaciones de entrada/salida en el sistema de archivos distribuido. Por lo tanto, logra un mayor rendimiento que Hadoop. Otro *framework open source* que implementa MR es MR-MPI [23], que permite un mayor control de la plataforma paralela permitiendo mejorar la performance del ancho de banda y reducir los costos de latencia. El paradigma MR, también, contribuye a construir nuevos modelos de optimización y de *machine learning*, en particular algoritmos metaheurísticos escalables, como *solvers* de problemas de optimización combinatoria, que despiertan un gran interés en la comunidad científica e industrial. En la literatura, muchos investigadores informan de algoritmos metaheurísticos programados en Hadoop ([24], [25], [26], [27], [28]) y Spark ([29], [30], [31]), y algunos bajo MR-MPI [32], según el conocimiento de los autores.

En la actualidad se tiende a abordar el estudio de problemas de *Big Data Optimization* con el uso de *solvers*, basados en metaheurísticas, secuenciales y distribuidos bajo *frameworks* de programación de alto nivel como los que incorporan el paradigma MapReduce. Dentro de los problemas de interés actual, que en el LISI venimos abordando, se encuentran los de diseño óptimo de redes de agua ([33], [34], [35]) y de sensores en plantas industriales ([36], [37]), como también la planificación de tareas [38], entre otros. Generalmente, los *solvers* se prueban con problemas académicos, con el objetivo de analizar su desempeño. Esto permite estudiar las características, parámetros, comportamiento en cuanto a calidad de soluciones, entre otras métricas. También se aprovechan estas pruebas para evaluar la escalabilidad de los algoritmos propuestos. Los problemas usualmente considerados para estos tests son: Knapsack ([39], [40]), OneMax [41], MaxCut [42], entre otros.

Línea de investigación y desarrollo

Los problemas de optimización dinámicos y, también, los que incluyan un gran número de variables (alta dimensionalidad) forman parte de lo que se conoce como *Big Data Optimization*. Estos se pueden resolver diseñando algoritmos metaheurísticos secuenciales y distribuidos (denominados *solvers*) bajo *frameworks* de programación de alto nivel como los que incorporan el paradigma MapReduce para el manejo de *Big Data*. Dichos *solvers*, en principio, serán diseñados y testeados con problemas académicos, con el objetivo de estudiar las características, parámetros, comportamiento en cuanto a calidad de soluciones y la escalabilidad de los algoritmos propuestos. Posteriormente, estos *solvers* serán adaptados para dar solución a problemas de diseño de redes de distribución de agua y de sensores en plantas industriales, además de problemas de planificación. Para una mejor comprensión del problema en su totalidad, y de su grado de complejidad, se formulan las siguientes preguntas:

1: ¿Cuál sería el diseño de *solvers* eficientes para resolver los problemas de *Big Data Optimization* planteados?

2: ¿Cuál de los *frameworks* disponibles actualmente permite a los *solvers* desarrollados alcanzar su mejor rendimiento, en cuanto a eficacia y eficiencia, para escalar a casos de estudio de alta complejidad y dimensionalidad?

3: ¿Cómo escalan estos *solvers* al considerar el incremento de los recursos computacionales para ejecutarlos?

4: ¿Cómo varían los tiempos de respuesta de los *solvers* al utilizar computación distribuida?

5: ¿Cuál es la eficiencia de los *solvers* desarrollados en comparación con los propuestos en la literatura para abordar los problemas de *Big Data Optimization* tratados?

El objetivo general de esta línea de investigación es evaluar la eficiencia de los *solvers* diseñados para resolver los problemas de *Big Data Optimization* académicos y los relacionados con el diseño óptimo de redes y de planificación, haciendo uso de *frameworks* de programación de alto nivel como los que incorporan el paradigma MapReduce.

Para alcanzar el objetivo general propuesto se establecen los siguientes objetivos específicos a seguir durante los próximos cuatro años:

1: Diseñar e implementar los *solvers* para resolver los problemas de *Big Data Optimization* planteados usando *frameworks* para el tratamiento de *Big Data*.

2: Realizar la experimentación para poder determinar el *framework* más adecuado y así mejorar el rendimiento, en cuanto a eficacia y eficiencia, de estos *solvers*.

3: Analizar la escalabilidad de los *solvers* en la ejecución al incrementar la disponibilidad de recursos computacionales (cantidad de nodos, capacidad de procesamiento, memoria, entre otros).

4: Estudiar los tiempos de respuesta de los *solvers* al utilizar computación distribuida.

5: Comparar los *solvers* desarrollados con los propuestos en la literatura para abordar los problemas de *Big Data Optimization* planteados desde los puntos de vista de eficiencia y eficacia.

Resultados esperados

Con este proyecto se espera que los *solvers* propuestos para la resolución de problemas de *Big Data Optimization* proporcionen soluciones eficientes de alta calidad a los casos de gran complejidad y alta dimensionalidad. También se espera aportar diseños de *solvers* innovadores y eficientes en el campo del diseño de redes de distribución de agua y de sensores en plantas industriales y de planificación.

Formación de recursos humanos

Cada año se incorporan al proyecto alumnos avanzados en la carrera Ingeniería en Sistemas y becarios de investigación. La tarea que se les asigna está relacionada a la solución de problemas de optimización usando técnicas inteligentes, con el objeto de guiarlos en el desarrollo de sus tesis de grado y, también, de formar futuros investigadores científicos. Por otra parte, los docentes-investigadores que integran el proyecto realizan diversos cursos de posgrado relacionados con la temática del proyecto, con el objetivo de sumar los créditos necesarios para cursar carreras de posgrado.

REFERENCES

- [1] E. Mezura-Montes and C. A. Coello Coello, "Constraint-handling in nature-inspired numerical optimization: Past, present and future," *Swarm and Evolutionary Computation*, vol. 1, no. 4, pp. 173–194, 2011.
- [2] T. T. Nguyen and X. Yao, "Continuous dynamic constrained optimization—the challenges," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 6, pp. 769–786, 2012.
- [3] T. T. Nguyen and X. Yao, "Evolutionary optimization on continuous dynamic constrained problems - an analysis," in *Evolutionary Computation for Dynamic Optimization Problems*, S. Yang and X. Yao, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 193–217.
- [4] M. Mavrouniotis and S. Yang, "Ant colony optimization for dynamic combinatorial optimization problems," pp. 121–142, 2018.
- [5] M. Ameca-Alducin, E. Mezura-Montes, and N. Cruz-Ramirez, "Dynamic differential evolution with combined variants and a repair method to solve dynamic constrained optimization problems: An empirical study," *Soft Comput.*, vol. 22, no. 2, p. 541–570, 2018.
- [6] J. K. Kordestani and M. R. Meybodi, *Application of Sub-Population Scheduling Algorithm in Multi-Population Evolutionary Dynamic Optimization*. John Wiley Sons, Ltd, 2020, ch. 7, pp. 169–211.
- [7] T. T. Nguyen and Xin Yao, "Benchmarking and solving dynamic constrained problems," in *2009 IEEE Congress on Evolutionary Computation*, 2009, pp. 690–697.
- [8] H. K. Singh, A. Isaacs, T. T. Nguyen, T. Ray, and Xin Yao, "Performance of infeasibility driven evolutionary algorithm (idea) on constrained dynamic single objective optimization problems," in *2009 IEEE Congress on Evolutionary Computation*, 2009, pp. 3127–3134.
- [9] S. Zeng, R. Jiao, C. Li, X. Li, and J. S. Alkassabeh, "A general framework of dynamic constrained multiobjective evolutionary algorithms for constrained optimization," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2678–2688, 2017.
- [10] M. Andrews and A. Tuson, "Dynamic optimization: An investigation of practitioners requirements," *Proceedings of the 24th Annual Workshop of the UK Planning and Scheduling Special Interest Group*, 2005.
- [11] Y. Wang and M. Wineberg, "Estimation of evolvability genetic algorithm and dynamic environments," *Genetic Programming and Evolvable Machines*, vol. 7, no. 4, p. 355–382, Dec. 2006. [Online]. Available: <https://doi.org/10.1007/s10710-006-9015-5>
- [12] D. M. Prata, E. L. Lima, and J. C. Pinto, "Simultaneous data reconciliation and parameter estimation in bulk polypropylene polymerizations in real time," *Macromolecular Symposia*, vol. 243, no. 1, pp. 91–103, 2006.
- [13] L. Araujo and J. J. Merelo, "A genetic algorithm for dynamic modelling and prediction of activity in document streams," in *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 1896–1903. [Online]. Available: <https://doi.org/10.1145/1276958.1277340>
- [14] P. Tawdross, S. K. Lakshmanan, and A. König, "Intrinsic evolution of predictable behavior evolvable hardware in dynamic environment," in *2006 Sixth International Conference on Hybrid Intelligent Systems (HIS'06)*, 2006, pp. 60–60.

- [15] M. Rocha, J. Neves, A. Veloso, E. Ferreira, and I. Rocha, *Evolutionary Algorithms for Static and Dynamic Optimization of Fed-batch Fermentation Processes*, 01 2005, pp. 288–291.
- [16] K. Deb, U. N. and K. Sindhya, “Dynamic multi-objective optimization and decision-making using modified nsga-ii: A case study on hydro-thermal power scheduling,” 05 2007, pp. 803–817.
- [17] P. Ioannou, A. Chassiakos, H. Jula, and R. Unglaub, “Dynamic optimization of cargo movement by trucks in metropolitan areas with adjacent ports,” 2002. [Online]. Available: <https://rosap.ntl.bts.gov/view/dot/15509>
- [18] K. Mertens, T. Holvoet, and Y. Berbers, “The dyncooa algorithm for dynamic constraint optimization problems,” in *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, ser. AAMAS '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 1421–1423. [Online]. Available: <https://doi.org/10.1145/1160633.1160898>
- [19] J. Dean and S. Ghemawat, “Mapreduce: Simplified data processing on large clusters,” in *OSDI*, 2004.
- [20] A. Cano, C. García-Martínez, and S. Ventura, “Extremely high-dimensional optimization with mapreduce: Scaling functions and algorithm,” *Information Sciences*, vol. 415–416, pp. 110–127, 11 2017.
- [21] T. White, *Hadoop: The Definitive Guide*, 1st ed. O’Reilly Media, Inc., 2009.
- [22] I. Hashem, N. Anuar, A. Gani, I. Yaqoob, F. Xia, and S. Khan, “Mapreduce: Review and open challenges,” *Scientometrics*, vol. 109, 04 2016.
- [23] S. J. Plimpton and K. D. Devine, “Mapreduce in mpi for large-scale graph algorithms,” *Parallel Computing*, vol. 37, no. 9, pp. 610–632, 2011, emerging Programming Paradigms for Large-Scale Scientific Computing.
- [24] A. Cano, C. García-Martínez, and S. Ventura, “Extremely high-dimensional optimization with mapreduce: Scaling functions and algorithm,” *Information Sciences*, vol. 415–416, pp. 110–127, 2017.
- [25] F. Ferrucci, P. Salza, and F. Sarro, “Using hadoop mapreduce for parallel genetic algorithms: A comparison of the global, grid and island models,” *Evolutionary Computation*, vol. 26, no. 4, pp. 535–567, 2018.
- [26] L. Di Geronimo, F. Ferrucci, A. Murolo, and F. Sarro, “A parallel genetic algorithm based on hadoop mapreduce for the automatic generation of junit test suites,” in *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, 2012, pp. 785–793.
- [27] A. Verma, X. Llorà, D. E. Goldberg, and R. H. Campbell, “Scaling genetic algorithms using mapreduce,” in *2009 Ninth International Conference on Intelligent Systems Design and Applications*, 2009, pp. 13–18.
- [28] A. Verma, X. Llorà, S. Venkataraman, D. E. Goldberg, and R. H. Campbell, “Scaling ecga model building via data-intensive computing,” in *IEEE Congress on Evolutionary Computation*, 2010, pp. 1–8.
- [29] C. Hu, G. Ren, C. Liu, M. Li, and W. Jie, “A spark-based genetic algorithm for sensor placement in large scale drinking water distribution systems,” *Cluster Computing*, vol. 20, 06 2017.
- [30] C. Paduraru, M.-C. Melemciuc, and A. Stefanescu, “A distributed implementation using apache spark of a genetic algorithm applied to test data generation,” in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, ser. GECCO '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 1857–1863. [Online]. Available: <https://doi.org/10.1145/3067695.3084219>
- [31] R. Qi, Z.-J. Wang, and S.-Y. Li, “A parallel genetic algorithm based on spark for pairwise test suite generation,” *Journal of Computer Science and Technology*, vol. 31, pp. 417–427, 03 2016.
- [32] C. Salto, G. Minetti, E. Alba, and G. Luque, *Developing Genetic Algorithms Using Different MapReduce Frameworks: MPI vs. Hadoop: 18th Conference of the Spanish Association for Artificial Intelligence, CAEPIA 2018, Granada, Spain, October 23–26, 2018, Proceedings*, 01 2018, pp. 262–272.
- [33] C. Bermudez, G. Minetti, and C. Salto, “SA to optimize the multi-period water distribution network design,” in *XXIX Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, 2018, pp. 12–21.
- [34] C. Bermudez, C. Salto, and G. Minetti, “Designing a multi-period water distribution network with a hybrid simulated annealing,” in *XLVIII JAIIO: XX Simposio Argentino de Inteligencia Artificial (ASAI 2019)*, 2019, pp. 39–52.
- [35] H. Alfonso, C. Bermudez, G. Minetti, and C. Salto, “A real case of multi-period water distribution network design solved by a hybrid SA,” in *XXVI Congreso Argentino de Ciencias de la Computación – CACIC 2020*, no. 1-12, 2020.
- [36] J. Hernandez, C. Salto, G. Minetti, M. Carnero, and M. C. Sanchez, “Hybrid simulated annealing for optimal cost instrumentation in chemical plants,” *Chemical Engineering Transactions*, vol. 74, pp. 709–714, May 2019. [Online]. Available: <https://www.cetjournal.it/index.php/cet/article/view/CET1974119>
- [37] J. Hernandez, C. Salto, G. Minetti, M. Carnero, C. Bermudez, and M. Sanchez, “Optimal instrumentation: Adjustment and hybridization of a simulated annealing based technique,” in *XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019)*, Oct. 2019, pp. –.
- [38] F. Morero, C. Salto, and C. Bermudez, “Parallelism and hybridization in differential evolution to solve the flexible job shop scheduling problem,” *Journal of Computer Science & Technology*, vol. 20, no. 1, pp. 33–42, 2020.
- [39] K. Klamroth and M. M. Wiecek, “Time-dependent capital budgeting with multiple criteria,” in *Research and Practice in Multiple Criteria Decision Making*, Y. Y. Haimes and R. E. Steuer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 421–432.
- [40] H. Kellerer, U. Pferschy, and D. Pisinger, *Introduction to NP-Completeness of Knapsack Problems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 483–493.
- [41] J. Schaffer and L. Eshelman, “On crossover as an evolutionarily viable strategy,” in *ICGA*, 1991.
- [42] M. X. Goemans and D. P. Williamson, “Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming,” *J. ACM*, vol. 42, no. 6, p. 1115–1145, Nov. 1995. [Online]. Available: <https://doi.org/10.1145/227683.227684>

CIUDADES INTELIGENTES Y SOSTENIBLES: INICIATIVAS Y DESAFÍOS

Villagra A.¹, Errecalde M.^{1,2}, Pandolfi D.¹, Mercado V.¹, Torres M.¹, Molina D.¹, Valdéz J.¹,
López M.¹

¹Laboratorio de Tecnologías Emergentes (LabTEM)
Instituto de Tecnología Aplicada (ITA) - Unidad Académica Caleta Olivia
Universidad Nacional de la Patagonia Austral

²Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)
Departamento de Informática - Universidad Nacional de San Luis

avillagra@uaco.unpa.edu.ar, merreca@unsl.edu.ar, {dpandolfi,
vmercado}@uaco.unpa.edu.ar, marianagalos@gmail.com, {dmolina, [jvaldez,
milopez](mailto:jvaldez.milopez@uaco.unpa.edu.ar)}@uaco.unpa.edu.ar

RESUMEN

Uno de los grandes retos que los países van a enfrentar en este siglo es la planificación, administración y gobernanza de las ciudades de forma sostenible, maximizando las oportunidades económicas y minimizando los daños medioambientales.

Todas las aglomeraciones urbanas presentan diversos retos. Cada vez más, las grandes ciudades son vistas como sistemas complejos con conexiones entre sus diferentes ambientes e individuos. Por ello, son muy importantes la planificación urbana y el desarrollo de mecanismos de decisión dinámicos que tomen en cuenta el crecimiento y la inclusión de procesos de participación ciudadana. Los problemas en las ciudades inteligentes son variados (economía, movilidad, gobernanza, personas, vida y entorno), multidisciplinares por naturaleza, aparentemente inconexos y requieren un vasto aparato científico y tecnológico.

Esta línea de trabajo se presenta una propuesta de investigación enfocada en los desafíos relacionados con la movilidad, medioambiente y

gobernanza inteligente. Se hace desde la perspectiva de construir nuevos prototipos basados en sistemas inteligentes, mejorados con metodología y tecnologías diferentes con el fin de exhibir "inteligencia holística".

Palabras clave: Ciudades Inteligentes, Optimización, Sistemas Inteligentes, Aprendizaje

CONTEXTO

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/B273 "Ciudades inteligentes y sostenibles: iniciativas y desafíos". Este proyecto se desarrolla en cooperación con el LIDIC de la UNSL, y el Grupo NEO de la UMA (España).

1. INTRODUCCIÓN

En la actualidad, debido a factores como la globalización, el crecimiento de-

mográfico y la evidente migración de los entornos rurales a los urbanos, se ha generado en las ciudades la necesidad de enfrentar nuevos retos relacionados con los procesos de planificación territorial urbana, la optimización de recursos, mejora en las condiciones económicas, gestión de la prestación de servicios y sostenibilidad medioambiental.

Los países enfrentan una serie de retos para atender las necesidades de las poblaciones en crecimiento, comenzando con elementos básicos como infraestructura, saneamiento, transporte, energía, vivienda, seguridad, empleo, salud y educación, y pasando por otros también fundamentales como comunicación y esparcimiento.

Existen diversas definiciones de una Ciudad Inteligente (CI) [CA19] [AEN17], se puede decir que es la visión holística de una ciudad que aplica las Tecnologías de la Información y Comunicación (TIC) para la mejora de la calidad de vida y la accesibilidad de sus habitantes y asegura un desarrollo sostenible económico, social y ambiental en mejora permanente.

El proceso para generar una CI demanda una serie de proyectos con una visión holística de todas las necesidades de la ciudad y, por lo tanto, de todo lo que se puede ofrecer en este contexto con la utilización de diversas tecnologías como Internet de las cosas (en inglés *Internet of Things*, IoT) [ZB+14], Big Data y Datos abiertos [DH+17], [BK20], Sistemas distribuidos a gran escala [VP+12], Cloud Computing [KA+13], Sistemas Inteligentes [SA14a], [VA+20] y la Inteligencia Artificial son algunas de las herramientas para resolver problemas complejos en todos los ámbitos de trabajo de la ciudad, tanto en contextos públicos como privados, y en contacto con sistemas ciberfísicos [L08] basados en IoT ([B06], [SR+10], [LT20], [W10]).

Según [MC+14] una CI es la que realiza actividades en al menos una de las seis áreas de acción inteligente: Economía inteligente, Sociedad inteligente, Gobernanza inteligente, Movilidad inteligente, Medioambiente inteligente y Modo de vida inteligente. Para abordar cualquiera de las áreas de acción inteligentes, y desarrollar sistemas inteligentes será necesario la aplicación de técnicas de vanguardia de diferentes dominios que permitan flexibilidad, autoadaptabilidad, robustez, alta dimensionalidad (escalabilidad) y eficiencia.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se describen las líneas de investigación que se llevan a cabo en el proyecto.

Las áreas de interés en este proyecto de investigación son: (a) Movilidad: tránsito verde; (b) Medioambiente: gestión inteligente de residuos; energía inteligente y (c) Gobernanza: chatbots aplicados a trabajo no formal colaborativo y medioambiente urbano.

(a) Movilidad: tránsito verde: En ciudades grandes son objetivos comunes controlar, disciplinar el tránsito y reducir accidentes en la ciudad, invirtiendo en sistemas de monitoreo y administración de tráfico. Los resultados que deben lograrse incluyen, por ejemplo, el uso de radares de velocidad, la programación adaptativa y en tiempo real de los semáforos, tomando en cuenta, entre otros factores, la concentración y el flujo de vehículos (dando prioridad a las ambulancias y los vehículos policiales y un corredor preferencial para colectivos), la concentración de peatones y la velocidad de los vehículos. Otra fuente de preocupación común es la oferta de sistemas de transporte público más eficiente, adecuado al desarrollo urbano y a la equidad social en relación con los des-

plazamientos. Muchas de las soluciones tienen el objetivo de preparar a la ciudad para la implementación, en el futuro, de un sistema multimodal de transporte, que incluya diferentes medios (bicicleta, subtes, franjas exclusivas para colectivos, vehículos livianos sobre rieles), y contribuya así a la reducción del consumo de combustibles, la emisión de gases y el tiempo de traslado, así como al mejoramiento de la calidad del aire.

La optimización de la movilidad inteligente surgió para reducir la contaminación generada por el tráfico [SR+03], con excelentes resultados (hasta un 50% de reducción). Una manera de reducir el impacto negativo del tráfico consiste en optimizar las rutas. Cada ruta tiene costos asociados: tiempo, dinero, contaminación. En consecuencia, se debe optimizar simultáneamente más de un objetivo (por ejemplo, tiempo versus contaminación) ([AR+05], [SA14a], [SA14b], [VA+20]).

Además, los datos utilizados en la optimización no son precisos (contienen errores) y varían durante un viaje. Existen trabajos que se centran en proporcionar rutas personalizadas a las necesidades de los ciudadanos y modificarlas según el estado actual de las rutas y el tráfico [P95].

(b) Medio ambiente inteligente: Los recursos como el agua y la energía son cada vez más escasos. Por ello, se requiere su uso racional e inteligente. Esto se refiere no solo a aumentar la eficiencia durante el consumo, sino también a preservar los manantiales, a utilizar fuentes renovables e, inclusive, a recolectar y destinar los residuos de manera adecuada.

El modelo de generación-distribución-uso de la energía eléctrica es actualmente el mismo que existía hace décadas siguiendo un esquema unidireccional de información en el que no hay realimentación del consumo de energía por parte de los hogares.

Los desequilibrios entre capacidad de generación y consumo tienen como consecuencia que parte de la energía no se aproveche. En este sentido surge la necesidad de desplegar redes inteligentes (*Smart Energy Grids*) para la gestión de la energía, que integren las fuentes de energías renovables en las actuales redes eléctricas [JL+15].

La utilización eficiente de la energía juega un papel muy importante para el desarrollo de redes inteligentes en el sistema eléctrico. Por lo tanto, la supervisión y el control adecuados del consumo de energía es una de las principales prioridades de la red inteligente [T11]. Las redes inteligentes facilitan a los clientes instrumentos que les permiten optimizar su propio consumo eléctrico y mejorar el funcionamiento del sistema global [HA+10]. Para que estas aplicaciones sean posibles se hace necesaria una interacción entre las infraestructuras de comunicación avanzadas, las tecnologías de la información y el uso de inteligencia computacional [FD+18], [BY+18].

Por otro lado, la gestión adecuada de los residuos urbanos es otro tema de creciente preocupación para los agentes públicos, con impactos directos sobre la salud, el medioambiente y la calidad de vida de la población ([AZ+17], [HA+18], [FA19]).

(c) Finalmente, la gobernanza inteligente figura en la mayoría de los estudios sobre CI como uno de los ejes más importantes para el desarrollo en un futuro próximo [BA+12]. Tomar decisiones basadas en la analítica de datos y con el objetivo de optimizar los recursos urbanos es el único camino a la excelencia. Los grandes volúmenes de datos disponibles en el sector público pueden ser utilizados por la ciudadanía tanto para generar nuevas informaciones y servicios, como para facilitar la participación ciudadana en las Administraciones públicas (tanto

en la toma de decisiones como en el análisis de los datos para incrementar la transparencia pública o fortalecer la integridad en las Administraciones públicas). Por esta razón, es importante contar con métodos efectivos para la extracción de conocimiento desde datos textuales, sino también proveer medios para que los ciudadanos se puedan comunicar con los sistemas de computación de una manera sencilla e intuitiva como cuando se comunica con otra persona utilizando el lenguaje natural.

En este eje, las innumerables historias de éxito observadas en su aplicación en empresas del área privada, ha llevado a un creciente interés en el uso de sistemas inteligentes conversacionales (SIC) tales como chatbots, asistentes virtuales y sistemas de respuestas automáticas en diversas áreas de gobierno como la salud y la seguridad social, el área impositiva, la educación, el turismo y la administración de justicia, entre otros [AK+19].

Existe un déficit de profesionales con experiencia en el desarrollo de SICs debido a los desafíos que plantea el desarrollo de sistemas que pueden interactuar en lenguaje natural convirtiéndola en un área de investigación sumamente activa e interesante tanto desde el punto de vista científico como por sus aplicaciones en el sector público y privado. En este contexto, el desarrollo de nuevos enfoques neuronales de aprendizaje profundo está revolucionando el procesamiento del lenguaje natural, como es el caso de aquellos basados en Transformers como BERT [DC+19] y/o basados en generación del lenguaje natural como GPT-3 [BM+20]. Estos métodos comienzan a ser integrados por las grandes empresas tecnológicas en herramientas para el desarrollo de chatbots como es el caso de Google con Dialogflow (<https://cloud.google.com/dialogflow/docs/>), Amazon con Lex (<https://aws.amazon.com/es/lex/>) e IBM con su Watson Natural Language Understanding (<https://www.ibm.com/ca-en/marketplace/natural-language-understanding>), pero también en plataformas de código abierto como Rasa (<https://rasa.com/>) o DeepPavlov (<https://deeppavlov.ai/>). Esto está abriendo un sinnúmero de oportunidades para la aplicación de estos avances en SICs y su uso en dominios como las ciudades y la gobernanza inteligentes.

3. RESULTADOS OBTENIDOS/ ESPERADOS

3. RESULTADOS OBTENIDOS/ ESPERADOS

En particular, en cuanto a movilidad hemos abordado el problema de programación de semáforos con dos versiones de un Algoritmo Genético Celular, cGA (síncrono y asíncrono) para resolver instancias grandes y reales. Nuestros algoritmos superan las técnicas de vanguardia y las configuraciones de expertos. Se llevaron a cabo varios análisis en profundidad de los resultados, estudio genotípico y fenotípico (para mayor detalle ver [VA+20]). La investigación futura se centrará en el número de evaluaciones con el objetivo de reducir el esfuerzo computacional en términos de tiempo de procesamiento. Además, planeamos definir una función de aptitud más precisa que explore el espacio de búsqueda de manera más eficiente.

Las otras dos áreas son nuevas en este proyecto y los objetivos esperados son (i) Estudio de la situación actual de la ciudad, identificación de problemas en las áreas de movilidad, medioambiente y gobernanza particularizando en los problemas enunciados en los objetivos. Planteamiento de posibles soluciones. (ii) Estudio de técnicas y herramientas (software y hardware) principalmente Metaheurísticas, IoT, Big Data, Aprendizaje de máquina, Aprendizaje profundo, aplicadas a solucionar los problemas identificados.

Creación de instancias y búsqueda de benchmark. (iii) Estudio y aplicación de técnicas y herramientas a problemas en redes eléctricas, tránsito y gobernanza. (iv) Diseño, construcción, evaluación y análisis de prototipos.

Finalmente, se pretende colaborar con los gobiernos locales para la implementación de políticas y acciones inteligentes y sostenibles que impacten en la calidad de vida de los ciudadanos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por dos Doctores y dos Magisters en Ciencias de la Computación, un Magister en Matemática Avanzada, dos Ingenieros en Sistemas y un estudiante de la Carrera Ingeniería en Sistemas de la UNPA.











Este proyecto de investigación proporcionará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes. De igual forma, será un ámbito adecuado para la realización de tesis de grado. Actualmente un integrante está desarrollando su tesis de Maestría y un integrante su tesis de Doctorado. Además, se cuenta un becario de grado.

5. BIBLIOGRAFIA

- [AEN17] AENOR/CTN 178. UNE 178201 - Smart Cities. Definition, attributes and requirements. [Online] Available [Accessed:28-Jun-2017]: <http://www.aenor.es>
- [AK+19] Androutsopoulou, A., Karacapilidis, N., Loukis, E., Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots, *Government Information Quarterly*, 36, (2),2019, 358-367.
- [AR+05] Arnott, R., Rave, T., & Schöb, R. (2005). *Alleviating urban traffic congestion*. MIT Press Books, 1.
- [AZ+17] Anagnostopoulos, T., Zaslavsky, A., Kolomvatsos, K., Medvedev, A., Amirian, P., Morley, J., & Hadjieftymiades, S. (2017). Challenges and opportunities of waste management in IoT-enabled smart cities: A survey. *IEEE Transactions on Sustainable Computing*, 2(3), 275-289.
- [B06] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- [BA+12] Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., ... & Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics*, 214(1), 481-518.
- [BK20] Bibri, S. E., & Krogstie, J. (2020). The emerging data-driven Smart City and its innovative applied solutions for sustainability: the cases of London and Barcelona. *Energy Informatics*, 3(1), 1-42.
- [BM+20] Brown, T., Mann, B., Ryder, N. et al. Language Models are Few-Shot Learners (2020). <https://arxiv.org/abs/2005.14165>. May 2020.
- [BY+18] Barman, B. K., Yadav, S. N., Kumar, S., & Gope, S. (2018, June). IOT based smart energy meter for efficient energy utilization in smart grid. In 2018 2nd International Conference on Power, Energy and Environment: Towards Smart Technology (ICEPE) (pp. 1-5). IEEE
- [CA19] Camero, A., & Alba, E. (2019). Smart City and information technology: A review. *cities*, 93, 84-94.
- [DC+19] Devlin, J., Chang, M., Lee, K. and Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (2019). *Proc. of NAACL-HLT 2019*, 4171-4186, ACL.
- [DH+17] Daki, H., El Hannani, A., Aqqal, A., Haidine, A., & Dahbi, A. (2017). Big Data management in smart grid: concepts, requirements and implementation. *Journal of Big Data*, 4(1), 1-19.
- [FA19] Ferrer, J., & Alba, E. (2019). BIN-CT: Urban waste collection based on predicting the container fill level. *Biosystems*, 186, 103962.
- [FD+18] Fallah, S. N., Deo, R. C., Shojafar, M., Conti, M., & Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: state of the art, future challenges, and research directions. *Energies*, 11(3), 596.
- [HA+10] Hutterer, S., Auinger, F., Affenzeller, M., & Steinmaurer, G. (2010). Overview: a simulation based metaheuristic optimization approach to optimal power dispatch related to a smart electric grid. In *Life System Modeling and Intelligent Computing* (pp. 368-378). Springer, Berlin, Heidelberg.
- [HA+18] Hannan, M. A., Akhtar, M., Begum, R. A., Basri, H., Hussain, A., & Scavino, E. (2018). Capacitated vehicle-routing problem model for scheduled solid waste collection and route optimization

- using PSO algorithm. *Waste management*, 71, 31-41.
- [JL+15] Jenkins, N., Long, C., & Wu, J. (2015). An overview of the smart grid in Great Britain. *Eng.*, 1(4), 413-421.
- [KA+13] Khan, Z., Anjum, A., & Kiani, S. L. (2013, December). Cloud based big data analytics for smart future cities. In 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing (pp. 381-386). IEEE.
- [LA08] Lee, E. A. (2008, May). Cyber physical systems: Design challenges. In 2008 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC) (pp. 363-369). IEEE.
- [LT20] Lebrusán, I., & Toutouh, J. (2020). Using Smart City Tools to Evaluate the Effectiveness of a Low Emissions Zone in Spain: Madrid Central. *Smart Cities*, 3(2), 456-478.
- [MC+14] Manville, C., Cochrane, G., Cave, J., Millard, J., Pederson, J. K., Thaarup, R. K., ... & Kottterink, B. (2014). Mapping smart cities in the EU.
- [P95] Psaraftis, H. N. (1995). Dynamic vehicle routing: Status and prospects. *Annals of operations research*, 61(1), 143-164.
- [SA14a] Stolfi, D. H., & Alba, E. (2014). Red Swarm: Reducing travel times in smart cities by using bio-inspired algorithms. *Applied Soft Computing*, 24, 181-195.
- [SA14b] Stolfi, D. H., & Alba, E. (2014, July). Eco-friendly reduction of travel times in european smart cities. In *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation* (pp. 1207-1214).
- [SR+03] Shaheen, S., Rodier, C. J., & Finson, R. S. (2003). University of California, Davis Long-Range Development Plan: A Davis Smart Mobility Model. California Partners for Advanced Transit and Highways (PATH).
- [SR+10] Santacana, E., Rackliffe, G., Tang, L., & Feng, X. (2010). Getting smart. *IEEE power and energy magazine*, 8(2), 41-48.
- [T11] Telefónica, F. (2011). *Smart Cities: un primer paso hacia la Internet de las Cosas* (Vol. 16). Fundación Telefónica.
- [VA+20] Villagra, A., Alba, E., & Luque, G. (2020). A better understanding on traffic light scheduling: New cellular GAs and new in-depth analysis of solutions. *Journal of Computational Science*, 41, 101085.
- [VP+12] Van Steen, M., Pierre, G., & Voulgaris, S. (2012). Challenges in very large distributed systems. *Journal of Internet Services and Applications*, 3(1), 59-66.
- [W10] Wong, A. K. (2010). The near-me area network. *IEEE internet computing*, 14(2), 74-77.
- [ZB+14] Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things journal*, 1(1), 22-32.

Deep Learning. Aplicaciones en Reconocimiento de Lengua de Señas, Generación de Lenguaje Natural e Imágenes Astronómicas.

F. Ronchetti^{1,2,3} , F. Quiroga^{1,2,4} , G. Rios^{1,4} , P. Dal Bianco¹ , I. Mindlin¹ , L. Lanzarini^{1,2} , A. Rosete⁶ , R. Gamen⁷ , D. Puig Valls⁹ , J. Torrents-Barrena⁹ , Y. Aidelman^{7,8}, C. Escudero⁸, A. Granada¹⁰.

¹ Instituto de Investigación en Informática LIDI, Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina.*

² Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina

³ Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)

⁴ Becario postgrado UNLP ⁵ Becario CIN

⁶ Universidad Tecnológica de La Habana “José Antonio Echeverría” (CUJAE), La Habana, Cuba

⁷ Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, La Plata, Argentina

⁸ Instituto de Astrofísica de La Plata (IALP CONICET), La Plata, Argentina

⁹ Universitat Rovira i Virgili (URV), Tarragona, España

¹⁰ Universidad Nacional de Río Negro (UNRN), Río Negro, Argentina

* 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 las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto F025 “Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data” perteneciente al Programa de Incentivos (2018-2021).

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de reconocimiento de patrones en imágenes, video y texto, utilizando técnicas de Aprendizaje Automático clásicas, junto con Redes Neuronales Convolucionales y Aprendizaje profundo. El trabajo presentado describe diferentes casos de aplicación en visión por computadora y procesamiento de lenguaje natural.

Una de las líneas de investigación principales que se continúa desarrollando es el reconocimiento de lengua de señas. Este es un problema complejo y multidisciplinar, que presenta diversos subproblemas a resolver como el reconocimiento del intérprete, la segmentación de manos, la clasificación de diferentes configuraciones y de un gesto dinámico, entre otros.

En esta área se está estudiando la forma de reconocer formas de mano de la Lengua de Señas con conjuntos de datos de tamaño reducido, dada la falta de datos de entrenamiento para este dominio. Además, se están utilizando Redes Recurrentes para reconocer señas dinámicas, utilizando la base de datos LSA64 de Lengua de Señas Argentina [2].

Por último, se están utilizando Redes Generativas Adversarias (GANs) para aumentar bases de datos de formas de mano, con el objetivo de complementar desde otro enfoque el entrenamiento de modelos para su clasificación.

Por otro lado, se está estudiando la forma en que las redes neuronales codifican la invarianza a las transformaciones y otras propiedades transformacionales, con el objetivo de poder analizar y comparar estos modelos. De esta forma se espera poder mejorar los modelos de clasificación de objetos transformados, en particular, de formas de mano.

Siguiendo con la línea de reconocimiento de patrones en imágenes, se está llevando a cabo una colaboración con investigadores de la Facultad de Astronomía y Geofísica de la UNLP para crear modelos de clasificación de imágenes de objetos celestes. Además, se está

desarrollando un sistema para recuperar la información de placas.

Por último, se creó un modelo neuronal capaz de generar texto artificial que se adapte al género *freestyle*.

Palabras clave: Redes Neuronales, Redes Convolucionales, Redes Recurrentes, Visión por Computadoras, Lengua de Señas, Redes Generativas Adversarias, Invarianza, Equivarianza, Imágenes Astronómicas.

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

Como resultado de estas investigaciones se han 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. Librería para experimentación con Formas de Mano de Lengua de Señas

Las lenguas de señas utilizan un conjunto finito de formas de mano que, en combinación con movimientos de las manos y el cuerpo, y expresiones faciales, se utilizan para señalar.

Una etapa fundamental en el reconocimiento de la lengua de señas es la clasificación de estas formas de mano, y por ende un área de investigación prioritaria para mejorar el reconocimiento.

En un estudio previo, se realizaron experimentos comparando diversas arquitecturas basadas en redes neuronales para clasificar formas de manos [1] pero advirtiendo que el desempeño de los modelos utilizados estuvo limitado por la falta de datos.

Por este motivo, se desarrolló una librería para facilitar la descarga y combinación de distintas bases de datos de formas de mano con el objetivo de mejorar los sistemas reconocedores de lengua de señas.



Figura 1. Ejemplos de las bases de datos de formas de mano RWTH y LSA16.

Dicha librería, de código abierto y disponible públicamente¹, permite la descarga de diez conjuntos de datos de imágenes de formas de mano, e incluye metadata que indica la correspondencia entre las formas de mano de las distintas bases de datos. De este modo, se logra simplificar y unificar las tareas de preprocesamiento y la comparación de distintas bases de datos, así como la transferencia de aprendizaje entre ellas.

1.2. Redes Generativas Adversarias para la Lengua de Señas.

Recientemente, las Redes Generativas Adversarias han mostrado resultados satisfactorios al crear imágenes artificiales a partir de datos aleatorios, o de una imagen para conseguir una transformación de la misma [6].

Por otro lado, las bases de datos de Lengua de Señas, dada su compleja naturaleza, o bien poseen pocos datos o los datos que las forman no presentan la suficiente diversidad. Además, la creación de datos artificiales no es trivial como en otros dominios, lo que hace muy difícil aplicar técnicas clásicas como *Data Augmentation*.

Por tal motivo, en el III-LIDI se están comenzando a utilizar Redes Generativas Adversarias (GANs) para generar imágenes artificiales relacionadas con la lengua de señas. Este tipo de redes permitirá aumentar

¹https://github.com/midusi/handshape_datasets

las bases de datos de formas de mano, con el objetivo de complementar desde otro enfoque el estudio de modelos y algoritmos de clasificación para bases de datos con pocos datos etiquetados. Estas investigaciones son llevadas a cabo en marco de una tesis doctoral financiada por la UNLP a través de una beca de postgrado.

1.3. Reconocimiento de Lengua de Señas con Redes Recurrentes y Convolucionales

En esta línea de investigación, el objetivo es la clasificación de gestos en videos de Lengua de Señas. En particular, estamos trabajando con señas de la base de datos LSA64. Esta consiste en un registro de 64 señas de la Lengua de Señas Argentina. Para la tarea se compararon diversas técnicas en el estado del arte del Aprendizaje Automático basadas en Redes Neuronales. Específicamente, se compararon tanto arquitecturas basadas en Redes Recurrentes y Convolucionales, como distintas estrategias de preprocesamiento para optimizar la calidad del reconocimiento. También se están analizando los modelos entrenados para comprender mejor el impacto de estas estrategias de preprocesamiento y de la forma de representación lograda por cada tipo de arquitectura.



Figura 2: Activaciones de las capas convolucionales de una Red Recurrente para reconocer señas en video.

1.4. Métricas de Equivarianza

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.

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.

En los últimos años, varios modelos fueron propuestos para añadir invarianza a la rotación y otras transformaciones en CNNs [3]. No obstante, no está claro como estos modelos impactan en el aprendizaje usual de los pesos de la red.

Por este motivo, se continua con la utilización de las métricas previamente definidas [4] para caracterizar modelos de redes neuronales típicos, ya sea desde capas muy utilizadas como Batch Normalization, Dropout, Max-Pooling, arquitecturas completas como Residual Networks, VGG o AllConvolutional, y arquitecturas especializadas como TI-Pooling.

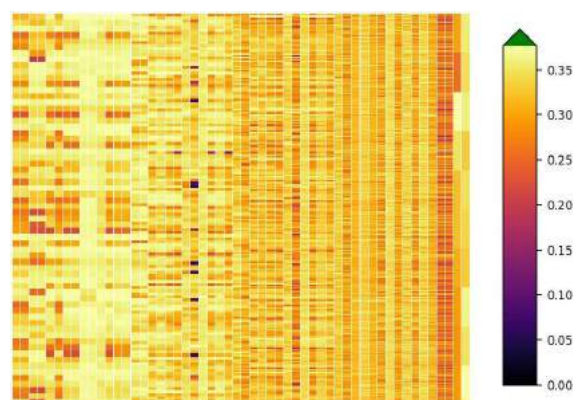


Figura 3. Invarianza por capas y unidades de una CNN con arquitectura ResNet.

1.5. Análisis de Imágenes Astronómicas

En los últimos años, la cantidad de información astronómica disponible se ha multiplicado de forma exponencial. En consecuencia, diversas tareas que anteriormente se realizaban de forma manual o semi-manual deben ahora automatizarse aún más.

En este ámbito, en el III-LIDI se están desarrollando dos proyectos. El primero consiste en determinar modelos de

clasificación a partir de información fotométrica de estrellas Be [5]. Estas estrellas son estrellas no-supergigantes cuyo espectro exhibe emisión en línea $H\alpha$. Para el análisis de muchos objetos astronómicos, los datos fotométricos son relativamente más fáciles de obtener debido al menor tiempo de uso del telescopio. Por lo tanto, existe una necesidad creciente de utilizar información fotométrica para identificar automáticamente objetos para estudios más detallados, especialmente estrellas con líneas de emisión $H\alpha$ como las estrellas Be. En este trabajo se evaluó el uso de redes neuronales para identificar candidatos a estrella Be a partir de un conjunto de estrellas tipo OB. Las redes se entrenaron utilizando un subconjunto etiquetado de las bases de datos VPHAS + y 2MASS, con filtros u, g, r, $H\alpha$, i, J, H y K. Para evitar el efecto del enrojecimiento, se propuso y se evaluó el uso de índices Q para mejorar la generalización del modelo a otras bases de datos. Para validar el enfoque, se etiquetó manualmente un subconjunto de la base de datos y se usó para evaluar modelos de identificación de candidatos. También se etiquetó un conjunto independiente de datos utilizando validación cruzada [7]. Los modelos entrenados lograron obtener un 25% de Sensibilidad (*Recall*) fijando la precisión (*Precision*) al 99%. Actualmente se está trabajando en agregar nuevas bases de datos de otros investigadores y así como otras áreas del cielo.

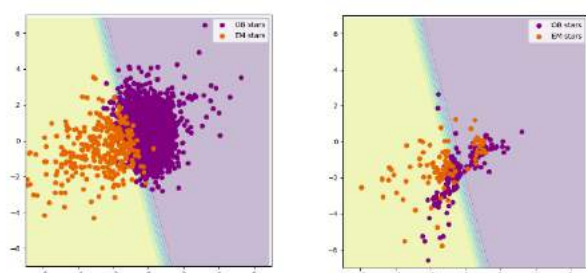


Figura 4. Reducción a 2 features para los conjuntos de datos de Mohr-Smith (izquierda) y Liu (derecha). Las líneas de decisión corresponden a una función sigmoidea de un regresor logístico entrenado en Mohr-Smith.

1.6. Generación de letras de rap con Redes Neuronales Recurrentes

En el marco del proyecto de investigación, desarrollo e innovación “Generación de

contenido multimedia mediante IA para entornos interactivos” evaluado y financiado por la Facultad de Informática de la UNLP y continuando la línea de trabajos anteriores de reconocimiento de patrones utilizando Deep Learning, se realizó un proyecto de generación automática de letras de rap con estilo freestyle.

Para llevar esto a cabo, se analizaron diferentes modelos de generación de texto basados en redes recurrentes y generativas [8][9]. Se confeccionaron dos bases de datos específicas de este tipo de letras en español, inexistentes hasta la fecha. Se analizaron diferentes sistemas de preprocesamiento del texto incluyendo tokenización a nivel de palabra, de sílaba y el algoritmo BPE. Se realizaron diversos experimentos con redes recurrentes bidireccionales tipo LSTM y se realizó un diccionario específico de rimas en español para controlar cómo el texto las genera.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Redes neuronales profundas, convolucionales y recurrentes.
- Invarianza y auto-equivarianza en redes neuronales.
- Reconocimiento de lenguaje de señas.
- Generación de imágenes con GANs.
- Análisis de datos astronómicos.
- Generación de lenguaje natural.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Comparación de modelos especializados para bases de datos con pocas muestras para la clasificación de formas de mano.
- Desarrollo de métricas de invarianza y auto-equivarianza para redes neuronales.
- Análisis de modelos para el reconocimiento de lengua de señas en video.

- Librería para facilitar la experimentación con imágenes de formas de manos para el lenguaje de señas.
- Redes generativas para la creación de datos artificiales en la Lengua de Señas.
- Modelo de clasificación de estrellas Be generalizable a distintos conjuntos de datos.
- Modelo generativo de texto artificial para batallas de *freestyle*.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 1 profesor con dedicación exclusiva, 1 investigador CIC-PBA, 2 becarios de posgrado de la UNLP con dedicación docente, 1 becario CIC, 1 becario CIN, 6 tesistas, 3 profesores extranjeros, y 3 investigadores externos.

Dentro de los temas involucrados en esta línea de investigación, en los últimos dos años se han finalizado 2 tesis de doctorado, 2 tesis de especialización, y 5 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 2 tesis de doctorado, 1 tesis de especialista, 5 tesinas de grado de Licenciatura y 4 trabajos finales de Ingeniería en Computación. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Quiroga, F., Antonio, R., Ronchetti, R., Lanzarini, L., Rosete, A. A Study of Convolutional Architectures for Handshape Recognition applied to Sign Language, publicado en el XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017) (pp. 13-22).
- [2] Cornejo Fandos, U., Rios, G., Ronchetti, F., Quiroga, F., Hasperué, W., Lanzarini, L. Recognizing Handshapes using Small Datasets, publicado en el XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019, Rio Cuarto) (pp. 105-114).
- [3] Quiroga F., Ronchetti F., Lanzarini L., Fernandez-Bariviera A. Revisiting Data Augmentation for Rotational Invariance in Convolutional Neural Networks. International Conference on Modeling and Simulation in Engineering, Economics and Management (MS'2018 GIRONA) .
- [4] Quiroga, F., Torrents-Barrena, J., Lanzarini, L., & Puig, D. (2019, June). Measuring (in) variances in Convolutional Networks. In Conference on Cloud Computing and Big Data (pp. 98-109). Springer, Cham.
- [5] Aidelman Y., Escudero C., Ronchetti F., Quiroga F., Lanzarini L. Reddening-Free Q Indices to Identify Be Star Candidates. Cloud Computing, Big Data & Emerging Topics. pp 111-123. Springer International Publishing. 2020.
- [6] Goodfellow I. J., Pouget-Abadie j., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative Adversarial Networks. NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems. v2. pp 2672-2680. 2014.
- [7] Jaschek M., Slettebak A., Jaschek C. *Be star terminology*. Be Star Newsletter. 1981.
- [8] Potash, P., Romanov, A., and Rumshisky, A. Ghostwriter: Using an lstm for automatic rap lyric generation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 1919-1924. 2015.
- [9] Young, T., Hazarika, D., Poria, S., and Cambria, E. Recent trends in deep learning based natural language processing. *iee Computational intelligenCe magazine*, 13(3):55. 2018.

Desarrollo de sistemas de scheduling de producción en el contexto de la industria 4.0

Daniel Díaz, Francisco Ibañez, Sandra Oviedo, Juan Cuneo, María Becerra
Fernando Guardia

LabIAI- Instituto de Informática – Dpto. de Informática
FCEFyN - Universidad Nacional de San Juan
CUIM – Av. Ignacio de la Roza 590 (O), Rivadavia – J5402DCS San Juan
{ fibanez, ddiaz, soviedo, jcuneo, mbecerra,fguardia}@iinfo.unsj.edu.ar}

Resumen

La nueva revolución industrial conocida como industria 4.0 va en camino a crear productos inteligentes y conectados que ofrecen posibilidades de expansión exponenciales hacia nuevas funcionalidades con mayor confiabilidad, mayor utilidad y capacidades que atraviesan y trascienden los límites de los productos tradicionales. Al interior de una planta industrial producir estos productos requiere de una transformación de los sistemas de producción, una componente importante de este proceso son los sistemas de planificación y scheduling de la producción. Es por ello que este proyecto propone investigar sobre el desarrollo de sistemas de scheduling de producción en el contexto de la industria 4.0.

Palabras clave:

Planificación de la producción, industria 4.0, fabricación inteligente

Contexto

El concepto de industria 4.0 se refiere a la integración de diferentes tecnologías tales como big data, edge computing, inteligencia artificial, servicios en la

nube, internet industrial de las cosas [1], sistemas ciberfísicos [2] entre otras.

El objetivo de este proyecto es investigar sobre tecnologías de la industria 4.0 plausibles de ser incorporadas en un sistema de scheduling de producción.

El proyecto se lleva a cabo en el Laboratorio de Informática Aplicada a la Innovación (LabIAI) del Instituto de informática de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan, el mismo está inserto en la línea de investigación sobre sistemas de scheduling que tiene el LabIAI y se presenta como una continuidad del último proyecto ejecutado bajo esta línea de investigación [3, 4].

Introducción

Desde el surgimiento de la informática los sistemas de producción han evolucionado conforme las tecnologías de la información y comunicación lo han hecho. Se pueden identificar 3 etapas en esta evolución. La etapa de digitalización, que va desde mediados del siglo XX hasta mediados de la década de 1990, en esta etapa se informatiza la industria mediante aplicaciones de gestión y dispositivos de control tales los PLC (Programmable Logic Controllers). Desde mediados de

los 1990 comienza la etapa de red, que surge con el advenimiento de internet, en esta etapa todo se conecta, los sistemas y los dispositivos electrónicos dejan de trabajar aislados y pasan a formar parte de la gran red. Hoy en día, un clúster de avances tecnológicos tales como big data, edge computing, inteligencia artificial, servicios en la nube, internet industrial de las cosas [1], sistemas ciberfísicos [2] entre otras, están dando surgimiento a la etapa de la inteligencia, también denominada la nueva generación de manufactura inteligente [5]. Estos nuevos avances han dado lugar al desarrollo de nuevas estrategias de manufactura tales como industria 4.0 en Alemania, Internet Industrial en EEUU [6] y Made In China 2025 en China [7]. Esta cuarta revolución industrial ya ha llegado a nuestro país, así lo demuestran las últimas conferencias de la Unión Industrial Argentina (UIA) [8-10], donde diversas temáticas relacionadas con la industria 4.0 han sido tratadas.

En este contexto la Industria 4.0 presenta nuevos desafíos y oportunidades para los sistemas de planificación y scheduling de la producción. Es por ello que esta propuesta de proyecto propone investigar sobre el desarrollo de sistemas de scheduling de producción en el contexto de la industria 4.0.

Industria 4.0

Según Hermann y otros [11] “la Industria 4.0 es un término colectivo para las tecnologías y conceptos de organización de la cadena de valor. Dentro de las fábricas inteligentes modulares y estructuradas de la Industria 4.0, los sistemas ciberfísicos (CPS) monitorean los procesos físicos, crean una copia virtual del mundo físico y toman decisiones descentralizadas. A través de Internet de las cosas (IoT), CPS se

comunica y coopera entre sí y con los humanos en tiempo real. A través de Internet de Servicios (IoS), los participantes de la cadena de valor ofrecen y utilizan servicios internos y de organización cruzada”

Sistemas ciberfísicos (CPS)

Los CPS pueden considerarse sistemas que unen el mundo físico y el virtual. Más precisamente, “los sistemas ciberfísicos son integraciones de procesos computacionales con procesos físicos. Dispositivos computarizados y redes integradas monitorean y controlan los procesos físicos, generalmente con bucles de retroalimentación donde los procesos físicos afectan los cálculos y viceversa. En el contexto de la manufactura, esto significa que la información relacionada con lo físico (el taller, la fábrica) y el espacio virtual de cómputo están altamente sincronizados. Esto permite un nuevo grado de control, vigilancia, transparencia y eficiencia en el proceso de producción. Con respecto a su estructura, CPS tiene dos redes paralelas para controlar, a saber, la red física de componentes interconectados de la infraestructura y la red cibernética compuesta por controladores inteligentes y los enlaces de comunicación entre ellos. CPS realiza la integración de estas redes mediante el uso de múltiples sensores, actuadores, unidades de procesamiento de control y dispositivos de comunicación.

Internet de las cosas (IoT)

El termino no tiene una sola definición, tal esta expresado en [12], la definición del Estandar de IEEE dice "Una red de elementos, cada uno integrado con sensores, que están conectados a Internet". Porter [13] “son los productos inteligentes y conectados que ofrecen

posibilidades de expansión exponenciales para nuevas funcionalidades, mayor confiabilidad, mayor utilidad del producto y capacidades que atraviesan y trascienden los límites de los productos tradicionales “. En resumen podría decirse que la IoT puede hacer que todas las cosas físicas puedan convertirse en cosas inteligentes conectadas a internet.

Internet de servicios (IoS)

De manera similar al IoT, está surgiendo una Internet de servicios (IoS), basada en la idea de que los servicios se ponen a disposición fácilmente a través de tecnologías web, lo que permite a las empresas y a los usuarios privados combinar, crear y ofrecer un nuevo tipo de servicios de valor agregado [14]. Los servicios de la nube, así como las aplicaciones basadas en arquitectura orientada servicios o en arquitectura de microservicios son parte del internet de servicios.

Fabricación inteligente

Los conceptos de CPS, IoT e IoS son los componentes principales de Industria 4.0. Cabe señalar que estos "conceptos" están estrechamente relacionados entre sí, ya que CPS se comunica a través de IoT e IoS, lo que permite la llamada "fábrica inteligente", que se basa en la idea de un sistema de producción descentralizado, en el que “Los seres humanos, las máquinas y los recursos se comunican entre ellos de manera tan natural como en una red social. En [5] se discuten en profundidad todos los conceptos de fabricación inteligente.

Sistemas de scheduling en dominios industriales

Los problemas de Scheduling aparecen en diferentes dominios, este proyecto trata con problemas de scheduling que surgen en el dominio industrial y que en la industria que se denominan production scheduling o manufacturing scheduling.

La definición más clásica de la palabra scheduling dice : “Scheduling es el problema de asignar recursos limitados a tareas en el tiempo con el objeto de optimizar uno o más objetivos” [15]. En la industria, un sistema de scheduling es el corazón del sistema de planificación y control de la producción. Normalmente el sistema de scheduling interactúa con otros sistemas de fabricación, tales como ERP (Enterprise Resource Planning) y MES (Manufacturing Execution System).

Existe una literatura muy extensa sobre scheduling y se ha escrito mucho sobre modelos y algoritmos que resuelven problemas de scheduling. Sin embargo, poco son los artículos que tratan de cómo traer estos modelos y algoritmos a implementaciones reales. Esto se conoce como el “hueco o gap” entre la teoría y la práctica de scheduling [16]. Para cerrar este hueco entre la teoría y la práctica los modelos y algoritmos de scheduling se deben implementar en una pieza de software que satisfaga las necesidades que tiene una empresa en el área de scheduling. Esto implica llevar a cabo un proceso de desarrollo de software para obtener un producto final, es decir un sistema de scheduling y que ayude a la toma de decisión. Según Yen y Pinedo [17] un sistema de scheduling se compone tres módulos: (1) módulos de base datos y base de conocimiento, (2) módulos del motor de scheduling y (3) módulo de interface de usuario. La parte esencial de todo sistema de scheduling es

el desarrollo del motor de scheduling, es donde se vinculan las necesidades que tiene la empresa con los modelos y algoritmos de scheduling.

Nuevas necesidades por parte de la empresa debido a los avances tecnológicos y evolución del mercado hacen que las técnicas aplicadas para desarrollar los sistemas de scheduling mencionados no sean aplicables al nuevo entorno industrial. McKay et al. [18] Describen un conjunto de necesidades o requisitos que debe reunir un sistema de scheduling, algunos de los cuales se refieren a los algoritmos de secuenciación y al uso del sistema. Con respecto a los algoritmos de secuenciación, éstos deben ser flexibles y configurables. Deben ser capaces de manejar diversos criterios (scheduling multicriterio). A estos requisitos se les debe agregar otros tales como facilitar el mantenimiento y la reducción de costos de desarrollo.

Por último, [19] intenta integrar esta definiciones definiendo a un manufacturing scheduling como el proceso de toma de decisión que consiste de asignar un conjunto de operaciones/tareas requeridas para manufactura un conjunto de productos con los recursos existentes en una planta, como así también los tiempos necesario para iniciar estas operaciones o tareas. Un Schedule o plan se define como un conjunto específico de asignaciones de operaciones o tareas a los recursos sobre una escala temporal.

Resultados y Objetivos

El objetivo es investigar sobre métodos y técnicas en pos de integrar las más recientes tecnologías en un marco de trabajo que permita construir sistemas de scheduling de producción para la industria 4.0.

En cuanto a resultados, se está trabajando en el desarrollo de chatbots que asistan en el proceso de planificación de la producción, más específicamente que ayude al planificador en la carga e interpretación de los datos de entrada y salida de un scheduling de producción

Formación de Recursos Humanos

El equipo de trabajo que lleva adelante este proyecto se compone de

- 4 docentes investigadores,
- 1 tesistas de grado en período iniciación.
- 2 alumnos avanzados de iniciación a la investigación
- 2 tesistas de posgrado (maestría)

Referencias

- [1] Z. Bi, L. Da Xu, and C. Wang, "Internet of Things for Enterprise Systems of Modern Manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 10, pp. 1537-1546, 2014 2014.
- [2] N. Jazdi, "Cyber physical systems in the context of Industry 4.0," *2014 IEEE International Conference on Automation, Quality and Testing, Robotics*, 2014 2014.
- [3] D. Díaz, S. Oviedo, F. Guardia, J. M. Cuneo, and A. Otazú, "Hacia un marco de trabajo para desarrollar sistemas de programación de la producción utilizando servicios en la nube.," presented at the Congreso argentino de ciencias de la computación, Universidad Nacional de Rio Cuarto - Cordoba, 2019.
- [4] D. Díaz, S. Oviedo, F. Guardia, J. Cuneo, F. Ibañez, and N. Alonso, "Desarrollo sistemas de programación de producción mediante servicios en la nube.,"

- presented at the X ENCUENTRO DE INVESTIGADORES Y DOCENTES DE INGENIERÍA, Los Reyunos, San Rafael, Mendoza, 2019.
- [5] J. Zhou, P. Li, Y. Zhou, B. Wang, J. Zang, and L. Meng, "Toward new-generation intelligent manufacturing," *Engineering*, vol. 4, pp. 11-20, 2018.
- [6] I. Industrial. (2014, 2019/12/3). *The Industrial Internet Consortium: A Global Nonprofit Partnership Of Industry, Government And Academia, March 2014*. Available: <https://www.iiconsortium.org/about-us.htmPartnePartnershiprship>
- [7] L. Li, "China's manufacturing locus in 2025: With a comparison of "Made-in-China 2025" and "Industry 4.0"," *Technological Forecasting and Social Change*, vol. 135, pp. 66-74, 2018.
- [8] UIA2018, "24° Conferencia Industrial (2018)," <https://www.uia.org.ar/conferenciaindustrial/conferencia/24>, 2018.
- [9] UIA2017, "23° Conferencia Industrial (2017)," <https://www.uia.org.ar/conferenciaindustrial/conferencia/23>, 2017.
- [10] UIA2019, "25° Conferencia Industrial (2019)," 2019.
- [11] M. Hermann, T. Pentek, and B. Otto, "Design principles for industrie 4.0 scenarios," in *2016 49th Hawaii international conference on system sciences (HICSS)*, 2016, pp. 3928-3937.
- [12] R. Minerva, A. Biru, and D. Rotondi, "Towards a definition of the Internet of Things (IoT)," *IEEE Internet Initiative*, vol. 1, pp. 1-86, 2015.
- [13] M. E. Porter and J. E. Heppelmann, "How smart, connected products are transforming competition," *Harvard business review*, vol. 92, pp. 64-88, 2014.
- [14] W. Wahlster, H.-J. Grallert, S. Wess, H. Friedrich, and T. Widenka, *Towards the internet of services: The THESEUS research program*: Springer, 2014.
- [15] K. R. Barker, *Elements of sequencing and scheduling*. New York: John Wiley and Sons, 1974.
- [16] B. L. MacCarthy and J. Liu, "Addressing the gap in scheduling research: A review of optimization and heuristic methods in production scheduling," *International Journal of Production Research*, vol. 31, pp. 59-79, 1993.
- [17] B. P.-C. Yen and M. Pinedo, "On the design and development of scheduling systems," in *Fourth International Conference on Computer Integrated Manufacturing and Automation Technology*, 1994, pp. 197 - 204.
- [18] K. McKay, M. L. Pinedo, and S. Webster, "Practice-Focused Research issues for scheduling systems," *Production and Operations Management*, vol. 11, pp. 249-258, 2002.
- [19] J. M. Framinan, R. Leisten, and R. R. García, "Manufacturing scheduling systems," *An integrated view on Models, Methods and Tools*, pp. 51-63, 2014.

Detección de patrones y tendencias en estudiantes universitarios de carreras de ingeniería para determinar el éxito académico aplicando Machine Learning

José Federico Medrano¹, Octavio Daniel Coro¹, Valeria Barriento²

jfmedrano@fi.unju.edu.ar, odcoro@fi.unju.edu.ar, barriento.valeria@gmail.com

¹VRAln / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy (UNJu) - Ítalo Palanca 10, +54 (388) 4221587

² Escuela Superior de Ciencias Jurídicas y Políticas / Universidad Nacional de Jujuy (UNJu)

RESUMEN

Uno de los indicadores por el cual son medidas y evaluadas las universidades es el número o cantidad de egresados por año. Si bien uno de los tantos objetivos de las instituciones académicas es la formación de profesionales, en las universidades se evidencia que no existe una relación directa entre el número de ingresantes y el número de egresados, es decir, la cantidad de egresados no necesariamente se incrementa al aumentar la cantidad de ingresantes. Esto representa un gran problema puesto que los presupuestos con los que cuentan las universidades, sobre todo las de gestión pública, muchas veces se ven reducidos, es más, cuando un alumno logra finalizar sus estudios, el costo de la inversión para formarlo es altísimo. Claro está que obtener una mayor cantidad de egresados no solo reduciría la inversión mencionada sino que mejoraría los indicadores empleados y posicionaría mejor a cualquier universidad. Lograr identificar algún patrón o tendencia en el modo en que los alumnos cursan su respectiva carrera, o identificar las variables que influyen tanto en el éxito como en la deserción académica, sería beneficioso para que las autoridades puedan tomar las

decisiones correctas. En este proyecto se tomarán las historias académicas de estudiantes de carreras de ingeniería de la Universidad Nacional de Jujuy, para identificar patrones o tendencias en el cursado que permitan determinar las variables que influyen para que un estudiante logre concluir la carrera. Se empleará *Machine Learning* como técnica para conducir este trabajo, encarándolo como un problema de clasificación y/o predicción.

Palabras clave: *Machine Learning; Detección de patrones; Tendencias; Rendimiento académico; Deserción*

CONTEXTO

La línea de investigación aquí presentada se encuadra dentro del Proyecto Bianual 2020-2021 denominado “*Detección de patrones y tendencias en estudiantes universitarios para identificar causas de deserción y éxito académico empleando Machine Learning*”, aprobado y financiado por la Secretaría de Ciencia y Técnica de la Facultad de Ingeniería

de la Universidad Nacional de Jujuy, aprobado por resolución CAFI 449/19 .

1. INTRODUCCIÓN

La permanencia y la graduación de los estudiantes suele ser una problemática relevante compartida por la mayoría de los sistemas universitarios e instituciones que lo componen, y de acuerdo al nivel alcanzado en sus indicadores se constituye en condición de existencia o desaparición de aquellas universidades que dependen en forma casi exclusiva de su matrícula, como es el caso de las universidades privadas, o de condicionamiento presupuestario en el caso de las universidades estatales (Lattuada, 2017).

El rendimiento académico es un fenómeno complejo que suele abordarse en las universidades desde distintos indicadores. Con frecuencia se consideran las calificaciones promedio obtenidas en distintos períodos de la estancia en la universidad, tasas de aprobación de materias y créditos, tasas de retención-deserción, tasas de graduación, períodos de graduación respecto de planes de estudio (Sánchez & Chinchilla Brenes, 2005). Estos indicadores son relativamente sencillos de calcular y medir, dado que se basan en recuentos estadísticos, y los informes y reportes que se arman en base a estos números solo reflejan una parte de la historia, la que la estadística descriptiva puede analizar (Fernández, Sánchez, Córdoba, & Largo, 2002).

Ahora bien, si el número de estudiantes que egresan cada año de una universidad fuera el indicador más relevante para estimar la calidad de una universidad, sería deseable no solo identificar las causas que influyen en la deserción de los estudiantes sino también identificar las causas que determinan que a un estudiante le demande mayor o menor

cantidad de años finalizar dicha carrera, ya sea una carrera de grado o pregrado. Por ello existen modelos, técnicas y herramientas avanzadas que emplean Inteligencia Artificial para poder identificar patrones y tendencias en los datos analizados (Wang, Rudin, Wagner, & Sevieri, 2015), del mismo modo estos modelos pueden realizar predicciones o pronósticos partiendo del análisis datos históricos, lo que se conoce comúnmente como Aprendizaje Automático o *Machine Learning* (Witten, Frank, Hall, & Pal, 2016).

El mecanismo que emplea el *Machine Learning* es muy sencillo, a partir de un conjunto de datos se realiza una división del mismo en tres grupos, uno de entrenamiento, uno de prueba y uno de evaluación. El conjunto de datos de entrenamiento es empleado por el modelo/algoritmo que se desee entrenar, esto es, el modelo aprende la lógica de los datos para realizar predicciones, para clasificar o agrupar datos, el modelo se ajusta a los datos. El conjunto de prueba sirve para testear y ajustar el modelo diseñado. Por último el conjunto de datos de validación es empleado para obtener indicadores de la eficiencia del modelo planteado. Una vez ajustado y validado el modelo, el mismo es utilizado con datos nuevos y desconocidos para obtener “nuevo conocimiento”.

Dentro de esta área del conocimiento (la Inteligencia Artificial), el *Machine Learning* ha recibido muchísima atención en esta última década, debido principalmente a la reducción de costos de los equipos de cómputo y a la proliferación de librerías y herramientas que automatizan gran parte del trabajo. En este sentido se ha aplicado efectivamente al ámbito académico para analizar la retención de estudiantes (Nandeshwar, Menzies, & Nelson, 2011) o para predecir la deserción

de estos (Kuna, García-Martínez, & Villatoro, 2010; Solis, Moreira, Gonzalez, Fernandez, & Hernandez, 2018; Bowen & Agustín, 2016).

Una ventaja adicional de este tipo de análisis serviría para producir visualizaciones y representaciones de información con técnicas novedosas que permiten encontrar nuevas relaciones e información a partir de dichas representaciones. Para ello se emplearán técnicas de Visualización de Información (InfoVis) que tienen por objeto presentar la información visualmente, en esencia, para reducir la carga de trabajo cognitivo al sistema perceptivo visual humano (Ware, 2004; Ware, 2008).

Otro estudio por ejemplo se ha centrado en la construcción de un modelo predictivo a través de técnicas de minería de datos que permita, por un lado, determinar la cantidad de alumnos que pueden graduarse en carreras de ingenierías de UTN Facultad Regional San Francisco, y por otro lado, identificar patrones que puedan incidir en la graduación (Carrizo, 2019). Del mismo modo, en (Istvan & Lasagna, 2019) se plantea la identificación de los diferentes factores que influyen en la deserción estudiantil, con el fin de caracterizar y extraer los perfiles de los estudiantes desertores empleando minería de datos como una alternativa útil para encontrar información derivada a partir de la detección de patrones de atributos individuales.

La Universidad Nacional de Jujuy posee un Sistema de Gestión Académica llamado SIU Guaraní, cada facultad posee una configuración e instalación individual, es decir, el sistema no es centralizado. Cada instancia en cada facultad tiene una base de datos única en donde se gestiona y almacena la “historia” académica de los alumnos, desde el momento en que ingresan a la facultad hasta que egresan de la misma. Partiendo de

esto, se plantea en este proyecto analizar la base de datos del sistema SIU Guaraní de la Facultad de Ingeniería para identificar las variables y causas que influyen tanto en el éxito como en el fracaso académico.

El presente proyecto tiene como objetivo general la búsqueda e identificación de patrones en la conducta de cursado de los estudiantes universitarios a lo largo de su vida académica.

De este objetivo general se desprenden los siguientes Objetivos Particulares:

- Analizar la base de datos actual del sistema SIU Guaraní.
- Determinar los atributos y variables relevantes para el estudio.
- Determinar la muestra que servirá para el análisis.
- Determinar el conjunto de datos que se utilizará para entrenamiento, prueba y validación.
- Verificar el balance de las clases a emplear en los modelos.
- Diseñar modelos de clasificación y agrupación de información mediante la aplicación de técnicas de Machine Learning.
- Representar las distintas relaciones y/o patrones entre los datos procesados mediante el empleo de técnicas de visualización de información.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación principal se enmarca en un proyecto de la Facultad de Ingeniería de la UNJu que aborda el estudio de las variables que influyen tanto en el éxito como en la deserción académica. Mediante el empleo de *Machine Learning* pretende predecir cuales

son los atributos que influyen para que un estudiante logre finalizar la carrera. De este modo, con el nuevo conocimiento adquirido, se podrán tomar distintas decisiones para atacar el problema y garantizar bajo ciertos parámetros una cursada exitosa.

El equipo de Investigación que llevará a cabo este trabajo centra sus líneas de investigación en la Inteligencia Artificial, Minería de Datos/ Textos, y Recuperación de Información.

3. RESULTADOS OBTENIDOS/ESPERADOS

A partir de los objetivos planteados en este trabajo, se espera diseñar ciertos modelos de aprendizaje automático que sean capaces, luego del correcto entrenamiento, ajuste y validación, predecir cuales son las variables y atributos del conjunto de datos (historia académica de los alumnos) que determinan el éxito en la cursada o por el contrario, detectar los factores que influyen en el re-cursado o abandono de la carrera.

Luego de este análisis, se espera conformar un conjunto de visualizaciones que ayuden a tener una visión más genérica de los resultados hallados (Lanzarini, Charnelli, Baldino, & Díaz, 2015).

4. FORMACIÓN DE RECURSOS HUMANOS

El Equipo de Trabajo está conformado por docentes investigadores de la Universidad Nacional de Jujuy. Los mismos llevan adelante esta línea de investigación desde hace años. Cada año se incorporan al proyecto alumnos avanzados de distintas carreras, quienes trabajan en temas relacionados con las temáticas planteadas. Del mismo modo, los integrantes del equipo participan en el

dictado de asignaturas/cursos de grado y postgrado de la UNLP, UNJu y UCSEDASS.

5. BIBLIOGRAFÍA

- Bowen, O., & Agustín, A. (2016). *Análisis de la deserción y permanencia académica en la educación superior aplicando minería de datos*. Tesis Doctoral, Universidad Nacional de Colombia-Sede Bogotá.
- Carrizo, C. J. (2019). Modelo predictivo para determinar la graduación de alumnos de carreras de ingenierías aplicando técnicas de minería de datos. *I Congreso Internacional de Ingeniería de Sistemas*.
- Fernández, S. F., Sánchez, J. M., Córdoba, A., & Largo, A. C. (2002). *Estadística descriptiva*. Esic Editorial.
- Istvan, R. M., & Lasagna, V. (2019). Aplicación de minería de datos en la extracción de perfiles de deserción estudiantil sobre ingresantes de la UTN FRLP. *XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)*. San Juan.
- Kuna, H., García-Martínez, R. V., & Villatoro, F. (2010). Pattern Discovery in University Students Desertion Based on Data Mining. *IV Meeting on Dynamics of Social and Economic Systems, 2*, pp. 275-285.
- Lanzarini, L. C., Charnelli, M. E., Baldino, G., & Díaz, F. J. (2015). Selección de atributos representativos del avance académico de los alumnos universitarios usando técnicas de visualización. *TE & ET(15)*, 42-50.

- Lattuada, M. (2017). Deserción y retención en las unidades académicas de educación superior. Una aproximación a las causas, instrumentos y estrategias que contribuyen a conocer y morigerar su impacto. *Debate Universitario*, 5(10), 100-113.
- Nandeshwar, A., Menzies, T., & Nelson, A. (2011). Learning patterns of university student retention. *Expert Systems with Applications*, 38(12), 14984-14996.
- Sánchez, E. G., & Chinchilla Brenes, S. (2005). Detección de estudiantes en riesgo académico en el Instituto Tecnológico de Costa Rica. *Revista Educación*, 29(2), 123-138.
- Solis, M., Moreira, T., Gonzalez, R., Fernandez, T., & Hernandez, M. (2018). Perspectives to Predict Dropout in University Students with Machine Learning. *IEEE International Work Conference on Bioinspired Intelligence (IWOBI)*, (pp. 1-6).
- Wang, T., Rudin, C., Wagner, D., & Sevieri, R. (2015). Finding patterns with a rotten core: Data mining for crime series with cores. *Big Data*, 3(1), 3-21.
- Ware, C. (2004). *Information Visualization - Perception for Design*. Morgan-Kaufmann.
- Ware, C. (2008). *Visual Thinking for Design*. Morgan Kaufman/Elsevier.
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Experticia. Un modelo de Sistema Experto aplicado al Poder Judicial

Oswaldo Sposito¹, Luis Busnelli², Viviana Ledesma¹, Laura Conti², Sergio García³, Gastón Procopio¹, Gustavo Pérez Villar⁴, Frega, Gerardo¹, Julio Bossero¹, Saizar, Victoria¹, Quintana, Fabio¹

¹ Universidad Nacional de La Matanza. Departamento de ingeniería e Investigación Tecnológicas. Florencio Varela 1903. San Justo. La Matanza.

{sposito, vledesma, gprocopio, gfrega, jbossero, fquintana, vsaizar}@unlam.edu.ar

² Universidad Nacional de La Matanza. Departamento Derecho y Ciencia Política. {lbusnell, lconti}@unlam.edu.ar

³ Palacio de Tribunales. Departamento Judicial de Morón. Alte. Brown. Piso 4. Morón. sergiogabriel.garcia@pjba.gov.ar

⁴ Subsecretaría de Tecnología Informática del Poder Judicial de la Provincia de Buenos Aires. Palacio de Justicia, avenida 13 entre 47 y 48, primer piso (La Plata). Argentina. gperez@scba.gov.ar

RESUMEN

Se presenta una línea de investigación y desarrollo, que tiene por objeto estudiar temas relacionados con Inteligencia Artificial (IA), Árboles binarios e informática jurídica, orientada al desarrollo de un modelo de Sistema Experto (SE) para la resolución de dictámenes jurídicos, con el fin de brindar a los profesionales del derecho una herramienta que les permita acelerar los tiempos de procesamiento de expedientes, minimizando posible de errores en la carga de datos.

Palabras clave: Informática Jurídica, Justicia Digital, Inteligencia Artificial, Sistemas Expertos, Sistemas Expertos Legales.

CONTEXTO

El trabajo aquí presentado está enmarcado en el proyecto de investigación “*Diseño e Implementación de un Sistema Experto como Apoyo al Proceso de Despacho de Trámites de un Organismo Judicial*”, dentro del Programa de Incentivos para Docentes Investigadores de la Secretaría de Políticas Universitarias (PROINCE) 2020-2021. Este proyecto es financiado

por la Universidad Nacional de La Matanza (UNLaM), y se lleva a cabo por investigadores de dos departamentos, el Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT) y el Departamento de Derecho y Ciencias Políticas. A su vez el mismo se realiza con la estrecha colaboración del Juzgado de Ejecución N° 2 del Departamento Judicial Morón. El mismo se desarrolla en el Centro de Desarrollo e Investigaciones tecnológicas (CeDIT) del Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT) de la Universidad Nacional de La Matanza (UNLaM).

INTRODUCCIÓN

Los sistemas expertos forman parte de la ciencia de la computación y constituyen una de las ramas de la inteligencia artificial. Estos sistemas son aplicaciones que usan el conocimiento y los procedimientos de inferencia para resolver problemas que son lo suficientemente difíciles como para requerir significativa experiencia humana para su solución [1] y son desarrollados para dominar un conocimiento en particular.

Las áreas específicas de interés para las cuales podemos aplicar Inteligencia Artificial son llamadas dominios. Sería imposible crear un programa que cubriera todos los dominios imaginables, el conocimiento requerido sería infinito. El camino que se sigue es el de confinarse a las áreas de las que se tiene suficiente información y que pueden ajustarse a un programa de computadora. Se puede decir que un sistema de Inteligencia Artificial creado para resolver problemas en un dominio particular es llamado Sistema Experto basado en conocimiento [2].

Características de los SE

Los SE están compuestos por dos partes principales: el ambiente de desarrollo y el ambiente de consulta. El ambiente de desarrollo es utilizado por el constructor para crear los componentes e introducir conocimiento en la base de conocimiento. El ambiente de consulta es utilizado por los no-expertos para obtener conocimiento experto y consejos [3].

Los componentes básicos de un SE son [2]:

✓ **Subsistema de adquisición de conocimiento:** Es la acumulación, transferencia y transformación de la experiencia para resolver problemas de una fuente de conocimiento a un programa de computadora para construir o expandir la base de conocimiento.

✓ **Base de conocimiento:** Contiene el conocimiento necesario para comprender, formular y resolver problemas.

✓ **Base de hechos:** Es una memoria de trabajo que contiene los hechos sobre un problema determinado.

✓ **Motor de inferencia:** Es el cerebro del SE, es el interpretador de reglas. Este componente es un programa que provee metodologías para el razonamiento de información en la base de

conocimiento. Este componente controla los pasos para resolver el problema cuando se realiza una consulta.

✓ **Subsistema de justificación:** Se encarga de explicar el comportamiento del SE al encontrar una solución. Permite al usuario hacer preguntas al sistema para poder entender las líneas de razonamiento que este siguió. Resulta especialmente beneficioso para usuarios no expertos que buscan aprender a realizar algún tipo de tarea.

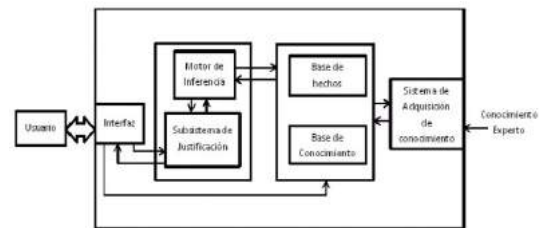


Figura 1: Estructura de un SE [2].

1. LÍNEAS de INVESTIGACIÓN y DESARROLLO

La línea de trabajo principal de este proyecto de investigación es el estudio de los Sistemas Expertos, sus características principales, su vinculación con la inteligencia artificial, las diferentes tecnologías asociadas y su posible aplicación en el poder judicial.

Además, entre las líneas de investigación a considerar en este proyecto se pueden mencionar:

- Diferentes técnicas aplicadas a Sistemas Expertos.
- Algoritmos utilizados en la Minería de Datos y en Machine Learning para el descubrimiento de patrones.
- Confeccionar los Árboles de decisión binarios.

Dado que el presente proyecto tiene como uno de sus objetivos el desarrollo de un prototipo informático se decidió utilizar la metodología de gestión de proyectos Scrum.

Este proyecto se aboca a la construcción de un prototipo de un Sistema Experto referido al dominio jurídico, el mismo ha sido denominado "Experticia". Para

contextualizar los procesos en cualquier fuero, ha de tenerse en cuenta que la función de los jueces implica la toma de decisiones ante situaciones determinadas relacionadas a los procesos a su cargo, estas se expresan a través de sentencias o resoluciones, para ello cuenta con la colaboración de distintos colaboradores a su cargo. Los procesos para realizar son muy estructurados y representa un gran caudal de trabajo diario.

Experticia está basado en la confección, por parte del usuario, de árboles binarios de decisión, los cuales recibirán como parámetros de entrada referencias almacenadas en la base de conocimiento del sistema. Luego, mediante la asistencia del Sistema Experto se completará información relacionada con la causa. Entregando como resultado uno o más trámites con los documentos, permitiendo automatizar el proceso de decisión en función del estado y datos del expediente electrónico.

Por lo tanto, como ya se mencionó, el propósito del presente proyecto se orienta al estudio de los elementos requeridos para el diseño y desarrollo de un prototipo de SE que ayude con la resolución de expedientes judiciales en el Poder Judicial de la Provincia de Buenos Aires.

2. RESULTADOS OBTENIDOS/ESPERADOS

En el Poder Judicial de la Provincia de Buenos Aires se utiliza el Sistema Informático de Gestión Asistida Multi-fuero y Multi-Instancia (GAM), más conocido como Augusta¹. Este sistema es utilizado en toda la provincia, como única herramienta informática, para la asistencia integral en la gestión de causas de los organismos jurisdiccionales de diferentes

¹

<http://www.scba.gov.ar/subinformacion/augusta.asp>

instancias y fueros. Experticia pretende dar soporte a las decisiones que permitan la resolución de una causa.

Por tal motivo, el proyecto fue presentado el año pasado a los responsables de la Subsecretaría de Tecnología Informática de la Suprema Corte de Justicia de la Provincia de Buenos Aires. Sector responsable del desarrollo de Augusta. Como resultado de esta presentación la Suprema Corte de Justicia de la Pcia. de Buenos Aires (SCJPBA) y la Universidad Nacional de La Matanza firmaron un convenio de Colaboración Recíproca con la idea de que Experticia funcionará como complemento de Augusta, y facilitará la automatización de diferentes procesos dentro de un expediente judicial.

De acuerdo al convenio firmado, ambas instituciones “intercambiarán entre sí todo tipo de datos, observaciones, memorias, publicaciones y toda otra documentación que pudiera estar relacionada con el desarrollo del SE. El resultado del trabajo será propiedad intelectual de la Suprema Corte mientras que la universidad tendrá el derecho de difundir y publicar los resultados [4], [5].

En la actualidad un primer prototipo de la aplicación Experticia, está siendo utilizado en el Juzgado de Ejecución Nro. 2 del Departamento Judicial Morón. La misma contempla:

1. La definición y gestión de los modelos de proceso a aplicar en cada tipo de trámite. La resolución por adoptar dependerá de características propias de cada causa. Inicialmente, como se mencionó antes, el algoritmo de decisión es definido a partir de árboles binarios. Se programa la secuencia de preguntas, conexiones entre árboles y los documentos de salida (Ver figura 3).

El grupo de investigación, presentó, en el año 2020, las siguientes publicaciones:

1. “*Predicción del riesgo de abandono universitario utilizando métodos supervisados*”. XXII Workshop de Investigadores en Ciencias de la Computación (WIIC 2020). Se realizó en forma de “Exposición Virtual”.
2. “*Sistema Experto para Apoyo del Proceso de Despacho de Trámites de un Organismo Judicial*”. Jornadas Argentinas de Informática (JAIIO 2020). Se realizó en forma de “Exposición Virtual”.
3. “*Metodológica para evaluar un modelo de Justicia Predictiva*”. Trabajo presentado en el Workshop del VIII Congreso Nacional de Ingeniería en Informática/Sistemas de Información. CONAISI 2020. Se realizó en forma de “Exposición Virtual”.

Así también el proyecto ha sido difundido en otros eventos:

- La Dra. Laura Conti ha participado de la Mesa de Análisis titulada *Digitalización Judicial en América Latina* organizada por la Academia Mexicana de Derecho en julio de 2020. Disponible en: <https://www.youtube.com/watch?v=K3s5EWZLSaE>
- El equipo de investigación fue invitado a exponer el proyecto en el *IV Encuentro del Programa MEP - Mejora de las Estrategias Pedagógicas*, organizado por el Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM realizado en diciembre de 2020. Los expositores en el encuentro fueron Sposito, Ledesma, Conti y García.

3. FORMACIÓN DE RECURSOS

HUMANOS

El equipo de trabajo está integrado por tres docentes de la carrera de Ingeniería en Informática, tres docentes de la carrera de Derecho y tres alumnos de la carrera de ingeniería, dos personas están como Asesores-Especialistas Externo. -

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

Por otra parte, los docentes investigadores que integran el proyecto realizaron diversos cursos relacionados con la temática del proyecto, entre estos, cabe mencionar los cursos de “Inteligencia Artificial y Derecho”, “Data Analytics and Visualization”, así como otros referidos a temas específicos de programación. El objetivo buscado es sumar saberes que luego serán convertidos en cursos para la comunidad educativa.

4. BIBLIOGRAFÍA

- [1] Fuentes Covarrubias, R.; Fuentes Covarrubias, A. (2013). Desarrollo de un sistema experto para el diagnóstico de fallas automotrices. Facultad de Ingeniería Mecánica y Eléctrica, Universidad de Colima, Colima, México. En línea en: http://sedici.unlp.edu.ar/bitstream/handle/10915/32398/Documento_completo.pdf?sequence=1&isAllowed=y. Último acceso 14/02/2021.
- [2] Badaró, S; Ibañez, L y Agüero, J. (2013). *Sistemas Expertos: Fundamentos, Metodologías y Aplicaciones*. Universidad de Palermo Ciencia y Tecnología, Facultad de Ingeniería. En línea en: https://www.palermo.edu/ingenieria/pdf2014/13/CyT_13_24.pdf. Último acceso 14/02/2021.
- [3] Turban, E. (1995). *Decision Support and Expert Systems* (4ta edición). EE.UU. Prentice-Hall.
- [4] Acta de Implementación Específica N° 1. En línea en: <http://www.scba.gov.ar/includes/descarga.asp?id=46049&n=Acta.pdf>. Último acceso 14/02/2021.
- [5] Convenio Marco de Colaboración Recíproca. En línea en: <http://www.scba.gov.ar/includes/descarga.asp?id=46049&n=>

Convenio%20marco.pdf. Último acceso
14/02/2021

LOGISTICA VERDE DE RESIDUOS TECNOLÓGICOS UNA RECOLECCIÓN INTELIGENTE Y SOSTENIBLE

D. Pandolfi¹, J. Rasjido¹, A. Villagra¹, S. Orozco¹, D. Perez¹, V. Varas, M. Bilbao, F. Brandan
G. Leguizamon¹²

{dpandolfi, jrasjido, avillagra, sorozco, dperez, vvaras, mbilbao, fbrandan}@uaco.unpa.edu.ar, legui@unsl.edu.ar

¹Laboratorio de Tecnologías Emergentes (LabTEM)
Instituto de Tecnología Aplicada (ITA) - Unidad Académica Caleta Olivia
Universidad Nacional de la Patagonia Austral

²Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)
Departamento de Informática - Universidad Nacional de San Luis

RESUMEN

Una ciudad inteligente es una ciudad donde el desempeño de sus componentes fundamentales (economía, movilidad, medioambiente, salud, gobernanza, etc.), están construidas sobre la base de nuevas capacidades donde sus ciudadanos pueden afrontar de manera inteligente los desafíos futuros de sostenibilidad urbana. La creciente demanda en el uso las tecnologías de la información y comunicación (TIC) han revolucionado nuestra vida cotidiana, sin embargo, han contribuido a aumentar las preocupaciones ambientales y sociales.

La recuperación de residuos tecnológicos requiere de estrategias inteligentes de recolección que minimicen los costos asociados, maximicen el reuso de materias primas y la reparabilidad de los bienes para la prolongación de su vida útil. Este trabajo consiste en generar rutas recolección de residuos tecnológicos dado una cantidad de clientes por atender, un conjunto de vehículos de recogida, permitiendo minimizar ciertos factores que ayuden a la empresa a obtener beneficios.

Palabras clave: Ciudades Inteligentes, Residuos Tecnológicos, Problema de enrutamiento de vehículos, Inteligencia Computacional.

CONTEXTO

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/B252 “Rutas inteligentes en la recolección de residuos tecnológicos”. Este proyecto se desarrolla en cooperación con el LIDIC de la UNSL, y el Grupo NEO de la UMA (España).

1. INTRODUCCIÓN

El paradigma de la economía circular plantea que los residuos al final del ciclo de vida útil de un producto pueden ser materias primas idóneas para unos nuevos productos. *Great Recovery* es un proyecto con el apoyo de *Innovate UK* (Londres, Reino Unido) que busca trabajar con la basura y los residuos que se generan a diario, para generar nuevos procesos de economía circular. Los equipos eléctricos y electrónicos (EEE) se convierten en uno de los grupos más importantes de residuos que contienen materiales fáciles de reciclar. Los desechos electrónicos son conocidos como RAEE (en español, Residuos de Aparatos Eléctricos y Electrónicos), o *e-waste*, o bien WEEE (en inglés, *Waste Electrical and Electronic Equipment*).

Los WEEE, contienen muchas sustancias que son tóxicas y potencialmente peligrosas para el medio ambiente y la salud humana (Oguchi et al., 2012). La reducción, reciclaje y eliminación de los desechos finales de los WEEE se ha convertido en uno de los principales objetivos para el desarrollo sostenible de la creciente industria tecnológica. La tasa de WEEE está creciendo a un ritmo alarmante, especialmente en los países desarrollados, donde los mercados están saturados con grandes cantidades de nuevos productos electrónicos. En 2013, un informe de SEPI (*Solving the e-waste Problem Initiative*), una iniciativa financiada por Naciones Unidas alertó que la cifra de WEEE se había elevado hasta cerca de los 49 millones de toneladas, un promedio de 7 kilos por cada uno de los 7.000 millones de habitantes del planeta Tierra. A la luz de estos datos, no resulta muy difícil hacerse una idea de la importancia que adquiere el reciclaje de este tipo de residuos.

Las empresas de recolección de residuos aplican diversos métodos de recolección, y pueden dividirse en estacionarios y móviles (Baxter et al., 2016). El objetivo tradicional de la gestión de distribución o recolección de productos consiste en minimizar los costos de todo el sistema. *Green Logistics* (GL) se ha convertido en una tendencia en la gestión de la distribución de bienes y la recolección de productos al final de su vida útil. Con su enfoque en maximizar el valor económico y ambiental mediante el reciclaje y el control de emisiones, GL contribuye al desarrollo sostenible de la industria, pero también requiere un esquema de transporte más completo cuando se realizan servicios de logística (Lin et al. 2014 (a) (b)).

Las metaheurísticas (MHs) son métodos que integran de diversas maneras, procedimientos de mejora local y estrategias de alto nivel para crear un proceso capaz de escapar de óptimos locales y realizar una búsqueda robusta en el espacio de búsqueda. En su evolución, estos métodos han incorporado diferentes estrategias para evitar la convergencia a óptimos locales,

especialmente en espacios de búsqueda complejos. En otras palabras, las MHs proveen de un marco general que permite crear nuevos híbridos a través de la combinación de conceptos derivados de: heurísticas clásicas, inteligencia artificial, evolución biológica, sistemas naturales, mecánica estadística, etc.

Boussaïd et al. 2013, desarrollan un estudio de varios algoritmos de optimización inspirados en las metáforas del comportamiento de enjambre en la naturaleza, conocida como Inteligencia Colectiva, (SI, *Swarm Intelligence*). Ejemplo de SI son Optimización de Colonias de Hormigas (ACO, *Ant Colony Optimization*), la Optimización de Cúmulo de Partículas (PSO, *Particle Swarm Optimization*), la Optimización de Forrajeo Bacteriano (BFOA, *Bacterial Foraging Optimization Algorithm*), Optimización de Colonias de Abejas (BCO, *Bee Colony Optimization*), Sistemas Inmunes Artificiales (AIS, *Artificial Immune Systems*) y Optimización Basada en Biogeografía (BBO, *Biogeography-Based Optimization*).

En particular, ACO fue propuesta fue propuesta por Dorigo et al. (1996, 1991) como una metaheurística inspirada en el comportamiento de una colonia de hormiga en el proceso de forrajeo y aplicada principalmente para la solución de problemas de optimización combinatoria.

Los algoritmos Genéticos Celulares (cGAs) son una subclase de un Algoritmo Genético con una población estructurada espacialmente, es decir, los individuos de la población pueden aparearse solo con sus vecinos. Los cGAs, se diseñaron inicialmente para trabajar en máquinas paralelas, formadas por muchos procesadores que ejecutaban simultáneamente las mismas instrucciones sobre diferentes datos (máquinas SIMD - El primer modelo de cGA fue propuesto por Robertson en 1987 (Alba y Dorronsoro 2009)).

El problema de recolección estacionaria se lleva a cabo en puntos de recolección específicos de desechos, ubicados en lugares popula-

res o en los alrededores de los lugares de residencia. La elección de la ubicación de dichos centros y la elección de un recorrido óptimo representa distintos problemas de decisión. Así mismo, la recolección móvil, propone nuevos medios de recolección de WEEE a pedido y principalmente está asociado a residuos tecnológicos medianos o grandes (Nowakowski et al., 2017).

Las distintas estrategias de recolección de e-waste pueden asociarse al conocido problema de enrutamiento de vehículos (VRP, *Vehicle Routing Problem*). Dantzig y Ramser (1959), formalizaron el VRP junto con sus numerosas variantes, y posteriormente Lenstra y Kan (1981) demostraron que dicho problema combinatorio es NP-Hard. En general, VRP consiste en asignar una cantidad de vehículos homogéneos a varios clientes, donde cada cliente tiene una cierta ubicación y demanda de productos (homogéneos). El objetivo de la optimización es entregar los bienes demandados a todos los clientes al tiempo que minimiza la suma de los costos (longitudes) de las rutas de los vehículos.

Existen diferentes variantes del VRP (Toth y Vigo, 2014), que incluyen restricciones adicionales y la incorporación de múltiples variables. Algunas restricciones más importantes son: a) cada vehículo tiene capacidad limitada (CVRP), b) todos los clientes deben ser atendidos dentro de una cierta ventana de tiempo (VRPTW), c) el proveedor utiliza muchos depósitos para suministrar a los clientes (*Multiple Depot* MDVRP) d) los clientes pueden devolver algunos productos (VRPPD) e) los clientes pueden recibir servicio por diferentes vehículos (SDVRP) f) algunos valores (como el número de clientes, sus demandas, el tiempo de servicio o el tiempo de viaje) son arbitrarios (VRP estocástico - SVRP) g) las entregas se pueden realizar en algunos días (VRP periódico - PVRP).

En la actualidad el problema de recolección o distribución de productos a partir de un depó-

sito original (punto de origen) y una cantidad de clientes con una demanda por atender, juega un papel importante en empresas logísticas ya que deben planificar inteligentemente las rutas de manera tal que signifiquen considerables ahorros en costos tales como: el consumo de combustible, horas hombre, entre otros; que ayudarán a una mejor rentabilidad para los negocios hoy en día. En relación con una creciente sensibilidad ambiental los problemas de enrutamiento de vehículos presentan una extensa disponibilidad bibliográfica de los problemas de enrutamiento verde de vehículos (GVRP). Los diferentes tipos de problemas de enrutamiento de VRP en logística inversa, describen perfectamente la complejidad en los casos del mundo real.

En este artículo trabajamos la variante de VRP con capacidad limitada (CVRP), en el que cada vehículo tiene una capacidad uniforme de un único artículo. Definimos el CVRP sobre un grafo no dirigido $G = (V, E)$ donde $V = \{v_0, v_1, \dots, v_n\}$ es un conjunto de vértices y $E = \{(v_i, v_j) / v_i, v_j \in V, i < j\}$ es un conjunto de ejes.

El vértice v_0 es el depósito, y es desde donde m vehículos de capacidad Q deben abastecer a todas las ciudades o clientes, representados por un conjunto de n vértices $\{v_1, \dots, v_n\}$.

Definimos en E una matriz $C = (c_{ij})$ de costo, distancia o tiempo de viaje no negativos entre los clientes v_i y v_j . Cada cliente v_i tiene una demanda no negativa de artículos q_i y tiempos de entrega δ_i (tiempo necesario para descargar todos los artículos). Siendo v_1, \dots, v_m una partición de V , una ruta R_i es una permutación de los clientes en V_i especificando el orden en el que se visitan, comenzando y terminado en el depósito v_0 . El costo de una ruta dada $R_i = \{v_0, v_1, \dots, v_{k+1}\}$, donde $v_j \in V$ y $v_0 = v_{k+1} = 0$ (0 indican el depósito), viene dada por la ecuación (1):

$$Cost(R_i) = \sum_{j=0}^k c_{i,j+1} + \sum_{j=0}^k \delta_j q_j \quad (1)$$

y el costo de la solución al problema (S) es:

$$FCVRP(S) = \sum_{i=1}^m Cost(R_i) \quad (2)$$

El CVRP consiste en determinar un conjunto de m rutas (i) de costo total mínimo - como especifica la ecuación (2); (ii) empezando y terminando en el depósito v_0 ; de forma que (iii) cada cliente es visitado una sola vez por un sólo vehículo, sujeto a las restricciones (iv) de que la demanda total de cualquier ruta no exceda Q_i ; y (v) la duración total de cualquier ruta no supera el límite preseleccionado D_i . Todos los vehículos tienen la misma capacidad y transportan el mismo tipo de artículo. El número de vehículos puede ser un valor de entrada o una variable de decisión. En este estudio, la longitud de las rutas se minimiza independientemente del número de vehículos utilizados.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de trabajo se enfoca en la generación de rutas recolección que permitan minimizar ciertos factores económicos y ambientales y maximizar los beneficios referidos a al reciclaje y tratamiento de WEEE. De los objetivos económicos podemos mencionar: minimizar el tiempo de recolección, kilómetros recorridos, maximizar el ahorro de combustible en los vehículos, minimizar la cantidad de vehículos, todo lo cual llevaría a obtener menores costos. Un objetivo económico adicional está dado por la valoración de la recuperación las materias primas tales como material ferroso, cobre, y metales preciosos, etc. De los beneficios ambientales podemos mencionar: la disminución en la emisión de gases efecto invernadero (CO_2 , CH_4 y N_2O) y los gases de contaminación ambiental (CO , SO_2 , PM , NO_x). Tanto los objetivos económicos como ambientales promueven una mejor calidad de servicio e imagen, para las empresas y los gobiernos locales

3. RESULTADOS OBTENIDOS / ESPERADOS

Como objetivo general se espera modelar, resolver e implementar un amplio y variado conjunto de servicios inteligentes de la ciudad y producir un mayor impacto en Ciencia e Industria.

Como objetivos específicos se pretende: (a) Estudiar el estado de tecnologías y técnicas de Inteligencia Computacional para resolver problemas de VRP. (b) Proponer instancias de problemas relacionados a recolección de WEEE. (c) Proponer y seleccionar elementos de medición relacionados a los problemas identificados. (d) Diseñar y construir prototipos de software relacionados a los problemas.

Actualmente se han generado cuatro instancias de pruebas para para 75, 105, 150 y 210 locaciones de recogida de WEEE de cuatro localidades (Caleta Olivia, Comodoro Rivadavia, Trelew y Río Gallegos). Las instancias se han probado con dos algoritmos de inteligencia computacional (ACO y cGA) los resultados preliminares fueron publicados en Pandolfi et al. 2019.

Para elegir los vehículos más adecuados para la colección, analizamos las siguientes variables: total de km recorridos, Costo total de uso vehículo, Total de Combustible consumido, Total de CO_2 emitido y Tiempo Total de horas de Trabajo.

En general, con las instancias iniciales y teniendo en cuenta la cantidad de km recorridos y la cantidad de horas de trabajo, los resultados indican que los vehículos de alta carga con un mayor proceso de automatización tienen un mejor rendimiento. Para la variable de consumo de combustible y, por lo tanto, la variable de emisión de CO_2 , los vehículos más pequeños sin automatización han mostrado mejores resultados. Finalmente, solo para la variable de costo del vehículo, el vehículo con una capacidad de 22 m3 de carga es la mejor opción.

En trabajos futuros analizaremos en distintas ciudades y se propondrán nuevas funcio-

nes de optimización teniendo en cuenta las variables consideradas en este estudio.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por cuatro investigadores con distintos niveles de posgrado, tres Doctores en Ciencias de la Computación, dos Magíster en Ciencias de la Computación, tres Ingenieros en Sistemas y un estudiante de la Carrera Ingeniería en Sistemas.

Esta línea de investigación proporcionará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes. De igual forma, será un ámbito adecuado para la realización de tesis de grado. En ese sentido, dos integrantes de este proyecto de investigación está desarrollando su Tesis de Maestría en temáticas afines y un integrante está desarrollando su Tesis de doctorado. Además, se cuenta con un becario alumno de la carrera de Ingeniería en Sistemas.

BIBLIOGRAFÍA

- E. ALBA, y B. DORRONSORO. (2009). *Cellular genetic algorithms* (Vol. 42). Springer Science & Business Media.
- J. BAXTER, K.A., LYNG, C. ASKHAM, O.J. HANSEN (2016). "High-quality collection and disposal of WEEE: environmental impacts and resultant" issues. *Waste Manage.* 57, 17–26.
- I BOUSSAÏD, J. LEPAGNOT, P. SIARRY, A survey on optimization metaheuristics, *Information Sciences*, Volume 237, 2013, Pages 82–117, ISSN 0020v-0255.
- G.B. DANTZIG, J. RAMSER, (1959). The truck dispatching problem, *Manage. Sci.* 6 (1) (1959) 80–91.
- M. DORIGO, V. MANIEZZO, A. COLORNI, (1991) Positive Feedback as a Search Strategy, Technical Report 91-016, Dipartimento di Elettronica, Politecnico di Milano, Milan, Italy, 1991.
- M. DORIGO, V. MANIEZZO, A. COLORNI, (1996) The ant system: optimization by a colony of cooperating agents, *IEEE Transactions on Systems, Man, and Cybernetics Part B* 26 29–41.
- C. LIN, K.L. CHOY, G.T.S. HO, T.W. Ng (a), A Genetic Algorithm-based optimization model for supporting green transportation operations, *Expert Systems with Applications*, Volume 41, Issue 7, 2014, Pages 3284-3296, ISSN 0957-4174.
- C. LIN, K.L. CHOY, G.T.S. HO, S.H. CHUNG, H.Y. LAM (b), Survey of Green Vehicle Routing Problem: Past and future trends, *Expert Systems with Applications*, Volume 41, Issue 4, Part 1, 2014, Pages 1118-1138, ISSN 0957-4174.
- P. NOWAKOWSKI, A. KRÓL, B. MRÓWCZNSKA (2017), Supporting mobile WEEE collection on demand: A method for multi-criteria vehicle routing, loading and cost optimisation, *Waste Management*, Volume 69, 2017, Pages 377-392, ISSN 0956-053X.
- M. OGUCHI, H. SAKANAKURA, A. TERAZONO (2012). Toxic metals in WEEE: characterization and substance flow analysis in waste treatment processes. *Sci. Total, Environ.*
- D. PANDOLFI, J. RASJIDO, A. VILLAGRA, G. LEGUIZAMON (2019), Following the green footprint of technological waste: A smart and sustainable collection, *RPIC - Reunión de Trabajo en Procesamiento de la Información y Control*. ISBN 9789871648.
- P. TOTH, D. VIGO (Eds), (2014). *Vehicle Routing: Problems, Methods, and Applications*, Society for Industrial and Applied Mathematics.

Metaheurísticas, búsqueda estocástica y cómputo eficiente en optimización aplicada

Tomás Tetzlaff¹, Adriana Gaudiani¹, Andrés Rojas Paredes¹, Diego Encinas², Esteban Fassio¹, Mariano Trigila³, Rodrigo González¹, Daniel Bertaccini¹

¹Instituto de Ciencias, Universidad Nacional de General Sarmiento

²Proyecto SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

³Facultad de Ingeniería y Ciencias Agrarias, Pontificia Universidad Católica Argentina-CABA

{tetzlaff, agaudiani, arojas, efassio, dbertacc, rgonzalez}@campus.ungs.edu.ar, dencinas@unaj.edu.ar, mariano_trigila@uca.edu.ar

Resumen

Las metaheurísticas son técnicas de optimización y resolución de problemas computacionales que toman inicialmente una solución factible, la cual es luego mejorada usando procedimientos heurísticos conocidos, como recocido simulado, algoritmos genéticos, búsqueda tabú y redes neuronales. La búsqueda estocástica está presente en estos métodos y su importancia reside en ser una herramienta general de optimización cuyo estudio puede aportar mejoras para las metaheurísticas y desarrollar variantes de ellas. Este proyecto propone estudiar propiedades teóricas y prácticas de estas técnicas y su aplicación en los problemas en los que trabajan sus integrantes, brindando metodologías para incrementar la eficiencia de los algoritmos involucrados y la confiabilidad de los resultados que producen. Estas metodologías aprovechan los avances en los métodos y técnicas de la computación eficiente y del cómputo en paralelo para desarrollar los algoritmos necesarios para validar las nuevas propuestas mediante casos de experimentación. La implementación de los algoritmos que permiten realizar las experiencias requiere la utilización de técnicas de las ciencias de los datos, del cómputo de alto rendimiento y métodos del campo de la optimización.

Palabras Claves: *metaheurísticas, técnicas de optimización, estimación de parámetros, soluciones factibles, eficiencia.*

Contexto

La línea de Investigación que se presenta corresponde al Proyecto: 30/1147 - Metaheurísticas, búsqueda estocástica y aplicaciones con cómputo eficiente, del Instituto de Ciencias de la Universidad Nacional de General Sarmiento (UNGS). (2020-2022).

Existe cooperación con grupos de investigación en otras Universidades Argentinas y del exterior, con las cuales los integrantes del proyecto han participado en investigaciones en conjunto.

Introducción

Un método heurístico es un procedimiento que, teniendo conocimiento de un problema y de las técnicas aplicables, aporta soluciones o se acerca a ellas usando una cantidad de recursos razonable. Por lo general, este recurso es el tiempo de cómputo que se emplea en la búsqueda de la solución. Se trata de procedimientos inteligentes en el sentido de realizar una tarea que no es producto de un riguroso análisis formal ya conocido y aplicable al tema, que es a veces ineficiente en cuanto al tiempo de cómputo, sino de algoritmos de simulación, de búsqueda o evolutivos que utilizan y también aportan al conocimiento experto sobre la tarea. Una solución heurística de un problema es aquella proporcionada por un método heurístico, es decir, aquella solución sobre la que se tiene cierta confianza de que alcanza algún grado de optimalidad y/o factibilidad [1].

Las metaheurísticas (también llamadas heurísticas modernas) se han desarrollado principalmente a partir de la década de 1980. En general, las metaheurísticas toman inicialmente una solución factible, la cual es luego mejorada usando heurísticas de mejoramiento embebidas en una estructura más general, i.e., Recocido Simulado (*Simulated Annealing*, SA), Algoritmos Genéticos (*Genetic Algorithms*, GA), Búsqueda Tabú (*Tabu Search*, TS), y Redes Neuronales (*Neural Networks*, NN). La característica común de estos enfoques es el uso de mecanismos para encontrar en lo posible soluciones óptimas evadiendo óptimos locales [2].

Las metaheurísticas pueden manejar variables discretas y reales, siendo aplicadas en la actualidad a una amplia gama de problemas de optimización de manera efectiva. Básicamente, los enfoques metaheurísticos basados tanto en la trayectoria (sucesión de posibles soluciones) como en la población (conjunto adaptable de posibles soluciones) apuntan a ubicar el óptimo global en el espacio de la solución a través de movimientos aleatorios. La diferencia clave entre las metaheurísticas está en la forma en que proponen el siguiente movimiento en el espacio de la solución [3]

El avance constante de las tecnologías de cómputo provee recursos computacionales con más rapidez de procesamiento, lo cual permite a investigadores y profesionales de la ciencia y la ingeniería resolver problemas combinatorios complejos y de gran escala. Esto genera un interés creciente en resolver problemas complejos del mundo real utilizando métodos metaheurísticos implementados con técnicas que mejoran la eficiencia, como el cómputo paralelo que permite reducir los tiempos de procesamiento y obtener la mayor eficiencia de las plataformas de hardware heterogéneas mediante el diseño y desarrollo de algoritmos paralelos adecuados para lograrlo [4].

En particular la Ingeniería del Software, a pesar de ser un área reciente, es en la actualidad una importante fuente de problemas de optimización. Los ingenieros y gestores de proyectos de software se enfrentan diariamente a diversos problemas de optimización en los que las técnicas exactas no tienen cabida por el corto intervalo de tiempo en que se requiere una respuesta [5]. En Ingeniería de requerimientos el procesamiento de texto y la extracción automática de contenido es un tema en investigación en rápido desarrollo [6]. Durante el preprocesado del texto se obtienen características decisivas para la clasificación y priorización de contenidos. Luego se aplican técnicas avanzadas de procesamiento y análisis de lenguaje natural [7] [8]. Las técnicas heurísticas ayudan a la estimación de parámetros de los métodos predictivos que permiten acceder al conocimiento latente en el texto analizado.

Cabe la posibilidad, por tanto, de aplicar a estos problemas, algoritmos metaheurísticos que ofrezcan al ingeniero una solución de cierta calidad en un breve periodo de tiempo: un compromiso entre calidad de la solución y la rapidez en encontrarla. En éstos y otros campos de investigación a menudo es esencial modelar y resolver tareas de optimización, de aprendizaje o de investigación para aplicaciones que no admiten una fácil formulación.

Líneas de Investigación y Desarrollo

Se propone abordar las siguientes líneas de I+D desde la perspectiva del estudio de las metaheurísticas y del uso del cómputo de alto rendimiento (HPC), ambos como ejes centrales:

- Evaluación de escalabilidad y rendimiento mediante balance de carga, mapping y scheduling de procesos, métricas de prestación y de consumo.
- Paralelización de aplicaciones (numéricas y no numéricas) con alta demanda computacional y/o grandes volúmenes de

datos sobre arquitecturas multiprocesador distribuidas (puras e híbridas).

- Desarrollo de soluciones aplicadas a problemas de impacto social con requerimiento de HPC.
- Aplicaciones de la teoría de probabilidad. Propiedades y simulación de procesos estocásticos. Búsqueda estocástica. Metaheurísticas.
- Estadística matemática y sus aplicaciones en espacios de búsqueda. Métodos de Monte Carlo.

Algunas de los estudios que se abordan en este trabajo ya vienen desarrollándose en la UNGS y en los grupos externos que colaboran con este proyecto, otros se inician a partir de cooperaciones con otras Universidades o para formalizar trabajos finales de grado o postgrado.

Se ha iniciado una cooperación en el área de Simulación para investigación y cursos de postgrado con el Proyecto SimHPC de la Universidad Nacional Arturo Jauretche y para formalizar trabajos finales. Entre otros temas el Proyecto SimHPC enfoca al desarrollo de modelos y simuladores para el área de Salud como así también para el tratamiento de grandes volúmenes de datos.

Existe una fuerte colaboración, con el grupo HPC4EAS (*High Performance Computing for Efficient Applications and Simulation*) del Departamento de Arquitectura de Computadores y Sistemas Operativos de la Universidad Autónoma de Barcelona en la dirección de tesis de postgrado. Se espera consolidar a corto plazo un grupo de investigación, desarrollo y transferencia que generará oportunidades de formación de recursos humanos, proveerá de servicios a la comunidad en el área de estudio y potenciará los vínculos de cooperación con otras instituciones.

Resultados y Objetivos

Este proyecto tiene como objetivo principal el estudio y análisis de la aplicación de metaheurísticas y técnicas de búsqueda estocástica, que permitan optimizar el funcionamiento de programas utilizados en diversas áreas de desarrollo. Es también un objetivo en todos los problemas de aplicación incrementar la eficiencia de los algoritmos involucrados y la confiabilidad de los resultados que producen. En particular nos referiremos a los siguientes problemas:

- **Optimización de la simulación de sistemas físicos** En la Ingeniería Hidrológica, una amplia gama de aplicaciones con diferentes y novedosas metodologías y técnicas metaheurísticas son creadas para abordar una variedad de problemas del modelado hidrológico [9] [10]. Se busca mejorar la simulación mediante métodos de búsqueda de un conjunto óptimo de parámetros del modelo, interpretando como óptimo a la mejor aproximación a la realidad que se puede hacer con ese modelo. Se continúan los trabajos de investigación en curso aplicando heurísticas para estimar los parámetros de ajuste que disminuyen la complejidad y la probabilidad de errores en un simulador del cauce del Río Paraná y contribuir a mejorar su predicción con el menor costo computacional posible [11] [12][13].
- **Métodos de aceleración del procesamiento de lenguaje natural.** Los usuarios de aplicaciones móviles generan a través de las *app stores* comentarios y calificaciones para describir su experiencia con las aplicaciones que descargan y usan. Estos comentarios, también llamados *reviews*, tienen información muy valiosa para retroalimentar el proceso de desarrollo de *apps*, por ejemplo, se pueden encontrar comentarios en forma de pedidos de nuevos features y reporte de errores. Este tipo de información es relevante para el desarrollo y la evolución de las aplicaciones móviles [14]. La enorme expansión de las aplicaciones móviles y el constante crecimiento de la cantidad de *reviews* de usuarios provocan un cuello de botella que surge del procesamiento de

grandes volúmenes de datos. Se espera aplicar metaheurísticas para encontrar de manera eficiente los parámetros de los métodos predictivos que buscan el conocimiento latente en el texto. Se propone crear un pipeline de preprocesamiento de las *reviews* en modo secuencial y compararlo contra una implementación con HPC. Se aplicarán las técnicas clásicas de paralelismo: paralelismo de datos o paralelismo de tareas, y combinación de ambas técnicas. Se propone explorar diferentes tecnologías como multicore, GPGPU, cluster standard y cluster de Raspberry. Adicionalmente nos preguntamos sobre la eficiencia energética de nuestra solución. En este punto nos proponemos realizar una implementación sobre un cluster de Raspberry que es la alternativa más adecuada en términos energéticos. Esperamos cuantificar el gasto energético de nuestra solución.

- **Fundamentos teóricos de la búsqueda estocástica:** La búsqueda estocástica permite encontrar soluciones satisfactorias a un problema en algunos casos en tiempos muy reducidos, aun careciendo de la fundamentación sobre porqué se llegó a las mismas. Nos proponemos seguir avanzando con la teoría matemática que fundamenta las propiedades y la eficiencia de estas metodologías [15][16]. Asimismo, esperamos realizar aportes rigurosos en este campo para los métodos usados en aplicaciones biológicas [17].

Se espera consolidar un grupo de investigación, desarrollo y transferencia a la comunidad con capacidad para abordar estos tipos de problemas.

Formación de Recursos Humanos

Tres miembros del equipo poseen formación de postgrado a nivel de doctorado, uno de ellos es especialista en Cómputo de Altas Prestaciones. Uno de los investigadores está en su etapa final para obtener el grado de Doctor en Ciencias Informáticas, dirigido por uno de los miembros del proyecto y un investigador de la UAB

(España). Otro investigador está comenzando una Maestría en Ciencia de los Datos, de interés para este proyecto. Un área afín a los del proyecto (evolución por deriva estocástica en poblaciones biológicas) es el tema de una tesis en curso (Doctorado en Ciencia y Tecnología, UNGS), dirigida por un integrante del proyecto. Uno de los integrantes coordina una Diplomatura en Ciencias de Datos.

Bibliografía

1. Reeves C.R. Fitness Landscapes. In: Burke E., Kendall G. (eds) Search Methodologies, pp 681-705. Springer, Boston, MA. (2014) https://doi.org/10.1007/978-1-4614-6940-7_22
2. Du K.L., Swamy M.N.S.: Genetic Algorithms. In: Search and Optimization by Metaheuristics. Birkhäuser, Cham, (2016)
3. Piotrowski, A.P., Napiorkowski, J.J. Some metaheuristics should be simplified. Inf. Sci. (Ny) **427**, 32–62, (2018)
4. H Hennessy, J. L., Patterson, and D. A., Computer Architecture, Fourth Edition: A Quantitative Approach. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. (2006)
5. Tekiner F., Tsuruoka Y., Tsujii J., Ananiadou S., Keane J. Parallel Text Mining for Large Text Processing, pp. 348-353 In: Proceedings of IEEE CSNDSP 2010, 21-23 July, Newcastle, UK (2010)
6. E. Guzman, R. Alkadhi, and N. Seyff, “A Needle in a Haystack: What Do Twitter Users Say about Software? In: Proc. of the International Requirements Engineering Conference, pp. 96–105. (2016)
7. E. Oehri and E. Guzman. Same Same but Different: Finding Similar User Feedback Across Multiple Platforms and Languages, 2020 IEEE RE, Zurich, Switzerland, pp. 44-54. (2020) doi: 10.1109/RE48521.2020.00017
8. Guzman E. and W. Maalej. How do users like this feature? A fine-grained sentiment analysis of app reviews. In; 22nd IEEE International Requirements Engineering Conference (RE'14), p. 153–162, (2014).
9. Chau K.-W., Use of meta-heuristic techniques in rainfall-runoff modelling, *Water*, vol. 9, no. 3, pp. 186, (2017)

10. Fallah-Mehdipour E., Haddad O.B.: Application of Genetic Programming in Hydrology. In: Gandomi A., Alavi A., Ryan C. (eds) Handbook of Genetic Programming Applications. Springer, Cham, (2015).
11. Gaudiani, E. Luque, P. García, M. Re, A. De Giusti, M. Naiouf. Computing, a powerful tool for improving the parameters simulation quality in flood prediction. *Procedia Computer Science – Volume 29*, p. 299-309, (2014).
12. Gaudiani, E. Luque, P. García, M. Re, A. De Giusti, M. Naiouf. How a computational method can help to improve the quality of river flood prediction by simulation. In: *Advances and new Trends in Environmental Informatics*. Springer-Verlag Berlin Heidelberg – Alemania. Volume 5, p.337-351, (2015).
13. Trigila M., Gaudiani A., Luque E. (2018) Agile Tuning Method in Successive Steps for a River Flow Simulator. In: Shi Y. et al. (eds) *Computational Science – ICCS 2018, Lecture Notes in Computer Science*, Volume 10862. Springer, Cham, (2018).
14. E. Guzman and A. Paredes Rojas. Gender and User Feedback: An Exploratory Study, *IEEE 27th International Requirements Engineering Conference (RE)*, pp. 381-385. (2019) doi: 10.1109/RE.2019.00049.
15. Maffei A.L., Tetzlaff T. Propiedades de un procedimiento de búsqueda evolutiva con diversidad. LX Reunión Anual de Comunicaciones Científicas de la Unión Matemática Argentina. Tandil, Buenos Aires. (2010)
16. Maffei A.L., Tetzlaff T. A schema theorem for stochastic search with forced diversity. Seventh Regional Meeting on Probability and Mathematical Statistics. Santa Fe, Argentina, (2010).
17. Marengo J., Tetzlaff T., Cálculo recursivo de probabilidad de presencia de secuencias en cadenas markovianas, 13° Simposio Argentino de Investigación Operativa, 44 Jornadas Argentinas de Informática, Universidad Nacional de Rosario, Rosario, Argentina. (2015)

MODELO PRESCRIPTIVO DINÁMICO PARA UN SISTEMA DE EVENTOS COMPLEJO

Esteban Schab^(1,3), Carlos Casanova^(1,3) y Fabiana Piccoli^(1 y 2)

⁽¹⁾ Universidad Autónoma de Entre Ríos, Concepción del Uruguay

⁽²⁾ LIDIC- Univ. Nacional de San Luis, San Luis

⁽³⁾ Univ. Tecnológica Nacional, Facultad Regional Concepción del Uruguay, Concepción del Uruguay
Argentina

mpiccoli@unsl.edu.ar

{schabe, casanovac}@frcu.utn.edu.ar

RESUMEN

La toma de decisiones en contextos dominados por grandes volúmenes de datos debe conjugar dos atributos usualmente contrapuestos: calidad y velocidad. La disponibilidad de información generada por personas y dispositivos abre nuevos desafíos en el diseño de mecanismos que puedan aprovecharla, de tal manera de que sean capaces de determinar las decisiones de mayor utilidad sujetas a ventanas temporales que garanticen su factibilidad. Uno de estos mecanismos lo constituyen las analíticas en sus distintos tipos, las cuales buscan transformar los datos en información a través de técnicas diversas. En este trabajo proponemos una línea de investigación enfocada en la analítica prescriptiva, capaz de determinar acciones a ser ejecutadas en el momento (decisiones operativas) o en el futuro (decisiones tácticas para corto y mediano plazo, decisiones estratégicas para largo plazo) para lograr un objetivo deseado. Para componerlas se propone la utilización de desarrollos provenientes de la Inteligencia Computacional y de la Computación de Alto Desempeño con el fin de obtener, de forma colaborativa, calidad y velocidad en las decisiones.

Palabras clave: Inteligencia Computacional. Analíticas. Big Data. Computación de Alto Desempeño.

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro de los proyectos de investigación: “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” (LIDIC, UNSL), “Cómputo de Altas Prestaciones aplicado a la Solución de Grandes Problemas” (UADER) y “Descubrimiento de conocimiento en bases de datos” (GIBD, UTN).

1. INTRODUCCIÓN

La mejora continua y adaptativa de los procesos de negocio resulta clave para mantener la competitividad de las organizaciones. La digitalización de los procesos, así como el incremento en las tecnologías de monitoreo, han llevado a producir una enorme cantidad de datos, los cuales tienen un gran potencial para la mejora de los procesos conducida por analíticas [1] [2] [3].

Las analíticas buscan transformar los datos en conocimiento para la toma de decisiones [4], pudiendo distinguirse cuatro tipos de analíticas según el nivel de automatización del proceso [5]. Primero, la descriptiva intenta responder qué ha

pasado o está pasando y la diagnóstica por qué ha pasado o está pasando, analizando para ello datos históricos. Luego, la analítica predictiva busca responder qué sucederá, aplicando el conocimiento para predecir nuevos datos sobre el presente o el futuro (pronóstico). Cabe aclarar que ninguno de estos enfoques sugiere acciones concretas, sino que descansan en el juicio subjetivo y las habilidades analíticas del usuario para deducir acciones de mejora. En cambio, la analítica prescriptiva responde qué debería hacerse para lograr un objetivo: determina acciones a ser ejecutadas en el momento (decisiones operativas) o en el futuro (decisiones tácticas para corto y mediano plazo, decisiones estratégicas para largo plazo).

Si bien el volumen de datos generados por personas y dispositivos se encuentra en continuo crecimiento, y a pesar de los avances tecnológicos en general, las analíticas de procesos existentes dentro de la industria actual no aprovechan completamente el conocimiento oculto debido a las siguientes limitaciones [1]:

- a. No hacen uso de técnicas prescriptivas para transformar los resultados del análisis en acciones de mejora concretas, dejando este paso completamente a criterio del usuario.
- b. Hacen un uso intensivo de datos de sistemas en producción, generando un deterioro en el desempeño de las herramientas de software que soportan los procesos.
- c. La optimización es conducida *ex post*, después de completado el proceso, en contraste a la mejora proactiva durante la ejecución del proceso.

En el área de datos masivos (o Big Data), se identifica como área emergente el procesamiento de *datastreams*, también llamado *Data Stream Mining* [6] [7] [8]. Un *datastream* es una representación digital y transmisión continua de datos, los cuales describen una clase de eventos relacionada [9] [11]. Mediante el

procesamiento de estos *datastreams* se puede lograr la respuesta en tiempo real a los eventos en forma de toma de decisiones.

Los grandes volúmenes de datos generados pueden ser utilizados en algoritmos de diversa índole para la generación de analíticas, particularmente las descriptivas y predictivas, como por ejemplo los de aprendizaje supervisado o no supervisado. Existe, sin embargo, una dificultad en lo relativo a las analíticas prescriptivas: no cuentan con un “profesor” [10] que les diga qué acción tomar en cada circunstancia. Un tipo de aprendizaje que no necesita de un profesor es el llamado aprendizaje por refuerzo. 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 recompensas observadas para aprender una política óptima (o aproximadamente óptima) del entorno, sin asumir ningún conocimiento *a priori* [10]. 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, ya que ambos comparten el mismo objetivo.

La ejecución de los algoritmos para la generación de analíticas debe ser lo suficientemente rápida como para procesar los *datastreams* generados de manera continua por el sistema, lo que implica el uso de técnicas y herramientas de Computación de Alto Desempeño (HPC).

En la siguiente sección se describen con mayor detalle las características de esta línea de investigación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación se basa en la idea de la Optimización de procesos de negocio por recomendación. En este contexto, los *datastreams* abren nuevas y amplias oportunidades para la creación de valor en las organizaciones. Como se mencionó, mediante su procesamiento se puede conseguir respuesta en tiempo real a eventos en forma de toma de decisiones.

En [11], [12] y [13] encontramos distintos enfoques para la definición de eventos en un proceso de negocio, variando en su complejidad. Una secuencia o flujo de eventos (*Event Stream*) es una serie ordenada y potencialmente ilimitada de eventos [14]. Los flujos de eventos se generan y utilizan en muchos sistemas. Ejemplo de ellos son:

- Bancos: un caso de particular interés es el de los bancos y otros sistemas de atención al público [15]. En ellos la generación de *datastreams* es causada por los sistemas de gestión de atención implementados. Estos sistemas en general utilizan modelos relacionales de bases de datos y, si bien permiten la elaboración de analíticas, puede resultar inapropiada su implementación debido a que el desempeño del sistema de atención, especialmente en un contexto de *Software as a Service* (SaaS) [16], puede colapsar ante las continuas consultas para realizar el monitoreo. Por tal motivo, la generación y el procesamiento de los eventos en forma de *datastreams* como una componente paralela al sistema de atención puede ser la tecnología de base para permitir un monitoreo eficiente, sin deterioro del desempeño del sistema en general.
- Enrutamiento de vehículos: es otro caso de interés, conocido por su siglas VRP (*Vehicle Routing Problem*) [17] con suministro de

información y reencaminamiento en tiempo real, orientado a la búsqueda de un paradigma de movilidad inteligente [18]. Dentro de este problema se pueden estudiar de forma particular o en conjunto la logística urbana, el transporte de personas y los conductores individuales. En este caso los *datastreams* son generados de forma distribuida por cada agente involucrado y pueden ser procesados de forma centralizada o distribuida dependiendo del esquema elegido y los recursos disponibles.

Considerando estos problemas, esta línea propone la composición de modelos de analítica prescriptiva, para superar los inconvenientes descritos anteriormente. Estos modelos serán parte esencial de un proceso de mejora continua basado en la recomendación de acciones operativas y tácticas destinadas a mantener los indicadores de rendimiento del sistema dentro de valores deseados, en un contexto con grandes flujos de eventos y Computación de Alto Desempeño (HPC).

Para la construcción de estos modelos prescriptivos, cuya principal función es la determinación de las acciones a llevar a cabo, se debe hacer uso de modelos predictivos para explorar los futuros cercanos y modelos descriptivos para calcular la aptitud de dichos estados. Para ello se propone el uso de agentes basados en aprendizaje por refuerzo, técnicas provenientes de la Inteligencia Computacional: redes neuronales como modelos, teoría de conjuntos difusos como lenguaje de especificación, y métodos numéricos y metaheurísticos para el entrenamiento de tales modelos [19] [20] [21].

Además, ante la necesidad de dar rápida respuesta a los procesos de negocios dinámicos, y dadas las características propias de cada una de las técnicas de inteligencia computacional antes mencionadas, es mandatorio pensar en la

aplicación de modelos/paradigmas de computación de alto desempeño [22] [23] [24], principalmente en el entrenamiento de los modelos. Este proceso debe ser lo suficientemente rápido como para procesar los *datastreams* que el sistema genera de manera continua y brindar resultados en tiempo real.

3. RESULTADOS OBTENIDOS/ESPERADOS

El primer caso de aplicación que se encuentra en desarrollo es el de enrutamiento de vehículos, con suministro de información y re-encaminamiento en tiempo real. Este caso es un caso de logística urbana. Se espera construir un modelo prescriptivo dirigido por los datos, basado en la recomendación automática y proactiva de acciones operativas y tácticas destinadas a mantener los indicadores de rendimiento del sistema dentro de los valores deseados, considerando un contexto de gran volumen de eventos.

Para el procesamiento, los primeros desarrollos tienden a construir algoritmos aplicando HPC capaz de reducir los tiempos derivados de entrenar y ejecutar el modelo. Aunque uno de los objetivos es lograr una solución paralela portable, de costo predecible, capaz de explotar las ventajas de modernos ambientes HPC a través de herramientas y “frameworks de computación” de alto nivel [23] [24], los primeros desarrollos se están haciendo en GPU Nvidia con CUDA [25], se prevé el uso de otras tecnologías como OneAPI[26].

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados esperados respecto a la formación de recursos humanos son el desarrollo de 1 tesis de doctorado, 2 de maestría y de varias tesinas de grado en las universidades intervinientes.

5. BIBLIOGRAFÍA

- [1] Christoph Gröger, Holger Schwarz, and Bernhard Mitschang. “Prescriptive analytics for recommendation-based business process optimization”. In *International Conference on Business Information Systems*, pages 25–37. Springer, 2014.
- [2] Mandeep Kaur Saggi and Sushma Jain. “A survey towards an integration of big data analytics to big insights for value-creation”. *Information Processing & Management*, 54(5):758{790, 2018.
- [3] Usarat Thirathon, Bernhard Wieder, Zoltan Matolcsy, and Maria-Luise Ossimitz. “Impact of big data analytics on decision making and performance”. In *International Conference on Enterprise Systems, Accounting and Logistics*, 2017.
- [4] Clyde Holsapple, Anita Lee-Post, and Ram Pakath. “A unified foundation for business analytics. *Decision Support Systems*”, 64:130{141, 2014.
- [5] Michael Minelli, Michele Chambers, and Ambiga Dhiraj. “Big data, big analytics: emerging business intelligence and analytic trends for today’s businesses”. Volume 578. John Wiley & Sons, 2013.
- [6] Albert Bifet and Jesse Read. “Ubiquitous artificial intelligence and dynamic data streams”. In *Proceedings of the 12th ACM International Conference on Distributed and Event-Based Systems, DEBS ’18*, Pp. 1–6, Association for Computing Machinery. New York, USA, 2018.
- [7] Sergio Ramírez-Gallego, Bartosz Krawczyk, Salvador García, Michal Wozniak, and Francisco Herrera. “A survey on data preprocessing for data stream mining: Current status and future directions”. *Neurocomputing*, 239:39 – 57, 2017.
- [8] Taiwo Kolajo, Daramola Olawande, and Adebisi Ayodele. “Big data stream analysis: a systematic literature review”.

- Journal of Big Data, vol. 6, no 1, pp. 1-30, 2019.
- [9] Federico Pigni, Gabriele Piccoli, and Richard Watson. "Digital data streams: Creating value from the real-time flow of big data". *California Management Review*, 58(3):5–25, 2016.
- [10] Stuart J Russell and Peter Norvig. "Inteligencia Artificial: un enfoque moderno". 2004.
- [11] Chris Wrench, et al. "Data stream mining of event and complex event streams: A survey of existing and future technologies and applications in big data." *Enterprise Big Data Engineering, Analytics, and Management*. IGI Global, 2016. pp. 24-47.
- [12] Annika Hinze, Kai Sachs, and Alejandro Buchmann. "Event-based applications and enabling technologies." *Proceedings of the Third ACM International Conference on Distributed Event-Based Systems*. pp. 1-15. 2009.
- [13] Opher Etzion and Peter Niblett, "Event processing in action", Manning, 2011.
- [14] Thomas J. Owens, "Survey of event processing", Air Force Research Lab Rome NY Information Directorate, 2007.
- [15] T. Olanrewaju. "The rise of the digital bank." <https://www.mckinsey.com/business-functions/digital-mckinsey/ourinsights/the-rise-of-the-digital-bank/>. Accessed: 2020.
- [16] M. Turner, D. Budgen and P. Brereton, "Turning software into a service". In *Computer*, vol. 36, no. 10, pp. 38-44, Oct. 2003.
- [17] Kris Braekers, Katrien Ramaekers, and Inneke Van Nieuwenhuysse, "The vehicle routing problem: State of the art classification and review", *Computers & Industrial Engineering*, vol. 99, pp. 300-313, 2016.
- [18] Sandra Melo, Joaquim Macedo, and Patrícia Baptista, "Guiding cities to pursue a smart mobility paradigm: An example from vehicle routing guidance and its traffic and operational effects", *Research in transportation economics*, vol. 65, p. 24-33, 2017.
- [19] A. Ebrahimnejad and J. L. Verdegay. "Fuzzy sets-based methods and techniques for modern analytics". Springer International Publishing. 2018.
- [20] Nazmul Siddique and Hojjat Adeli. "Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing". John Wiley & Sons, 2013.
- [21] Lotfi A. Zadeh. "Fuzzy logic, neural networks, and soft computing". *Commun. ACM* 37, 3, 77–84. DOI: <https://doi.org/10.1145/175247.175255>. March 1994.
- [22] Mercedes Barrionuevo, Mariela Lopresti, Natalia Miranda, and María Fabiana Piccoli. "Solving a big-data problem with gpu: the network traffic analysis". *Journal of Computer Science and Technology*, 15(01): p.30–39, Apr. 2015.
- [23] S. Kurgalin and S. Borzunov, "A Practical Approach to High-Performance Computing". Springer. 2019.
- [24] Peter Pacheco. "An Introduction to Parallel Programming", 1st ed., San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.
- [25] Nvidia. "*CUDA C++ Programming Guide, Design Guide*". https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf. 2019.
- [26] OneAPI Specification, release 1.0-rev 3. <https://spec.oneapi.com/versions/1.0-rev-3/>. 2020.

Modelos de Pronósticos para Incendios Forestales en la Provincia de Córdoba

Julio Castillo¹, Marina Cardenas¹, Nicolás Hernández¹

¹ Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
{jotacastillo, ing.marinacardenas}@gmail.com

RESUMEN

En este artículo describimos el proyecto denominado Modelos de pronósticos para la detección de incendios forestales. Se describen los diferentes componentes que se encuentran operativos hasta el momento, sus objetivos y utilidades en lo que atañe a la gestión de incendios forestales.

Este proyecto tiene como objetivo principal elaborar un modelo que permita pronosticar la ocurrencia de incendios forestales en la Provincia de Córdoba, especialmente en las sierras de Córdoba y la región del Parque Chaqueño de la provincia.

Es parte de un desarrollo que involucra desde la recolección de la información meteorológica, su procesamiento, análisis, hasta su visualización.

Palabras clave: Pronósticos, modelo, incendios, forestales.

CONTEXTO

El proyecto presentado en este artículo se denomina *Modelos de pronósticos para la detección de incendios forestales*, siendo un proyecto homologado por rectorado de UTN y cuyo código de Proyecto es: SIUTNCO0005351.

Se encuentra consolidado dentro de la línea de investigación que se dedica a la resolución de problemas físicos, reales y sociales, a través del uso de herramientas computacionales basadas en aprendizaje automático.

En el mismo, se abordan temáticas de construcción de modelos de pronósticos y aplicación de diversas técnicas de inteligencia artificial. Físicamente, su desarrollo tiene lugar en el Laboratorio de Investigación de Software LIS¹ del Departamento de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba.

Formalmente, el proyecto se encuentra dentro del grupo UTN denominado GA2LA: Grupo de Aprendizaje Automático, Lenguajes y Autómatas y del LIS¹.

El grupo GA2LA [1] articula diversos proyectos de investigación que se encuentran en tres líneas de investigación diferentes, esto es, la relacionada a teoría de autómatas y lenguajes, otra relacionada a aprendizaje por computadora, y una última relacionada a la trazabilidad y

¹ www.investigacion.frc.utn.edu.ar/mslabs/

calidad en los proyectos de desarrollo de software. Adicionalmente, se mantiene una colaboración estrecha con expertos en ciencias sociales del IIGHI-CONICET, con los cuales se trabaja en el desarrollo conjunto de modelos teóricos-computacionales que permitan capturar la influencia del riesgo de la vivienda humana para la salud, basados en una perspectiva holística del riesgo.

Al grupo lo componen docentes-investigadores, doctorandos, pasantes y becarios.

1. INTRODUCCIÓN

Los incendios forestales son uno de los mayores problemas ambientales que se repiten año tras año produciendo graves daños en términos de recursos naturales, recursos económicos y muchas veces, en términos de pérdidas de vidas humanas [2] [3].

Por ello, un adecuado pronóstico que permita determinar la ocurrencia o no de incendios forestales en una fecha determinada, es un problema que, de ser resuelto, ayudaría a minimizar los efectos nocivos de los incendios forestales.

En este contexto, el proyecto de investigación aborda la problemática de la predicción de incendios forestales utilizando técnicas de aprendizaje automático por computadora [4] [5].

Existen muchas y variadas aproximaciones al problema de la predicción de fenómenos climatológicos. Entre las que destacamos aquellas basadas en sensores meteorológicos [6], otras basadas en satélites [7] y otras basadas en scanners de humo e infrarrojos [8].

En ese contexto, el problema es abordado desde la predicción de

incendios forestales mediante un modelo matemático-computacional que se nutre de información de diversas fuentes de datos [9].

Se trata de una tarea difícil, la cual abordamos desde un enfoque basado en aprendizaje automático, que nos provea de información útil para la planeación estratégica de recursos destinados a combatir los incendios forestales. Como resultado esperamos obtener un sistema que indique las áreas en las que ocurrirá un incendio y la cantidad de hectáreas que se verán afectadas. Se trata de un trabajo en progreso [10] [11] [12], y sus avances pueden observarse en la sección Resultados.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto involucra las áreas conocimientos de inteligencia artificial, en particular aquellos relacionados con aprendizaje automático por computadoras de tipo supervisado. A su vez, involucra la construcción de modelos de pronósticos. Se trata de un problema complejo y no lineal, el cual se ve afectado en gran medida por la mano del hombre, puesto que la mayor parte de los incendios son intencionales.

El proyecto actual se trata de la continuación de un proyecto previo, también homologado en UTN denominado “*Construcción de un modelo de pronósticos para predicción de incendios forestales en la provincia de Córdoba*”. Se ha continuado trabajando sobre los resultados de ese proyecto, lo cuales se pueden ver reflejados en los diferentes subsistemas descriptos en la siguiente sección.

El desarrollo del proyecto se mencionan como parte de los resultados alcanzados hasta el momento, y la innovación del proyecto radica en la articulación de diferentes fuentes de información para poder construir un modelo que mejor capture la ocurrencia de los incendios forestales.

3. RESULTADOS OBTENIDOS / ESPERADOS

Desde un enfoque sistémico, el proyecto se ha dividido en diferentes subsistemas que componen lo que hemos denominado Sistema de Gestión de información Meteorológica y de Predicción de Incendios Forestales.

Este sistema cuenta con varios subsistemas que se detallan a continuación:

- Subsistema de Recolección Automática: que permiten llevar a cabo un registro de información meteorológica, tanto manual como automáticamente. Incluye el registro diario de toda la información de las estaciones meteorológicas, así como el historial de las mismas.

- Subsistema de chat de consultas y de reporte de incendios: posibilita a usuarios realizar un reporte de un avistamiento de incendios al personal de bomberos.

Los anteriores se complementan con dos módulos adicionales que permiten:

- Subsistema de Reportes: se encarga de la generación de reportes de incendios para el personal decisor.

- Subsistema predicción de incendios forestales: se trata del principal módulo del sistema, puesto que permite realizar una predicción efectiva y se articula con los demás módulos para que el resultado de las predicciones puedan ser utilizadas por los tomadores de decisión. Las

predicciones son llevadas a cabo bajo el paradigma del aprendizaje automático supervisado, y el modelo permite predecir la ocurrencia y la cantidad de hectáreas que probablemente se vean afectadas.

- Subsistema de visualización: Este subsistema provee información visual acerca de los datos recolectados por el Subsistema de Recolección Automática de Información Meteorológica. Su visualización se basa en un cálculo de índices de riesgo. El índice de riesgo de incendio forestal, se clasifica en base a una escala de colores que varían desde rojo si el riesgo de incendio es extremo, anaranjado si es muy alto, amarillo si es alto, azul si es moderado y verde si es un riesgo de incendio forestal bajo. Estos índices son propios de cada estación meteorológica y son calculados en base a la información meteorológica.

Este sistema muestra mediante mapas electrónicos distintas situaciones: la predicción de incendios, los índices de riesgo, las mediciones atmosféricas, los siniestros, entre posibilidades de visualización de información.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación y desarrollo de software está formado principalmente por docentes investigadores de la Universidad Tecnológica Nacional, Facultad Regional Córdoba. Además, anualmente se incorporan al proyecto (por el lapso de un año), becarios alumnos, y eventualmente un becario de posgrado.

Adicionalmente, participan del proyecto alumnos de cuarto o quinto año

de la carrera de Ingeniería en Sistemas de Información de la misma universidad, en el contexto de su desarrollo de prácticas supervisadas que es parte de los requisitos para la obtención del grado de Ingeniero.

Los becarios alumnos que participan del proyecto aprenden diversas tareas entre las que destacamos "como trabajar en un proyecto de investigación", esto permite complementar su formación curricular desde el punto de vista científico.

Por otra parte, en el marco del proyecto se han desarrollado una tesina de especialista y una tesis de maestría en Ingeniería en sistemas de información en la UTN-FRC.

En síntesis, el equipo de investigación está conformado por:

- 4 docentes investigadores (incluye un doctor y 1 doctorando en ingeniería con mención en sistemas, y 1 maestrando en ingeniería en sistemas, todos en UTN-FRC).
- 2 alumnos auxiliares.
- 2 alumnos becarios.
- 1 alumnos practicantes.

5. BIBLIOGRAFÍA

- [1] Juan C Vázquez, Julio J Castillo, Leticia Constable, Marina E Cardenas. (2018). XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).
- [2] Calderón Marnez, J., A., Gavarrete Galdámez, O., A., & Guzmán Navarrete, R., S. (2013). La prevención y manejo de desastres ambientales que afectan el medio ambiente, el patrimonio y la vida de los salvadoreños. Tesis de Licenciatura. Universidad de El Salvador.
- [3] Cruz, A. (2005). Causa calentamiento global desastres. Diario: El Universal. México.
- [4] Castillo, J., Cardenas, M., Medel, R., Casco, O., Navarro, M. & Gutierrez, S. (2016). "Sistema para Predicción de Incendios Aplicado a la Provincia de Córdoba". XVII Workshop de Investigadores en Ciencias de la Computación, WICC'2016, Entre Ríos, Argentina.
- [5] Martín Navarro Mugas, Osvaldo Casco, Marina Cárdenas, Julio Castillo. (2017). "Uso de tecnologías web y móvil para alertas tempranas de incendios en la Provincia de Córdoba". 46 JAIHO-CLEI. EST 2017. ISSN: 2451-7615. UTN-FRC. Córdoba, Argentina. Septiembre de 2017.
- [6] Chandler, C., Cheney, P., Thomas, P., Trabaud, L., & Williams, D. (1983), Fire in Forestry, Vol I. Forest Fire Behaviour and Effects, John Wiley, New York, p. 450.
- [7] Rauste, Y. (1996) Forest Fire Detection with Satellites for Forest Fire Control. Archives of Photogrammetry and Remote Sensing, Vol.31, Part B7, Proc. XVIII Congress of the Int'l Soc. for Photogrammetry and Remote Sensing, ISPRS, 584–588.
- [8] Arrue, B., Ollero, A., & Matinez de Dios, J. (2000) An Intelligent System for False Alarm Reduction in Infrared Forest-Fire Detection. IEEE Intelligent Systems, 15(3):64–73.
- [9] Rauste, Y. (1996). "Forest Fire Detection with Satellites for Forest Fire Control," Int'l Archives of Photogrammetry and Remote Sensing, Vol.31, Part B7, Proc. XVIII Congress of the Int'l Soc. for Photogrammetry and Remote Sensing, ISPRS, pp. 584–588, 1996.
- [10] Cardenas Marina, Navarro Martín, Hernández Nicolás, Castillo Julio. (2018).

"Recolección Automática de Información Meteorológica". 6to. Congreso Nacional de Ingeniería Informática – Sistemas de Información - CONAIISI 2018. Mar del Plata, Buenos Aires. Noviembre de 2018.

[11] Cardenas Marina, Nicolás Hernandez, Martín Navarro Mugas, Castillo Julio. (2019). "Herramientas de Recolección, Visualización y Análisis de Incendios Forestales", 7mo CONAIISI. 14-15 de Nov. de 2019 en la Universidad

Nacional de La Matanza – San Justo, Buenos Aires, Argentina.

[12] Castillo Julio, Hernández Nicolás, Cardenas Marina. (2020). "Visualización de Información Meteorológica mediante Mapas Digitales Interactivos". 8vo. Congreso Nacional de Ingeniería Informática – Sistemas de Información - CONAIISI 2020. UTN San Francisco, Córdoba.

Optimización de Carteras de Inversión: Un Benchmark con Modelos Clásico, de Computación Cuántica y de Hibridación AI /QC

Juan Pablo Braña, Alejandra M.J. Litterio y Alejandro Fernández
Centro de Altos Estudios en Tecnología Informática- Facultad de Tecnología Informática -
U.A.I
Av. Montes de Oca 745 - (C1270AAH) Ciudad Autónoma de Buenos Aires,
República Argentina
{juan.brana, alejandra.litterio, alejandroa.fernandez}@uai.edu.ar

RESUMEN

La optimización de carteras de inversión representa un desafío para el inversor al momento de seleccionar la combinación correcta de acciones a efectos de maximizar los retornos esperados y minimizar los riesgos. Es así que en contraste con el paradigma de la computación tradicional, la computación cuántica no sólo acelera de manera sustancial el proceso de las computadoras tradicionales si no que optimiza la performance del método que determina cuáles acciones deben ser incluidas en las carteras de inversión y cuáles no. En este trabajo presentamos un estudio de comparación y referencia entre el resultado obtenido por tres diferentes metodologías de optimización de cartera: a) aproximación clásica, b) aproximación cuántica, c) un híbrido entre la aproximación cuántica incorporando criterios de Inteligencia Artificial. Para el primero de los casos utilizamos el modelo de Markowitz, el cual es un algoritmo clásico para determinar carteras eficientes, para nuestra aproximación cuántica hemos trabajado con un algoritmo de optimización cuántico llamado *Variational Quantum Eigensolver* (VQE) y por último hemos intentado mejorar este último criterio de optimización con un índice de

sentimiento calculado con procesamiento de lenguaje natural (NLP) y una métrica de *forecasting* multivariado basado en *Machine Learning*.

Palabras clave: Computación Cuántica, Procesamiento de Lenguaje Natural Cuántico, Optimización de Cartera de Inversión, Finanzas, Machine Learning, AI

CONTEXTO

El presente trabajo es parte del proyecto “Inteligencia Artificial y Computación Cuántica en Finanzas”, iniciado en Marzo 2020, el cual se desarrolla en el Centro de Altos Estudios en Tecnología Informática (CAETI) dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI). Se enmarca dentro de la línea de Algoritmos y Software y continúa las investigaciones iniciadas en el proyecto “Modelo de Sentiment Analysis para la clasificación de noticias en tiempo real”. Es financiado y evaluado por la Secretaría de Investigación de la Universidad. Cuenta con la participación de docentes y alumnos de la Maestría en Tecnología Informática y de la Diplomatura en Análisis de Datos para Negocios, Finanzas e Investigación de Mercados.

1. INTRODUCCION

El inversor racional siempre ha buscado dos metas cuando toma la decisión de conformar una cartera de inversión: a) maximizar los retornos, b) disminuir el riesgo. Para ello, diferentes metodologías han sido empleadas, desde un enfoque clásico como el Modelo de Markowitz [12], uno de mayor reconocimiento por la industria financiera, por un lado, y un enfoque moderno, por el otro lado, que nos presenta algoritmos de *Machine Learning*, tanto modelos supervisados como no-supervisados que permiten la selección de instrumentos financieros y su rebalanceo periódico de la cartera, incluso, híbridos con Aprendizaje por Refuerzo (*Reinforcement Learning*) los que han sido presentados con éxito [1]. En relación a lo mencionado y trabajos anteriores han realizado numerosos experimentos y benchmarking en un esfuerzo por evaluar y e identificar las ventajas en aplicaciones en el mundo real [6], [8]. De la misma manera, se ha observado que la Computación Cuántica (QC) cuenta con el potencial de resolver problemas de índole financiera como detección de fraude, *pricing* de derivados [6], [8], [14], [16], [18] riesgo crediticio [5], [19], entre otros. En particular, el caso que nos ocupa, la Computación Cuántica como mecanismo de optimización de carteras de inversión ha introducido un nuevo paradigma [10], [11], [13], [15], [17]. Además, se espera que la fusión entre la Computación Cuántica y la Inteligencia Artificial sea la próxima revolución que cambie el rumbo de la industria, no solo la financiera, sino cualquiera relacionada con grandes volúmenes de procesamiento, velocidad y seguridad en los datos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente estudio se enmarca en la investigación de aplicación de técnicas y algoritmos de minería de datos y pretende dar un enfoque y soporte académico a toda la comunidad sea que esté relacionada a la industria financiera o dentro del marco académico. En este estudio, damos continuidad a las estrategias desarrolladas para codificar algoritmos de *Machine Learning* en computadoras cuánticas para la administración de carteras de inversión y su relación con trading algorítmico.

3. RESULTADOS OBTENIDOS/ESPERADOS

Pregunta Problema:

Una aproximación híbrida entre la Computación Cuántica e Inteligencia Artificial podría conseguir mejores resultados para un modelo de optimización de cartera de inversión en comparación con los resultados obtenidos con una aproximación que solo utiliza Computación Cuántica.

Objetivo General:

Realizar un benchmark entre los resultados obtenidos por un modelo clásico, uno basado en Computación Cuántica, y otro híbrido entre Computación Cuántica e Inteligencia Artificial.

Objetivos Específicos:

- Crear una metodología que permita generar un modelo híbrido entre los resultados obtenidos por los modelos de *Machine Learning* de *forecasting*, nuestro indicador de Sentimiento y los resultados provistos por el algoritmo cuántico QVE.
- Testear nuestros desarrollos contra una de las computadoras cuánticas físicas de 5 qubits (no simuladores) que provee IBM.

Metodología de Trabajo

En nuestra investigación hemos creado un panel que permite:

- a. Seleccionar cinco instrumentos financieros y
- b. Correr: (i) En primera instancia: un modelo tradicional de Markowitz; (ii) en segunda instancia un algoritmo cuántico (VQE) desarrollado por la comunidad de Qiskit [21] y (iii) en tercera instancia un algoritmo desarrollado por nuestro equipo de investigadores que conjuga el algoritmo VQE del punto anterior, pero le agrega una capa de lógica y confirmación de la compra/venta de instrumentos con reglas de nuestro desarrollo de la fase anterior de índice de sentimiento más un índice de *forecasting* realizado con Modelos de *Machine Learning*.

Nuestra estrategia se basa en un criterio de asignar un instrumento financiero que tiene el potencial de conformar la cartera de inversión a un qubits de una computadora cuántica, cuando el estado de ese qubits se acerca a 0 una regla lógica nos indica que ese instrumento no ingresa a la cartera de inversión o lo hace en un porcentaje muy bajo, por el contrario, cuando el estado de un qubits se acerca a 1, ese instrumento ingresa a la cartera de inversión con un porcentaje alto. En este punto nos parece importante aclarar que la cantidad de cinco instrumentos a optimizar en este trabajo no es arbitraria, sino que se justifica teniendo en cuenta que la compañía IBM, con su servicio IBM-Q Experience brinda una serie de computadoras cuánticas físicas de 5 qubits de manera libre y gratuita, lo cual presenta un excelente entorno de desarrollo y testing.

Recordemos también que en trabajos anteriores hemos presentado diferentes algoritmos de NLP [3], [20], que permitían procesar noticias financieras en

tiempo real con nuestro lexicón FSAL [2] asignándoles un índice de sentimiento que a su vez se reforzaba con un indicador de predicción (*forecasting*) realizado con redes neuronales LSTM. En el presente trabajo, ambos indicadores brindan una capa de ajuste para determinar con que peso un instrumento ingresa a la cartera o no lo hace.

Resultados Obtenidos /Esperados

A. Resultados Obtenidos

Realizando un *backtesting* de los tres criterios de comparación descriptos previamente hemos obtenido los siguientes resultados en nuestra primera fase:

- Las carteras de inversión con los retornos más altos, a menor riesgo fueron obtenidos en un 92% por la APROXIMACIÓN CLÁSICA.
- El 8% de las veces, nuestro DESARROLLO HÍBRIDO entre la aproximación Cuántica con AI ha superado a la APROXIMACIÓN CLÁSICA.
- El 89% de las veces, nuestro desarrollo híbrido ha sido mejor que el netamente CUÁNTICO sin AI.

B. Resultados Esperados

Como se ha mencionado, solo hemos podido experimentar con la generación de carteras con hasta cinco (5) instrumentos financieros. Al momento de escribir este trabajo, IBM ha liberado computadoras cuánticas con hasta 16 qubits y se espera que pronto haga lo propio con 65 qubits, lo que nos permitirá ampliar nuestro universo de manera considerable. Cuando esto suceda, y cuando los algoritmos cuánticos, los cuales se encuentran en parmente desarrollo y actualización, esperamos que estos resultados obtenidos de manera provisoria se reviertan hasta el punto de que la aproximación híbrida Cuántica con *Machine Learning* se

acerque e incluso mejore la performance de la aproximación Clásica.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto consta de un equipo multidisciplinario conformado por docentes de la Diplomatura en Análisis de Datos para Negocios, Finanzas e Investigación de Mercado, y la Maestría en Tecnología Informática así como expertos del área de Lingüística y Finanzas. Por su parte, señalamos que el proyecto contará con la participación de alumnos avanzados de la Maestría en Tecnología Informática, quienes llevan a cabo su pasantía de investigación al tiempo que identifican temas en los que puedan desarrollar su tesis. Además se cuenta con la colaboración de alumnos de la mencionada Diplomatura.

5. BIBLIOGRAFÍA

- [1] Alcazar, J., Leyton-Ortega, V. and A. Perdomo-Ortiz (2020). Classical versus Quantum Models in Machine Learning: Insights from a Finance Application en *arXiv:1908.10778v2 [quant-ph]* 8 Jan 2020.
- [2] Braña, J.P., Litterio, A.M.J. and A. Fernández (2018). "FSAL: Lexicón financiero de sentimiento en español rioplatense diseñado para "Bolsas y Mercados Argentinos" (BYMA) en Revista Abierta de Informática Aplicada (RAIA), vol.2, Nro1., 2018. ISSN: 2591-5320. Disponible en: <http://portalreviscienc.uai.edu.ar/ojs/index.php/RAIA/article/view/159>
- [3] Braña, J.P., Litterio, A.M.J. and A. Fernández. (2020). "Inteligencia Artificial y Computación Cuántica en Finanzas". XXII Workshop de Investigadores en Ciencias de la Computación WICC 2020. Universidad Nacional de la Patagonia Austral (UNPA), El Calafate, Santa Cruz, Argentina. 25 de Mayo de 2020.
- [4] Dasgupta, S., Banerjee, A. (2019). Quantum Annealing Algorithm for Expected Shortfall based Dynamic Asset Allocation en *arXiv:1909.12904v1 [q-fin.RM]* 27 Sep 2019.
- [5] Egger, D., García Gutiérrez, R., Cahué Mestre, J. and S. Woerner (2019). Credit Risk Analysis using Quantum Computers en *arXiv:1907.03044v1 [quant-ph]* 5 Jul 2019.
- [6] Hao, W., Lefevre, C., Tamturk, M. and S. Utev, (2019). Quantum option pricing and data analysis en *Quantitative Finance and Economics*, QFE, 3(3): 490–507. DOI:10.3934/QFE.2019.3.490
- [7] Grant, E., Humble, T. and B. Stump (2021). Benchmarking Quantum Annealing Controls with Portfolio Optimization, in *Physical Review Applied* 15,014012. DOI: 10.1103/PhysRevApplied.15.014012
- [8] Gonzalez-Conde, J., Rodriguez –Rosas, A., Solano, E. and M. Sanz (2021). Pricing Financial Derivatives with Exponential Quantum Speedup in *Quantum Physics (quant-ph)*, *arXiv:2101.04023 [quant-ph]*
- [9] Hodson, M., Ruck, B., Ongy, H., Garvin, D., and S. Dulman. (2019). Portfolio rebalancing experiments using the Quantum Alternating Operator Ansatz en *arXiv: 1911.05296v1 [quant-ph]* 13 Nov 2019.
- [10] Kerenidis, I., Prakash, A. and D. Szilagy. (2019). Quantum Algorithms for Portfolio Optimization en *arXiv: 1908.08040v1 [math.OC]* 22 Aug 2019.
- [11] Mahajan, R.P. (2011). A Quantum Neural Network Approach for Portfolio Selection en *International Journal of Computer Applications*, Volume 29– No.4, September 2011.
- [12] Markowitz, H. (1952). Portfolio Selection in *The Journal of Finance* Vol. 7, No. 1 (Mar., 1952), pp. 77-91
- [13] Marzec, M. (2016). "Portfolio Optimization: Applications in Quantum Computing", en Ionut Florescu, Maria C. Mariani, H. Eugene Stanley, Frederi G. Viens (eds.) *Handbook of High-Frequency Trading and Modeling in Finance*, John Wiley & Sons, Inc.
- [14] Martin, A., Candelas, B., Rodríguez-Rozas, Martín-Guerrero, J. Chen, X., Lamata, L., Orus, R., Solano, E. and M. Sanz (2019). Towards Pricing Financial Derivatives with an IBM Quantum Computer en *arXiv: 1904.05803v1 [quant-ph]* 11 Apr 2019.
- [15] Mugel, S., Kuchkovsky, C., Sánchez, E., Fernández-Lorenzo, S., Luis-Hita, J., Lizaso, E. and R. Orús (2020). Dynamic Portfolio Optimization with Real Datasets Using Quantum Processors and Quantum-Inspired Tensor Networks, in *arXiv:2007.00017 [quant-ph]*
- [16] Orrell, D. (2020). A Quantum Walk Model of Financial Options (January 1, 2020). Available at SSRN: <https://ssrn.com/abstract=3512481> or <http://dx.doi.org/10.2139/ssrn.3512481>

- [17] Orus, R., Mugel, S. and E. Lizaso (2019). Quantum computing for finance: overview and prospects en *Reviews in Physics*, Volume 4, November 2019, 100028. <https://doi.org/10.1016/j.revip.2019.100028>
- [18] Stamatopoulos, N., Egger, D.J., Sun, Y., Zoufal, C., Iten, R., Shen, N. and S. Woerner. (2020). Option Pricing using Quantum Computers en *arXiv: 1905.02666 [quant-ph]*, 17 Feb 2020.
- [19] Woerner, S., Egger, D.J. Quantum risk analysis. En *NPJ Quantum Inf*, 5, 15 (2019). <https://doi.org/10.1038/s41534-019-0130-6>.
- [20] Zeng, W. and Coecke, B. (2016) Quantum Algorithms for Compositional Natural Language Processing en Kartsaklis, D., Lewis, M. and Rimell, L. (Eds) *2016 Workshop on Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science (SLPCS'16) EPTCS 221*, 2016, pp. 67–75, doi:10.4204/EPTCS.221.8
- [21] IBM Qiskit Portfolio Optimization Tutorial https://qiskit.org/documentation/tutorials/finance/01_portfolio_optimization.html

Red neuronal multiescala para clasificación de la calidad vocal

García Mario Alejandro ✉¹, Rosset Ana Lorena², Eduardo Destéfanis¹

¹Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN FRC)

²Universidad Nacional de Córdoba (UNC)

mgarcia@frc.utn.edu.ar

RESUMEN

La valoración de la calidad vocal mediante el análisis audio-perceptual es parte de la rutina clínica de evaluación de pacientes con trastornos de la voz. La debilidad de este método reside en la subjetividad y en la necesidad de que sea realizada por oyentes experimentados. Este proyecto tiene como objetivo la realización de una clasificación automática de la calidad vocal, valuada en la escala GRBAS, mediante la aplicación de técnicas de aprendizaje profundo sobre voces grabadas. Particularmente, en este trabajo se muestran los resultados del diseño de una red neuronal multiescala para la clasificación de la calidad vocal.

Palabras clave: *red neuronal multiescala, calidad vocal, aprendizaje profundo*

CONTEXTO

Este trabajo de investigación se desarrolla en el marco del proyecto “*Deep learning* para clasificación de señales vocales” (UTN5274) de la Universidad Tecnológica Nacional, Facultad Regional Córdoba y cuenta con la colaboración del Departamento de Investigación Científica, Extensión y Capacitación “Raquel Maurette”, Escuela de Fonoaudiología, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba.

1. INTRODUCCIÓN

Se intenta reconocer, de forma automática, características del análisis acústico de la voz que permitan clasificar muestras de audio. El estudio se enfoca en la medición de la calidad vocal según la escala GRBAS. La clasificación se realiza aplicando modelos de aprendizaje profundo (*deep learning*), un subgrupo de técnicas del campo de aprendizaje automático (*machine learning*).

GRBAS: La escala GRBAS es un método de valoración perceptivo-auditivo de la voz. Surge de la necesidad de estandarizar la valoración subjetiva y de interrelacionar los aspectos auditivos y fisiológicos de la producción vocal. Está basada en estudios del año 1966 de la *Japan Society of Logopedics and Phoniatics* [1] y posteriormente divulgada y descripta por Minoru Hirano en el año 1981 [2]. Consiste en la valoración de la fuente glótica a través de 5 parámetros que forman el acrónimo GRBAS:

G: (*Grade*) Grado general de disfonía.

R: (*Roughness*) Rugosidad, irregularidad de la onda glótica.

B: (*Breathiness*) Soplosidad, sensación de escape de aire en la voz.

A: (*Astheny*) Astenia, pérdida de potencia.

S: (*Strain*) Tensión, sensación de hiperfunción vocal.

Puede valorarse de dos maneras: a través de 4 grados, desde el 0 al 3 o mediante un valor en un rango continuo de 0 a 100. En ambas el 0 es ausencia de disfonía y el 3 o 100 implican disfonía severa. La escala fue mundialmente adoptada y validada en numerosos países [3-6]. Actualmente se utiliza en la investigación y de manera rutinaria en los consultorios de los profesionales que hacen clínica vocal. Sirve como metodología simple y al alcance de la mano para valorar la evolución pre-post tratamiento. La debilidad de este método reside en la subjetividad de la valoración de la voz y en la necesidad de que sea realizada por oyentes experimentados en la escucha y la disociación de los parámetros [7,8].

Estado del arte: La aplicación de técnicas de aprendizaje profundo es el estado del arte en el análisis automático de audio, con la detección de los fonemas pronunciados y la identificación de la persona que habla como objetivos principales [9-15], pero también utilizadas en detección de emociones, edad y género entre otros. Durante la ejecución de este proyecto se aplican técnicas de aprendizaje a la clasificación de la calidad de la voz.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación que se presenta en este trabajo se enfoca en el desarrollo de un clasificador neuronal multiescala para la clasificación de la calidad vocal.

La teoría del espacio escalar propone que el ojo humano es capaz de reconocer patrones en distintos niveles de suavizado de la imagen original y que los patrones reconocidos en diferentes niveles mejoran la clasificación general. Esta idea aplica con frecuencia en el reconocimiento de imágenes. Existen trabajos de reconocimiento de patrones en audio que

utilizan la misma técnica [16-19]. En estos se muestra que representando el audio en distintas escalas se puede mejorar la capacidad de clasificación.

El suavizado se logra convolucionando la señal con una ventana gaussiana. El efecto obtenido es la pérdida de detalles. Cuanto más grande sea la ventana, mayor será el efecto.

La incorporación de la clasificación multiescala en el proyecto se lleva a cabo mediante la adaptación de un modelo de red neuronal profunda (M_0) desarrollado en etapas anteriores.

El modelo M_0 realiza la clasificación de la calidad vocal partiendo desde el audio crudo (*raw* audio). El audio se divide en segmentos y se multiplica cada segmento por una ventana de pesos adaptables mediante una capa STHadamard [20]. Para cada segmento de *windowed* audio se calcula el power cepstrum con otras capas de la red neuronal, tal como se explica en [21]. Las salidas anteriores son las entradas de la última parte de la red, donde se realiza la extracción de características con capas de convolución y la clasificación con dos capas de neuronas densamente conectadas.

Los distintos niveles de suavizado se introducen entre la representación cepstral del audio y la extracción de características (Figura 1). En el modelo presentado en este trabajo se utilizan cuatro niveles o escalas. El suavizado se logra con una capa de convolución de tamaño n y desplazamiento 1, donde los pesos de cada kernel de convolución forman una campana gaussiana y no se modifican durante el entrenamiento. La salida de las cuatro capas de convolución se concatenan formando un vector de tres dimensiones.

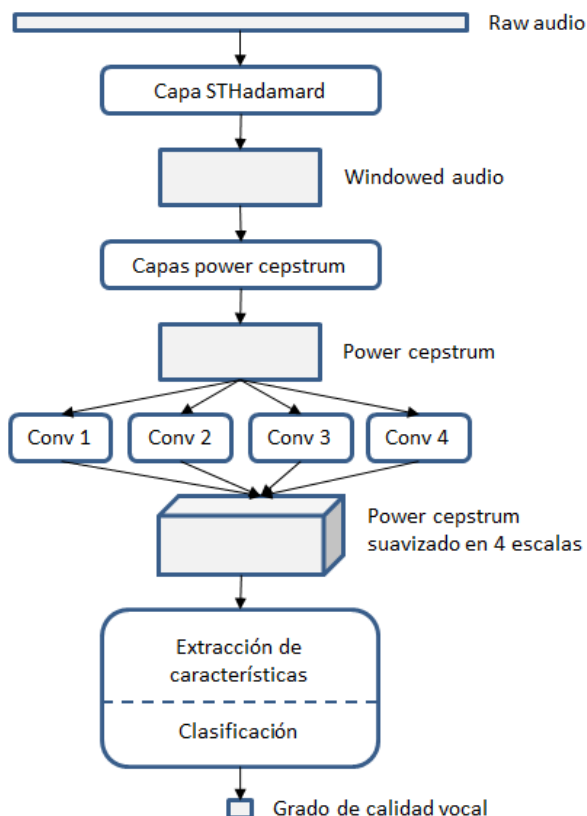


Figura 1. Modelo M_0 modificado con escalado múltiple de cuatro niveles.

Experimento:

Se utilizaron audios de la Voice Disorders Database (VDD) [22], grabados por la Universidad Politécnica de Madrid en colaboración con el Hospital Universitario Príncipe de Asturias. Estas grabaciones se realizaron sobre personas sanas y personas con patologías vocales pronunciando una vocal /a/ sostenida durante aproximadamente dos segundos. Para el experimento se extrajeron segmentos de 1 segundo y se aplicaron las técnicas de *data augmentation* explicadas en [23].

Se realizó una clasificación binaria del grado general de disfonía (G). Debido a la distribución de los valores de G, se crearon las categorías 0 para $G = 0$ y 1 para $G = 1, 2$ y 3.

Los tamaños de los kernels de convolución para suavizado fueron $n = 3, 11, 21$ y 31. En la Figura 2 se muestra la salida de las capas de suavizado.

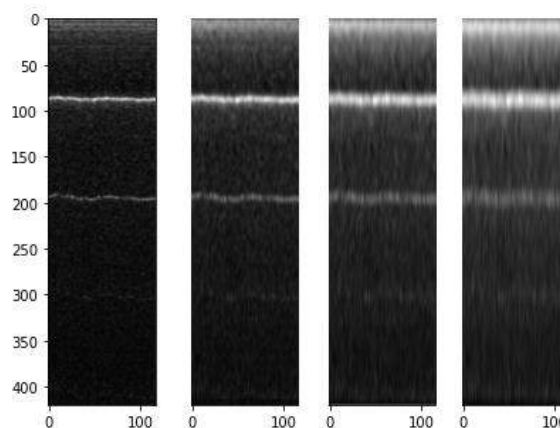


Figura 2. Salidas de las capas de suavizado para $n = 3, 11, 21, 31$ de izquierda a derecha.

3. RESULTADOS OBTENIDOS

A continuación se exponen los resultados de 50 entrenamientos del modelo propuesto y 50 entrenamientos del modelo M_0 . Los pesos se inicializaron con valores aleatorios entre -10^{-6} y 10^{-6} uniformemente distribuidos. Se utilizó el método de optimización Adam [24] con los parámetros provistos por los autores y las actualizaciones de los pesos se realizaron en batches de tamaño 5244 (la totalidad de datos de entrenamiento).

Los cálculos se realizaron sobre una GPU NVIDIA Titan Xp, donada a través del GPU Grant Program de NVIDIA.

El *accuracy* medio de validación alcanzado para el modelo M_0 fue 0.7698, mientras que para el modelo M_0 con suavizado multiescala se obtuvo 0.8069, lo que significa una mejora media del 4.8%.

Estos resultados indican que el modelo propuesto es capaz de mejorar la clasificación de la calidad vocal y que se puede integrar en una red neuronal profunda.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto está formado por un docente/investigador de la UTN FRC, dos docentes/investigadores de la UNC y cuatro alumnos de la carrera de grado de la UTN FRC.

Además de formación de los alumnos participantes, el conocimiento generado por el proyecto se incorporará a las cátedras de los docentes de la UTN y UNC.

5. REFERENCIAS

- [1] Isshiki, N., Yanagihara, N., & Morimoto, M. (1966). *Approach to the objective diagnosis of hoarseness*. *Folia Phoniatica et Logopaedica*, 18(6), 393-400.
- [2] Hirano, M. (1981). *Clinical examination of voice* (Vol. 5). Springer.
- [3] Yun, Y. S., Lee, E. K., Baek, C. H., & Son, Y. I. (2005). *The correlation of GRBAS scales and laryngeal stroboscopic findings for the assessment of voice therapy outcome in the patients with vocal nodules*. *Korean Journal of Otolaryngology-Head and Neck Surgery*, 48(12), 1501-1505.
- [4] Hui, H., Weijia, K., & Shusheng, G. (2007). *The Validation of Acoustic Analysis and Subjective Judgment Scales of Several Voice Disorders* [J]. *Journal of Audiology and Speech Pathology*, 3, 010.
- [5] Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). *Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders*. *Journal of Voice*, 21(5), 576-590.
- [6] Jesus, L. M., Barney, A., Couto, P. S., Vilarinho, H., & Correia, A. (2009, December). *Voice quality evaluation using cape-v and GRBAS in european Portuguese*. In *MAVEBA* (pp. 61-64).
- [7] Kreiman, J., & Gerratt, B. R. (2010). *Perceptual assessment of voice quality: past, present, and future*. *SIG 3 Perspectives on Voice and Voice Disorders*, 20(2), 62-67.
- [8] Núñez-Batalla et al (2012). El espectrograma de banda estrecha como ayuda para el aprendizaje del método GRABS de análisis perceptual de la disfonía. *Acta Otorrinolaringológica Española*, 63(3), 173-179.
- [9] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., Kingsbury, B. (2012): Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *Signal Processing Magazine*, vol. 29.6, 82-97. IEEE.
- [10] Mitra, V., Sivaraman, G., Nam, H., Espy-Wilson, C., Saltzman, E., Tiede, M. (2017) Hybrid convolutional neural networks for articulatory and acoustic information based speech recognition. *Speech Communication*, vol. 89. pp 103-112.
- [11] Collobert, R., Puhersch, C., Synnaeve, G. (2016) Wav2letter: an end-to-end convnet-based speech recognition system. *arXiv preprint arXiv:1609.03193*.
- [12] Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C., Chen, J. (2016) Deep speech 2: End-to-end speech recognition in english and mandarin. *International Conference on Machine Learning*. pp. 173-182.
- [13] Palaz, D., Collobert, R. (2015) Analysis of cnn-based speech recognition system

using raw speech as input (No. EPFL-REPORT-210039). Idiap.

[14] Sainath, T. N., Kingsbury, B., Mohamed, A. R., Ramabhadran, B. (2013) Learning filter banks within a deep neural network framework. IEEE Workshop on ASRU. pp 297-302. IEEE.

[15] Farrús, M. (2007) Jitter and shimmer measurements for speaker recognition. 8th Annual Conference of ISCA. pp. 778-781. (2007)

[16] Chi, T., Shamma, S. A. (2006). Spectrum restoration from multiscale auditory phase singularities by generalized projections. IEEE transactions on audio, speech, and language processing, 14(4), 1179-1192.

[17] Zhu, Z., Engel, J. H., Hannun, A. (2016). Learning Multiscale Features Directly from Waveforms. Interspeech 2016, 1305-1309.

[18] von Platen, P., Zhang, C., Woodland, P. C. (2019). Multi-Span Acoustic Modelling Using Raw Waveform Signals. Interspeech 2019, 1393-1397.

[19] Mesgarani, N., Slaney, M., & Shamma, S. A. (2006). Discrimination of speech from nonspeech based on multiscale spectro-temporal modulations. IEEE Transactions on Audio, Speech, and Language Processing, 14(3), 920-930.

[20] García, M. A., Destéfanis, E. A., Rosset, A. L. (2020, September). Trainable Windowing Coefficients in DNN for Raw Audio Classification. In Conference on Cloud Computing, Big Data & Emerging Topics (pp. 153-166). Springer, Cham.

[21] García, M. A., Destéfanis, E. A. (2019). Power cepstrum calculation with convolutional neural networks. Journal of Computer Science & Technology, 19.

[22] Arias-Londoño, J. D., Godino-Llorente, J. I., Markaki, M., Stylianou, Y.: On combining information from modulation spectra and mel-frequency cepstral coefficients for automatic detection of pathological voices. Logopedics Phoniatrics Vocology 36.2. 2011: 60-69

[23] García, M. A., Destéfanis, E. A. (2020). Data Augmentation para la Clasificación Automática de la Calidad Vocal. AJEA, (5).

[24] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

REDES NEURONALES PARA LA CLASIFICACIÓN DE PARTÍCULAS DE ARENA

Mg. Ing. Carlos Gustavo Rodriguez Medina ¹, Dr. Ing. Oscar Daniel Chuk ²,
Ing. Regina Bertero, Lic. Adriana Luna, Ing. Enrique Núñez. Sr. Darío Quinteros

Instituto de Investigaciones Mineras / Facultad de Ingeniería /
Universidad Nacional de San Juan

Av. Libertador Gral. San Martín 1109 oeste. San Juan
0264-4211700 (int. 285 ¹, int. 389 ^{1,2})

grodriguez@unsj.edu.ar ¹, dchuk@unsj.edu.ar ²

RESUMEN

El presente trabajo trata respecto a la utilización de redes neuronales para el reconocimiento de formas y patrones de imágenes digitales de partículas de arena, con destino al uso en procesos de fracturación hidráulica en explotación de hidrocarburos de manera no convencional (*Fracking*).

Si bien, la mayor parte de las redes neuronales permiten el reconocimiento de patrones y clasificación de objetos, hay un tipo particular de redes que en la actualidad han permitido un salto importante en el procesamiento de información e imágenes, y son las redes neuronales convolucionales. Estas permiten la extracción de características directamente desde las propias imágenes sin la necesidad de un trabajo de pre-procesado.

Es decir, para este caso, se usan como entradas a la red neuronal las imágenes digitales de partículas o granos arena, y no un conjunto de características extraídas de manera previa por parte del diseñador de la red neuronal. En cuanto a la clasificación de las partículas de arena, en este trabajo se refiere en poder determinar a través de la red neuronal convolucional características de Redondez y Esfericidad específicamente.

También se presenta una alternativa para generar el conjunto de datos para el entrenamiento y validación de la red neuronal,

mediante la elaboración de un algoritmo de procesamiento de imágenes que permite la generación artificial (gráfica) de partículas de arena, dado que no siempre se puede disponer de un banco de imágenes adecuado.

Palabras clave: Redes neuronales convolucionales, Conjunto de datos, Fracturación hidráulica, Partículas de arena.

CONTEXTO

La temática desarrollada representa parte de los avances y resultados del proyecto de Investigación y Creación (PIC – CICITCA) “*Clasificación de partículas de arena de fracturación mediante redes neuronales convolucionales*”. En ejecución durante el periodo 2020 – 2021.

Dicho proyecto se inscribe dentro de la línea de investigación que lleva adelante un grupo de trabajo en la temática de Inteligencia Artificial y Visión Artificial orientado al control automático de procesos mineros, que se viene desarrollando hace varios años en el Instituto de Investigaciones Mineras de la Facultad de Ingeniería de la UNSJ.

A su vez, el mencionado proyecto representa la continuidad del proyecto “*Visión artificial aplicada a la determinación de propiedades geométricas de arenas de fracturación para el análisis de calidad*”, desarrollado durante 2018 y 2019.

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 *Fracking* para actuar como apuntalantes o propantes, se reconocen como base las normas de referencia del *American Petroleum Institute* (API) [1]. Esta norma recomienda una serie de ensayos a realizar sobre las arenas con el fin de conocer la calidad del material a utilizarse en la fracturación hidráulica. Entre los ensayos a realizarse se encuentran los de redondez y esfericidad de las partículas de arena.

Tales ensayos padecen de una fuerte dependencia del criterio del observador, dado que el procedimiento se realiza mediante la utilización de un microscopio y las mediciones de esfericidad y redondez se realizan por comparación respecto a tablas que contienen una serie de formas provistas por la norma (cartillas de *Krumbein y Sloss* – figura 1). Esto deriva en resultados sustancialmente disímiles que invalidan el procedimiento, aunque es el que se sigue usando por norma.

El uso de técnicas de Visión Artificial se presenta como la alternativa tecnológica que puede permitir obtener medidas objetivas no dependientes del criterio de un operador [1].

Cada partícula de arena puede presentar diversas variaciones unas a otras, 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 [2].

Se sabe que las redes neuronales artificiales son modelos de procesamiento de información, inspirados por la forma en que el cerebro procesa información. Aunque son una simplificación de las redes neuronales biológicas, las redes neuronales artificiales son aptas para resolver problemas que la gente puede resolver, pero las computadoras

no pueden, o los algoritmos resultan muy complicados [3].

El conocimiento es adquirido por la red a través de un proceso que se denomina aprendizaje [4]; dicho conocimiento se almacena mediante la modificación de la fuerza o peso sináptico de las distintas uniones entre elementos de procesamiento [5]. Todo conocimiento de una red neuronal se encuentra distribuido en los pesos sinápticos de las neuronas. Una red neuronal artificial aprende a través de un proceso de ajuste de sus pesos sinápticos. El aprendizaje es el proceso por el cual una red neuronal modifica sus pesos sinápticos en respuesta a una entrada, para proporcionar la salida adecuada [3].

Se habla de conjunto de datos, al utilizado para llevar a cabo la tarea de aprendizaje de la red neuronal, como así también para realizar las pruebas o validación. Existe cierta relación entre la cantidad de datos para el entrenamiento y la capacidad que adquiere la red para generalizar. Por lo general se subdivide el conjunto original en un subconjunto de datos para el entrenamiento y otro subconjunto para las pruebas.

Al utilizar redes neuronales, si necesitaráramos operar con imágenes, y realizar una clasificación teniendo en cuenta alguna característica, primero habría que extraer tales características del conjunto de imágenes y armar un conjunto nuevo conteniendo dichas características, y estas serían las entradas a utilizar por la red. Esto implica tener que realizar un trabajo previo de extracción de características y así poder contar con este conjunto nuevo de datos (características de interés), siendo utilizadas luego como subconjuntos de entrenamiento y de prueba.

Existen otro tipo de redes neuronales, que permiten operar directamente sobre un conjunto de imágenes, sin necesidad de tener que realizar ningún pre-procesado, para la extracción de la o las características de interés, estas son las redes neuronales convolucionales [6].

Las redes neuronales convolucionales son especialmente útiles para localizar patrones en imágenes con el objetivo de reconocer

objetos, formas, etc. Aprenden directamente a partir de los datos de imágenes, utilizando patrones para clasificar las imágenes y eliminar la necesidad de una extracción manual de características [7].

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 componen de capas Convolucionales y de Pooling de forma alternada, Finalizando con una serie de capas de neuronas “completamente conectadas”.

En cuanto el proceso de extracción de petróleo mediante fracturación hidráulica, tal como se dijo en párrafos anteriores, la norma que se aplica para la determinación de los parámetros de calidad de las partículas de arena, utiliza las cartillas de *Krumbein y Sloss* para clasificarlas en cuanto a redondez y esfericidad, como se expone en la figura 1. Dicha tabla posee solamente 20 elementos gráficos de acuerdo a distintos valores de redondez y esfericidad. Tales elementos imágenes de partículas de arena de la tabla resultan extremadamente insuficientes si se quisieran utilizar para llevar a cabo el entrenamiento de la red neuronal. Además, la bibliografía consultada hace referencia que para el caso de las redes neuronales convolucionales se requieren un mayor número de elementos (imágenes) de entrenamiento que para otro tipo de Red Neuronal.

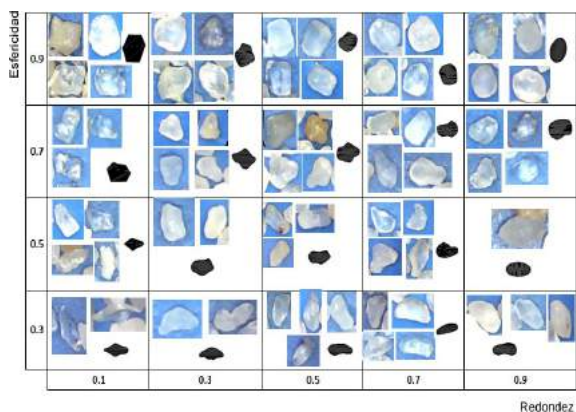


Figura 1. Cartilla de esfericidad y redondez de *Krumbein y Sloss*

Por lo antes enunciado, se requiere contar con un conjunto o banco de imágenes capturadas por microscopio electrónico, y a su vez que las mismas se encuentren clasificadas en cuanto a medidas de Redondez y Esfericidad. Se requeriría aproximadamente entre 1000 y 10.000 elementos de acuerdo a la bibliografía consultada [6].

Dado que en la práctica esto resulta difícil de concretar, se propone una alternativa para el armado del conjunto de imágenes de partículas de arena, mediante el software matemático Matlab [7], generando partículas de arena que no son imágenes tomadas mediante un microscopio (granos de arena reales), sino que las mismas son elaboradas mediante funciones matemáticas (elementos geométricos básicos tales como elipses y círculos), pudiéndose así hacer referencia como un conjunto de elementos de partículas “artificiales” de arena.

Tales imágenes generadas, tienen la particularidad que como están compuestas por formas básicas de elipses y círculos, entonces se puede aplicar con cierta facilidad la ecuación desarrollada por *Wadell* (1932) para el cálculo de la redondez, tal como se expone en la Ecuación 1. A la vez resulta de importancia aclarar que la cartilla de *Krumbein y Sloss* (1963) se basa en la aplicación de la ecuación de *Wadell*.

$$R = \frac{\sum_{i=1}^N r_i}{r_{\max-in}} \quad (\text{Ecuación 1.})$$

Donde:

N : Cantidad de círculos inscriptos.

r_i : Radios de los círculos.

$r_{\max-in}$: Radio del máximo círculo inscripto

Por lo tanto, a cada partícula generada gráficamente como una partícula artificial se le calcula la redondez mediante la ecuación 1, utilizando *Wadell*, y la correspondiente para el cálculo de esfericidad, por lo que en definitiva se estaría aplicando la referida norma API para la determinación de estas características, utilizando técnicas de

inteligencia artificial (procesamiento de imágenes, etc.)

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 para ser utilizadas en la extracción de petróleo mediante el método de *Fraking*.

En búsqueda de dar solución al tema de clasificación de partículas, se ha incorporado el trabajo con redes neuronales.

Otra de las líneas de investigación y desarrollo del grupo de trabajo, está orientado al de optimización. Concretamente, se está trabajando en la optimización multiobjetivo, del tipo predictivo, para la explotación de minas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Durante el año 2020 se ha trabajado en la elaboración del algoritmo y técnicas que permiten generar cada imagen de partícula de arena (artificial), para conformar el conjunto de entrenamiento y prueba de la Red Neuronal Convolutiva. Así también, se está trabajando en el diseño de la red neuronal y su implementación.

Tal como se observa de la figura 2, la entrada de la red neuronal se alimenta con las imágenes de las partículas de arena en forma directa (partículas gráficas generadas artificialmente para este caso particular). Esto trae como ventaja de que no deben extraerse

manualmente ni tampoco mediante algoritmos características de las imágenes para ser ingresadas a la red neuronal convolutiva. Posteriormente pasa a una capa de convolución y pooling, y así sucesivamente hasta extraer las características requeridas de cada imagen. Por último, se pasa a una red neuronal completamente conectada, ajustándose sus pesos sinápticos, durante la etapa de aprendizaje.

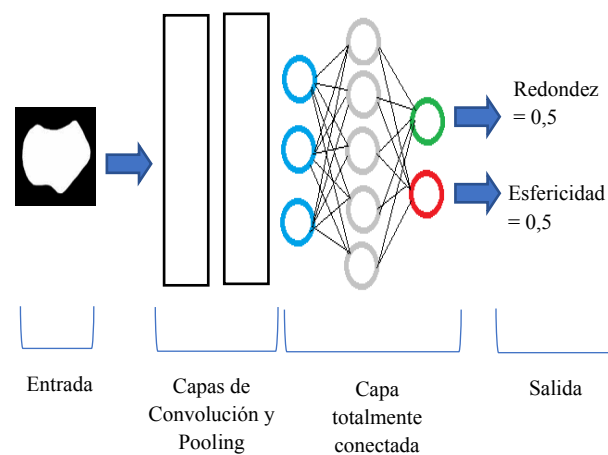


Figura 2. Diagrama de Red Neuronal Convolutiva para clasificar partículas de arena

Resulta importante que una vez entrenada y validada la red neuronal convolutiva, en la salida se obtienen los valores de redondez y esfericidad de cada imagen que se está analizando por parte de la red. Para el caso de ejemplo de la figura 1, para la partícula ingresada, si se recurre a la tabla 1 (cartilla de *Krumbein y Sloss*) se aprecia que una imagen de partícula de arena de aspecto similar efectivamente tiene un valor de redondez de 0,5, al igual que su esfericidad.

La figura 3 expone el algoritmo que elabora cada imagen artificial de partícula de arena. Sobre una imagen con fondo de color negro se grafica una serie de elipses y círculos, se determina la cantidad de elementos a emplear, relacionado a la cantidad de picos o esquinas de la partícula a generar. En la medida que más círculos o picos se le agreguen a la elipse graficada previamente, menos redondeada será la partícula artificial resultante, obsérvese

la figura 4-A. Luego se aplica un filtro de bordes, de *Canny* específicamente [8], para poder trabajar solamente con el contorno resultante, figura 4-B.

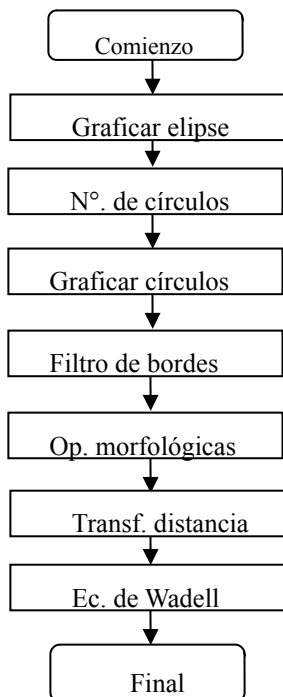


Figura 3. Algoritmo para graficar partículas artificiales de arena.

En la figura 4-C y 4-D se aplican operaciones morfológicas, esto se hace con el objeto de lograr un suavizado de la partícula obtenida, ya que hasta esa instancia es muy notoria la presencia de los círculos sobre la elipse.

Luego se realiza el cálculo de la transformada de distancia de la partícula obtenida lográndose el resultado de la figura 3-E.

Finalmente se calcula la redondez mediante la Ecuación 1 de *Wadell*, utilizando los radios de los círculos inscriptos en las esquinas de la partícula artificial, y tomando el valor de la transformada de distancia como el valor del radio del máximo círculo inscripto en la partícula. Se calcula también la esfericidad.

Como resultado preliminar, los cálculos obtenidos mediante la aplicación de la fórmula de *Wadell*, utilizando el algoritmo elaborado (figura 3), es coincidente con el que se obtiene de aplicar las cartillas de *Krumbein y Sloss* (figura 1) para la obtención de los valores de redondez y esfericidad (de acuerdo

a las normas API, universalmente utilizada y explicado en el apartado 1).

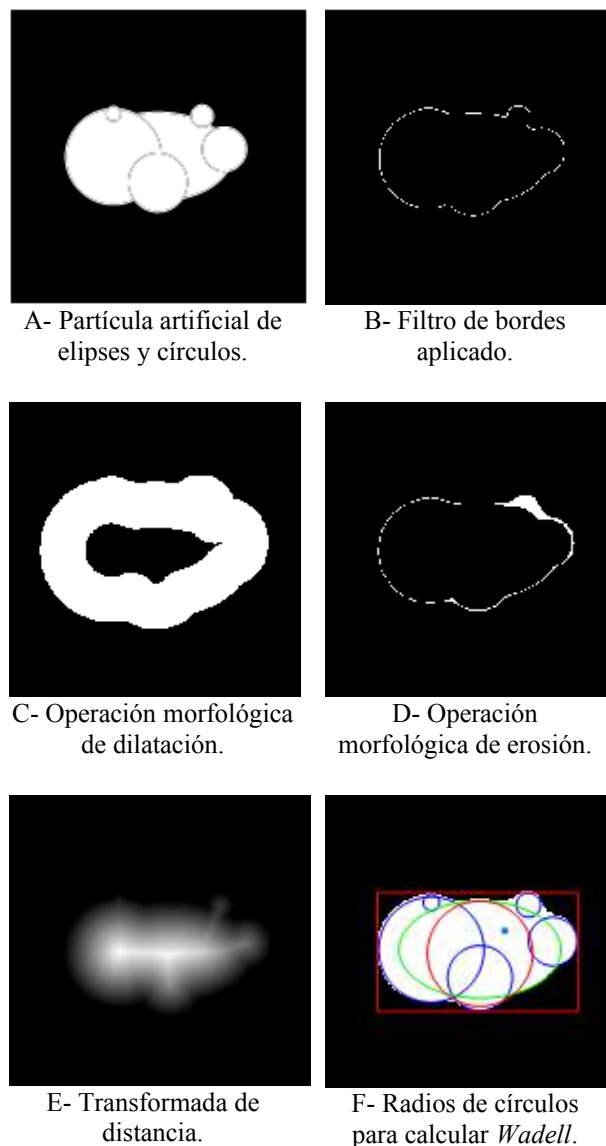


Figura 4. Resultados de aplicar el algoritmo para graficar partículas artificiales de arena.

Por lo antes expuesto, se ha desarrollado el algoritmo que permite generar partículas de arena, posibilitando la obtención de conjuntos de imágenes para el entrenamiento y testeado de la red neuronal convolucional. Así también se ha trabajado en el diseño de la red que se aplica para la clasificación de partículas de arenas con destino al *Fracking*.

Resta optimizar dicho diseño y llevar a cabo las pruebas necesarias para medir su desempeño, seguida de los correspondientes ajustes.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por Ingenieros Electrónicos, Electromecánicos, en Minas, Metalurgistas y Geólogos, con formación diversa en posgrado, tal como Doctorado en sistemas de control, Maestría en informática, Especialización en gestión y vinculación tecnológica, etc.

Actualmente, una de las integrantes del equipo de trabajo, se encuentra desarrollando su Tesis de Doctorado en Ingeniería en procesamiento de minerales, siendo su director de Tesis también otro de los integrantes del grupo de investigación.

Además, como inicio en la carrera de investigador, otra integrante comenzó en este grupo hace unos años participando de un proyecto PROJOVI (Proyecto de Jóvenes Investigadores), para luego continuar su participación en proyectos PIC (Proyecto de Investigación y Creación) en el periodo 2018 – 2019 y 2020-2021.

También, un alumno avanzado de la carrera de grado de Ingeniería en Electrónica, ha comenzado a participar del proyecto PIC en ejecución, periodo 2020 – 2021, llevando a cabo su tesina de fin de carrera.

5. BIBLIOGRAFIA

[1] Rodriguez Medina, C. G.; Chuk, O. D.; Bertero R.; Luna, A.; Núñez, E.; Trigo, P. 2019. Visión artificial aplicada a la determinación de propiedades geométricas de arenas de fracturación para el análisis de calidad. Universidad Nacional de San Juan. Informe de Final de Proyecto PIC.

[2] Jain, A. K. 2000. Statistical Pattern Recognition: a review. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol 22, Nro. 1.

[3] Castro G., J. F. 2006. Fundamentos para la implementación de red neuronal perceptron multicapa mediante software. Tesis de grado. Facultad de Ingeniería, Universidad de San Carlos de Guatemala.

[4] Haykin S. 1994. Neural Networks - A comprehensive foundation, IEEE Press - Macmillan College Publishing Company, Inc.

[5] Long, L. 1990. Introducción a las computadoras y al procesamiento de información. 2ª. Edición; México: Editorial Prentice All Hispanoamericana.

[6] Rodriguez Medina, C. G.; Chuk, O. D.; Bertero R.; Luna, A.; Núñez, E.; Quinteros, D. 2020. Clasificación de partículas de arena de fracturación mediante redes neuronales convolucionales. Universidad Nacional de San Juan. Informe de Avance de Proyecto PIC.

[7] Mathworks. (1994-2019). Image Processing Toolbox. https://es.mathworks.com/products/image.html?s_tid=srchtitle.

[8] Rodriguez Medina, C. G. y Navas, G. S. 2014. Aplicación del filtro de Canny en la esteganografía digital. XVI Workshop de Investigadores en Ciencias de la Computación - WICC 2014 Ushuaia. RedUNCI. p 806 - 811. ISBN 978-950-34-1084-4.

RIESGOS, DILEMAS ÉTICOS Y BUENAS PRÁCTICAS EN INTELIGENCIA ARTIFICIAL

Dra. Bibiana Beatriz Luz Clara; Lucía Rosario Malbernat
Grupo Ética-IA, Departamento de Sistemas, Universidad CAECE
bluzclara@ucaecemdp.edu.ar; lmalbernat@ucaecemdp.edu.ar

RESUMEN

Mientras que Europa publica su primera guía ética para el uso responsable de Inteligencia Artificial (IA), en la región hay necesidad de contar con una guía de pautas deseables y prácticas preferidas por las organizaciones que propicien la definición de buenas prácticas.

Por lo tanto, es necesario elaborar directrices éticas que se basen en el marco regulador existente y aporten a los desarrolladores, proveedores y usuarios de IA orientaciones marco.

Se comparte en este documento un Proyecto de investigación de reciente presentación que pone su foco en las prácticas que llevan a cabo las organizaciones de la región que aplican técnicas de IA en sus desarrollos o procedimientos, de modo de aportar conocimiento para minimizar sus riesgos éticos.

Palabras clave: Inteligencia Artificial, ética, Buenas prácticas

CONTEXTO

La investigación que se reporta está radicada en el Departamento de Sistemas de la Universidad CAECE, República Argentina, presentada para el período 2021-2022 con fecha de inicio 1° de abril de 2021.

Se trata de un Proyecto inter sedes ya que participarán investigadores radicados en la Sede Central y en la Subsede Mar del Plata que tendrá impacto en carreras de grado y posgrado del Departamento. En particular, investigadores del proyecto esperan poder volcar el conocimiento adquirido en las asignaturas Aspectos Profesionales para Ingeniería e Inteligencia de Negocios.

1. INTRODUCCIÓN

Si bien el concepto de Inteligencia artificial, acuñado en 1956 por John McCarthy en la que se considera primer conferencia dedicada al tema (Dartmouth College, Hanover, Nuevo Hampshire, Estados Unidos) no es nuevo y hace más de 30 años que la industria informática apuesta a la inclusión de técnicas propias de esta disciplina a sus desarrollos, en la última década empiezan la organizaciones en general a obtener provecho de sus ventajas y beneficios por lo que comienzan a impactar en la sociedad y a tomar cuerpo los dilemas éticos.

Para dar respuesta a ese contexto los países alrededor del mundo comienzan, en el último quinquenio, a definir sus estrategias nacionales o planes de desarrollo. Tal el caso de EEUU, que lo hizo en 2016, Canadá y China en 2017, Japón y Francia en 2018 [1] mientras que el gobierno de España presentó en 2020 el Consejo Asesor de Inteligencia Artificial, un órgano consultivo destinado a asesorar y a realizar recomendaciones independientes sobre las medidas adoptables para garantizar un uso seguro y ético de la IA. [2]

Para la Comisión Europea, la IA son “Sistemas de software (y posiblemente también de hardware) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno a través de la adquisición de datos, interpretando los datos estructurados o no estructurados recolectados, argumentando sobre el conocimiento o procesando la información, derivada de estos datos y decidiendo las mejores acciones para lograr el objetivo dado. Los sistemas de IA pueden usar reglas simbólicas o aprender un modelo numérico, y también pueden adaptar su comportamiento analizando cómo el entorno es afectado por sus acciones previas”. [3]

Esta definición ha sido sugerida recientemente por el Grupo de expertos de

alto nivel independiente sobre inteligencia artificial (AI HLEG) con el fin de aclarar aspectos de la IA como disciplina científica y como tecnología con el objetivo de evitar malentendidos, para lograr un conocimiento común compartido de la IA que pudiera ser utilizado de manera fructífera también por los no expertos en el tema y para proporcionar detalles útiles que se puedan utilizar en la discusión sobre sus pautas de ética y recomendaciones de políticas. [3]

Un informe del Centro Común de Investigación (JRC) al servicio de ciencia y conocimiento de la Comisión Europea del año 2020 da cuenta de la ausencia de una definición de IA formal acordada en común pero toma la definición anterior como referencia y la compara con una colección de definiciones desarrolladas desde 1955 hasta 2019 que incluye más de 50 versiones, al tiempo que aporta una visión general completa sobre el dominio de la IA y propone una taxonomía y palabras clave representativas. [4]

Los aportes valiosos de este trabajo son: la colección de definiciones desarrolladas entre 1955 y 2019; el resumen de las principales características del concepto de inteligencia artificial tal como se refleja en la literatura relevante; y el desarrollo de un proceso replicable que pueda proporcionar una definición dinámica y taxonomía de la IA.

Además de la Unión Europea, otras organizaciones internacionales como el BID, el Foro Económico Mundial (FEM), el IEEE, la OCDE, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) o el Instituto de Ética en la IA de la Universidad de Oxford, están trabajando en la formulación de lineamientos, guías y herramientas para que los países promuevan el desarrollo de sistemas autónomos centrados en los derechos humanos. Sin embargo, América Latina se encuentra aún en una etapa temprana en el desarrollo de guías, protocolos, evaluaciones, marcos regulatorios y procesos de vigilancia social en temas de ética de IA. [5].

Un Informe de fAIr LAC, red de profesionales y expertos académicos,

gubernamentales, de la sociedad civil, la industria y el sector emprendedor cuyo objetivo es promover la aplicación ética de la IA en América Latina y el Caribe, señala en sus conclusiones que los países estudiados (Argentina, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Paraguay, Perú e Uruguay) presentan esfuerzos incipientes en cuanto a conversaciones y debates sobre privacidad, rendición de cuentas, seguridad, transparencia y explicabilidad, justicia y discriminación, responsabilidad profesional, promoción de valores humanos e impacto ambiental.[5]

La inteligencia artificial es un campo de las ciencias de la computación que investiga y desarrolla sistemas capaces de realizar tareas inteligentes, es decir, aquellas que se suelen atribuir a la inteligencia cognitiva de los seres humanos, tales como aprender, resolver problemas y tomar decisiones. En la actualidad, a través del machine learning y el deep learning, se busca, además, que los sistemas realicen dichas tareas de forma autónoma [2].

El hombre de hoy está tomando, en el entorno de la nueva economía, un papel cada vez más activo, buscando nuevos desarrollos, basados en investigación, que muchas veces integran IA con la finalidad de lograr mayor eficacia y autonomía en los procesos, así como el mejoramiento de la calidad de vida de la sociedad. En ese contexto, es preciso analizar y estudiar el marco normativo para un desarrollo adecuado, que esté fundado en la ética y valores humanos, lo que tal vez implica poner algunos límites al avance de la IA.

De hecho, los avances en Inteligencia Artificial plantean un enorme desafío ya que esta disciplina toca muchos aspectos de la ciencia, la tecnología y, especialmente, por lo dicho, de la sociedad, por lo que su estudio requiere un enfoque interdisciplinario y su aplicación una visión ética en todas sus perspectivas donde el ser humano ocupe el lugar central.

El uso responsable de la IA va más allá de no ejercer prácticas ilegales mediante su uso; se trata de usar la IA de una manera que no

vulnere a minorías, que evite violaciones a derechos humanos y que no lleve al incremento de la brecha de desigualdad existente, ya sea de manera intencional o accidental. Dentro de las consecuencias que trae consigo el uso de la IA existen riesgos intrínsecos, directamente relacionados con datos, y extrínsecos, vinculados con la adopción de la IA en la sociedad. [6]

Un reporte desarrollado por la organización mexicana C Minds y comisionado por el Centro de Estudios en Tecnología y Sociedad (CETyS) de la Universidad argentina de San Andrés identifica 4 tipos de riesgo intrínsecos (que dependen exclusivamente de los datos) en las etapas de planeación, programación e implementación para un uso responsable de la IA: justicia e inclusividad; fiabilidad y seguridad del sistema; privacidad y seguridad de los datos del usuario y transparencia y rendición de cuentas. En relación a los riesgos extrínsecos, hace referencia a la automatización como causante de cambios en los puestos de trabajos. [6]

El informe “Artificial Intelligence Index Report 2019”, del Human-Centered Artificial Intelligence Institute de la Universidad de Stanford [7] señala que el 58% de las grandes empresas encuestadas informaron haber adoptado Inteligencia Artificial en, al menos, una función o unidad de negocio en 2019, frente al 47% en 2018. El informe aborda también la percepción de la sociedad y las preocupaciones existentes con respecto al avance de la inteligencia artificial y señala que los países están empezando a introducir la regulación de la inteligencia artificial en sus agendas pero no se menciona la aprobación de leyes o regulaciones concretas. Además, en cuanto a las preocupaciones sobre principios éticos en inteligencia artificial, indica que se destacan la equidad, la interpretabilidad y la explicabilidad.

El informe reporta que se revisaron 3.600 artículos de prensa a nivel mundial y que se encontró que las preocupaciones pivotan en torno a los marcos legislativos, la privacidad de los datos, el uso de reconocimiento facial, el sesgo de los algoritmos y el papel de las grandes tecnológicas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Es posible que la falta de regulación de la Inteligencia Artificial provoque demoras en la investigación y desarrollo. Es fundamental determinar parámetros de actuación y fijar límites, condiciones y responsabilidades dejando en claro si las máquinas inteligentes pueden considerarse jurídicamente responsables por sus acciones ya sea de forma parcial o total.

Se ha definido, en consecuencia, como Problema a investigar cómo se están adoptando tecnologías y técnicas propias de la Inteligencia Artificial en las organizaciones de la Región, los resguardos que se toman y el marco jurídico con el que cuentan para hacerlo.

Si bien existen normas jurídicas que tipifican los delitos informáticos, protegen datos o algoritmos y múltiples regulaciones locales, es necesario conocer la normativa que regula el desarrollo y utilización de actividades propias de la IA así como también las normas técnicas que configuran estándares para el mercado.

Las normas ISO sobre inteligencia artificial [8], [9], [10] la define de forma un poco vaga, con la finalidad de no limitarse, como: “la capacidad de una unidad funcional para realizar funciones que generalmente están asociadas con la inteligencia humana, como el razonamiento y el aprendizaje.” En el ámbito de la Unión Europea, se propone el cumplimiento de estos objetivos a través de la aplicación de normas internacionales, de la creación de un código de Conducta, de la revisión de las reglas de responsabilidad vigentes y de la evaluación del impacto social que la IA pueda llegar a tener. Se propone, como una de las primeras actividades del Proyecto, relevar las normativas vigentes en la región, que se vinculen con cualquier etapa de la producción y uso de aplicaciones.

Dentro de las ciencias de la computación, la IA, como se apuntó en la Introducción, constituye una categoría cuya finalidad es diseñar sistemas inteligentes y autónomos, basados en el conocimiento humano. El

concepto de IA se aplica a los sistemas que manifiestan un comportamiento inteligente, capaces de analizar su entorno y pasar a la acción con cierto grado de autonomía, con el fin de alcanzar objetivos específicos. En la región se realizan desarrollos y se utilizan aplicaciones con tales características. Se propone como otro de los objetivos del proyecto y primeras actividades, conocer el panorama de desarrollo y aplicación de IA en las organizaciones de la región, identificando quienes llevan a cabo el desarrollo, con qué fines y contraparte, quienes son los usuarios externos o internos, etc..

La incorporación de Inteligencia Artificial a los procesos organizacionales impulsa a dotar de valores y principios a la aplicación tecnológica y a comprender las implicaciones morales y éticas que conllevan. En este sentido, se plantea, también, que es necesario, además, relevar las preocupaciones de las organizaciones sobre principios éticos aplicables a la Inteligencia Artificial.

La Enciclopedia de IA de Shapiro [11] señala que la IA se ocupa del descubrimiento de modelos computacionales que se comportan de la misma manera que lo hacen los humanos y, por lo tanto, de las representaciones y los procesos que se definen utilizando las herramientas computacionales disponibles. Los riesgos e impactos que pueden derivar de la utilización de sistemas provistos de IA llevan a preguntarse por la responsabilidad civil contractual y extracontractual que debe considerarse dentro de diversos ámbitos y, en ese sentido, se estableció otro objetivo con el fin de conocer los resguardos que toman las empresas para evitar que la IA sea utilizada con fines ilícitos o contra la ética en todas las etapas de los procesos vinculados a su utilización y de tipificar tales riesgos.

En particular, los riesgos que aparecen ante la falta de explicabilidad o transparencia ya que la IA trae consigo riesgos que pueden derivarse de defectos, errores o sesgos en la programación que atenten contra la privacidad, intimidad, dignidad, salud o patrimonio de las personas y la responsabilidad eventualmente podría estar

distribuida entre distintos actores haciendo más complejo el problema. En este aspecto surge otro objetivo tendiente a averiguar si las organizaciones que producen las aplicaciones y programan los algoritmos, toman, también, recaudos, ya sean las mismas u otras distintas a las que las aplican.

Se llevará a cabo una investigación cualitativa, con una etapa exploratoria que tomará por unidad de análisis, primero, el corpus jurídico vinculado a la IA y luego las organizaciones que desarrollan o aplican IA en la región en función de los objetivos planteados y previamente descriptos en este trabajo.

Entre las actividades propuestas en el marco del proyecto está previsto el diseño de cuestionarios para la recolección de datos en organizaciones ya mencionadas, a ser aplicados entre usuarios dentro y fuera de las organizaciones (desarrolladores, usuarios, clientes, proveedores, etc.)

3. RESULTADOS ESPERADOS

En conformidad con lo planteado en los objetivos específicos del proyecto se espera:

- a) Identificar normas, protocolos, guías, etc que se estén aplicando en la región en relación a las prácticas vinculadas con el desarrollo y aplicación de IA
- b) Identificar el contexto
 - a. quiénes desarrollan IA,
 - b. qué técnicas aplican
 - c. con qué fines,
 - d. con qué contrapartes,
 - e. usuarios internos y externos alcanzados, etc.,
- c) recoger valores y principios de responsabilidad corporativa,
- d) reconocer normas de conducta que se aplican en torno al diseño, desarrollo y aplicación de las tecnologías IA
- e) e identificar dilemas éticos y prejuicios.

La información que se releve y el conocimiento que se genere podrá ser un insumo para el desarrollo de una guía de pautas deseables y prácticas preferidas por las organizaciones que permitan la definición de buenas prácticas orientadoras

Se espera que los logros alcanzados por este Proyecto, además de generar conocimiento empresarial útil para las organizaciones que se estudien y el mercado informático, impacten positivamente en la carrera de grado Ingeniería en Sistemas y demás carreras del Departamento, en especial en las asignaturas Inteligencia de Negocios y Aspectos profesionales para Ingeniería.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación Etica-IA está integrado por profesores del Departamento de Sistemas, docentes de la asignatura Inteligencia de Negocios y de Aspectos Profesionales para Ingeniería de la carrera de grado Ingeniería en Sistemas, asignaturas que toman como insumo académico lo producido en el marco de los Proyectos de Investigación que desarrollan.

Las investigaciones en curso del Departamento de Sistemas son propuestas cada año a los estudiantes como espacios de desarrollo de sus trabajos finales. Con este Proyecto en particular, se espera, además, aportar organizaciones dónde los estudiantes avanzados puedan realizar sus prácticas profesionales.

Las tareas propuestas en el proyecto para el primer trimestre incluyen la incentivación de estudiantes a integrar el Proyecto de investigación en carácter de investigador inicial.

5. BIBLIOGRAFÍA

[1] Mantilla, Sofia E. “Hacia una Estrategia Nacional de Inteligencia Artificial. Instituto de Estrategia Internacional, 2018.

[2] Ministerio de Asuntos Económicos y Transformación Digital, “ENIA, Estrategia Nacional de Inteligencia Artificial” Versión 1.0, 2020

[3] High-Level Expert Group on Artificial Intelligence “A Definition of AI: Main Capabilities and Disciplines” European Commission, 2019

[4] Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., and

Delipetrev, B., AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence, EUR 30117 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-17045-7, doi:10.2760/382730, JRC118163.

[5] Constanza Gómez Mont; Claudia May Del Pozo; Cristina Martínez Pinto; Ana Victoria Martín del Campo Alcocer. “La Inteligencia Artificial al Servicio del Bien Social en América Latina y el Caribe: Panorámica Regional e Instantáneas de doce Países”, 2020.

[6] Gómez Mont, C.; Del Pozo, C.; Martín del Campo, A. V. (2020). Economía de datos e inteligencia artificial en América Latina. Oportunidades y riesgos para un aprovechamiento responsable. En C. Aguerre, (Ed.). Inteligencia Artificial en América Latina y el Caribe. Ética, Gobernanza y Políticas. Buenos Aires: CETyS Universidad de San Andrés.

[7] Raymond Perrault, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles, “The AI Index 2019 Annual Report”, AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA, December 2019.

[8] ISO 10218–1:2011, Robots and robotic devices — Safety requirements for industrial robots.

[9] ISO 13855, Safety of machinery, Positioning of safeguards with respect to the approach speeds of parts of the human body.

[10] ISO/TS 15066:2016 Robots and robotic devices — Collaborative robots.

[11] Shapiro, Stuart C.: Encyclopedia of artificial intelligence (2nd Ed.). John Wiley & Sons, Inc., New York, 1992.

Subsistemas para Análisis de Textos

Marina Cardenas¹, Julio Castillo¹, Nicolas Hernandez¹

¹ Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
{jotacastillo, ing.marinacardenas}@gmail.com

Resumen

En este artículo se presenta el proyecto de investigación denominado *Desarrollo de Sistemas de Análisis de Texto*.

Este proyecto se centra en el estudio, análisis y procesamiento de información textual y en el desarrollo de herramientas software que permitan agilizar y facilitar su exploración e inferencia de información.

Se presenta desde un enfoque de sistemas y subsistemas, las partes principales proyecto y las herramientas que se están desarrollando para abordar el procesamiento de textos.

Las líneas de investigación en la que se encuadra el proyecto es dentro de las áreas de lingüística computacional y de aprendizaje automático, para atacar problemas con orígenes de texto con y sin estructura definida.

Palabras clave: análisis de texto, extracción de información, corpus, machine learning.

Contexto

En este artículo se describe el proyecto denominado Análisis de Texto (ADT), un proyecto homologado por la SCyT de la UTN [1], y que se enmarca dentro del área de lingüística computacional [2]. El proyecto físicamente se desarrolla en el Laboratorio de Investigación de Software LIS¹ del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC).

En el laboratorio convive el desarrollo de proyectos de muy variada envergadura,

entre los temas abordados resaltamos el estudio y diseño de compiladores y metalenguajes de programación, proyectos relacionados con análisis de imágenes, autómatas y modelos de pronósticos, y un proyecto de educación en tecnologías de información [3].

El LIS está compuesto por doctores, doctorandos, ingenieros, docentes-investigadores, pasantes, becarios alumnos y de posgrado. También se trabaja con expertos en ciencias sociales de CONICET, con los cuales se colabora para desarrollar modelos computacionales que permitan abordar la problemática de la salud humana siguiendo una perspectiva holística y en base a la interacción con su entorno [4].

Los proyectos desarrollados en el LIS se enmarcan en varias líneas de investigación y desarrollo (que se mencionan en la siguiente sección) y que dieron origen o colaboraron en la creación de varios grupos UTN de investigación.

1. Introducción

El proyecto denominado Desarrollo de sistemas de análisis de texto (ADT) aborda el problema de análisis y procesamientos de textos los cuales pueden provenir de orígenes muy variados y disímiles [5].

En ese contexto, una de las tareas que se desarrollan es la orientada a la construcción de corpus lingüísticos, cuya necesidad varía de acuerdo al problema que se intenta abordar.

Para llevar a cabo esta tarea se ha desarrollado un subsistema que facilite la clasificación manual y que sirva para disminuir la tasa de errores de anotación de

¹ www.investigacion.frc.utn.edu.ar/mslabs/

los etiquetadores humanos, el cual hemos denominado Subsistema de Asistente de Creación de Corpus (ACC).

Nos concentramos en una serie de problemas de procesamientos de textos que detallamos a continuación:

- el problema de la normalización de diversas fuentes de información en un repositorio común, al cual lo hemos aproximado mediante el desarrollo de un subsistema denominado Subsistema de Mapeo de Datos

- el problema de evaluación y selección de métricas en base a la elección objetiva de un criterio, a través del Subsistema de Banco de Prueba de Algoritmos de Semejanza.

- el problema de detección de similitud en textos se ha abordado mediante el desarrollo del Subsistema de Análisis de Similitud

- y el problema de visualización de diferencias en el texto, se ha atacado a través del desarrollo de un Subsistema de Visualización de Diferencias en Documentos.

Los resultados proporcionados por estos subsistemas se pueden utilizar en el contexto de varias tareas de análisis de texto, ya que pueden ser útiles para problemas de extracción de información y minería de datos en textos no estructurados [6,7]. Otros problemas, como búsquedas de implicaciones en textos [8], implicaciones textuales [9, 10], extracción de información [11,12] o minería de datos [13, 14] también pueden hacer uso de estos subsistemas.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de investigación en la que se enmarca este proyecto es computación lingüística desde el abordaje del aprendizaje por computadoras.

Esta línea abarca un campo científico interdisciplinar cuyo principal objetivo es el de desarrollar sistemas con la capacidad de reconocer y comprender el lenguaje natural

humano a través de modelos computacionales. Más aún, esto da origen a otras sublíneas de investigación en lingüística, como la denominada Lingüística de Corpus [15], que es aquella que aborda los problemas del lenguajes a partir de ejemplos reales de producciones lingüísticas (orales o escritas) que se almacenan en un computador y a partir de las cuales es posible inferir conocimientos.

A su vez, hay otras líneas de investigación, que se desarrollan en el mismo laboratorio (LIS), como lo son: la línea de investigación en teoría de autómatas y gramáticas formales, la línea de investigación de construcción de modelos de pronósticos, y otra de modelado de problemas de ciencias sociales utilizando técnicas de inteligencia artificial.

Con estas líneas de investigación, y en particular con los proyectos que los componen, se forma una interesante sinergia que permite fortalecer a cada línea y proyecto de investigación en base a la fructífera interacción y colaboración entre los investigadores que la integran.

Estas líneas de investigación se han plasmado concretamente en la creación de un grupo UTN de investigación.

3. Resultados

En el proyecto se han desarrollado varios sistemas de análisis y procesamiento de texto, entre los más importantes mencionaremos a un sistema Software de Asistente de Creación de Corpus (ACC), un Sistema de Mapeo de Datos (PMD), un Sistema de detección de similitudes en archivos de código fuente (SDS), un subsistema de Banco de Pruebas de Algoritmos de Semejanza (BPAS), y un subsistema de Visualización de Diferencias en Documentos (VDD). Estos sistemas se describen con más detalle a continuación.

El ACC permite realizar la clasificación de un conjunto de fenómenos de origen léxico, sintáctico, semánticos y morfológicos, mediante la selección de

subcadenas entre las que se sostiene un determinado fenómeno lingüístico. Esto es especialmente útil en tareas como la implicación de textos o en la detección de paráfrasis. Una funcionalidad destacable del ACC es permitir trabajar con subcadenas de texto, y anotar el fenómeno lingüístico que entre ellos se sostiene.

Este subsistema continúa en ampliación de sus funcionalidades, y se está trabajando con un módulo que genere reportes estadísticos que permitan calcular automáticamente la consistencia del material de entrenamiento, midiendo el acuerdo inter-anotador.

El subsistema de Mapeo de Datos (PMD) permite procesar información originada en múltiples fuentes de datos estructurados, tales como archivos de textos, tablas de bases de datos diferentes, y armar con ellos un repositorio de información centralizada, sobre la cual puedan aplicarse posteriormente técnicas de recuperación de información. La característica positiva es que concentra la información de manera centralizada, pero existe un aspecto negativo, ya que la información debe mantenerse actualizada periódicamente a los efectos de reflejar los últimos cambios de la información en origen. En este sentido, es similar a lo que ocurre con bases de datos multidimensionales y con cubos OLAP.

El subsistema de Banco de Pruebas de Algoritmos de Semejanza (BPAS), permite trabajar con archivos de configuraciones en los cuales se encuentran un conjunto de criterios y un peso relativo de importancia. Luego, en base a estos valores se emplea el método de Jerarquía Analítica (AHP), el cual permite encontrar la opción óptima en base a los criterios previamente establecidos en los archivos de configuraciones. El método AHP [16] es utilizado cuando se deben tomar decisiones en escenarios complejos [17] y provee una manera de encontrar la solución óptima al

problema de seleccionar una opción entre un conjunto de opciones posibles.

Esta herramienta dota de un criterio más preciso y riguroso para justificar la elección objetiva de una alternativa ante un conjunto de opciones.

El subsistema de Análisis de Similitud (SAS) realiza el procesamiento de archivos de códigos fuentes en varios lenguajes de programación e informa un valor porcentual con el grado de similitud de los mismos. Esta herramienta permite la comparación de un archivo contra un conjunto de archivos a los efectos de poder encontrar los k-archivos más similares. Es útil en la detección de patrones en códigos fuentes (que pueden estar escritos en varios lenguajes de programación), en la reutilización de código, entre otras potenciales aplicaciones.

Por último, el subsistema de Visualización de Diferencias en Documentos (VDD) cuenta con una interfaz gráfica que permite contrastar visualmente la diferencia entre dos archivos de texto. Para ello, se usan diferentes colores a los efectos de resaltar la información nueva, faltante o modificada, en relación a los documentos origen-destino.

Los subsistemas mencionados anteriormente se encuentran en etapa de desarrollo, ampliación e integración en un pipeline de trabajo común.

4. Formación de Recursos Humanos

El equipo de investigación y desarrollo de software, está formado por docentes investigadores de la Universidad Tecnológica Nacional, Facultad Regional Córdoba, que a continuación se detallan:

- Un doctor en ciencias de la computación, quién guía a becarios de grado y de posgrado, dirige prácticas profesionales supervisadas y pasantías.
- Un doctorando en ingeniería con mención en sistemas de información de la Universidad Tecnológica Nacional,

Facultad Regional Córdoba que trabaja en temáticas similares a las de este proyecto. Colabora también en la dirección de becarios de grado y posgrado realizadas en el proyecto.

- Un grupo de dos o tres becarios alumnos participan cada año realizando actividades de investigación en el proyecto, complementando así su formación curricular.
- Eventualmente se desarrollan prácticas supervisadas que constituyen uno de los requisitos para la obtención del grado de Ingeniero, y son realizadas en tareas específicas en el proyecto, acotadas en tiempo y con alcances claramente definidos.
- El proyecto además posee investigadores en formación y en proceso de categorización.
- Finalmente, se han realizado charlas de difusión y jornadas de capacitación a alumnos y a docentes de ingeniería en sistemas de información en las líneas temáticas enumeradas anteriormente.

5. BIBLIOGRAFÍA

- [1] Marina Cardenas, Julio Castillo, Martín Navarro, Nicolás Hernández, Melisa Velazco. (2019). "Herramientas para el desarrollo de sistemas de análisis de textos no estructurados". XXI Edición de Workshop de Investigadores en Ciencias de la Computación. WICC 2019, UN de San Juan, Argentina, 25 y 26 de abril de 2019.
- [2] Judith Klavans y Philip Resnik. *The Balancing Act. Combining Symbolic and Statistical Approaches to Language*. MIT Press, 1996.
- [3] Castillo, J., Cárdenas, M., Serrano, D. (2011). "Experiencias en el Desarrollo de Competencias de Programación en UTN-FRC". VI Congreso de Tecnología en Educación y Educación en Tecnología. Teyet 2011.
- [4] Rojas MC, Meichtry NC, Ciuffolini MB, Vásquez JC, Castillo J. (2008). Repensando de manera holística el riesgo de la vivienda urbana precaria para la salud: un análisis desde el enfoque e la vulnerabilidad sociodemográfica. *Salud Colectiva*. 2008; 4(2): 187- 201.
- [5] Castillo Julio J., Cardenas Marina E., Curti Adrián, Casco Osvaldo, Navarro Martín, Hernández Nicolás A., Velazco Melisa. Desarrollo de sistemas de análisis de texto. XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017). 2017.
- [6] I. Goodfellow, Y. Bengio y A. Courville. *Deep Learning*. MIT Press. 2016.
- [7] N. Buduma. *Fundamentals of Deep Learning: Designing Next-Generation Artificial Intelligence Algorithms*. O' Really book. 2015.
- [8] Castillo J. Sagan in TAC2009: Using Support Vector Machines in Recognizing Textual Entailment and TE Search Pilot task. TAC, 2009.
- [9] Castillo J., Cardenas M. Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. Iberamia 2010, LNCS, vol. 6433, pp. 366-375, 2010.
- [10] Castillo J. Using Machine Translation Systems to Expand a Corpus in Textual Entailment. Proceedings of the Iccetal 2010, LNCS, vol. 6233, pp.97-102, 2010.
- [11] FeldmanR. y Hirsh H.. Exploiting Background Information in Knowledge Discovery from Text. *Journal of Intelligent Information Systems*, 1996.
- [12] Lewis, D.. Evaluating and optimzing autonomous text classification systems. In Proceedings of SIGIR-95, 18th

ACM International Conference on Research and Development in Information Retrieval. Seattle, US, págs. 246-254, 1995.

[13] M. Craven y J. Shavlik. Using Neural Networks for Data Mining. *Future Generation Computer Systems*, 13, págs. 211-229, 1997.

[14] Gaikwad, S., Chaugule, A., & Patil, P.B. (2014). Text Mining Methods and Techniques. *International Journal of Computer Applications*, 85, 42-45.

[15] Stefan Th. (2006). Y Anatol Stefanowitsch. *Corpora in Cognitive Linguistics. CorpusBased Approaches to Syntax and Lexis*, Berlin: Mouton, pág. 1-17.

[16] Bhushan N., Kanwal R. (2004). "Strategic Decision Making: Applying the Analytic Hierarchy Process". London: Springer-Verlag.

[17] Jiandong Zhang, Jin Yang, Xuan Ma and Juanjuan Gong. (2013). "Research and Application on Improved group decision making AHP algorithm". *Applied Mechanics and Materials* Vols. 336-338.

Utilizando Argumentación Rebatible en la Detección de Intrusión en Sistemas Biométricos

Graciela R. Etchart¹, Juan C.L. Teze¹, Carlos E. Alvez¹, M. Vanina Martínez², Gerardo I. Simari³

¹*Facultad de Ciencias de la Administración, Universidad Nacional de Entre Ríos (UNER),*

²*Instituto de Investigación en Ciencias de la Computación, Universidad Nacional de Buenos Aires (UBA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET),*

³*Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET), Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS)*

¹{graciela.etchart, carlos.alvez, carlos.teze}@uner.edu.ar, ²mvmartinez@dc.uba.ar,

³gis@cs.uns.edu.ar

Resumen

Hoy en día, las tecnologías biométricas representan un componente integral en sistemas de gestión de la identidad y de control de acceso; sin embargo, los sistemas biométricos son vulnerables a ataques que pueden comprometer su seguridad y privacidad. Para aumentar su seguridad, las técnicas de detección de intrusión son considerablemente útiles. En este trabajo se busca realizar la detección de ataques dirigidos al canal de comunicación utilizando argumentación rebatible con el propósito de proporcionar una estructura que favorezca una toma de decisiones de seguridad más informada. Se considera que el enfoque formal que brinda la argumentación complementará sustancialmente los sistemas de seguridad existentes en sistemas biométricos.

Palabras clave: Argumentación Rebatible, Sistemas Biométricos, Seguridad, Detección de intrusión

Contexto

Este trabajo se da en el marco del Proyecto PID “Modelos de Machine Learning para la mejora de la precisión, seguridad y eficiencia en la gestión de datos biométricos”, que da continuidad a los Proyectos PID07/G035 “Identificación de personas mediante Sistemas

Biométricos. Estudio de factibilidad y su implementación en organismos estatales” y PID07/G044 “Gestión de datos biométricos en base de datos objeto - relacionales” [1, 2, 3].

Además, este trabajo se realiza en el marco del desarrollo de una tesis para la Maestría en Sistemas de la Información de la Universidad Nacional de Entre Ríos.

Introducción

En la actualidad, el acceso automático de las personas a los servicios es prácticamente una cuestión esencial. Esto ha dado lugar al desarrollo de diferentes métodos de autenticación. Un área tecnológica de relevancia en este sentido es conocida como reconocimiento biométrico o simplemente biometría [4]. El objetivo básico de la biometría es discriminar automáticamente entre sujetos - de forma fiable y según alguna aplicación- basándose en una o varias imágenes o señales derivadas de rasgos físicos o de comportamiento, como la huella dactilar, el rostro, el iris, la voz, la geometría de la mano o la firma escrita, entre otros.

Aunque el proceso biométrico presenta varias ventajas en la gestión de la identidad y en el control de acceso, es vulnerable a ataques que pueden disminuir su seguridad y comprometer la privacidad de los datos. Los sistemas de autenticación biométrica pueden recibir ataques

externos o sufrir una intrusión en la información privada del usuario [5], causando problemas graves y persistentes, ya que los datos biométricos son irremplazables. En la literatura han sido descriptos puntos de ataques potenciales o puntos vulnerables en los sistemas biométricos [6-8]. El mayor número de esos puntos de vulnerabilidad involucran el tráfico de datos a través del canal de comunicación del sistema biométrico.

En el contexto de la seguridad en redes, los sistemas de detección de intrusos (IDS, por sus siglas en inglés) son una herramienta de creciente preponderancia. Según los datos de la auditoría, la detección de intrusión puede clasificarse como basada en *host* o basada en red. Un IDS basado en red analiza el tráfico del canal de comunicación, mientras que uno basado en *host* utiliza en su análisis de registros del sistema operativo o de las aplicaciones. Por otro lado, según la técnica de detección empleada, un IDS puede ser clasificado como sistema de detección de anomalías o sistema de detección de mal uso. En relación a los procesos biométricos, la detección de intrusión puede ser utilizada como una capa de defensa contra los ataques que surgen en el canal de comunicación. Si se detecta la intrusión, se puede iniciar una advertencia para prevenir o minimizar los daños que pueda sufrir el sistema. En diferentes escenarios se han estudiado y aplicado con efectividad sistemas de detección de intrusión [9]. Sin embargo, la naturaleza dinámica del dominio donde se producen los ataques en ocasiones conduce a situaciones donde la información que se maneja es incompleta o potencialmente contradictoria. Este contexto constituye un escenario ideal para los sistemas argumentativos [10, 11]. El mecanismo de inferencia sobre el cual están basados, permite decidir entre conclusiones contradictorias y adaptarse fácilmente a entornos cambiantes.

En inteligencia artificial, el área de argumentación computacional se especializa en modelar el proceso de razonamiento humano de manera tal de establecer qué conclusiones son

aceptables en un contexto de desacuerdo. En el marco de este trabajo, este proceso de razonamiento permite filtrar selectivamente información para detectar amenazas a la seguridad de los sistemas biométricos, sugerir acciones y acelerar la respuesta ante acontecimientos complejos como la presencia de una posible intrusión. Este trabajo se centra en una arquitectura que extiende las capacidades de razonamiento de los sistemas biométricos incorporando argumentación al proceso de detección de intrusión. En la solución propuesta se utiliza el concepto formal de servidor de razonamiento en DeLP (DeLP-server) [12, 13], cuyo mecanismo de inferencia se base en el sistema argumentativo llamado Programación en Lógica Rebatible (DeLP, por sus siglas en inglés) [11]. Una consulta para un DeLP-server es un par (Co, L) donde L es una consulta DeLP y Co es el contexto para la consulta. El contexto puede ser cualquier programa DeLP. Un sistema basado en reglas como DeLP constituye una herramienta para el soporte a la toma de decisiones que brinda la posibilidad de explicar a los analistas humanos por qué se recomienda una acción y no otra. Además, este tipo de sistemas resulta adecuado para funcionar con un esquema “*human-in-the-loop*”, donde los analistas humanos brindan retroalimentación que permite validar o rectificar los resultados del sistema.

Línea de Investigación y Desarrollo

Esta línea de investigación se enfoca sobre la problemática involucrada en la utilización de argumentación rebatible para la detección de intrusión. Diversas técnicas de inteligencia artificial se han utilizado para la detección y/o prevención de intrusión en redes de computadoras. En este contexto, cabe mencionar algunos trabajos que aplican argumentación en cuestiones relacionadas con la seguridad informática. En [14-17] se utiliza argumentación para el desarrollo de

firewalls. En [18] los autores analizan la aplicación de un marco de argumentación abstracta para el análisis general de la seguridad de la red de un sistema. Por otra parte, en los trabajos [19, 20] se utiliza la argumentación para abordar el problema de la atribución cibernética. Una propuesta preliminar para el desarrollo de un sistema de detección de intrusión basado en representación de conocimiento y razonamiento rebatible para un entorno de red LAN, se encuentra en [21]. El trabajo presentado en [22] considera el uso de argumentación para la correlación de alerta y el análisis de intrusión. En el presente trabajo se consideran principalmente los aportes de [21, 22] para aplicar el razonamiento basado en argumentación en la detección de ataque al canal de comunicación de sistemas biométricos.

Resultados y objetivos

Para supervisar la seguridad de la red en un sistema biométrico, se propone un framework que utiliza un DeLP-server. Los primeros resultados de este trabajo fueron publicados recientemente en [23]. Como se mencionó anteriormente, los servicios de razonamiento basados en DeLP, tienen la capacidad de representar conocimiento y responder consultas contextuales. En el framework propuesto, el contexto es información sobre alertas de intrusión al canal de comunicación del sistema biométrico.

El enfoque formal proporcionado por la argumentación rebatible permite manejar la inconsistencia de los datos que deben utilizarse en la toma de decisiones y extraer un conjunto coherente de reglas que puedan aplicarse para llegar a una decisión. Además, los argumentos que se construyen permiten explicar al responsable de la toma de decisiones humana los resultados del razonamiento, de manera que se aclare la situación y se mejore la calidad de las decisiones. En el contexto de este trabajo, el proceso de argumentación llevado a cabo por el DeLP-server filtra selectivamente información para el diagnóstico de ataques al canal de

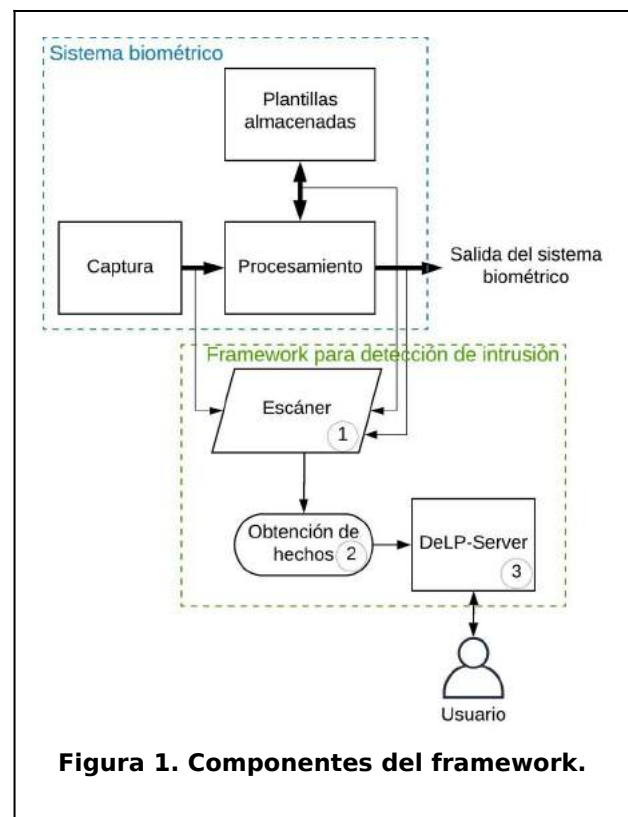


Figura 1. Componentes del framework.

comunicación del sistema biométrico y para proporcionar una estructura que informe al analista acerca de una intrusión y posibles contramedidas a adoptar.

El framework propuesto consta de tres componentes (Figura 1). Uno de los componentes es el escáner (1 en la Figura 1); encargado de obtener y registrar información sobre el tráfico del canal de comunicación, analizando paquetes capturados en segmentos de red que conforman el sistema biométrico. Para la implementación de este componente se utiliza la herramienta open source *Snort*¹, la cual además de su tipo de licencia de uso, presenta como ventaja la posibilidad de configuración para obtener información ampliada de las alertas. Otro de los componentes del framework es el módulo para la obtención de hechos (2 en la Figura 1), que consiste en un módulo computacional específico para generar hechos que expresan observaciones de los datos registrados por el escáner. Estos hechos resumen los eventos

¹ <https://www.snort.org/>

anómalos que se producen en el canal de comunicación en un período de tiempo determinado. En el framework propuesto, este conjunto de hechos luego conformará el conocimiento para contextualizar el pedido de recomendación que reciba el DeLP-server (3 en la Figura 1). En esta propuesta, el conocimiento público del DeLP-server está representado en una base de conocimiento mediante un programa DeLP con hechos y reglas que permiten detectar posibles ataques e indicar la contramedida que puede adoptarse frente a ellos. Para el armado de la base de conocimiento es necesaria la participación del experto humano en seguridad para identificar características de los ataques que pueden realizarse sobre el canal de comunicación del sistema biométrico, y para considerar políticas generales de seguridad.

Actualmente, se están desarrollando algoritmos que permitan implementar el módulo computacional para la generación de los hechos que denoten observaciones sobre los datos registrados por el escáner.

Formación de Recursos Humanos

En la presente línea de investigación se enmarca el desarrollo de una tesis para la Maestría en Sistemas de la Información de la Universidad Nacional de Entre Ríos.

Referencias

- 1 Alvez C., Etchart G., Ruiz S., Miranda E. and Aguirre J., “Extensión de una base de datos Objeto-Relacional para el soporte de datos de iris”. XXIII Congreso Argentino de Ciencias de la Computación. Universidad Nacional de La Plata– Argentina, 2017.
- 2 Ruiz S., Etchart G., Alvez C., Miranda E., Benedetto M. and Aguirre J., “Iris Information Management in Object-Relational Databases”. XXI Congreso Argentino de Ciencias de la Computación. Universidad Nacional del Noroeste de la Provincia de Buenos Aires, Junín - Argentina, 2015.
- 3 Etchart G., Luna L., Leal R., Benedetto M. and Alvez C., “Sistema adecuado a estándares de reconocimiento de personas mediante el iris”. CGIV - XIV Workshop de Investigadores en Ciencias de la Computación, Universidad Nacional de Entre Ríos, Concordia - Argentina, 2012.
- 4 Alvez C., Benedetto M., Berón G., Etchart G., Luna L. y Leal C., “Desarrollo de un sistema multi-biométrico mediante reconocimiento de iris y voz, adecuado a estándares, para su aplicación en organismos públicos”. SIE 2011 – Simposio de Informática en el Estado. Córdoba, 31 de agosto, 01 y 02 de septiembre de 2011. 40° JAIIO. pp. 206 - 220.
- 5 Rui, Z., and Yan, Z., “A Survey on Biometric Authentication: Towards Secure and Privacy-Preserving Identification”. IEEE Access, 2019, 7, pp. 5994 – 6009.
- 6 Marcel, S., Nixon, M. S., and Li, S. Z., Eds., “Handbook of Biometric Anti-Spoofing - Trusted Biometrics under Spoofing Attacks”. ser. Advances in Computer Vision and Pattern Recognition. Springer, 2014.
- 7 Galbally, J., Cappelli, R., Lumini, A., Gonzalez-de-Rivera, G., Maltoni, D., Fierrez, J., Ortega-Garcia, J., Maio, D., “An evaluation of direct attacks using fake fingers generated from ISO templates”. Pattern Recognition Letters, Vol. 31, Issue 8, 2010, pp. 725-732.
- 8 Ratha, N., Connell, J., Bolle, R. “An analysis of minutiae matching strength”. In Proc. AVBPA. LNCS, Vol. 2091. Springer, 2001, pp. 223–228.
- 9 Hamed, T., Ernst, J.B., Kremer, S.C. “A Survey and Taxonomy of Classifiers of Intrusion Detection Systems”. In Daimi K. (eds) Computer and Network Security Essentials. Springer, Cham. 2018.
- 10 Simari, G.R. and Loui, R., “A mathematical treatment of defeasible reasoning and its implementation”. Artificial Intelligence 53 (2–3). 1992, pp. 125–157.
- 11 García, A. and Simari, G.R., “Defeasible Logic Programming: An Argumentative Approach”. Theory and Practice of Logic Programming 4(1). 2004, pp. 95–138.
- 12 García, A., Rotstein, N., Tucac, M. and Simari, G.R. “An argumentative reasoning service for deliberative agents”. In KSEM, 2007, pp. 128–139.
- 13 García, A. and Simari, G.R., “Defeasible logic programming: Delpservers, contextual queries,

- and explanations for answers”. *Argument & Computation* 5. 2014, pp. 63-88.
- 14 Applebaum, A., Levitt, K., Rowe, J., and Parsons, S., “Arguing about firewall policy in COMMA”. ser. *Frontiers in Artificial Intelligence and Applications*, Verheij, B., Szeider, S., and Woltran, S. (eds), Vol. 245. IOS Press, 2012, pp. 91–102.
 - 15 Bandara, A., Kakas, A., Lupu, E., and Russo, A., “Using argumentation logic for firewall policy specification and analysis”, in *DSOM*, ser. LNCS, R. State, S. van der Meer, D. O’Sullivan, and T. Pfeifer, Eds., Vol. 4269. Springer, 2006, pp. 185–196.
 - 16 Bandara, A., Kakas, A., Lupu, E., and Russo, A., “Using argumentation logic for firewall configuration management”. In *Integrated Network Management*. IEEE, 2009, pp. 180–187.
 - 17 Rajkhowa, P., Hazarika, S.M., Simari, G.R. “An Application of Defeasible Logic Programming for Firewall Verification and Reconfiguration”. In Singh, K., Awasthi, A.K. (eds) *Quality, Reliability, Security and Robustness in Heterogeneous Networks*. QShine, Vol 115. Springer, Berlin, Heidelberg, 2013.
 - 18 Martinelli, F., Santini, F. and Yautsiukhin, A., “Network Security Supported by Arguments”. 2015.
 - 19 Shakarian, P., Simari, G.I., Moores, G., Parsons, S., and Falappa, M., “An Argumentation-based Framework to Address the Attribution Problem in Cyber-Warfare”. *Proceedings of the 3rd ASE International Conference on Cyber Security*, 2014.
 - 20 Nunes, E., Shakarian, P., Simari, G.I. and Ruef, A., “Argumentation models for cyber attribution”, *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, San Francisco, CA, 2016, pp. 837-844.
 - 21 Guasco, L., Echaiz, J., and Ardenghi, J., “Framework para detección de intrusos usando DeLP”. *IX Workshop de Investigadores en Ciencias de la Computación*, 2007.
 - 22 Applebaum, A., Levitt, K., Li, Z., Parsons, S., Rowe, J., Sklar, E., “Cyber reasoning with argumentation: Abstracting from incomplete and contradictory evidence”. *MILCOM 2015 - IEEE Military Communications Conference*, 2015, pp. 623-628.
 - 23 Etchart G., Teze J.C., Alvez C., Martínez M.V., Simari G.I., “Hacia la Detección de Intrusión en Sistemas Biométricos Utilizando Argumentación Rebatible”. *8º Congreso Nacional de Ingeniería Informática – Sistemas de Información (CoNaIISI)*. Universidad Tecnológica Nacional - Facultad Regional San Francisco – Argentina, noviembre de 2020.

Bases de Datos y Minería de Datos

ANÁLISIS DE CALIDAD DE VIDA DE LA POBLACIÓN SANJUANINA APLICANDO CIENCIA DE DATOS

Herrera, Myriam^{1,2}, Susana Ruiz³, Simón Pedro Gonzalez², María Gema Romagnano^{1,2,3}, María Inés Lund¹, Leonel Ganga¹, María del Valle Scheffer², Magalí Sosa³

¹Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan

²Instituto de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan

³Departamento de Geofísica y Astronomía, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan

{myriamhrrr, leonelganga, sbruizr, simon.pedro.g, magamsosa}@gmail.com
{mlund, maritaroma}@iinfo.unsj.edu.ar

RESUMEN

En el proceso de toma de decisiones resulta imprescindible la disponibilidad de información relevante, particularmente cuando las consecuencias de dichas decisiones tienen el potencial de impactar en la calidad de vida de un significativo número de personas. Para obtener dicha información, resulta importante contar con datos, recursos y un análisis adecuado. Por lo tanto resulta fundamental la aplicación de técnicas y herramientas basadas en la “Ciencia de Datos”.

Específicamente se propone la aplicación de metodologías de análisis exploratorio sobre datos de la población sanjuanina, para realizar descripciones mediante números reducidos y manejables de factores y el establecimiento de perfiles, lo que, sumado a otros métodos, permitan proveer información concisa, adecuada y entendible para la toma de decisiones a nivel de Estado Provincial y/o Nacional.

Se presenta un informe basado en un análisis realizado considerando datos socio-económicos sanjuaninos, provenientes del Instituto Nacional de Estadística y Censos de

la República Argentina (INDEC). Dicho análisis, posibilita el descubrimiento de estructuras internas en los datos y hacen posible el establecimiento de perfiles o tipologías que caracterizan a la población en relación a la calidad de vida

Palabras clave: Ciencia de Datos, Población, Calidad de Vida, Perfiles.

CONTEXTO

Este trabajo forma parte del proyecto “**Aplicación de la Ciencia de los Datos para determinar la Calidad de Vida de la Población del gran San Juan en relación con la Nación**” Este proyecto fue aprobado y acreditado por CICITCA Código: 21/E1122 (Código SIGEVA: 80020190100021SJ) que ha sido presentado en la Convocatoria 2019 de la Universidad Nacional de San Juan, para el período comprendido entre 01/01/2020 al 31/12/2021 y se desarrolla en el Gabinete de Matemática Aplicada del Instituto de Informática de la Facultad de Ciencias

Exactas, Físicas y Naturales de la Universidad Nacional de San Juan.

Dicho proyecto que es continuación de las líneas de investigaciones desarrolladas en proyectos anteriores, entre los que se mencionan:

- “Calidad Educativa Universitaria mediante Técnicas de Clasificación”. Acreditado por CICITCA. (2018-2019). Código: 21/E-1050
- “Técnicas de Clasificación aplicadas al rendimiento académico”. Acreditado por CICITCA. (2016-2017). Código: 21/E-1011.
- “Algoritmos de Clasificación de Procesos Multivariados utilizando Medidas de Asociación Espacial”. Acreditado por CICITCA. (2014-2015). Código: 21/E-948.
- “Determinación y Comparación de Perfiles Sociales y Culturales de Estudiantes Universitarios a través de Técnicas Estadísticas Multivariadas”. Acreditado por CICITCA. (2014-2015). Cód.: 21/F-982.
- “Clasificación Espacial Multivariada”. Acreditado por CICITCA. (2011-2013). Cód.: 21/E-878
- “Reducción y Selección de Variables en la Clasificación Digital”. Acreditado por CICITCA. (2008-2010). Cód.: 21/E-820.
-

1. INTRODUCCIÓN

En los tiempos que corren, los datos constituyen una nueva materia prima cuya explotación debe ser fuente de conocimiento para la comunidad y los organismos que la administran. Tradicionalmente, para realizar caracterizaciones socio-económicas de poblaciones se recurre al análisis de parámetros de tendencia, dispersión y distribuciones. Estos resultados, que sin duda son útiles, se pueden enriquecer en gran medida mediante la aplicación de técnicas de Ciencias de Datos que permitan analizar los datos de forma global, resumida y rápida, para definir la tipología de la población. En

particular, que posibiliten situar a los argentinos y en especial a los sanjuaninos, en temas de condiciones de vida. Por las razones mencionadas, mediante el proyecto de investigación, se propone como objetivo la utilización de herramientas y técnicas descriptivas y exploratorias propias de la Ciencia de Datos, campo interdisciplinario que agrupa Estadística, Inteligencia Artificial (incluyendo el Aprendizaje de Máquina), Minería de Datos, etc.[1], que permitan la extracción de información con un enfoque más holístico de los datos, para su posterior conversión a un mensaje manejable. Este mensaje debe ser entregado en forma de gráficas y tablas que resuman los datos de manera comprensible, de las que se pueden extraer conclusiones generales al servicio de la toma de decisiones en políticas públicas [2].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El trabajo que desarrolla este equipo de investigación, en base al objetivo propuesto, tiene como eje central a la “**Ciencia de Datos**” que consiste en la práctica de obtener información valiosa, para dar respuesta al afrontar los retos del tratamiento de grandes conjuntos de datos [3], [4], [5]. Por ello, se parte de un estudio bibliográfico donde se tienen en cuenta, diversas técnicas y teorías de muchos campos dentro de amplias áreas como la Estadística y las Tecnologías de la Información, incluyendo modelos probabilísticos, machine learning, lógica difusa, aprendizaje estadístico, programación, ingeniería de datos, reconocimiento de patrones, visualización, modelización de la incertidumbre, etc. Además, se analizan las posibilidades de aplicación de Python, SQL y R, conocidos como “los tres mosqueteros en la ciencia de datos” [6] para extraer e integrar datos (estructurados, no estructurados o semi-estructurados). Estos representan un recurso crítico en muchas organizaciones y por lo tanto, el acceso eficiente a estos, la extracción de información y el uso que se hace con ellos, se transforman en procesos muy importantes, que apoyan el entendimiento de grandes y

complejos conjuntos de datos, brindando valiosa información, común a diversos campos, tales como el de los negocios, de las ciencias sociales y la ingeniería.

Se observó que existen varios softwares de licencia paga con características útiles para el análisis de datos de tipo socio-demográfico, tal como por ejemplo “Coheris Analytics SPAD”[7]. Sin embargo, respondiendo a un espíritu de accesibilidad y democratización de la información, se buscaron alternativas libres, optándose finalmente, en el caso del ejemplo, por R [8] [9], un entorno y lenguaje de programación libre con enfoque estadístico, muy popular en las investigaciones biomédicas, bioinformáticas, financieras y demográficas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Para realizar descripciones de la población de estudio, mediante números reducidos y manejables de factores, para establecer perfiles que permitan proveer información concisa, adecuada y entendible, para la toma de decisiones a nivel de Estado Provincial y/o Nacional, se requiere de la aplicación adecuada de técnicas y recursos de Ciencia de los Datos [10].

De esta manera, este grupo de investigación, logra un primer avance, dentro de sus objetivos, que resulta de un análisis realizado considerando datos socio-económicos sanjuaninos, provenientes del Instituto Nacional de Estadística y Censos de la República Argentina (INDEC) [11]. Para dicho análisis se requirió de la aplicación de técnicas para el descubrimiento de estructuras internas en los datos que posibilitará establecer perfiles o tipologías que caracterizan a la población en relación a la calidad de vida.

Concretamente, se tomaron las bases de datos correspondientes a hogares e individuos, ambas resultantes de la Encuesta Permanente de Hogares del cuarto trimestre del año 2017. Se contaron con 58182 respuestas de la encuesta de individuos de Argentina (de las cuales 1957 pertenecen al aglomerado Gran San Juan) y 18793 respuestas de la encuesta

de hogares de Argentina (de las cuales 566 pertenecen al aglomerado Gran San Juan). Como las variables estudiadas resultaron mayoritariamente de tipo nominal, se consideró la aplicación del Análisis de Correspondencia Múltiple (ACM) [12], el cual es una buena opción. En nuestro caso, cada variable representa una pregunta de la encuesta y las posibles respuestas son las modalidades o categorías mutuamente excluyentes de cada variable. Al responder a la encuesta, un individuo es el objeto de estudio que asumió una categoría para una variable determinada. En la siguiente Tabla se muestran las variables relacionadas a los Individuos considerados en este avance.

Tabla: AMBITOS, VARIABLES DE LA BASE DE DATOS DE INDIVIDUOS

Ámbito	Variables
Características generales	Relación de Parentesco
	Sexo
	Edad
	Estado Civil
	Nivel Educativo
Características de Ocupación	Estado de Actividad
	Actividad Laboral
	Razón de Inactividad
	Busca Trabajo
	Tipo de Empresa (donde trabaja)

Una vez seleccionadas las variables, los datos se representaron como una nube de puntos en un espacio de dimensión NP (N-individuos, P-modalidades), donde fue posible identificar a los individuos (filas de la matriz) y las modalidades (columnas de la matriz) mediante coordenadas y calcular distancias, con la métrica de X^2 (Chi-cuadrado). La distancia entre individuos refleja qué tan semejantes son. Por otro lado, dos modalidades se consideraron próximas si fueron asumidas en gran medida por el mismo grupo de individuos. Cabe destacar que todo el análisis se realizó con el lenguaje de programación R.

Una de las informaciones que se pudieron obtener para dar una tipología de la provincia de San Juan queda manifiesto en los siguientes clusters.

CLUSTER 1: 42.57 % DE INDIVIDUOS DE SAN JUAN correspondiente a individuos ocupados, mayoritariamente varones, entre 26 y 49 años, de nivel educativo secundario o universitario, jefes/as de hogar. Podemos considerar que esta sección de la población constituye la fuerza de trabajo remunerada.

CLUSTER 2: 43.94 % DE INDIVIDUOS DE SAN JUAN corresponde a individuos inactivos, mayoritariamente mujeres, de nivel educativo más modesto, jubilados, estudiantes o amos/as de casa. Podemos ver cómo el 37,9% del cluster son estudiantes, el 38,95% jubilados y el 17% amas/os de casa. Esto causa que el rango etario varíe significativamente.

CLUSTER 3: 13.49 % DE INDIVIDUOS DE SAN JUAN corresponde a niños de entre 0 y 13 años. Son por lo general hijos o nietos del jefe del hogar.

En el presente año se espera trabajar con datos de texto, ya que en la actualidad es muy difícil contar con encuestas estructuradas y es factible contar con datos provenientes de diferentes dispositivos disponibles en redes. Esto nos llevará a profundizar en el estudio de técnicas cualitativas multidimensionales de textos [13][14] con el objetivo de establecer Nuevas variables influyentes para la determinación de la calidad de vida de los habitantes del Gran San Juan relacionados con Nación.

4. FORMACIÓN DE RECURSOS HUMANOS

En el grupo de trabajo se encuentran un doctorando, un docente estudiando una maestría, dos ayudantes alumnos de la carrera de grado de Licenciatura en Ciencias de la Computación.

5. BIBLIOGRAFÍA

[1] V. Dhar, “Data science and prediction”, 2013. Communications of the ACM. pp. 64-73, 56 (12). Retrieved from: <https://dl.acm.org/doi/abs/10.1145/2500499>

[2] Z. Engin and P. Treleaven, Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies”, in *The Computer Journal*, vol. 62, no. 3, pp. 448-460, March 2019, Retrieved from: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8852885>

[3] “Capítulo 1 ¿Qué es la ciencia de datos?,” in *Ciencia de Datos para Gente Sociable*. Available: https://bitsandbricks.github.io/ciencia_de_datos_gente_sociable/que-es-la-ciencia-de-datos.html [Accessed: 15-Jun-2020].

[4] F. Vivas, “Qué es la ciencia de datos y qué posibilidades abre en el mercado laboral,” *Rev. El Economista*, 2019. [Online]. Available: <https://www.economista.com.ar/2019-11-que-es-la-ciencia-de-datos-y-que-posibilidades-abren-en-el-mercado-laboral/>. [Accessed: 04-Mar-2020].

[5] J. P. Mora, *Que es la Ciencia de Datos, el aprendizaje automático (ML), el Big Data y cuales son sus usos?*.

[6] J. Jiménez, “Científico de datos: así es y así se forma uno en esta profesión cada vez más demanda”, 2018. [Online]. Available: <https://www.xataka.com/otros/de-profesion-cientifico-de-datos>. [Accessed: 04-Mar-2020]

[7] “¿Qué es Coheris Analytics SPAD?” [Online]. Available: <https://www.comparasoftware.com/coheris-analytics-spad-mineria-datos> [Accessed: 01-Feb-2021].

[8] “R: The R Project for Statistical Computing.” [Online]. Available: <https://www.r-project.org/>. [Accessed: 04-Mar-2020].

[9] “R: What is R?” [Online]. Available: <https://www.r-project.org/about.html>. [Accessed: 04-Mar-2020].

[10] M. Herrera, M. Romagnano, L. Ganga, “Aporte de la Ciencia de Datos a la toma de decisiones institucionales”. II Congreso de Matemática Aplicada, Computacional e Industrial, (MACI2019), 7 al 9 de mayo 2018, Río Cuarto, Córdoba. Retrieved from: <https://amcaonline.org.ar/maci/index.php/maci2019/maci/paper/viewFile/5112/344>

[11] Instituto Nacional de Estadística y Censos de la República Argentina, Encuesta Permanente de hogares. Retrieved from:

<https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos> [Accessed: 01-Mar-2020]

[12] D. Peña, *Análisis de datos multivariantes*. McGraw-Hill/Interamericana, 2002.

[13] Nora Moscoloni, “*El análisis multidimensional de datos como herramienta metodológica en ciencias sociales*”. VI Jornadas de Sociología. Facultad de Ciencias Sociales, Universidad de Buenos Aires, Buenos Aires, 2004.

[14] Schettini, P., & Cortazzo, I. , “*Análisis de datos cualitativos en la investigación social*”. La Plata: edulp, 2015. Retrieved from: http://sedici.unlp.edu.ar/bitstream/handle/10915/49017/Documento_completo.pdf?sequence=1 [Accessed: 17-Feb-2020]

ANÁLISIS DE DATOS SIMBÓLICOS PARA DATA SCIENCE

Adriana Mallea¹, Jorgelina Carrizo¹, Leonel Ganga², Cecilia Martínez²,
Andrea Salas¹

¹Departamento de Matemática, FFHA, Universidad Nacional de San Juan

²Departamento de Informática, FCFN, Universidad Nacional de San Juan

lamallea@ffha.unsj.edu.ar

RESUMEN

La ciencia de datos, considerada como una ciencia en sí misma, es en términos generales, la extracción de conocimiento de los datos. Data Mining es una poderosa tecnología con gran potencial para extraer tal conocimiento. Sin embargo, desde el punto de vista estadístico, sus herramientas sólo han sido desarrolladas para trabajar con matrices de datos clásicas, es decir, donde cada unidad es individual y las variables toman un único valor para cada individuo. El análisis de datos simbólicos (SDA, por sus siglas en inglés) brinda una nueva forma de pensar en Data Science al extender la entrada estándar a un conjunto de clases de entidades individuales. Por lo tanto, las clases de una población dada se consideran unidades de una población de nivel superior a estudiar. Tales clases a menudo representan las unidades reales de interés. Para tener en cuenta la variabilidad entre los miembros de cada clase, las clases se describen por intervalos, distribuciones, conjunto de categorías o números que a veces se ponderan. De esa manera, obtenemos nuevos tipos de datos, llamados "simbólicos", ya que no se pueden reducir a números sin perder información sobre la variabilidad interna. SDA es un nuevo paradigma que abre un vasto dominio de investigación y aplicaciones al

proporcionar resultados complementarios a los métodos clásicos aplicados a los datos estándar.

A lo largo de las últimas tres décadas se han extendido distintos métodos del análisis clásico de datos al simbólico, la mayoría de ellos descriptivos. Esto fundamenta la necesidad de continuar investigando sobre la modelización e inferencia en el contexto de datos de naturaleza simbólica. El presente proyecto pretende responder a esta necesidad. Las metodologías se aplicarán a problemas reales o simulados.

Palabras clave: Data Science, Simbólico, Análisis

CONTEXTO

El proyecto *Análisis de Datos Simbólicos para Data Science* propone continuar con la investigación y desarrollo de nuevas metodologías, cuyo estudio se inició en el proyecto CICITCA 2018-2019, 21/F1085, sobre todo referidas a modelación e inferencia en el marco del SDA, que permitan la extracción de conocimientos. Es un proyecto cuyo tipo de actividad de I+D es investigación básica, presentado para su acreditación en diciembre de 2019, inició en 2020 y es de carácter bi-anual, financia la UNSJ. Tiene como unidad ejecutora el Departamento de Matemática de la FFHA

y sus integrantes desarrollan sus tareas de docencia e investigación en las áreas de matemática e informática. La línea de investigación corresponde a minería de datos, en un marco más general de Ciencia de los Datos.

1. INTRODUCCIÓN

El volumen de datos que circula en la Web, o almacenado por las empresas, está creciendo constantemente. Para explotar esta riqueza, es necesario extraer conocimiento de grandes volúmenes de información. El dominio que apunta a resolver esta problemática es la ciencia de los datos (Data Science). Data Science tiene como objetivo extraer conocimiento de todo tipo de datos (estructurados o no, de fuentes homogéneas o heterogéneas, etc.). Representa la intersección de varias disciplinas como estadística, matemática, inteligencia artificial y minería de datos (Data Mining). Data Mining ofrece métodos de análisis muy útiles para extraer conocimiento. Sin embargo, desde el punto de vista estadístico, sus herramientas sólo han sido desarrolladas para trabajar con matrices de datos clásicas, es decir, donde cada unidad es individual y las variables toman un único valor para cada individuo. Para que el estudio de los datos sea accesible en varios niveles de agregación, ha surgido el campo del análisis de datos simbólicos (SDA). Desde entonces, este campo se ha desarrollado proponiendo varios métodos específicos de análisis de datos simbólicos. Estos métodos se han implementado en herramientas como Sodas [S3], Syr [S1] y bibliotecas de R [S5, S4, S2] que permiten sus pruebas y sus aplicaciones en nuevas bases de datos. SDA en esencia se construye sobre la noción de que las inferencias estadísticas se requieren comúnmente a nivel de grupo en lugar de a un nivel individual

(Billard, 2011, Billard y Diday, 2006). Por ejemplo, en los exámenes de prueba estandarizados, el rendimiento de la escuela, y las unidades de nivel superior suelen ser de interés más que el rendimiento de los estudiantes individuales. SDA adopta explícitamente esta idea al agregar datos de nivel individual (los microdatos) en resúmenes distribucionales a nivel grupo (es decir, los símbolos), y luego construir modelos para inferencia directamente a nivel de grupo basado en estos resúmenes (Billard, 2011, Billard y Diday, 2006). La elección más común de estos resúmenes es el intervalo aleatorio (o el equivalente d-dimensional: el hipercubo aleatorio) pero existen otro tipo de símbolos que incluyen histogramas aleatorios (Dias y Brito, 2015, Le-Rademacher y Billard, 2013) y variables categóricas multivaluadas (Billard y Diday, 2006).

Este enfoque es extremadamente atractivo dadas las tendencias tecnológicas actuales que requieren análisis de conjuntos de datos cada vez más grandes y complejos. Al agregar los microdatos a un número mucho menor de símbolos a nivel de grupo, los análisis de “big data” se pueden realizar de manera económica y efectiva en los dispositivos informáticos de baja gama. Más allá de la agregación de datos, las observaciones de valor distribucional pueden surgir naturalmente a través del proceso de registro de datos.

Desde sus inicios en la década de los ochenta, se han desarrollado muchas técnicas del SDA para analizar variables aleatorias con valores de distribución, incluyendo modelos de regresión (Irpino y Verde, 2015, Dias y Brito, 2015, Giordani, 2015), análisis de componentes principales (Kosmelj et al., 2014, Le-Rademacher y Billard, 2013, Ichino, 2011), series de tiempo (Lin y Gonzalez-Rivera, 2016, Wang et al., 2016, Arroyo et al., 2011), clustering (Brito et al.,

2015), análisis discriminante (Silva y Brito, 2015) y modelado jerárquico bayesiano (Lin et al., 2017). Sin embargo, si bien ha habido muchos avances en el análisis de datos simbólicos, desde una perspectiva estadística, la mayoría de las técnicas de SDA son descriptivas y no permiten la inferencia estadística sobre los parámetros del modelo, lo que impide que sus métodos realicen su potencial en el conjunto de herramientas del estadístico moderno.

Existen pocos trabajos en la literatura que emplean inferencia: la inferencia basada en la verosimilitud fue introducida por Le-Rademacher y Billard (2011), con mayor desarrollo y aplicación por Brito y Duarte Silva (2012). Los trabajos más recientes en este sentido se deben a Zhang y Sisson (2016) y B. Beranger, H. Lin and S. A. Sisson (2018).

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación, tal como se mencionan en el resumen, se enmarcan dentro de Data Science y Data Mining. Debido a que las herramientas desarrolladas en esta última sólo sirven para trabajar con matrices de datos clásicas, ha surgido en la década de 1980 el análisis de datos simbólicos que brinda una nueva forma de pensar en Data Science, al extender la entrada estándar a un conjunto de clases de entidades individuales. En muchas ocasiones tales clases son el objeto de estudio y para tener en cuenta la variabilidad entre los miembros de cada clase, las mismas se describen por datos distribucionales. De esa manera, obtenemos nuevos tipos de datos, llamados "simbólicos", ya que no se pueden reducir a números sin perder mucha información. El primer paso en SDA es construir la tabla de datos simbólicos donde las filas son clases y las

columnas son variables que pueden tomar valores simbólicos. El segundo paso es estudiar y extraer nuevos conocimientos de estos nuevos tipos de datos mediante al menos una extensión de Estadística Computacional y Data Mining a datos simbólicos.

SDA es un nuevo paradigma que abre un vasto dominio de investigación y aplicaciones al proporcionar resultados complementarios a los métodos clásicos aplicados a los datos estándar. SDA también brinda respuestas a los grandes volúmenes de datos (big data) y datos complejos, ya que los primeros se pueden reducir y resumir por clases y los datos complejos, con múltiples tablas de datos no estructurados y las variables no apareadas se pueden transformar en una tabla de datos estructurada con variables apareadas de valores simbólicos.

En este proyecto trabajamos con ambos enfoques, Data Mining y SDA para la extracción de conocimientos.

3. RESULTADOS OBTENIDOS/ESPERADOS

El presente proyecto tiene como finalidad investigar sobre metodologías del Análisis de Datos Simbólicos en el contexto de Data Science. El mismo sigue la línea de investigación del proyecto acreditado desarrollado en el período 2018-2019. En él se ha trabajado fundamentalmente con el estudio de Series Simbólicas de intervalo y de Regresión Lineal Simbólica. En el proyecto actual se ha profundizado, en el primer año de trabajo, el estudio y aplicación de técnicas del SDA referentes a Clustering, Regresión y Series Temporales. Algunas de las metodologías propuestas en regresión simbólica de intervalo se han aplicado a datos en un

contexto biométrico. Por otra parte se ha trabajado con datos de COVID-19 publicados en el sitio <https://github.com/owid/covid-19-data>.

Para estos datos se han empleado técnicas del SDA para describir los países de América respecto a características de la evolución de la pandemia y posteriormente hacer una clasificación supervisada que evidencia el posicionamiento de cada país frente a la pandemia, de acuerdo a variables tales como los valores de los casos confirmados acumulados, el nuevo aumento diario de casos confirmados y los relativos por millón de habitantes. Los resultados obtenidos se han presentado y publicado en congresos nacionales e internacionales. Además se comenzó el estudio de papers recientemente publicados que abordan el problema de la modelización e inferencia en datos de naturaleza simbólica.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado por docentes investigadores de dos facultades de la UNSJ, algunos de ellos son jóvenes investigadores. En el marco del proyecto desarrolla su segundo año de beca de iniciación a la investigación una egresada de Licenciatura en Matemática, que actualmente cursa las últimas materias en la carrera Maestría en Matemática de la Universidad Nacional de San Luis y está escribiendo su trabajo de tesis sobre Series Simbólicas de Intervalo. Entre los integrantes del proyecto hay además dos maestrandos, que aplicarán en sus trabajos de tesis las herramientas objeto de la presente investigación.

5. BIBLIOGRAFÍA

Books

- [B1] Afonso, F., Diday, E., Toque, C. (2019) "Data Science par Analyse Des Données Symboliques". Editions TECHNIP
- [B2] Arroyo, J. (2008) "Métodos de predicción para series temporales de intervalos e histogramas". Ph. D. Dissertation, Universidad Pontificia Comillas, Madrid.
- [B3] Billard, L., Diday, E. (2006). Symbolic Data Analysis: Conceptual Statistics and Data Mining. Wiley.
- [B4] Diday, E. and Noirhomme-Fraiture, M. (2008). Symbolic Data Analysis and the SODAS Software. Wiley.

Papers

- [P1] Arroyo, J., R. Espínola, and C. Maté (2008). "Diferent approaches to forecast interval time series: a comparison in finance." *Computation Statistics and Data Analysis* (submitted).
- [P2] Arroyo, J. and C. Maté (2008). "Forecasting time series of observed distributions with smoothing methods based on the barycentric histogram". In *Computational Intelligence in Decision and Control. Proceedings of the 8th International FLINS Conference*, pp. 61-66. World Scientific.
- [P3] Berenger B., Lin H., Sisson S.A. (2018) "New models for symbolic data analysis". arXiv: 1809.03659v1[stat. CO]
- [P4] Billard, L. (2011). "Brief overview of symbolic data and analytic issues".
- [P5] Brito, P., A. P. D. Silva, and J. G. Dias (2015). "Probabilistic clustering of interval data. *Intelligent Data Analysis*". 19, 293-313
- [P6] Brito, P., Duarte Silva, A. P. (2012): "Modelling Interval Data with Normal and Skew-Normal Distributions". *Journal of Applied Statistics*, Volume 39, Issue 1, 3-20.

- [P7] Brito, P. (2007): "Modelling and Analysing Interval Data". In: "Advances in Data Analysis", Decker, R., Lenz, H.-J. (Eds.), Series "Studies in Classification, Data Analysis and Knowledge Organization", Springer, Berlin, Heidelberg, New-York, 197-208.
- [P8] Giordani, P. (2015). "Lasso-constrained regression analysis for interval-valued data". *Advances in Data Analysis and Classification* 9, 5-19
- [P9] González-Rivera, G. and Arroyo, J. (2012). "Time series modeling of histogram-valued data: The daily histogram time series of S&P500 intradaily returns", *International Journal of Forecasting*, 28 (1), 20–33.
- [P10] Han, A., Hong, Y., Lai, K. and Wang, S. (2008). "Interval time series analysis with an application to the Sterling-Dollar exchange rate", *Journal of Systems Science and Complexity*, 21 (4), 558-573.
- [P11] Irpino, A. and R. Verde (2015). "Linear regression for numeric symbolic variables: a least squares approach based on Wasserstein distance". *Advances in Data Analysis and Classification* 9, 81-106.
- [P12] Kosmelj, K., J. Le-Rademacher, and L. Billard (2014). "Symbolic covariance matrix for interval-valued variables and its application to principal component analysis: a case study". *Metod. Zvezki* 11, 1-20.
- [P13] Le-Rademacher, J. and L. Billard (2011). "Likelihood functions and some maximum likelihood estimators for symbolic data". *Journal of Statistical Planning and Inference* 141, 1593-1602.
- [P14] Lin, H., M. J. Caley, and S. A. Sisson (2017). "Estimating global species richness using symbolic data meta-analysis". arXiv:1711.03202
- [P15] Lin, W. and G. González-Rivera (2016). "Interval-valued time series models: Estimation based on order statistics exploring the Agriculture Marketing Service data". *Comp. Stat. Dat. An* 100, 694-711.
- [P16] Teles, P. and Brito, P. (2015). "Modelling Interval Time Series with Space-Time processes", *Communications in Statistics: Theory and Methods*, Volume 44, Issue 17. DOI: 10.1080/03610926.2013.782200
- [P17] Wang, X., Z. Zhang, and S. Li (2016). "Set-valued and interval-valued stationary time series". *Journal of Multivariate Analysis* 145, 208-223.
- [P18] Zhang, X. and S. A. Sisson (2016). "Constructing likelihood functions for interval-valued random variables". arXiv:1608.00945 .

Software

- [S1] F. Afonso, E. Diday, and R. Haddad (2012) Latest developments of the syr software for symbolic data analysis of complex data. 3rd Workshop on Symbolic Data Analysis.
- [S2] V. Batagelj and N. Kejzar (2010) Clamix clustering symbolic objects. Computer software R program. Vienna : R Foundation for Statistical Computing. Available : <https://r-forge.r-project.org/projects/clamix>.
- [S3] E. Diday, M. Noirhomme-Fraiture, et al.(2008) Symbolic data analysis and the SODAS software. J.Wiley & Sons.
- [S4] A. Irpino and R. Verde (2015) Basic statistics for distributional symbolic variables : a new metric-based approach. *Advances in Data Analysis and Classification*, 9(2) :143–175.
- [S5] R. Oldemar, J. Murillo, and J. Villalobos (2012) An r package for symbolic data analysis. XVIII SIMMAC.

Aplicaciones de Bases de Datos Espaciales y Espacio Temporales en el ámbito agropecuario.

Edilma Olinda Gagliardi, Maria Gisela Dorzán, Maria Teresa Taranilla,

Pablo Rafael Palmero y Carlos Andrés Casanova

Departamento de Informática

Facultad de Ciencias Físico Matemáticas y Naturales

Universidad Nacional de San Luis, Argentina

{oli, mgdorzan, tarani, prpalmero, cacasanova}@unsl.edu.ar

RESUMEN

El proyecto *Tecnologías Avanzadas de Bases de Datos*, de la Universidad Nacional de San Luis, tiene como objetivo principal el estudio de bases de datos no convencionales, donde se involucra el diseño y desarrollo de herramientas para administrar eficientemente sistemas de bases de datos no estructurados. En la línea de investigación *Bases de Datos Espaciales y Espacio Temporales* se exploran dominios de aplicación de bases de datos espaciales y espacio temporales, empleando técnicas y herramientas de apoyo procedentes de diferentes disciplinas en la resolución de los problemas.

Palabras clave: Bases de Datos, Bases de Datos Espaciales y Espacio Temporales, Geometría Computacional, Metaheurísticas, AgroTIC.

CONTEXTO

El presente trabajo se enmarca en la línea de investigación *Bases de Datos Espaciales y Espacio Temporales*, perteneciente al Proyecto Consolidado *Tecnologías Avanzadas de Bases de Datos* (PROICO 03-2218) de la Universidad Nacional de San Luis, incluido en el Programa de Incentivos (Código 22/F814).

El objetivo del proyecto es el desarrollo de nuevos modelos para administrar y recuperar información almacenada en bases de datos no estructurados, que requieren modelos no tradicionales tales como las bases de datos espaciales y espacio temporales, bases de datos de imágenes, bases de datos de texto, espacios métricos, entre otros.

En particular, las aplicaciones que utilizan bases de datos espaciales y espacio temporales precisan almacenar y consultar información actual e histórica respecto de posiciones espacialmente referenciadas y cambios de forma de los objetos de estudio en diferentes escenarios a lo largo del tiempo. En este contexto, es apropiado utilizar métodos y herramientas que provienen de diferentes disciplinas en la resolución de problemas en diversos dominios de aplicación y en la búsqueda de soluciones para problemas complejos orientados a optimización. Con este fin en la línea de investigación *Bases de Datos Espaciales y Espacio Temporales* se vinculan las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas.

1. INTRODUCCIÓN

El avance en las Tecnologías de la Información y la Comunicación (TIC) ha permitido una evolución en las actividades

humanas, promoviendo cambios sociales, culturales y económicos transformándose en procesos clave para accionar en el tiempo presente y proyectar el futuro. Los beneficios aportados por las TIC en diferentes situaciones del mundo real promueven avances en tópicos que refieren a calidad, gestión, economía, entre otros. En particular, la actividad agropecuaria se perfila como un dominio de aplicación de las TIC, donde se originan nuevas experiencias que permiten innovar en prácticas de producción y formas de comunicación para optimizar la organización y gestión de establecimientos agropecuarios.

Con el objetivo general de promover un ámbito de investigación, intercambio y desarrollo de TIC, así como para estudiar su potencialidad y alcance real en el ámbito de la producción agropecuaria se articula el proyecto *Campo Conectado*. El mencionado proyecto surge de un convenio entre la Universidad Nacional de San Luis y el Instituto Nacional de Tecnología Agropecuaria (INTA-Estación Experimental Agropecuaria San Luis). En este marco se realiza un trabajo de colaboración multidisciplinar entre proyectos de investigación de UNSL, especialistas del INTA y productores agropecuarios de la región [1] [2].

En el contexto del ámbito agropecuario, para algunas de las aplicaciones de bases de datos se manifiesta la necesidad de representar y manipular tipos de datos tradicionales y también, tipos de datos complejos. En este sentido, las Bases de Datos Espaciales permiten representar y almacenar objetos junto con sus características espaciales, así como expresar consultas que involucran sus propiedades espaciales. Las Bases de Datos Temporales involucran algún aspecto de tiempo para organizar su información y almacenan datos históricos junto con los datos actuales. Las Bases de Datos Espacio Temporales concentran características de ambas y permiten observar la evolución de objetos espaciales en el tiempo [4] [8] [10]. En esta línea de trabajo se exploran dominios de aplicación para bases de datos espaciales y espacio temporales, utilizando diferentes

técnicas y herramientas en la resolución de problemas. Se plantea el diseño y aplicación de índices espacio temporales, aplicables en diversos escenarios de movimiento y se propone aplicación de técnicas de Geometría Computacional en la solución de los problemas involucrados [3].

2. LÍNEA DE INVESTIGACIÓN

Entre los tópicos de estudio de la línea de investigación se destaca:

- Diseño y aplicación de índices espaciales y espacio temporales, en diversos escenarios de movimiento.
- Desarrollo de herramientas y aplicaciones vinculadas con bases de datos espaciales y espacio temporales.
- Aplicación de la Geometría Computacional y su marco disciplinar en aspectos particulares de los problemas estudiados.
- Optimización de estructuras geométricas aplicando técnicas metaheurísticas.

Como objetivos específicos se propone:

- Estudio de la indexación espacial y espacio temporal sobre objetos en movimiento para diversos escenarios. Desarrollo de las estructuras de almacenamiento, algoritmos de consulta y evaluación experimental.
- Desarrollo de aplicaciones que utilizan bases de datos espaciales y espacio temporales, con apoyo de métodos provistos por la Geometría Computacional en la resolución de situaciones problemáticas.
- Diseño e implementación de herramientas para la visualización de información de interés vinculada a las bases de datos mencionadas.

3. RESULTADOS OBTENIDOS

En el marco del proyecto *Campo Conectado*, desde la línea de investigación se proponen algunos desarrollos tecnológicos aplicados en

la gestión de la producción agropecuaria en sistemas reales de producción.

Se trabaja en el desarrollo de una plataforma soporte de eventos y sistemas de información orientada al seguimiento espacio temporal de individuos y la gestión espacial y tradicional de los elementos que integran un establecimiento agropecuario. Entre las características principales de la plataforma se menciona la posibilidad de la captura de datos desde diferentes fuentes, la gestión de su almacenamiento, explotación y visualización de información relevante con capacidad para la generación de reportes, estadísticas, gráficas etc. La plataforma está disponible en la web con accesibilidad mediante tecnología móvil y de escritorio [5].

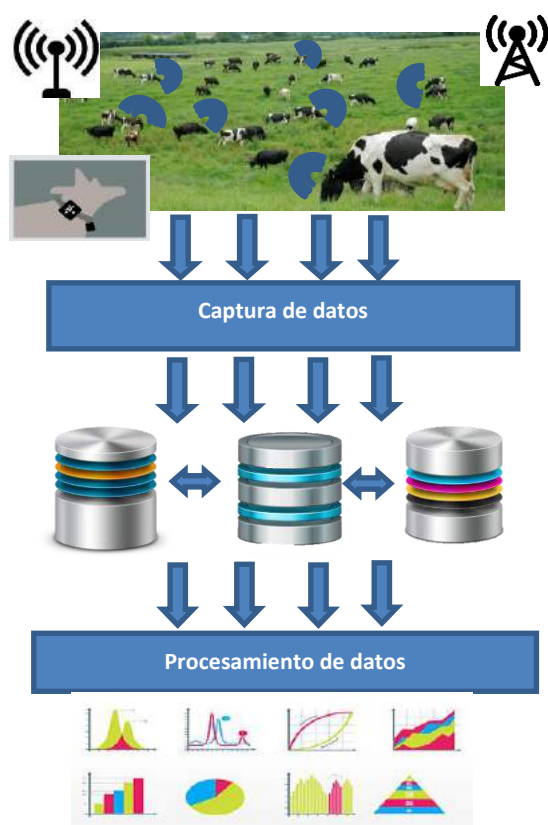


Figura 1.

En el diagrama de la Figura 1 se muestra el proceso de obtención de la información que comienza con la captura de datos desde sensores dispuestos en el campo. Luego, el proceso continúa con el envío de los datos

recolectados a repositorios de datos donde se almacenan y se analizan para generar información, en forma de indicadores y recomendaciones, que se disponen en forma inmediata para que el productor y personal autorizado pueda observarlos desde dispositivos móviles o desde sistemas de escritorio.

En la producción ganadera, algunos de los requerimientos refieren al seguimiento de indicadores productivos y ambientales, la trazabilidad animal, caracterización y seguimiento de un rodeo, control de alimentación, detección de la actividad de celo, monitoreo de la condición corporal, entre otros, como herramientas en la ganadería de precisión. En la plataforma integral se incluyen una herramienta para el seguimiento espacial y temporal de rodeos ganaderos y una herramienta de apoyo para la gestión y administración de rodeos de cría. Las mismas permiten acoplar la sensorización instalada en un establecimiento productivo y los individuos de un rodeo para garantizar la recolección de los datos desde fuentes varias, la gestión del almacenamiento de los datos y la obtención información relevante en el proceso de toma de decisiones.

La herramienta para el seguimiento espacio temporal de rodeos manipula las componentes espaciales y tradicionales de los datos, implementa consultas espaciales y espacio temporales sobre los objetos de tratamiento, proporciona comunicación con las bases de datos y visualización del geoseguimiento del rodeo accesible por medio de tecnología móvil [6] [9].

La herramienta para la gestión de rodeos de cría incluye funcionalidades para el análisis de datos relevados, seguimiento de indicadores como la condición corporal de los individuos del rodeo y la obtención de conocimiento de apoyo con su correspondiente visualización disponible con tecnologías móviles [7].

Como visión de futuro, se continuará trabajando en la incorporación de nuevas funcionalidades, en la detección y estudio de nuevos indicadores de gestión que aporten información de interés y en el análisis de

resultados obtenidos. El objetivo es proseguir con actividades que involucran la pertinencia y transferencia de proyectos I+D+i en las áreas de intervención seleccionadas, afianzando el ámbito interdisciplinario de trabajo y el fortalecimiento de competencias.

4. FORMACIÓN DE RECURSOS HUMANOS

La formación de recursos humanos se refleja en tesis doctorales, tesis de maestría y trabajos finales de Licenciatura en Ciencias de la Computación. Los integrantes de la línea realizan actividades de continua formación académica que incluyen el dictado y realización de cursos de posgrado y de especialización, actividades de divulgación científica, publicaciones en congresos entre otras actividades académico-científicas. Por otra parte, el grupo de trabajo fortalece su formación con tareas de cooperación mutua e intercambio recíproco de información científica, tecnología y desarrollo de nuevos conocimientos con investigadores locales y de otras universidades.

5. BIBLIOGRAFÍA

- [1] *Agrotic: Revolución Digital, Desarrollo y Producción*.
<https://inta.gob.ar/noticias/agrotic-revolucion-digital-desarrollo-y-produccion> (2016)
- [2] *Campo Conectado*, proyecto interinstitucional
<http://inta.gob.ar/noticias/campo-conectado-un-nuevo-proyecto-interinstitucional> (2017)
- [3] de Berg, M., Cheong O., van Kreveld, M., Overmars, M. : *Computational Geometry: Algorithms and Applications.*, Springer-Verlag, Heidelberg (2008)
- [4] Elmasri, R. Navathe, S.B.: *Fundamentals of Database Systems*. 7^{ma} Edition. Pearson (2015)
- [5] Gagliardi, E., Dorzán, M. G., Taranilla, M. T., Palmero, P., Casanova, C.: *Propuesta de plataforma para la integración de TIC orientadas al Agro* Congreso de AgroInformática CAI 2017, 46JAIIO (2017)
- [6] 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, 47JAIIO (2018).
- [7] Gagliardi, E.; Dorzán, M.G.; Taranilla, M.T.; Palmero, P.; Casanova, C.: *Diseño de una herramienta de apoyo en la gestión de rodeos de cría*. Congreso de AgroInformática CAI 2018, 47JAIIO (2018).
- [8] Manolopoulos, Y., Papadopoulos, A., Vassilakopoulos, M.: *Spatial Databases: Technologies, Techniques and Trends*. Idea Group (2005)
- [9] Palmero, P.; Gagliardi, E.; Dorzán, M.G.; *Seguimiento de rodeos en establecimientos agropecuarios*. CACIC: XXVI Congreso Argentino de Ciencias de la Computación (2020)
- [10] Shekhar, S.; Chawla, S.: *Spatial databases: a tour*. Prentice Hall (2003)

Aplicaciones de Procesamiento de Lenguaje Natural y Ciencia de Datos

Anabella De Battista, Soledad Retamar, Patricia Cristaldo, Esteban Schab, Carlos Casanova, Lautaro Ramos, Ramiro Rivera, Cristhian Richard, Lucas La Pietra, Lautaro Retamar, Franco Miret, Rodrigo Mignola, Juan Manuel Franzante

Grupo de Investigación en Bases de Datos, Departamento Ingeniería en Sistemas de Información,
Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional,
Entre Ríos, Argentina

{debattistaa, retamars, schabe, ramosl, riverar, richardc, cristaldop, lapietral, retamarl, miretf, mignolar, franzantejm}@frcu.utn.edu.ar

Norma Edith Herrera

Departamento de Informática, Universidad Nacional de San Luis, San Luis, Argentina
{nherrera}@unsl.edu.ar

Resumen

Cada día se generan grandes cantidades de datos de diversos tipos y vinculadas a diversas áreas de conocimiento (textos, imágenes, audios, videos, entre otros). La posibilidad de analizar estos enormes volúmenes de datos permite generar conocimiento que sirva como fundamento para la toma de decisiones en ámbitos tan diversos como la salud, la administración de distintos recursos, el deporte, las actividades turísticas, entre otros. En este contexto resultan de gran utilidad los modelos desarrollados en áreas de conocimiento como la Ciencia de Datos, el Aprendizaje Automático, la Minería de Textos por citar algunas. En el ámbito de análisis automático de textos (para aplicaciones de análisis de sentimientos, minería de opinión, entre otras) la mayoría de los algoritmos, herramientas y recursos disponibles han sido desarrollados para el idioma inglés por lo que no pueden aplicarse a textos escritos en idioma español.

En este artículo se presentan soluciones desarrolladas en el proyecto *Descubrimiento de Conocimiento en Bases de Datos*, basadas en aplicación de técnicas de Procesamiento de Lenguaje Natural y Ciencia de Datos. En particular, se presentan aplicaciones web para seguimiento de la evolución de casos de Dengue y COVID-19 en Argentina, otra aplicación que presenta una aplicación de Topic Modelling a noticias en lenguaje español, otro caso de aplicación de Análisis de

Sentimientos a entidades vinculadas al turismo y finalmente un caso de generación de estadísticas avanzadas a partir del procesamiento de conjuntos de datos históricos de básquet.

Palabras clave: ciencia de datos, minería de textos, bases de datos, descubrimiento de conocimiento, idioma español.

Contexto

Las aplicaciones que se presentan en este trabajo se desarrollan en el ámbito del proyecto de investigación *Descubrimiento de conocimiento en Bases de Datos (Código 5109)* 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

En los últimos años han ocurrido importantes avances en lo que respecta al desarrollo de soluciones basadas en aplicaciones de Ciencia de Datos. Las organizaciones en general han adoptado este tipo de herramientas para mejorar sus procesos de toma de decisiones y disminuir la incertidumbre en base al descubrimiento de patrones o reglas que, de otro modo, permanecerían ocultas en sus bases de datos. El Descubrimiento de Conocimiento en Bases de Datos consiste en el proceso de identificar

patrones válidos, novedosos y potencialmente útiles y comprensibles a partir de los datos, e involucra áreas de conocimiento como inteligencia artificial, aprendizaje automático, estadística, sistemas de gestión de base de datos, técnicas de visualización de datos y medios que apoyan toma de decisiones.

La Ciencia de Datos es un área de estudio multidisciplinar que integra campos de conocimiento variados como estadística, aprendizaje automático y análisis de datos, y que se puede aplicar en muchos dominios de conocimiento [1].

La Minería de Textos tiene como objetivo la extracción de patrones relevantes a partir de un gran conjunto de textos con el propósito de obtener conocimiento. Abarca una serie de tareas que se enfocan en distintos aspectos del análisis de textos. Algunas de las más relevantes:

- Recuperación de información (*Information Retrieval, IR*): es la tarea de encontrar material de naturaleza no estructurada (generalmente textos) proveniente de grandes colecciones que satisfagan determinadas necesidades de información [2]. Una tarea crucial para un sistema de IR es indexar la colección de documentos para hacer que sus contenidos sean accesibles de manera eficiente. Generalmente la indexación se realiza sobre una representación lógica del documento, que puede consistir en un conjunto de palabras clave o términos relevantes que aparezcan en el texto [3].

- Procesamiento del Lenguaje Natural: es un campo de las ciencias de la computación que combina Inteligencia Artificial y conceptos lingüísticos con el fin de hacer que oraciones o palabras escritas en lenguaje natural puedan ser interpretados por programas de computadoras [1, 4].

- Extracción de Información (*Information Extraction, IE*): es una subdisciplina de la Inteligencia Artificial que se aboca a la identificación, y consecuente clasificación y estructuración en grupos semánticos, de información específica que se encuentran en fuentes de datos no estructurados, como el texto en lenguaje natural, lo que hace que la información sea más adecuada para las tareas

de procesamiento de la información [5].

- Resumen de textos (*Text Summarization*): es la tarea de producir un resumen conciso y fluido preservando el contenido clave de la información y el significado general de una colección de textos [6].

- Métodos de Aprendizaje Supervisado y No Supervisado: los métodos de aprendizaje supervisado son técnicas de aprendizaje automático relacionadas con entrenar un modelo, por ejemplo, de clasificación, utilizando un conjunto de datos de entrenamiento para realizar predicciones sobre datos desconocidos de antemano. Existe una amplia gama de métodos supervisados, como clasificadores de vecinos más cercanos, árboles de decisión, clasificadores basados en reglas y clasificadores probabilísticos.

Los trabajos sobre minería de textos en español que se presentan en la actualidad se enfocan principalmente en Análisis de Sentimientos o Minería de Opinión, en los cuales se evidencian dos enfoques: uno basado en el empleo de lexemas, y otro en técnicas de *Machine Learning*, para identificar los sentimientos expresados en los textos. En la gran mayoría de estos trabajos se utilizan recursos traducidos de forma automática generados para otros idiomas, como el inglés, lo cual manifiesta una escasez de recursos genuinos para el lenguaje español.

Existen también trabajos sobre perfilado de autor, en los que se menciona la dificultad de encontrar colecciones de textos adecuadamente etiquetados y con poco ruido [7]. A partir de eso se han producido trabajos tendientes al desarrollo de conjuntos de textos en español específicos para esta tarea [8,9].

Las técnicas de minería de datos de texto se han propuesto para tareas de estudios bibliográficos, simplificación de textos y etiquetado semántico. Se ha propuesto la aplicación de algoritmos de clasificación [10] para identificar automáticamente el dominio disciplinar de un nuevo texto académico en un repositorio bibliográfico mediante la construcción de lexemas compartidos en cada disciplina.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de trabajo principal de nuestro proyecto de investigación es el estudio, análisis y comparación de técnicas de minería de datos para el tratamiento de textos en español, el análisis de su desempeño en distintos escenarios y la propuesta de modificaciones o mejoras a las técnicas de minería de textos existentes para incrementar la calidad de los resultados en el tratamiento de textos en español. También está previsto realizar una evaluación del funcionamiento de técnicas de minería de datos sobre conjuntos "tradicionales" enriquecidos con atributos provenientes de textos relacionados. Durante el año 2020 se realizaron también trabajo en el ámbito de Analítica de Datos, aplicados al seguimiento de la pandemia de COVID-19 y a la epidemia de Dengue, lo que incrementó la visibilidad del trabajo realizado por el grupo de investigación y permitió concretar acuerdos para trabajar, por ejemplo, en analítica de datos aplicada al deporte.

2.1. Topic Modelling

El término Agenda Setting hace referencia a la influencia que tienen los medios de comunicación en la fijación de temas en la opinión pública [14]. Se desarrolló una aplicación web que despliega la aplicación de Topic Modeling a un corpus de texto en español, generado a partir de la recolección de noticias sobre COVID-19 del mes de mayo de 2020 de los periódicos digitales argentinos Clarín, Infobae y La Nación. El objetivo propuesto fue investigar los patrones subyacentes en los documentos y determinar las temáticas principales de las noticias publicadas por los medios durante esa etapa de la pandemia. Se realizó una comparativa de dos herramientas que implementan el algoritmo LDA y se realizaron las pruebas con la herramienta que obtuvo mejores resultados. Se desarrolló además una aplicación web en la que se presentan los hallazgos del proyecto a través de diferentes visualizaciones. Para la publicación de resultados se desarrolló una aplicación web accesible en la URL

<https://nlp-noticiascorona.herokuapp.com> y que permite visualizar distintos aspectos resultantes del procesamiento del conjunto de noticias como la cantidad de noticias clasificadas recolectadas por cada diario, la cantidad de noticias clasificadas en cada tópico, los bi-gramas y n-gramas de palabras que aparecen con más frecuencia en las noticias.

2.2. Análisis de Sentimientos

Mediante un convenio con la Dirección de Producción del Municipio de Concepción del Uruguay, se trabaja en un estudio consistente en la aplicación de técnicas de minería de textos sobre las opiniones que turistas que visitan la ciudad vierten en la plataforma digital de turismo TripAdvisor (www.tripadvisor.com.ar). Para realizar el análisis se implementó una tarea de categorización de textos conocida como Análisis de Sentimientos que consiste en la determinación de la orientación positiva o negativa que un escritor expresa hacia algún objeto [15]. Estas técnicas pueden aplicarse a revisiones de películas, libros, productos o servicios en la web, textos editoriales o políticos, entre otros. El resultado de su aplicación permite identificar el sentimiento hacia el objeto o sujeto bajo análisis, lo que las convierte en aplicaciones de gran relevancia en ámbitos como experiencia de usuarios, marketing o política. La aplicación de Análisis de Sentimientos a comentarios de turistas que han visitado hoteles, restaurantes y lugares turísticos de la ciudad de Concepción del Uruguay, permite conocer la percepción sobre los lugares visitados y obtener información detallada que sirva como soporte para la toma de decisiones en aspectos relacionados con el turismo como el diseño de estrategias de marketing orientadas según el tipo de público que visita la ciudad, acciones de fortalecimiento orientadas a determinados rubros, entre otras.

2.3. Visualización de datos

La generación y el almacenamiento de grandes volúmenes de información hacen que el mismo pase desapercibido y muchas veces se pierda la oportunidad de encontrar valor en ella. La visualización de datos es el proceso de representación de datos, en formato gráfico, de una manera clara y eficaz. Se convierte en una herramienta poderosa para el análisis e interpretación de datos grandes y complejos, volviéndose un medio eficiente en la transmisión de conceptos en un formato universal [16, 17].

Se desarrollaron dos aplicaciones que, mediante la aplicación de técnicas de analítica de datos, permiten visualizar la evolución de la pandemia de COVID-19 en Sudamérica, y con mayor granularidad en Argentina y sus provincias, y otra aplicación que permite comparar la evolución de la epidemia de Dengue en Argentina en los últimos años.

La aplicación para seguimiento de COVID-19 se desarrolló con el objetivo de aportar información sobre la evolución de la pandemia. Se seleccionaron fuentes de datos oficiales y mediante la aplicación de técnicas de analítica de datos y visualización de información, se presentan distintos análisis que permiten comprender la evolución de casos. La aplicación web está accesible en la URL <https://gibd.github.io/covid> [18].

En el caso de la aplicación de seguimiento de Dengue se procedió inicialmente a la identificación de fuentes oficiales de información, accediendo a sitios web de los Ministerios de Salud de las distintas provincias argentinas. El principal inconveniente encontrado fue que no todas las provincias publicaban información sobre esta epidemia, por lo que no era posible generar un conjunto de datos homogéneo. Se realizó un análisis de los datos publicados por el Ministerio de Salud, específicamente en el Boletín Integrado de Vigilancia (BIV), y se decidió tomar esa información para generar un conjunto de datos apropiado para aplicar análisis estadístico. Para realizar el análisis de la evolución de casos en Argentina en los últimos años se trabajó con la información publicada en el sitio

web de la Organización Panamericana de la Salud. Posteriormente se trabajó en la generación de datasets a partir de la información recolectada de las fuentes antes mencionadas se generaron dos conjuntos de datos. Finalmente se realizó el análisis exploratorio de datos: en esta etapa se trabajó en la generación de visualizaciones empleando la librería Plotly de Python. Se adoptaron distintos enfoques para el análisis desde el punto de vista geográfico, generando gráficas a nivel país, a nivel provincia y a nivel región. Para el análisis a nivel país se realizó una comparativa de cantidad de casos y fallecimientos para el período 2014-2020. Para el análisis de la evolución de casos durante el año 2020 se realizaron visualizaciones de: Casos autóctonos, Casos importados, Casos en Investigación, Notificados y Fallecimientos, a nivel Provincia, Región y País. Además, se generó un mapa con la Cantidad Acumulada de casos por provincia hasta la última Semana Epidemiológica para la cual se publicaron datos. La aplicación web está accesible en la URL <http://dengueibd.herokuapp.com> [19].

2.4. Aplicación de metodología para la gestión de proyectos de Minería de Datos

En la gestión de las actividades de cada una de las líneas de investigación y desarrollo del proyecto se emplean fundamentos de metodologías ágiles. Partiendo de la propuesta metodológica de CRISP-DM [25] se realizó una adaptación empleando dichos fundamentos ágiles. Se formalizó dicha adaptación como una propuesta de metodología ágil para la gestión de proyectos de ciencia de datos [26, 27]. Actualmente se trabaja en definir y desarrollar métricas e indicadores para la evaluación transversal de metodologías de gestión de proyectos, a partir de expresiones textuales de la descripción de un proyecto. Se busca especificar un marco de medición que permita, a través de las métricas consignadas, evaluar las metodologías, obteniendo un resultado métrico que exprese el grado de aplicabilidad, eficacia y eficiencia en

diferentes contextos. Por otro lado, se busca medir la eficiencia de las metodologías en gestión de proyectos a partir de la descripción del proyecto en base a los hallazgos estadísticos, y aprendizaje automático con las métricas e indicadores propuestos.

3. Resultados obtenidos y esperados

Con este proyecto se espera lograr aplicaciones novedosas de técnicas y herramientas de minería de textos para textos en español, en particular en áreas de estudio como bibliometría y la teoría de establecimiento de agenda. Además, se continúa trabajando en analítica de datos aplicada al básquet. Estas iniciativas se desarrollan mediante la aplicación de la metodología ágil para proyectos de ciencias de datos propuesta.

4. Formación de Recursos Humanos

En el marco del proyecto se está desarrollando una tesis de maestría y unade las investigadoras está cursando el Doctorado en Informática. Se cuenta con un becario graduado con beca de iniciación a la investigación y dos becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación. Está prevista la realización de al menos una Práctica Supervisada de Ingeniería en Sistemas de Información en el marco del proyecto.

5. Referencias

[1] Wang, R., Hu, G., Jiang, C., Lu, H., & Zhang, Y. (2020). Data Analytics for the COVID-19 Epidemic. 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), 1261–1266. <https://doi.org/10.1109/COMPSAC48688.2020.00-83>

[2] Stefan Büttcher, Charles L.A. Clarke, Gordon V. Cormack. Information Retrieval: Implementing and Evaluating Search Engines. Published July 23rd 2010 by MIT Press (MA)

[3] Allahyari, Mehdi et al. A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. 2017. arXiv:1707.02919v2

[4] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to information retrieval. Cambridge University Press, 2008.

[5] S. Ceri, A. Bozzon, M. Brambilla, E. Della Valle, P. Fraternali, and S. Quarteroni, The Information Retrieval Process. In Web Information Retrieval, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 13–26.

[6] D. Khurana, A. Koli, K. Khatter, and S. Singh, “Natural Language Processing: State of The Art, Current Trends and Challenges,” Aug. 2017.

[7] Moens, Marie-Francine. Information Extraction: Algorithms and Prospects in a Retrieval Context. Springer Netherlands, 2006.

[8] M. Allahyari et al. Text Summarization Techniques: A Brief Survey. Jul. 2017.

[9] M. J. Garciarena Ucelay, M. P. Villegas, L. Cagnina, and M. L. Errecalde. Cross domain author profiling task in spanish language: an experimental study. *JCS&T*, vol. 15, no. 02, pp. p. 122-128. Nov. 2015.

[10] M. P. Villegas, M. J. Garciarena Ucelay, M. L. Errecalde, and L. Cagnina. A Spanish text corpus for the author profiling task. 2014.

[11] M. J. Garciarena Ucelay and M. P. Villegas. Determinación del Perfil del Autor de Documentos en español. Universidad Nacional de San Luis, 2015.

[12] R. Venegas. Clasificación de textos académicos en función de su contenido léxico-semántico. *Rev. signos*, vol. 40, no. 63, pp. 239–271, 2007.

[13] SCOPUS. <http://www.scopus.com> Accedido 02/2021.

[14] Scopus API. <https://goo.gl/mqpFpA> Accedido 03/2021.

[15] Jurafsky, Daniel and Martin, James H. Speech and Language Processing. 2019. <https://web.stanford.edu/~jurafsky>

[16] M. McCombs and D. Shaw. The agenda-setting function of mass media. Public opinion

quarterly, 36(2):176–187, 1972.

[17] Yeoul Kim, Suin Kim, Alejandro Jaimes, and Alice Oh. A computational analysis of agenda setting. In Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion), 323-324, 2014.

[18] Grupo de Investigación en Base de Datos. Aplicación para seguimiento de COVID-19. <https://gibd.github.io/covid>. Accedido en 03/2021

[19] Grupo de Investigación en Base de Datos. Aplicación para seguimiento de Dengue. <http://denguegibd.herokuapp.com> Accedido en 03/2021

[25] Ken Schwaber and Jeff Sutherland. The scrum guide. Scrum Alliance, 2011, vol. 21.

[26] Cristaldo, Richard, Rivera, Schab, Anabella De Battista. Propuesta Metodológica de Enfoque “Híbrido” para la Gestión de Proyectos de Minería de Datos. SABTIC 2018. ISSN 2237-2970

[27] Cristaldo, Schab, Richard, Rivera, Anabella De Battista, Retamar, Herrera. Adecuación de una Propuesta Metodológica de Enfoque “Híbrido” para la Gestión de Proyectos de Ciencia de Datos. 6to CoNaIISI. 29 y 30 de noviembre de 2018 – Universidad CAECE – Mar del Plata, Bs. As., Argentina.

Bases de Datos no Convencionales: Índices y Operaciones

M.D. Alba, J. Arroyuelo, Maria E. Di Genaro, A. Grosso, M. Jofré, V. Ludueña, N. Reyes
Dpto. de Informática, Fac. de Cs. Físico-Matemáticas y Naturales, Universidad Nacional de San Luis
{bjarroyu, digeme, agrosso, vlud, nreyes}@unsl.edu.ar, {mdaniela.alba, monicajofre}@gmail.com

Edgar Chávez

Centro de Investigación Científica y de Educación Superior de Ensenada, México

elchavez@cicese.mx

Karina Figueroa

Fac. de Cs. Físico-Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, México

karina@fisimat.umich.mx

Rodrigo Paredes

Dpto. de Cs. de la Computación, Fac. de Ingeniería, Universidad de Talca, Chile

raparede@utalca.cl

Resumen

Debido a que en la actualidad se generan gran cantidad de datos digitales, desde fuentes muy disímiles, los repositorios especializados en datos no estructurados se vuelven cada vez más necesarios. Por este motivo, los mismos deben adaptarse rápidamente, para administrar de manera eficiente el gran volumen de datos generados, al igual que el tipo de requerimientos al que son sometidos los mismos; éstos pueden ser tan dispares como los tipos de datos que puede ser necesario administrar, dado que pertenecen a campos muy diferentes.

Para ello, se investigan distintos aspectos relacionados con este tipo de bases de datos, como la administración del espacio disponible, que se vuelve crucial debido a la gran cantidad de datos que se debe manipular; formas más sofisticadas de búsqueda sobre las mismas, que permitan enfrentar tales requerimientos; optimización de estos depósitos, o desarrollo de nuevos, considerando incluso la arquitectura del procesador.

Un modelo de base de datos no convencionales que se adapta a tales requerimientos, en el cual se pueden utilizar métodos de acceso que contemplen estos aspectos, son las *Bases de Datos Métricas*. Esta investigación pretende contribuir a la madurez de este nuevo modelo de bases de datos considerando distintas perspectivas.

Palabras Claves: bases de datos métricas, índices, búsquedas por similitud.

Contexto

Esta investigación está enfocada en lograr la consolidación del modelo de Bases de Datos Métricas para soportar distintos tipos de bases de datos no convencionales. Se espera contribuir a este modelo obteniendo índices dinámicos, con E/S eficiente,

capaces de manejar grandes volúmenes de datos (escalables) y que resulten más eficientes para memorias jerárquicas. Además, con el fin de obtener mejoras en los administradores de estas bases de datos a un muy bajo nivel, se analizan nuevas arquitecturas del procesador, buscando así contribuir en diferentes campos de aplicación: sistemas de información geográfica, robótica, visión artificial, diseño asistido por computadora, computación móvil, entre otros.

El presente trabajo se realiza en el marco del Proyecto Consolidado *Tecnologías Avanzadas de Bases de Datos*, dentro de la línea *Bases de Datos no Convencionales*, (Cód. 03-2218 y en Programa de Incentivos 22-F814) de la Universidad Nacional de San Luis. Colaboran investigadores de otros grupos de: Universidad de Talca (Chile), Universidad Michoacana de San Nicolás de Hidalgo y Centro de Investigación Científica y de Educación Superior de Ensenada (México).

Introducción

El uso masivo de las redes y la disponibilidad de dispositivos electrónicos en diversos ámbitos, como el laboral, educativo, productivo, de la salud, recreativo, científico, etc., ha provocado la aparición de una significativa cantidad y variedad de datos, de diferentes tipos y procedencia. De esta manera, el volumen de datos generados y almacenados, ha crecido de manera exponencial.

Como respuesta a este fenómeno, las bases de datos han debido adaptarse rápidamente, tanto a la cantidad de datos que deben administrar, como a su va-

riedad y disimilitud. El tipo de requerimientos sobre estos datos también es diferente; por ejemplo, para encontrar las huellas digitales más similares a una dada, las búsquedas tradicionales (exactas) carecen de sentido, en cambio las *búsquedas por similitud* resultan más apropiadas.

Un modelo que resulta adecuado, por englobar muchas de las características compartidas por estas aplicaciones tan diversas, es el de *espacios métricos*. Este modelo es determinado por un universo de objetos y una función de distancia definida entre ellos, que mide cuán diferentes son estos objetos. Sin embargo, para responder eficientemente a los requerimientos que se realizan sobre los objetos no estructurados almacenados en estas bases de datos, son necesarios los llamados *Métodos de Acceso Métricos* (MAMs), que permiten resolver este tipo de búsquedas, sin realizar una examinación secuencial del conjunto de datos. La diversidad de ámbitos en los que se aplica el modelo, vuelve esencial la actualización y optimización de los MAMs, logrando su adaptación a cada caso particular y afrontando retos como el soporte de conjuntos masivos de datos, el permitir actualizaciones (inserciones/eliminaciones) y la resolución de búsquedas más complejas.

Por otro lado, pensando en mejorar el desempeño de los administradores de bases de datos (DBMS), también a bajo nivel, se busca caracterizar nuevas arquitecturas que permitan reducir el flujo de bits entre el procesador y la memoria en relación a la cantidad de datos utilizados por cada programa.

Estos avances se reflejan en áreas como: reconocimiento de voz, reconocimiento facial, bases de datos médicas, minería de datos, biología computacional, entre otros.

Líneas de Investigación y Desarrollo

Bases de Datos Métricas

Como se mencionó, aquellas bases de datos capaces de administrar secuencias de ADN o de proteínas, imágenes, videos, texto libre, audio, etc., comúnmente llamadas *bases de datos no convencionales*, serán modelizadas utilizando espacios métricos. En este modelo, debido a lo costoso que suele resultar calcular la distancia entre dos objetos, el número de cálculos realizados durante alguna operación se usa habitualmente como medida de complejidad. Calcular distancias es necesario ya sea para crear un índice o para realizar búsquedas sobre él.

Así, para evitar comparaciones exhaustivas con la base de datos durante una consulta por similitud y

responder más eficientemente a las mismas, se hará uso de MAMs. Por ello, el objetivo es optimizar los MAMs, analizando aquellos que han mostrado buen desempeño en las búsquedas para reducir su complejidad considerando, cuando sea necesario, la jerarquía de memorias y/o priorizando además su dinamismo y escalabilidad, cuando sea posible. En general, dado un espacio métrico cuyo universo de objetos es \mathbb{U} y una función de distancia $d : \mathbb{U} \times \mathbb{U} \mapsto \mathbb{R}^+$, una base de datos $X \subseteq \mathbb{U}$ y una consulta $q \in \mathbb{U}$, las consultas por similitud son de dos tipos: *por rango* (q, r) o de *k-vecinos más cercanos* (k -NN).

Métodos de Acceso Métricos

En el contexto del diseño y adaptación de los MAM's para ser útiles en un mayor espectro de aplicaciones reales, se deben considerar necesariamente características tales como soportar dinamismo y en muchos casos adaptarse adecuadamente a ser almacenados en memoria secundaria, optimizando en ese caso no sólo el número de cálculos de distancia que se realizan tanto en la construcción del MAM como en las búsquedas, sino también el número de operaciones de E/S necesarias.

Considerando la necesidad de dinamismo en los MAM's, es que surge el desarrollo del *Árbol de Aproximación Espacial Dinámico* (DSAT) [11], basado en el *Árbol de Aproximación Espacial* (SAT), que a pesar de ser uno de los índices de mejor desempeño en espacios de mediana a alta dimensión, era totalmente estático. El DSAT permite realizar inserciones y eliminaciones, manteniendo el buen desempeño en las búsquedas, aunque agrega un parámetro a sintonizar. Una variante también estática del SAT, es el *Árbol de Aproximación Espacial Distal* (DiSAT) [6], éste logra optimizar las búsquedas respecto de ambos (SAT y DSAT) y además no necesita ningún parámetro. Por ello, se ha propuesto una optimización del mismo, la *Foresta de Aproximación Espacial Distal* (DiSAF) [3], que es dinámica, para memoria principal y que para lograr mejorar al máximo su desempeño, aplica la técnica de dinamización de Bentley y Saxe al DiSAT y aprovecha el profundo conocimiento que se tiene sobre la aproximación espacial. Sin embargo, los costos de construcción son altos debido a la necesidad de reconstruir subárboles luego de cada inserción. Por ello, se está desarrollando una versión que utiliza inserción perezosa para amortizar los costos de reconstrucción, mientras mantiene buen desempeño en las búsquedas.

Cada vez es más habitual la necesidad de diseñar

índices que se almacenen en memoria secundaria. Esto se debe a que la cantidad masiva de datos a almacenar; por ejemplo, los provenientes de redes sociales, o el tamaño de algunos de los tipos de datos almacenados, como las imágenes satelitales, provocan que los índices no quepan en memoria principal. Por ello, se está trabajando en lograr una versión del *DiSAT* dinámica para memoria secundaria, que además de amortizar los costos de reconstrucción entre varias inserciones y mantener un buen desempeño en las búsquedas, se adapte a memoria secundaria realizando un buen uso de las páginas de disco para minimizar el número de operaciones de E/S. Para ello, se debe considerar no sólo que logre un desempeño comparable al de la versión de memoria principal en cantidad de cálculos de distancia, sino que las páginas en disco mantengan una buena ocupación, que la cantidad de operaciones de E/S necesarias sea baja y que se mantenga la localidad en los accesos.

Cabe aclarar que muchos de estos índices se basan en “agrupar elementos” en conjuntos de objetos similares. Por ello, es posible utilizar estrategias de optimización para analizar cuán buenos son los agrupamientos que logran, y tratar de mejorarlos. Por ello, se está buscando optimizar una versión del *DSAT* para memoria secundaria que considere además de la propia agrupación de elementos que realiza el índice, la posibilidad de adicionar en cada nodo un clúster de elementos cercanos.

Una faceta tan importante como las anteriores, involucra los requerimientos de aplicaciones que priorizan la rapidez en las respuestas, a costa de perder algunos elementos: se intercambia precisión (devolviendo sólo algunos objetos relevantes) por velocidad en la respuesta. Este tipo de búsquedas se denominan *aproximadas*. Para conjuntos de datos masivos, las búsquedas por similitud aproximadas permiten obtener un buen balance entre el costo de las búsquedas y la calidad de la respuesta obtenida. Uno de los mejores representantes de este tipo de consultas es el *Algoritmo Basado en Permutaciones (PBA)* [2], que logra una respuesta de alta calidad a un bajo costo. Por esta razón, se ha utilizado como base del diseño de la *Lista Dinámica de Permutaciones Agrupadas (DLCP)* [9], que además de ser dinámica es para memoria secundaria. Este índice, que combina *LC* con *PBA*, agrupa por distancia entre las *permutaciones* de los objetos, en lugar de hacerlo por distancia entre objetos, y además se le puede indicar cuántos cálculos de distancia y/o operaciones de

E/S utilizar, para obtener una respuesta rápida, aunque menos precisa.

Grafo de los k Vecinos

Encontrar los restaurantes más cercanos a un lugar en particular, es una búsqueda que puede resolverse en espacios métricos mediante una consulta k -NN. Una variación de esta consulta por similitud, la *All- k -NN*, resulta tan útil como la anterior. La misma permite obtener los k -vecinos más cercanos de *todos* los elementos de la base de datos; es decir, relaciona cada elemento $u \in X$, con los k objetos en $X - \{u\}$ que tengan la menor distancia a él. La solución ingenua tiene una complejidad de n^2 cálculos de distancia ($|X| = n$), y consiste en comparar cada objeto en la base de datos con todos los demás, devolviendo los k más cercanos a él. Evidentemente, resulta mucho más eficiente preprocesar la base de datos construyendo un índice, y luego buscar en él los k -NN de cada elemento del conjunto.

No obstante, al enfrentar una base de datos masiva, o cuando la función de distancia utilizada es demasiado costosa de calcular, o aún si se trabaja con espacios métricos de alta dimensión, esta solución puede resultar tan ineficiente como la ingenua, ya que estas situaciones pueden requerir revisar la base de datos completa, sin importar la estrategia implementada. Además, es necesario considerar los requerimientos de algunas aplicaciones particulares, que priorizan la velocidad de respuesta sobre la precisión de la misma [13, 7, 14, 8]. Atendiendo a todas estas circunstancias es que se han considerado las llamadas *búsquedas por similitud aproximadas*, las cuales mejoran su complejidad aceptando algunos “errores” en la respuesta.

Una estructura que permite indexar un espacio métrico, y resolver las consultas por similitud es el *Grafo de los k -vecinos más cercanos (k NNG)*[12]. Este grafo asocia cada elemento de la base de datos a sus k vecinos más cercanos y resulta ser un índice eficiente, con respecto a algunas de las técnicas clásicas. Dado que la solución al problema *All- k -NN* permite construir el k NNG, se han propuesto nuevas técnicas, que resuelven el problema de *All- k -NN*, y permiten computar una aproximación del k NNG. Éstas, *sin utilizar ningún índice* para buscar en él, conectan cada objeto u de la base de datos con k vecinos *cercanos*, relajando la condición que exige que no haya, en toda la base de datos, algún objeto más cercano a u que los k vecinos devueltos. Aunque estas técnicas pueden perder algún objeto muy cercano a u y en su lugar devolver otro un poco más

lejano obtienen, a cambio, una respuesta más rápida. Al grafo construido de esta manera se lo denominó *Grafo de vecinos cercanos* ($knNG$) [5].

Una primera aproximación, aprovecha el profundo conocimiento que se tiene del *DiSAT* y plantea un enfoque ingenioso para un caso particular del problema, en el que $k = 1$. Esta técnica utiliza información provista por la *construcción* del *DiSAT* para obtener el $1nNG$, conectando a cada objeto con un elemento *cercano* de la base de datos, que puede ser, o no, su vecino más cercano [5]. Esta propuesta permite recuperar el $1nNG$ con bajo costo, logrando muy buena precisión, un error bajo y por ende un buen compromiso calidad/tiempo, y *sin realizar ninguna búsqueda en el índice*.

Las demás propuestas abordadas no se apoyan en ningún índice y se enfocan en el problema general, resolviendo la consulta *All-k-nN* y computando el $knNG$. La base de estos desarrollos es aprovechar de manera ingeniosa las propiedades de la *función de distancia*. En ellos se sugieren distintas maneras de seleccionar muestras de la base de datos, a partir de las cuales se obtiene un conjunto de distancias que serán el punto de partida de este proceso. Se analizan diferentes maneras de utilizar la información conseguida, para calcular, en algunos casos los vecinos exactos [4] y en otros los aproximados, para todos los objetos de la base de datos, utilizando propiedades como la simetría o la desigualdad triangular. Los resultados obtenidos hasta el momento, de esta propuesta, se muestran muy prometedores.

Arquitecturas de Procesadores Orientadas a Bases de Datos

Algunos autores realizan una distinción entre arquitectura, implementación y realización, según la cual la arquitectura es la interfaz entre el software y el hardware de una computadora, y es lo que define el conjunto mínimo de propiedades que determinan qué programas correrán y qué resultados producirán sobre el procesador; la organización básica del flujo de datos y el control, conforma la implementación; y la estructura física que comprende la implementación, conforma la realización [1].

En la actualidad, la investigación sobre arquitecturas de procesadores ha sido desplazada por investigación sobre la implementación de procesadores. La mayoría de los trabajos de investigación están dedicados a mejorar técnicas de predicción (tanto de control como de datos), técnicas para sincronizar y comunicar procesadores (núcleos) a través de mensajes y/o memoria compartida. Muchas de estas

técnicas de implementación surgieron en los años 60 y hoy se han incorporado a los diseños de microprocesadores actuales. No obstante, estas técnicas de implementación se pueden aplicar a todo tipo de arquitecturas, desde una arquitectura RISC¹ a una arquitectura que se aleje del hardware e intente disminuir el tráfico de bits entre procesador y memoria.

Si bien las arquitecturas RISC compitieron en desempeño con las arquitecturas CISC², las mismas poseen un alto tráfico de bits entre el procesador y la memoria para una determinada traza de ejecución. Esto finalmente favoreció a las CISC sobre las RISC, una vez que las CISC mejoraron sus técnicas de implementación. El objetivo de las investigaciones en esta línea de trabajo es plantear nuevas arquitecturas que minimicen el tráfico de bits entre el procesador y la memoria. Para esto se está realizando la construcción de un simulador del set de instrucción AMD-64 o x86-64 con el fin de evaluar el tráfico de bits para benchmarks como Specint y Specfp para la arquitectura x86. El próximo paso sería evaluar el tráfico de bits para la arquitectura propuesta sobre los mismos benchmarks, lo cual implica construir no sólo el simulador de la arquitectura sino también el compilador C para la misma. Finalmente, se pretende aprovechar el conocimiento adquirido para, desde bajo nivel, mejorar el desempeño de los DBMSs.

Resultados y Objetivos

Las investigaciones llevadas a cabo en este proyecto han permitido mejorar el desempeño de algunos MAMs estudiados. Los resultados obtenidos animan a probar su aplicación en otros métodos de acceso [4, 5, 10, 6, 11, 3].

Se espera brindar nuevas herramientas, para administrar bases de datos métricas, cuya eficiencia acerque a estas bases de datos a la madurez de los modelos tradicionales. Con esto en mente, se continuará profundizando en el estudio de nuevos diseños de estructuras de datos, procurando incrementar su eficiencia en tiempo y espacio, logrando una mejor adaptación al nivel de la jerarquía de memorias donde se almacenarán y a las características de los datos a ser indexados. Del mismo modo, se continuará indagando sobre técnicas innovadoras que, sin utilizar índices, permitan resolver consultas eficientemente. Además, se espera mejorar el desempeño de las operaciones de bajo nivel en los DBMS, mediante una nueva arquitectura del procesador.

¹acrónimo del inglés "Reduced Instruction Set Computer".

²acrónimo del inglés "Complex Instruction Set Computer".

Actividades de Formación

Los docentes-investigadores que se forman dentro de esta línea de investigación están involucrados en: **Doctorado en Cs. de la Computación**: una tesis sobre expresividad de lenguajes lógicos de consulta.

Maestría en Cs. de la Computación: una tesis sobre una versión dinámica eficiente del *DiSAT*.

Maestría en Informática (UNSJ): una sobre la evaluación del *k*nNG para búsquedas por similitud y otra sobre una versión dinámica y para memoria secundaria del *DiSAT*.

Referencias

- [1] G. Blaauw and F. Brooks, Jr. *Computer Architecture: Concepts and Evolution*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1997.
- [2] E. Chávez, K. Figueroa, and G. Navarro. Effective proximity retrieval by ordering permutations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(9):1647–1658, Sept 2008.
- [3] E. Chávez, M. Di Genaro, N. Reyes, and P. Roggero. Decomposability of *disat* for index dynamization. *Computer Science & Technology*, pages 110–116, 2017.
- [4] E. Chávez, V. Ludueña, and N. Reyes. Solving all-*k*-nearest neighbor problem without an index. In *Procs. del XXV Congreso Argentino de Ciencias de la Computación CACIC 2019*, pages 567–576. UniRío editora, 2019.
- [5] E. Chávez, V. Ludueña, N. Reyes, and F. Kasián. All near neighbor graph without searching. *Computer Science & Technology*, 18:61–67, 2018.
- [6] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal {SAT}. *Information Systems*, 59:15 – 47, 2016.
- [7] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [8] P. Ciaccia and M. Patella. Approximate and probabilistic methods. *SIGSPATIAL Special*, 2(2):16–19, 2010.
- [9] K. Figueroa, C. Martínez, R. Paredes, N. Reyes, and P. Roggero. Dynamic list of clustered permutations on disk. In *Computer Science and Technology Series: XXI Argentine Congress of Computer Science Selected Papers*, pages 201–211. EDULP, 2016.
- [10] K. Figueroa, N. Reyes, A. Camarena-Ibarrola, and L. Valero-Elizondo. Improving the list of clustered permutation on metric spaces for similarity searching on secondary memory. In *10th Mexican Conference on Pattern Recognition (MCPR2018)*, volume 10880, pages 82–92, 2018.
- [11] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.
- [12] R. Paredes, E. Chávez, K. Figueroa, and G. Navarro. Practical construction of *k*-nearest neighbor graphs in metric spaces. In *Proc. 5th Workshop on Efficient and Experimental Algorithms (WEA)*, LNCS 4007, pages 85–97, 2006.
- [13] H. Samet. *Foundations of Multidimensional and Metric Data Structures (The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2006.
- [14] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach (Advances in Database Systems)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005. XVIII, 220 p., Hardcover ISBN: 0-387-29146-6.

CIENCIA DE DATOS APLICADA AL ESTUDIO LA FAUNA ÍCTICA EN LA ZONA DEL RÍO PARANÁ.

Cinthia A. Cuba L.^{(1,2,*), Paola V. Britos^(3,*) y Gladis G. Garrido^(1,2,*)}

⁽¹⁾ Proyecto Biología Pesquera – Instituto de Biología Subtropical (IBS).
Rivadavia 2307, N3300LQH, Posadas - Misiones, Argentina.

⁽²⁾ Facultad de Ciencias Exactas, Químicas y Naturales (FCEQyN). Universidad Nacional de Misiones (UNaM). Félix de Azara 1552, N3300LQH, Posadas - Misiones, Argentina.

⁽³⁾ Universidad Nacional de Río Negro (UNRN). Laboratorio de Informática Aplicada. Río Negro, Argentina.

*E-mail: cinthiacuba@gmail.com, pbritos@unrn.edu.ar, gladysgarrido@gmail.com

RESUMEN

El constante crecimiento del volumen de información es transversal a todos los ámbitos de nuestra vida cotidiana y hacen necesaria la utilización de nuevos métodos y enfoques analíticos que nos permitan navegar entre el volumen de información que se genera día a día y nos brinde la posibilidad de extraer conocimiento útil u oculto en esa gran cantidad de datos. La ciencia de datos, minería de datos o descubrimiento de conocimiento, son términos que tienen una amplia trayectoria y una gran cantidad de estudios realizados. En contraparte, en el ámbito biológico, existe una gran cantidad de estudios sobre la comunidad y dinámica de la fauna íctica; pero no se ha localizado ningún estudio específico que unifique las dos áreas. El presente trabajo propone la aplicación de técnicas de Ciencia de Datos a un set de información que describe el monitoreo de la fauna íctica en un tramo del Río Paraná.

El objetivo es poder determinar y entender las relaciones entre variables biológicas y ambientales e intentar descubrir comportamientos no apreciables a simple vista.

El resultado de este trabajo será un aporte al Proyecto de Biología Pesquera Regional y a través de ellos, a las entidades que toman decisiones y gestionan los recursos para la conservación y preservación de la fauna íctica en el río Paraná.

Palabras clave: ciencia de datos, minería de datos, descubrimiento de conocimiento en bases de datos, ictiofauna de agua dulce.

CONTEXTO

Este trabajo se lleva a cabo dentro del Proyecto de Biología Pesquera en el Laboratorio del Instituto de Biología Subtropical (IBS) de la Facultad de Ciencias Exactas, Químicas y Naturales (FCEQYN), de la Universidad Nacional de Misiones (UNaM), en el marco de un plan de trabajo final de la Maestría en Tecnologías de Información de la FCEQYN - UNaM

1. INTRODUCCION

El propósito de este trabajo es el desarrollo y aplicación de técnicas de Ciencia de Datos que permitan el descubrimiento de patrones y/o regularidades sobre la dinámica de la fauna íctica en un tramo del río Paraná. De esta manera se espera entender las relaciones entre variables biológicas y ambientales y que éstas sirvan de soporte a la toma de decisiones sobre la conservación y estudio de comunidades ícticas.

Los sistemas acuáticos en particular, se caracterizan por presentar gran variabilidad espacial y temporal en sus componentes físicos y químicos [1,2], los cuales

determinan la dinámica de las diferentes comunidades. Para analizar la variación espacio-temporal de la comunidad de peces, evaluar el rendimiento pesquero en término de biomasa, abundancia, riqueza y su relación con variables ambientales (como temperatura del agua y del ambiente, turbidez del agua, oxigenación, entre otros) se tomará como caso de estudio, datos de monitoreos sobre la fauna íctica que el Proyecto de Biología Pesquera Regional de la UNaM ha realizado en un tramo del río Paraná que abarca casi 200 km y comprende cinco sitios, siendo éstos: Toma de Agua Eriday (TAE), Puerto Garapé (GPE), Puerto Nemesio Parma - Candelaria (NPM), Puerto Maní (PMI) y Arroyo Yabebiry (YBY). Este monitoreo se realiza desde el año 1993 a la actualidad.

Para el presente trabajo se acota la información al rango de fechas que fue establecida después de fijar la cota definitiva en el río, y esto va desde 2010 a 2020.

Si bien, desde el punto de vista ictiológico existen diversos estudios relacionados a la biología reproductiva y rendimientos pesqueros [3], no se han expuesto patrones o características que expliquen el porqué de esta dinámica. Para dar respuesta a lo expuesto, este trabajo propone la aplicación de técnicas de Ciencias de Datos que permitan descubrir patrones y/o regularidades sobre la dinámica de la fauna íctica en el tramo del río Paraná antes mencionado.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Toda la información recolectada del 1993 a la actualidad se encuentra almacenada en una base de datos relacional y puesto que una de las actividades del proyecto es brindar información a través de distintos informes, el análisis de ese gran volumen de datos dio pie

a la creación de un almacén de datos, que junto a un sistema OLAP, forman parte de las nuevas herramientas de trabajo del proyecto. Si bien este set tecnológico aporta al rápido análisis de información desde distintas perspectivas, aún queda pendiente explicar e interpretar la dinámica de la comunidad íctica. Se espera responder a esta problemática aplicando técnicas de Ciencia de Datos que permitan descubrir patrones sobre la dinámica de la fauna íctica.

Finalmente, plasmar los hallazgos de manera simple, resumida y comprensible para el usuario.

3. RESULTADOS OBTENIDOS / ESPERADOS

Para comenzar con el proceso de extracción de conocimiento se han estudiado y comparado diferentes metodologías, como los expuestos en [4], [5] y [6] sobre Explotación de Información y Minería de Datos; luego las metodologías emergentes para Ciencia de Datos [7] o la adaptación de alguna existente [8] y [9].

Se decidió realizar un análisis comparativo entre las metodologías ASUM y Educación de Requisitos para Proyectos de Explotación de Información. Ambas se basan en la metodología CRISP-DM y la finalidad del análisis entre ambas es tomar sus mejores cualidades. ASUM, planteada como un Método Unificado de Soluciones Analíticas se presenta como un Manual de Usuario en línea que permite navegar entre cada fase o proceso planteado e indica las tareas a realizar en cada una de ellas (orientada a su herramienta comercial SPSS). Educación de Requisitos realizada como una tesis doctoral, tiene como puntos fuertes una serie de plantillas entregables que son fáciles y comprensibles de presentar al cliente; además de mostrar ejemplos concretos sobre la aplicación de la misma.

Del análisis de metodologías se obtienen las siguientes cinco fases: 1- Definición del Proyecto. 2- Educción de Procesos de Negocio. 3- Educción de Datos de Procesos de Negocio. 4- Conceptualización del Negocio. 5- Especificación de Procesos de Explotación de Información.

De la aplicación de la fase 1 se obtuvieron cinco objetivos generales. A continuación se describen las tareas realizadas para resolver la premisa “Determinar e identificar características sobre Estructura de Peces y su relación con variables ambientales”.

Para realizar las tareas prácticas se utiliza la herramienta RapidMiner Studio bajo una licencia “*Educational Edition*” en su versión 9.8.

En primer lugar, se tomó una muestra correspondiente a 1 (un) año de monitoreo con el fin de realizar las tareas de pre-procesamiento, los cuales incluyen limpieza de los datos, tratamiento de valores nulos, faltantes y calidad en general de los datos. Con la muestra tomada se trabajaron 25 variables tanto cualitativas como cuantitativas, de las cuales 8 representan datos ambientales y 17 datos bilógicos de los peces. En total la muestra se representa por 5380 registros.

Para comprender qué ocurre con los datos y cómo se relacionan entre sí, es necesario analizar una porción representativa del problema; para ello se ejecutó el proceso “Descubrimiento de Grupos” [10] con el algoritmo *K-means*.

Una vez que la herramienta formó los grupos, se utilizó una matriz de correlación para identificar qué tan fuertemente correlacionadas se encontraban las variables y con qué atributos. Además se analizaron los pesos de cada variable para tener una noción sobre cuáles son las más representativas. Como resultado se pudo observar que las variables con mayor peso al

momento de definir grupos fueron: peso y largo del pez (como principales atributos biológicos ícticos); temperatura, oxígeno y transparencia del agua (como principales atributos ambientales). Por otra parte, los pares de variables mayormente correlacionadas fueron altura-peso, altura-largo, largo-peso y temp_ambiente-temp_agua.

Teniendo este panorama general, se procedió con el análisis de cada partición (verificando si la correlación de los datos en general es correcta.). Por cada grupo conformado se identifica el atributo cluster - con el que fue catalogado en el paso anterior - y en base a este atributo se aplica el algoritmo de inducción para obtener “Reglas de pertenencia a cada grupo”. Esta tarea se ejecutó con el algortimo *Decision Tree*.

Como resultado se obtuvo que las variables que definieron el agrupamiento fueron tres: peso, largo y temperatura del agua, en el orden indicado.

Actualmente la investigación para este trabajo se encuentra en la tercera fase de la metodología, donde se realizan tareas asociadas a la Educción de Datos de Procesos de Negocio. Estas tareas se subdividen en dos grandes actividades, por un lado el relevamiento de los datos del negocio y por el otro, el análisis de los repositorios de datos. En esta etapa se está analizando la incorporación de información a través de la transformación de datos existentes, por ejemplo: a partir de la fecha de pesca, transformar este dato a estaciones del año y con este nuevo valor, generar otras salidas que sirvan al cumplimiento de los objetivos del trabajo.

4. FORMACION DE RECURSOS HUMANOS

El presente se lleva a cabo como Trabajo Final de la Maestría en Tecnologías de Información de la FCEQyN –UNaM y se conforma por tres integrantes: el maestrando, Lic. Cinthia Cuba de la FCEQyN – UnaM; la Dra. Paola Britos, perteneciente al Laboratorio de Informática Aplicada de la UNRN; y la Mgter. Gladys G. Garrido, Docente de categoría III (Sistema Nacional de Incentivo a la Investigación) perteneciente al Departamento de Biología de la FCEQYN – UNAM.

5. BIBLIOGRAFIA

[1] Fontoura, Nelson Ferreira, Ceni, Gianfranco, Braun, Aloisio Sirangelo, & Marques, Camilla da Silva. (2018). “*Defining the reproductive period of freshwater fish species using the Gonadosomatic Index: a proposed protocol applied to ten species of the Patos Lagoon basin*”. *Neotropical Ichthyology*, 16 (2), e170006. Epub July 16, 2018.

[2] Stebniki, Samanta; González, Iván; D’Anatro, Alejandro y Teixeira de Mello, Franco. (2016) “*Relaciones entre variables ambientales y la comunidad de peces en el río Uruguay bajo*” Uruguay. *Aqua-LAC* - Vol. 8 - N° 1 - Mar. 2016. pp. 62- 67.

[3] Torruco, Daniel, González-Solis, Alicia y Torruco-González, Ángel D. (2017). “*Diversidad y distribución de peces y su relación con variables ambientales, en el sur del Golfo de México.*” *Revista de Biología Tropical*, vol. 66, no. 1, 2018, p. 438.

[4] Pollo Cattaneo, M. Florencia (2017). “*Modelo de proceso para elicitación de requerimientos en Proyectos de Explotación de Información*” (Tesis doctoral). Universidad Nacional de La Plata. Argentina.

[5] M. T. Rodríguez Montequín, J. V. Álvarez Cabal, J. M. Mesa Fernández and A. González Valdés. (2005) “*Metodologías para la realización de proyectos de Data Mining*”, in VII Congreso Internacional de Ingeniería de Proyectos, pp. 257-265.

[6] J. M. Moine, S. Gordillo and A. S. Haedo, (2011) “*Análisis comparativo de metodologías para la gestión de proyectos de minería de datos*”, in XVII Congreso Argentino de Ciencias de la Computación (CACIC), pp. 931-938.

[7] Angée Santiago - Lozano, Silvia, Montoya-Munera, Edwin - Ospina Arango, Juan - Tabares, Marta. (2018). “*Towards an Improved ASUM-DM Process Methodology for Cross-Disciplinary Multi-organization Big Data & Analytics Projects*”: 13th International Conference, KMO 2018, Žilina, Slovakia, August 6–10, 2018, Proceedings. 10.1007/978-3-319-95204-8_51

[8] IBM – (2015) “*Analytics Solutions Unified Method (ASUM)*”, Publicado en: http://gforge.icesi.edu.co/ASUM-DM_External/index.htm#cognos.external.asum-DM_Teaser/deliveryprocesses/ASUM-DM_8A5C87D5.html <Último acceso: Febrero 2021>

[9] Fois, Giuliana, Agüero, Gustavo, Britos, Paola. (2020). “*Comparación de metodologías ágiles para Ciencia de Datos*” – Publicación InGenio 2020

[10] Britos, Paola. (2008). “*Procesos de explotación de información basados en sistemas inteligentes*”. (Tesis doctoral). Universidad Nacional de La Plata. Argentina.

CIENCIA DE DATOS APLICADA. ESTUDIO DE CASOS EN DIVERSAS ÁREAS, FOCALIZADO EN LA INDUSTRIA AGROPECUARIA DE LA PROVINCIA DE RIO NEGRO

Paola Britos¹, Paola Pizzingrilli^{1,6}, Lina María Montoya Suárez³, Pamela Britos⁴, Gastón Di Bonis², Melisa Isaja², Giuliana Fois¹, Gustavo Agüero¹, Pablo Enrique Argañarás^{1,5}, Martín René Vilugrón^{1,5}, Maximiliano Donadio¹

¹ Universidad Nacional de Río Negro. Laboratorio de Informática Aplicada. Río Negro. Argentina. pbritos@unrn.edu.ar, ppizzingrilli@unrn.edu.ar, gfois@unrn.edu.ar, gustavo54ar@gmail.com, parganaras@unrn.edu.ar, maxdonadio21@gmail.com.

² Universidad Nacional de Río Negro. Río Negro. Argentina. melisa.isaja@gmail.com, gdibonis@unrn.edu.ar

³ Universidad Católica Luis Amigó, Grupo de Investigación SISCO
Medellín, Colombia lina.montoyasu@amigo.edu.co

⁴ Sede Regional Sur: Metán – Rosario de la Frontera, Universidad Nacional de Salta. Argentina. pamebritos@gmail.com

⁵ Comisión Nacional de Energía Atómica, Centro Atómico Bariloche, banda@cab.cnea.gov.ar, martinvilu@cab.cnea.gov.ar

⁶ Universidad Nacional de Río Negro. Instituto de Investigaciones de Recursos Naturales, Agroecología y Desarrollo Rural (IRNAD). Río Negro. Argentina. ppizzingrilli@unrn.edu.ar

RESUMEN

Las organizaciones están buscando una nueva manera de producir optimizando y potenciando sus recursos, es por ello que están incorporando tecnologías que conectan el mundo físico con el virtual a través de diversos dispositivos que le permitan acceder a grandes cantidades de datos en tiempo real y tomar mejores decisiones en todo el proceso productivo. Por otra parte, la Ciencia de Datos es un conjunto de principios fundamentales que apoyan y guían la extracción de información y conocimiento a partir de los datos; incluye diversas metodologías, técnicas, algoritmos y herramientas que facilitan el procesamiento avanzado y automático de los mismos; permitiendo identificar información relevante y estratégica, que a simple vista no es detectada. Este proyecto tiene por *objetivo articular integradamente mecanismos de proceso de interpretación de grandes masas de información en diversas industrias, (con hincapié en la agropecuaria de la provincia de Río Negro), a través diversas técnicas de tratamiento de datos y su visualización brindada por la Ciencia de Datos.*

Palabras claves: Data Science aplicada, explotación de información, aplicaciones agropecuarias.

CONTEXTO

Este proyecto de I&D se encuentra radicado en la Universidad Nacional de Río Negro. Cuenta con los siguientes antecedentes:

- Proyectos de investigación:
 - Estudio y Evaluación de Tecnologías de la Información y la Comunicación para el desarrollo de ciudades inteligentes en Río Negro (PI 40-C-270).
 - Explotación y visualización de información. Su aplicación a las industrias rionegrinas (PI 40-C-542).

1. RELEVANCIA DEL PROBLEMA

Las organizaciones están buscando una nueva manera de producir optimizando y potenciando sus recursos, es por ello que están incorporando tecnologías que conectan el mundo físico con el virtual a través de diversos dispositivos que le permitan acceder a grandes cantidades de datos en tiempo real y tomar mejores decisiones en todo el proceso productivo: desde la concepción del producto o servicio hasta y comercialización del mismo, el estudio de los datos de las organizaciones es realizado por la Ciencia de Datos.

Por otra parte, de acuerdo al Boletín Oficial de la República Argentina [1] “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 la ciudad por las que el mundo la ha premiado de manera regular como una ciudad innovadora, sustentable e inteligente [2] y que a través de diversas actividades en la región [3], [4] y [5] y en conjunto con diversas organizaciones, como por ejemplo Instituto Nacional de Tecnología Agropecuaria (INTA Río Negro) se han realizado innovaciones al respecto.

Por lo anteriormente citado, es que el objetivo es *articular integradamente mecanismos de proceso de interpretación de grandes masas de información de diversos tipos de industria, haciendo hincapié en la agronómica, a través diversas técnicas de tratamiento de datos y su visualización, las cuales son brindada por la Ciencia de Datos.*

1. INTRODUCCIÓN

1.1. Cambio en las Organizaciones

De acuerdo a lo informado por [1] uno de los ejes estratégicos del Gobierno Nacional Argentino es impulsar procesos de transformación productiva que estimulen el desarrollo económico, promoviendo las capacidades tecnológicas y productivas de las empresas, la creación de empleos de calidad, la participación en los mercados internacionales, la mejora de la productividad y la incorporación y el desarrollo de las tecnologías emergentes.

El surgimiento de nuevos desafíos en el ámbito de las organizaciones está siendo afectado transversalmente por transformaciones tecnológicas y productivas que tienen impacto en la dinámica económica de los países que resultan determinantes para su desarrollo, planteando nuevos retos y demandas productivas [6], [7]. Las tecnologías que se insertan hoy en las organizaciones tienen que ver con la automatización, la inteligencia artificial, el *uso de ciencia de datos*, entre otras [4], [5], [3].

Ciencia de Datos

Un estudio realizado en [8] menciona que cada vez son más las empresas que se basan en datos, las cuales han mejorado su desempeño en medidas objetivas de los resultados financieros y operativos. En un entorno tan competitivo como el actual, las organizaciones, de cualquier tipo, requieren estrategias para poder adaptarse a los cambios y generar conocimiento a partir de los datos disponibles, es una muy buena estrategia [9]. Debido a la evolución de las TICs, los sistemas de almacenamiento pueden acopiar gran cantidad de datos, de diferente índole, a un menudo coste; lo que conlleva a que estos crezcan de manera exponencial en pequeños intervalos de tiempo. Los datos no son recogidos únicamente para el mantenimiento de registros o para validar las hipótesis generadas por los recursos humanos, sino que a medida que el volumen de datos crece, existe la posibilidad de probar las hipótesis que aún no se habían previsto [10]. El crecimiento en cantidad y diversidad de los datos ha llevado a que las herramientas convencionales de administración práctica, no puedan manejar estos conjuntos de datos [11].

Para poder tomar decisiones pertinentes y administrar estos conjuntos de datos potencialmente invaluable, es necesario realizar un análisis exhaustivo sobre los mismos. La Ciencia de Datos o Data Science (DS) es un conjunto de principios fundamentales que apoyan y guían la extracción de información y conocimiento a partir de los datos; incluye diversas metodologías, técnicas, algoritmos y herramientas que facilitan el procesamiento avanzado y automático de los mismos; permitiendo identificar información relevante y estratégica, que a simple vista no es detectada [9], [11], [12]. Fois, Aguero y Britos [13] desde la perspectiva de la ciencia de datos realizan una comparación de metodologías ágiles para aplicar en proyectos de ciencias de datos, del que se han tomado las actividades más inherentes para planificar cómo abordar el problema y organizar las actividades para abordar este caso de estudio. En síntesis, puede definirse la DS, como la aplicación de métodos cualitativos y cuantitativos para predecir resultados y solucionar problemas relevantes las cuales debido a la creciente e inmensa cantidad de datos que existe actualmente, el conocimiento y el análisis del dominio no pueden estar separadas. Existen muchos

algoritmos para el procesamiento y extracción de datos, así como una gran cantidad de detalles sobre los métodos aplicados en este campo; pero DS implica mucho más que solo algoritmos de Data Mining. DS proporciona a los profesionales del área una estructura y un conjunto principios que brindan un marco para tratar sistemáticamente los problemas de extracción de conocimiento útil de los datos; donde los métodos y la metodología para tratar los datos y llevar adelante este tipo de proyectos son trascendentales [11], [12]. Dentro de la ingeniería de sistemas, los proyectos de DS son un tipo especial de proyectos; cuyo objetivo es extraer conocimiento relevante, idóneo y original de los datos [14], [15].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto tiene por objetivo principal articular integradamente mecanismos de proceso de interpretación de grandes masas de información, de diversos tipos de industria (con hincapié en la industria agropecuaria y la educación de la provincia de Río Negro – Argentina, y Medellín - Colombia), a través diversas técnicas de tratamiento de datos y su visualización brindada por la Ciencia de Datos.

Específicamente tiene por objetivo: (a) Identificar modelos de comportamiento de crecimiento en diversos entornos, teniendo presente las características del mismo, y (b) Implementar nuevos modelos de visualización y estudio de datos brindados con la Industria y los estudios científicos de diversos centros académicos, en diferentes temáticas.

3. RESULTADOS OBTENIDOS/ ESPERADOS

A continuación, se indican los resultados obtenidos hasta el momento:

- Melisa Isaja, Paola Pizzingrilli, Pamela Britos, Maximiliano Donadio, Giuliana Fois, Gustavo Agüero, Pablo Enrique Argañarás, Martín René Vilugrón, Lina María Montoya Suárez, Paola Britos y Gastón Di Bonis. 2020. Estudio de comportamiento de cultivo de mostaza blanca (*Sinapis alba L.*). Un caso de aplicación con Big Data. XI Congreso

Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2020). ISBN 978-84-18254-84-0.

- Viviana Moscher, Paola Britos. Desarrollo de un modelo predictivo para descubrir factores que inciden en la deserción de alumno en la Facultad de Humanidades, Universidad Nacional del Nordeste. XI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2020). ISBN 978-84-18254-84-0
- Caleb Faillance, Florencia Pollo, Britos. Análisis De Sentimientos: Detección de Sarcasmo en Español a Través de Emoticones En Twitter. XI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2020). ISBN 978-84-18254-84-0.
- Victor Rojo, Florencia Pollo, Britos. Desarrollo de un recurso léxico de palabras informales en español de Argentina para el análisis de sentimientos en Twitter. XI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2020). ISBN 978-84-18254-84-0.
- Melisa Isaja, Paola Pizzingrilli, Pamela Britos, Maximiliano Donadio, Giuliana Fois, Gustavo Agüero, Pablo Enrique Argañarás, Martín René Vilugrón, Lina María Montoya Suárez, Paola Britos y Gastón Di Bonis. 2020. BIG DATA, aplicado a la Agricultura. IX Semana de la Tecnología, Creatividad e Innovación. Universidad Luis Amigo. Medellín. Colombia.
- Britos, P; Fois, Guiliana; Agüero Cobella, G. 2020. Evaluación comparativa de las metodologías Team Data Science Process TDSP y Analytics Solutions Unified Method for Data Mining ASUM-DM desde la perspectiva de la ciencia de datos. Investigación Formativa en Ingeniería. 264-270. ISBN 978-958-52-3335. Medellín. Colombia.
- Britos, Paola; Montoya Suarez, Lina María; Vilugron, M.; Donadio, Maximiliano. Ciencia de Datos en

Ingeniería de Sistemas Virtual y Distancia de la Uniremington. XXIII Encuentro Nacional de Investigación 2020. Universidad Luis Amigo. Medellín. Colombia.

4. FORMACION DE RECURSOS HUMANOS

Este proyecto tiene por objetivo la formación de: 3 tesis de maestría, 2 tesis de grado, 4 investigadores juniors.

5. BIBLIOGRAFÍA

- [1] M. de P. y Trabajo, “Boletín Oficial de la República Argentina,” *RESFC-2019-1-APN-MPYT*. Feb. 22, 2021, [Online]. Available: <https://www.boletinoficial.gob.ar/#!DetalleNorma/208007/20190522>.
- [2] “Medellín, el milagro de una Ciudad Inteligente,” *PanamericanWorld*. Feb. 22, 2014, [Online]. Available: <https://panamericanworld.com/revista/economia/medellin-el-milagro-de-una-ciudad-inteligente/>.
- [3] P. Britos, “Proyecto de I&D: ‘Explotación y visualización de información. Su aplicación a las industrias rionegrinas.’” 2017.
- [4] P. Britos and H. 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.’” Universidad Nacional de Río Negro, 2015.
- [5] H. Vivas, “Proyecto de I&D: ‘Herramientas Informáticas para el Desarrollo de Servicios Digitales Innovadores para las Comunidades urbanas y Rurales en el Marco de Ciudades y Regiones Inteligentes.’” Universidad Nacional de Río Negro, 2017.
- [6] L. F. O. Clavijo, J. D. F. Ledesma, S. C. Nieto, and C. J. G. Duque, “Computación en la Nube, estudio de herramientas orientadas a la Industria 4.0,” *Lámpsakos*, no. 20, pp. 68–75, Feb. 2018, [Online]. Available: <https://dialnet.unirioja.es/servlet/articulo?codigo=6771316>.
- [7] N. Velásquez, E. C. Estévez, and P. M. Pesado, “Cloud Computing, Big Data and the Industry 4.0 Reference Architectures,” in *VI Jornadas de Cloud Computing & Big Data - JCC & BD 2018 (La Plata, junio 2018)*, Feb. 2018, vol. 18, n, [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/71687>.
- [8] E. McAfee, A., & Brynjolfsson, “Big Data: The Management Revolution | Request PDF,” *ResearchGate*. Feb. 22, 2012, [Online]. Available: https://www.researchgate.net/publication/232279314_Big_Data_The_Management_Revolution.
- [9] K. B. Eckert and P. V. Britos, “Modelo basado en la Toma Decisiones con Criterios Múltiples para la elección de metodologías de Data Science,” *XX Workshop de Investigación en Ciencias de la Computación*, p. 5, 2018.
- [10] R. Agarwal and V. Dhar, “Editorial-Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research,” *Inf. Syst. Res.*, vol. 25, no. 3, pp. 443–448, Feb. 2014, doi: 10.1287/isre.2014.0546.
- [11] M. A. Waller and S. E. Fawcett, “Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management,” *J. Bus. Logist.*, vol. 34, no. 2, pp. 77–84, Feb. 2013, doi: <https://doi.org/10.1111/jbl.12010>.
- [12] F. Provost and T. Fawcett, “Data Science and its Relationship to Big Data and Data-Driven Decision Making,” *Big Data*, vol. 1, no. 1, pp. 51–59, Feb. 2013, doi: 10.1089/big.2013.1508.
- [13] P. Fois, G.; Agüero, G.; Britos, “Evaluación comparativa de las metodologías Team Data Science Process TDSP y Analytics Solutions Unified Method for Data Mining ASUM-DM desde la perspectiva de la ciencia de datos,” *Investigación Formativa en Ingeniería*, vol. 4ta, pp. 264–270, 2020, doi: <https://zenodo.org/record/4031253#.YDTxX-gzbiU>.
- [14] P. Pytel, P. Britos, and R. García-Martínez, “A Proposal of Effort Estimation Method for Information Mining Projects Oriented to SMEs,” in *Enterprise Information Systems of the Future*, vol. 139, G. Poels, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 58–74.
- [15] J. Schiefer, J. Jeng, S. Kapoor, and P. Chowdhary, *Process Information Factory: A Data Management Approach for Enhancing Business Process Intelligence*.

CIENCIA DE DATOS COMO HERRAMIENTA DE SOPORTE EN LA GESTIÓN PÚBLICA DE CALIDAD DEL AGUA

J. J. Bolano, M. G. Rey, U. Ramirez, J. G. A. Pautsch, E. Zamudio, H. D. Kuna

Instituto de Investigación Desarrollo e Innovación en Informática (IIDII)
Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones

hdkuna@gmail.com

RESUMEN

El agua es un recurso vital y como tal requiere una adecuada gestión de su calidad para la subsistencia de la población en todas sus dimensiones. En la provincia de Misiones, existen varias organizaciones que intervienen de alguna manera en la gestión de la calidad del agua, aunque se encuentran limitados respecto de la generación de información adecuada para la gestión, pese a esfuerzos individuales de algunas de estas organizaciones.

La gestión de calidad del agua es un proceso que requiere datos. Estos datos son generados a un alto costo, ya que cada estudio implica un traslado al lugar donde se realizan las muestras, gastos en recursos humanos, y gastos en el análisis de las propiedades del agua para determinar su calidad, entre otros.

En esta línea de investigación se aborda la aplicación de procesos de ciencia de datos que podrían contribuir en la generación de información dirigida a la gestión de calidad del agua.

A través de un análisis descriptivo se determinaron parámetros asociados a la calidad del agua, destacándose la temperatura, turbiedad, conductividad, pH, P-ortoFosfatos, solidos disueltos totales, oxígeno disuelto, sólidos suspendidos totales y coliformes fecales.

Asimismo, se evidencia la necesidad de ampliar el desarrollo de conjuntos de datos

que integren variables relevantes para la gestión de la calidad del agua.

CONTEXTO

Esta línea de investigación se desarrolla en el ámbito Programa de Investigación en Computación (PICom), perteneciente al Instituto de Investigación Desarrollo e Innovación en Informática de la Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones (IIDII, FCEQyN, UNaM). El PICom desarrolla líneas de investigación relacionadas con la explotación de información y la robótica.

El grupo de trabajo aborda principalmente las áreas relacionadas con el tratamiento de datos y el descubrimiento de información a partir de éstos mediante técnicas de ciencia de datos, en distintos campos de aplicación. Este contexto ha permitido contribuir con estrategias para la detección de datos anómalos, expansión de consultas, algoritmos de ranking, generación automática de perfiles y desambiguación de entidades, recomendación y selección de grupos de expertos.

1. INTRODUCCION

Una adecuada gestión de calidad del agua es determinante para la subsistencia de la población en todas sus dimensiones. Esto es evidente, en general debido a la dependencia del recurso para la vida, y en particular, debido a los usos del recurso en diversas aplicaciones como el consumo humano, el desarrollo agrícola, industrial, y recreativo.

En la Argentina, existen varios organismos que intervienen de alguna manera en la gestión de la calidad del agua, algunos inclusive con presencia en la Provincia de Misiones. Estos organismos incluyen al Instituto Nacional del Agua (INA), la Comisión Mixta del Río Paraná (COMIP), la Entidad Binacional Yacyretá (EBY), la Universidad Nacional de Misiones (UNaM), y varias otras dependencias del estado y organizaciones privadas, entre otros. En particular, la EBY ha desarrollado un Sistema de Información Geográfica (GIS) llamado GNOMA, con el objetivo de mantener datos de calidad del agua proveniente de diversas organizaciones. Por el momento, GNOMA carece de herramientas que permitan la generación de conocimiento a partir de sus datos contenidos para su uso en la gestión de la organización. Algunos ejemplos de posibles aplicaciones de generación de conocimiento, incluyen la posibilidad de clasificar zonas en forma automática que permitan indicar el uso potencial del agua en dicha zona.

Las características geográficas de provincia de Misiones, en relación a su carácter mesopotámico definido por sus ríos, arroyos, reservas, así como a sus límites geográficos nacionales e internacionales, establecen un especial interés por la gestión de la calidad del agua.

Mantener una adecuada gestión de calidad del agua implica entre otras importantes actividades, disponer de información relevante, confiable y oportuna. Sin embargo, la diversidad de organizaciones y la falta de información consolidada y disponible en relación a la calidad del agua dificultan la tarea de la gestión.

Salvo excepciones, los esfuerzos individuales de las organizaciones en relación a la recolección y mantenimiento de datos de calidad del agua, difícilmente son traducidas en información útil para la gestión. Más aún,

la gran cantidad de datos relacionados con la gestión de calidad del agua mantenidos por estas organizaciones, junto con la diversidad de sus estrategias de almacenamiento, dificultan la tarea de producir información a partir de dichos datos.

En la actualidad se evidencia una tendencia respecto del uso de la ciencia de datos aplicada a la gestión en diversos ámbitos. En particular, el crecimiento de datos disponibles y los costos en el poder de cómputo, permitieron un gran desarrollo de estrategias de ciencia de datos. Estas estrategias de ciencia de datos, permiten la generación de conocimiento a partir del descubrimiento de patrones en grandes conjuntos de datos. Asimismo, estas estrategias permiten el desarrollo de modelos de predictivos con precisiones que superan al humano en varias tareas específicas, como el reconocimiento de patrones o la predicción de variables cuantitativas.

En resumen, una adecuada gestión de calidad del agua requiere de estrategias que permitan generar conocimiento a partir de datos caracterizados por un gran volumen y heterogeneidad.

Como consecuencia, resulta necesario desarrollar estrategias que permitan generar conocimiento a partir de datos asociados a la gestión de calidad del agua. Sin embargo, hasta el momento, organizaciones relevantes en la provincia de Misiones aún no disponen de recursos avanzados para la aplicación de procesos relacionados con ciencia de datos que le permitan un aprovechamiento de los datos asociados a la calidad del agua, los cuales son particularmente costosos de generar.

2. LINEAS DE INVESTIGACION Y DESARROLLO

En la actualidad se evidencia una tendencia respecto del uso de la ciencia de datos

aplicada a la gestión en diversos ámbitos. Estas aplicaciones incluyen áreas de importancia como la agricultura [1].

El crecimiento de datos disponibles y los costos en el poder de cómputo, permitieron un gran desarrollo de estrategias de ciencia de datos. Estas estrategias de ciencia de datos, permiten la generación de conocimiento a partir del descubrimiento de patrones en grandes conjuntos de datos. Asimismo, estas estrategias permiten el desarrollo de modelos de predictivos con precisiones que superan al humano en varias tareas específicas, como el reconocimiento de patrones o la predicción de variables cuantitativas.

Recientemente se conoció la creación de un emprendimiento a nivel global denominado Aquagenity [2], el cual busca proveer servicios de información relacionados con la calidad del agua. Por otra parte, en el ámbito del aprendizaje automático, la compañía SomData ha desarrollado una inteligencia artificial para la detección de microorganismos en el agua [8]. Otras organizaciones como el Group on Earth Observations (GEO) desarrollan trabajos en relación al monitoreo, gestión y toma de decisiones relacionados con la calidad del agua [3].

La disponibilidad de datos es esencial para proyectos de ciencia de datos. En este sentido, existen esfuerzos internacionales para la creación, mantenimiento y disposición pública de datos relacionados con calidad del agua como es el caso del paquete de datos Global Open Data for Agriculture and Nutrition [4].

Actualmente es frecuente el uso de grandes cantidades de datos (Big Data) en varios dominios de aplicación. Esto también ocurre en el ámbito de la gestión del agua, como en el análisis de datos para la gestión de aguas urbanas [5].

Esta línea de trabajo aborda el desarrollo de procesos de ciencia de datos que permitan la identificación de patrones y generación de información dirigida a la toma de decisiones en relación a la gestión de la calidad del agua.

Esta línea de trabajo incluye los siguientes objetivos:

-Identificar las necesidades específicas en relación a la gestión pública de la calidad del agua en el ámbito de la provincia de Misiones.

-Determinar las fuentes y las características de los datos asociados a la gestión pública de la calidad del agua.

-Desarrollar mecanismos de procesamiento de datos disponibles que permitan su uso en modelos de ciencia de datos dirigidos a la gestión de la calidad del agua.

-Desarrollar estrategias de ciencia de datos, incluyendo el entrenamiento de modelos de aprendizaje automático, para dar soporte a los requerimientos de información de la gestión de la calidad del agua.

-Evaluar el impacto de la implementación de los procesos desarrollados de ciencia de datos en organizaciones intervinientes en la gestión de calidad del agua.

3. RESULTADOS OBTENIDOS/ESPERADOS

Hasta el momento se trabajó con el procesamiento de un conjunto de datos de calidad del agua del Río Paraná censados en el intervalo de tiempo para los años 1980 a 1990. A través de un análisis descriptivo se determinaron parámetros asociados a la calidad del agua, destacándose la temperatura, turbiedad, conductividad, pH, P-ortoFosfatos, solidos disueltos totales, oxígeno disuelto,

sólidos suspendidos totales y coliformes fecales.

Un análisis de la distribución de frecuencia de los valores avistados con los expertos en la calidad del agua determinó que no siguen una distribución normal estándar y no existe un marco de referencia sobre el cual poder contrastar los datos actualmente, dado que la naturaleza del Río Paraná es única en su contexto geográfico e hidrológico.

La ausencia de un marco de referencia para poder analizar el comportamiento de los datos incentivó a tomar la iniciativa en la reconstrucción de los datos históricos del Río Paraná a través de la articulación de los datos disponibles que disponen las organizaciones que se encargan de medir la calidad del agua del río.

Además de organizar los datos disponibles que se encuentran dispersados en varias entidades, otro desafío que enfrenta este proyecto es la propia naturaleza del Río Paraná, que posee un gran nivel de auto-depuración debido a su tamaño y fluidez. Un mismo parámetro de calidad del agua puede variar excesivamente en su medición si se realiza en un instante posterior y a una escasa distancia de la medición original.

Para definir el proceso de ciencia de datos a aplicar, [6] realiza un mapeo sobre la literatura relacionada con la predicción de la calidad del agua haciendo uso de técnicas de inteligencia computacional (IC), donde se detecta que la mayor proporción de trabajos estudian métodos híbridos que involucran más de una tecnología de IC para la construcción de un modelo predictivo de calidad de agua. En el mismo estudio se detecta la tendencia a la investigación de técnicas de Redes Neuronales Artificiales (RNA) para la construcción de modelos predictivos.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se compone de un equipo interdisciplinario, el cual desde la especialidad de las áreas de trabajo de sus integrantes contribuyen a la generación de conocimiento en el área de ciencia de datos por un lado, y en la gestión de calidad del agua por otro. Asimismo, integrantes del equipo de trabajo han colaborado en forma conjunta en objetivos comunes relacionados con el área de calidad del agua.

En particular a la gestión de calidad del agua, el grupo ha desarrollado actividades de colaboración con la Comisión Mixta del Río Paraná (COMIP), organismo del estado nacional, para procesos de digitalización de datos de calidad del agua a partir de informes históricos. A partir de dichas actividades, se desarrollaron procesos de control de calidad de datos dirigidos a la identificación desvíos, datos faltantes y anómalos.

Estas actividades implicaron el desarrollo de procesos específicos para garantizar la compatibilidad y calidad de los datos para su exportación a un Sistema de Información Geográfica gestionado por la Entidad Binacional Yaciretá (EBY).

Las líneas de investigación presentadas cuentan con doce integrantes relacionados con las carreras de Ciencias de la Computación de la UNaM. El grupo de investigación desarrolla dos tesis de grado articulando sus trabajos con una beca de Estímulo a las Vocaciones Científicas del Consejo InterUniversitario Nacional (CIN) y una beca UNaM; dos tesis de maestría en curso y una finalizada. Asimismo, las líneas de investigación y sus integrantes se vinculan con grupos de la Universidad de Castilla-La Mancha, España y la Universidad de Sonora, México.

5. BIBLIOGRAFÍA

- [1] J. J. Dabrowski, A. Rahman, A. George, S. Arnold, and J. McCulloch, "State Space Models for Forecasting Water Quality Variables," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining - KDD '18*, 2018, pp. 177–185.
- [2] D. Lazovskam, "Aquagenuity, plataforma de calidad del agua – ExpokNews," 2019. [Online]. Available: <https://www.expoknews.com/aquagenuity-plataforma-de-calidad-del-agua/>. [Accessed: 08-Aug-2019].
- [3] G. El Serafy, A. Taylor, and S. Greb, "Water Quality Aquatic Services: Use of Data Science in the Aquawatch GEO Initiative," in *20th EGU General Assembly, EGU2018, Proceedings from the conference held 4-13 April, 2018 in Vienna, Austria, p.19859*, 2018, vol. 20, p. 19859.
- [4] Global Open Data for Agriculture and Nutrition, "Introducing the Agricultural Open Data Package: BETA Version | GODAN," 2016. [Online]. Available: <https://www.godan.info/news/introducing-agricultural-open-data-package-beta-version>. [Accessed: 08-Aug-2019].
- [5] R. Harsh, G. Acharya, and S. Chaudhary, "Scope of Big Data Analytics in Bikaner Urban Water Management," *Int. J. Comput. Intell. IoT*, vol. 2, no. 3, 2018.
- [6] I. D. Lopez, A. Figueroa, and J. C. Corrales, "Un mapeo sistemático sobre predicción de calidad del agua mediante técnicas de inteligencia computacional," *Rev. Ing. la Univ. Medellín*, vol. 15, no. 28, pp. 35–52, 2016.

Clasificación automática de correos electrónicos

Juan M. Fernandez¹, Nicolás Cavasín¹, Agustín Rodríguez¹, Marcelo Errecalde²

{jmfernandez, ncavasin, arodriguez}@unlu.edu.ar, merreca@unsl.edu.ar

¹Departamento de Ciencias Básicas, Universidad Nacional de Luján

²LIDIC, Universidad Nacional de San Luis

Resumen

El correo electrónico es una de las herramientas de comunicación asincrónica más extendidas en la actualidad, habiendo desplazado a los canales más clásicos de comunicación debido a su alta eficiencia, costo extremadamente bajo y compatibilidad con muchos tipos diferentes de información [30].

En el último tiempo, con el objetivo de mejorar su uso y aprovechar a los correos electrónicos como fuente de conocimiento, se han aplicado diversas técnicas de aprendizaje automático a este tipo de información. En este sentido, existe un área particular del aprendizaje automático, denominada minería de textos, donde el conocimiento es generado a partir de la adopción de bases de datos exclusivamente textuales como fuente de datos [32]. A su vez, el correo electrónico posee características particulares respecto de otros elementos de texto que hace que existan diferencias y problemáticas particulares entre la minería de textos tradicional y la minería de correos electrónicos, conocida como *email mining* [5].

En este trabajo, se describen las acciones abordadas en el proyecto de investigación “Clasificación automática de correos electrónicos”, así como las líneas de I+D comprendidas en el mismo.

Contexto

Este trabajo se encuentra en el marco del Proyecto de Investigación “Clasificación automática de correos electrónicos”, aprobado por Disposición CDD-CB N°086/2020, del Departamento de Ciencias Básicas de la Universidad Nacional de Luján.

El objetivo general consiste en estudiar y analizar el conocimiento existente sobre técnicas de aprendizaje automático aplicadas a la clasificación automática de textos, particularmente de correos electrónicos, y generar modelos que aborden problemas concretos.

Asimismo, se está trabajando en conjunto con el Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) de la Universidad Nacional de San Luis a efectos de integrar los resultados encontrados en este proyecto con desarrollos de ese Laboratorio en torno a la clasificación de preguntas para sistemas conversacionales (*chatbots*).

Introducción

Existen trabajos que han recogido estimaciones respecto de la utilización mundial del correo electrónico donde se afirma que actualmente existen más de 3930 millones de usuarios y se proyectan 4371 millones para el año 2023 [11], alcanzando el tráfico actual de 293.6 billones de correos enviados diariamente [28].

Muchos de estos correos electrónicos son enviados a centros de contacto de organizaciones públicas y privadas, debido a que este medio se ha constituido en un canal de comunicación estándar [27].

Sin embargo, este canal de comunicación requiere una importante afectación de recursos humanos. A efectos de graficar este costo, algunos autores han relevado este aspecto a través de estudios de casos; por ejemplo, se demostró que responder un correo electrónico de un ciudadano enviado a la Agencia de Pensiones de Suecia lleva unos 10 minutos y, por lo tanto, los 99000 mensajes que reciben por año pueden necesitar hasta 10 empleados de tiempo completo para responderlos [13].

Con el fin de mejorar su uso y aprovechar a los correos electrónicos como fuente de conocimiento se han aplicado diversas técnicas de minería de textos, donde el conocimiento es generado mediante la adopción de bases de datos exclusivamente textuales como fuentes de datos [32]. Estas técnicas se enfrentan a problemáticas muy complejas dentro del área de la ciencia de la computación, debido principalmente a la dificultad del análisis del lenguaje (derivada de su ambigüedad), fundamentalmente en la etapa de análisis semántico, como así también, a los relativamente escasos materiales de entrenamiento y a la capacidad de cómputo necesaria para correr determinados algoritmos muy demandantes en recursos de hardware [7].

A su vez, el correo electrónico como fuente de datos posee un conjunto de características particulares respecto de otros elementos de texto que hace que existan diferencias y problemáticas particulares entre la minería de textos tradicional y la minería de correos electrónicos, conocida como *email mining* [5]. Por un lado, los correos electrónicos poseen información adicional en el encabezado que pueden ser explotadas para la obtención de conocimiento. Asimismo,

poseen una extensión reducida que hace que muchas técnicas de minería de textos sean ineficientes para estas fuentes de datos. A su vez, este tipo de comunicaciones muchas veces se da en un contexto informal o inmersos en una cultura organizacional particular, por lo tanto, los errores ortográficos y gramaticales, así como las abreviaturas o acrónimos aparecen con frecuencia. Por otro lado, además de los datos textuales, los correos electrónicos pueden contener tipos más ricos de datos, como enlaces URL, marcas HTML e imágenes. Aprovechar al máximo esos datos no textuales existentes en los correos electrónicos representan un problema interesante y desafiante para abordar [30].

Antecedentes

En lo relativo a clasificación de correos electrónicos, existen abordajes desde el procesamiento y generación de resúmenes [31], utilización de redes neuronales [4], clasificación para respuesta automática de correos [26], aplicación de técnicas basadas en máquinas vector-soporte, Naive Bayes, clasificadores TF-IDF [30] y utilización de multi-view y semi-supervised learning [17, 35], entre otras.

Algunos autores que abordaron la clasificación de correos electrónicos para la respuesta automática [26] categorizan las técnicas de acuerdo a, básicamente, tres enfoques de recuperación de texto: categorización de texto por aprendizaje automático, cálculo de similitud estadística de texto y coincidencia de patrones de texto y plantillas.

En relación a la categorización de texto mediante técnicas de aprendizaje automático, existen trabajos [6, 24] que desarrollaron modelos utilizando las técnicas de KNN, Naive Bayes, RIPPER y SVM, encontrando que SVM fue la técnica que demostró mejor performance. En el mismo sentido, existen trabajos [9] en los cuales se compara la precisión de técnicas como K-means++, KNN y Naive Bayes, alcanzando niveles de precisión muy altos, por encima del 96 %, para K-means++. En otras experiencias se realizan comparaciones entre los métodos de clasificación de Naive Bayes, SMO, J48 y Random Forest [23], observando que el algoritmo Random Forest fue el que obtuvo la mejor precisión, siendo esta de un 95.5 % mientras que el algoritmo Naive Bayes fue el más veloz en la construcción del clasificador. Siguiendo la misma línea, en otra experiencia [30] se comparó la precisión de diferentes algoritmos de clasificación, árboles de decisión, redes neuronales, Naive Bayes, K-Nearest Neighbor y SVM. Se utilizaron datos académicos para predecir la performance de los alumnos encontrando que los árboles de decisión y redes neuronales fueron los que mejor performance obtuvieron.

Otros abordajes a partir de la clasificación de correos electrónicos mediante el cálculo de similitud es-

tadística también obtuvieron resultados alentadores [3]. En estos casos, el modelo mantiene respuestas estándar asociadas a una variedad de preguntas etiquetadas como preguntas frecuentes. Cuando llega un correo electrónico de consulta, el sistema hace coincidir las oraciones en la consulta con las preguntas de la etiqueta considerando la distancia entre conceptos en las oraciones utilizando WordNet.

Un enfoque alternativo es el basado en coincidencia de patrones de texto y plantillas [2], donde el sistema mantiene un diccionario que contiene palabras y la probabilidad de que una palabra aparezca en un mensaje de una determinada categoría de texto, categorizando los mensajes en base a esa probabilidad junto con información adicional que toma de los mensajes de consulta.

También existen técnicas de clasificación de correos electrónicos utilizando un enfoque denominado de múltiples vistas o *multi-view* [35]; lo cual implica generar múltiples grupos de características de los correos y aprovechar los algoritmos de *Disagreement-based Semi-Supervised Learning* que proporcionan herramientas para ser entrenados en diferentes vistas. En algunas experiencias [17, 35], se generaron dos grupos de características de los correos, internas y externas, donde las primeras explotan el cuerpo del correo y las últimas aprovechan otras como el asunto y los destinatarios y luego se utilizó *Disagreement-based Semi-Supervised Learning* para generar varios modelos a múltiples vistas y permitirles colaborar para explotar ejemplos no etiquetados.

Por último y de forma más reciente, surgen los abordajes basados en *Deep Learning* [9] que implementan una red neuronal basada en un modelo *Long-Short-Term-Memory* para clasificar correos no deseados. Para resolver el problema de la gran cantidad de datos etiquetados necesarios para los métodos de *Deep Learning*, utilizaron un método de aprendizaje activo. Este método selecciona diferentes muestras y sólo entrena esas, buscando disminuir el costo del etiquetado manual de los datos. Este modelo demostró una mejor performance con respecto a los tradicionales CNN y RNN.

Líneas de I+D

Las principales líneas de investigación en las cuales se trabaja actualmente en el marco de este proyecto consisten en:

- Análisis y aplicación de diferentes estrategias de etiquetado de correos electrónicos.
- Implementación y comparación de estrategias de representación de documentos convencionales a correos electrónicos.
- Evaluación de técnicas de aprendizaje automático para la clasificación de correos electrónicos.

- Construcción de un clasificador automático como solución a la respuesta de consultas académicas.

A continuación se presenta una descripción breve de las líneas de I+D previamente enunciadas.

a. Estrategias de etiquetado

Una de las primeras tareas que se deben llevar a cabo para la construcción de un clasificador automático es la clasificación inicial de un conjunto de documentos que luego serán utilizados como conjuntos de entrenamiento y prueba para el entrenamiento y validación del clasificador. La estrategia tradicional para el etiquetado de documentos consiste en que esta tarea sea realizada por un humano, de forma manual. En muchas ocasiones, este etiquetado manual debe ser realizado por expertos en el tema que forma parte del problema que se desea abordar.

Si bien estas etiquetas de expertos proporcionan la piedra angular tradicional para evaluar los modelos de aprendizaje automático, el acceso limitado o costoso a los expertos representa un cuello de botella. A su vez, para caracterizar con precisión la efectividad de un sistema, la experiencia ha demostrado que los sistemas de IR (recuperación de información) deben evaluarse a la escala operativa en la que se utilizarán en la práctica, lo cual resulta una limitación para esta metodología puesto que debido a que los tamaños de las colecciones de datos han crecido rápidamente en los últimos años, se ha vuelto cada vez menos factible etiquetar manualmente tantos ejemplos usando el etiquetado experto tradicional [16].

En este sentido, han surgido metodologías alternativas que aportan mayor escalabilidad como la inferencia automática de etiquetas en función del comportamiento de los usuarios [14], la “supervisión distante”, en la que los datos de entrenamiento son etiquetados a partir de algunas características del texto, como tags, emoticones y otros metadatos [10] o el etiquetado de palabras representativas [19].

En el marco de este proyecto, se indagan las estrategias existentes, así como se exploran nuevas soluciones, que resulten escalables y efectivas para la clasificación de correos electrónicos.

b. Representación de correos

Aunque un documento de texto expresa una gran variedad de información, lamentablemente carece de la estructura impuesta en una base de datos tradicional; por lo tanto, los datos no estructurados deben transformarse en datos estructurados previo a la aplicación de técnicas de aprendizaje automático. Después de convertir datos no estructurados en datos estructurados, necesitamos tener un modelo de representación de documentos efectivo para construir

un sistema de clasificación eficiente [12]. En el marco de este proyecto, se evalúan y aplican diferentes estrategias de representación de documentos usualmente utilizadas como *bag of words* [18], *topic modeling* [34], *embeddings* [33] y BERT [29]. El objetivo de esta línea de investigación es evaluar las diferentes técnicas aplicadas a la clasificación de correos electrónicos.

c. Evaluación de algoritmos de clasificación

Luego de obtener los datos, realizar el preprocesamiento de los mismos para la extracción de características, realizar el etiquetado y avanzar en un esquema de representación, se entrena el clasificador utilizando distintos enfoques o algoritmos [22], como el aprendizaje bayesiano [21], la regresión logística, redes neuronales, árboles de decisión y máquinas de vectores soporte [15].

El modelo generado a partir del entrenamiento debe ser capaz de capturar las características distintivas de los documentos del conjunto de entrenamiento para luego poder analizar otros textos no observados previamente, lográndose así la capacidad de generalización del clasificador que se suele evaluar sobre otro conjunto de prueba separado [20].

A la fecha, y debido a la cantidad de algoritmos de aprendizaje existentes, resulta muy complejo sistematizar todos los abordajes posibles. De acuerdo a los antecedentes estudiados, los algoritmos escogidos que se están utilizando en el marco de esta investigación son el clasificador de Naive Bayes [21], la regresión logística [25], las máquinas de vector soporte [15], las redes neuronales recurrentes LSTM (*Long short-term memory*) [1] y XGBoost [8].

Las aspiraciones en torno a esta línea de estudio son encontrar las características de los problemas que hacen más acorde la utilización de un algoritmo de aprendizaje automático por encima del resto, sin descuidar las estrategias de representación de documentos ni las características del proceso de preprocesamiento realizado previamente.

d. Clasificador automático para la respuesta de consultas

La Universidad Nacional de Luján cuenta con un sistema informático propio para llevar adelante la gestión académica así como los trámites que de éstas se desprenden. Este sistema posee una funcionalidad que permite a los estudiantes realizar consultas vía correo electrónico al staff administrativo. Al cuerpo de ese correo, además del texto escrito por el estudiante, se agregan datos académicos y de la persona. Ante la llegada de un correo electrónico, el personal administrativo debe dar respuesta dentro de las 48 horas de realizada la solicitud. Como política de res-

guardo de la información, la Universidad Nacional de Luján realiza periódicamente una copia de seguridad con estas consultas y las respuestas brindadas a los estudiantes en cada caso, llegando a un total actual de 24700 correos electrónicos.

Utilizando esa base de conocimiento, se aborda el desafío de generar un modelo mediante técnicas de aprendizaje automático para clasificar cual es el tema de cada consulta realizada en función del contenido de los mensajes enviados y así poder responder las consultas de forma focalizada.

Objetivos

Los objetivos específicos que se persiguen en el marco de este proyecto se plantean a continuación:

1. Evaluar alternativas de pre-procesamiento de correos electrónicos y enriquecimiento de la representación de los mismos.
2. Evaluar nuevas técnicas de aprendizaje automático que permitan mejorar la precisión en problemas de clasificación de correos electrónicos.
3. Generar modelos que permitan clasificar correos electrónicos relacionados con casos concretos de acuerdo a los tópicos de los mismos.
4. Integrar y adaptar las técnicas y estrategias abordadas en este proyecto a sistemas de clasificación de preguntas para sistemas conversacionales (*chatbots*).

Complementariamente, su transferencia a la sociedad para resolver problemas del mundo real resulta de alto interés. Las herramientas a desarrollar pueden ser aplicadas a espacios tanto académicos como comerciales con lo que hay una oportunidad concreta de transferencia al sector público y privado.

Recursos Humanos

Se espera que este proyecto contribuya a consolidar un grupo de investigación en la temática y brindar un marco adecuado para la formación de recursos humanos en la Universidad Nacional de Luján a partir de la incorporación de saberes y competencias provenientes de la participación en actividades de investigación.

Concretamente, se ha incluido como integrantes del proyecto a dos docentes auxiliares del Departamento de Ciencias Básicas y un estudiante de la Carrera de Licenciatura en Sistemas de Información. A su vez, el director trabaja con estos temas en su tesis de Maestría, esperando culminarla en el marco del proyecto.

Por otro lado, se espera que este proyecto brinde la posibilidad a estudiantes de la Licenciatura en Sistemas de Información de la Universidad Nacional de Luján de realizar sus Tesinas de Grado en temas relacionados con la temática del proyecto.

Referencias

- [1] AGGARWAL, C. C., ET AL. Neural networks and deep learning. *Springer 10* (2018), 978–3.
- [2] AL-ALWANI, A. Improving email response in an email management system using natural language processing based probabilistic methods. *Journal of Computer Science 11*, 1 (2015), 109.
- [3] ALFALAHI, A., ERIKSSON, G., AND SNEIDERS, E. Shadow answers as an intermediary in email answer retrieval. In *International Conference of the Cross-Language Evaluation Forum for European Languages* (2015), Springer, pp. 209–214.
- [4] ALGHOUL, A., AL AJRAMI, S., AL JAROUSHA, G., HARB, G., AND ABU-NASER, S. S. Email classification using artificial neural network. *ACM* (2018).
- [5] BOGAWAR, P. S., AND BHOYAR, K. K. Email mining: a review. *International Journal of Computer Science Issues(IJCSI) 9*, 1 (2012).
- [6] BUSEMANN, S., SCHMEIER, S., AND ARENS, R. G. Message classification in the call center. *arXiv preprint cs/0003060* (2000).
- [7] CARDENAS, M. E., CASTILLO, J. J., NAVARRO, M., HERNÁNDEZ, N., AND VELAZCO, M. Herramientas para el desarrollo de sistemas de análisis de textos no estructurados. In *XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)*. (2019).
- [8] CHEN, T., AND GUESTRIN, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (2016), pp. 785–794.
- [9] CHEN, Z., TAO, R., WU, X., WEI, Z., AND LUO, X. Active learning for spam email classification. In *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence* (2019), pp. 457–461.
- [10] GO, A., BHAYANI, R., AND HUANG, L. Twitter sentiment classification using distant supervision. *CS224N project report, Stanford 1*, 12 (2009), 2009.
- [11] GROUP, T. R. Email statistics report, 2019-2023. url: <http://www.radicati.com>, 2019.

- [12] HARISH, B. S., GURU, D. S., AND MANJUNATH, S. Representation and classification of text documents: A brief review. *IJCA, Special Issue on RTIPPR (2)* (2010), 110–119.
- [13] HENKEL, M., PERJONS, E., AND SNEIDERS, E. Examining the potential of language technologies in public organizations by means of a business and it architecture model. *International Journal of Information Management* 37, 1 (2017), 1507–1516.
- [14] JOACHIMS, T. Optimizing search engines using clickthrough data. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (2002), pp. 133–142.
- [15] JOACHIMS, T., ET AL. Transductive inference for text classification using support vector machines. In *Icml* (1999), vol. 99, pp. 200–209.
- [16] JUNG, H. J., AND LEASE, M. Evaluating classifiers without expert labels. *arXiv preprint arXiv:1212.0960* (2012).
- [17] LI, W., MENG, W., TAN, Z., AND XIANG, Y. Design of multi-view based email classification for iot systems via semi-supervised learning. *Journal of Network and Computer Applications* 128 (2019), 56–63.
- [18] LI, Z., XIONG, Z., ZHANG, Y., LIU, C., AND LI, K. Fast text categorization using concise semantic analysis. *Pattern Recognition Letters* 32, 3 (2011), 441–448.
- [19] LIU, B., LI, X., LEE, W. S., AND YU, P. S. Text classification by labeling words. In *AAAI* (2004), vol. 4, pp. 425–430.
- [20] MARIÑELARENA-DONDENA, L., ERRECALDE, M. L., AND SOLANO, A. C. Extracción de conocimiento con técnicas de minería de textos aplicadas a la psicología. *Revista Argentina de Ciencias del Comportamiento* 9, 2 (2017), 65–76.
- [21] MCCALLUM, A., NIGAM, K., ET AL. A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization* (1998), Citeseer, pp. 41–48.
- [22] RUSSELL STUART, J., AND NORVIG, P. *Artificial intelligence: a modern approach*. Prentice Hall, 2009.
- [23] SAHA, S., DASGUPTA, S., AND DAS, S. K. Spam mail detection using data mining: A comparative analysis. In *Smart Intelligent Computing and Applications*. Springer, 2019, pp. 571–580.
- [24] SCHEFFER, T. Email answering assistance by semi-supervised text classification. *Intelligent Data Analysis* 8, 5 (2004), 481–493.
- [25] SKIENA, S. S. *The data science design manual*. Springer, 2017.
- [26] SNEIDERS, E. Review of the main approaches to automated email answering. In *New advances in information systems and technologies*. Springer, 2016, pp. 135–144.
- [27] SNEIDERS, E., SJÖBERGH, J., AND ALFALAHI, A. Automated email answering by text-pattern matching: Performance and error analysis. *Expert Systems* 35, 1 (2018), e12251.
- [28] STATISTA. Most popular global mobile messenger apps as of july 2019, based on number of monthly active users (in millions). url: <http://www.statista.com/>, 2019.
- [29] SUN, C., QIU, X., XU, Y., AND HUANG, X. How to fine-tune bert for text classification? In *China National Conference on Chinese Computational Linguistics* (2019), Springer, pp. 194–206.
- [30] TANG, G., PEI, J., AND LUK, W.-S. Email mining: tasks, common techniques, and tools. *Knowledge and Information Systems* 41, 1 (2014), 1–31.
- [31] UPADHYAY, M., RADHAKRISHNAN, D., AND NATARAJAN, M. Summarization and processing of email on a client computing device based on content contribution to an email thread using weighting techniques, Oct. 16 2018. US Patent 10,102,192.
- [32] USAI, A., PIRONTI, M., MITAL, M., AND MEJRI, C. A. Knowledge discovery out of text data: a systematic review via text mining. *Journal of knowledge management* (2018).
- [33] WU, L., YEN, I. E., XU, K., XU, F., BALAKRISHNAN, A., CHEN, P.-Y., RAVIKUMAR, P., AND WITBROCK, M. J. Word mover’s embedding: From word2vec to document embedding. *arXiv preprint arXiv:1811.01713* (2018).
- [34] YILDIRIM, S., AND YILDIZ, T. A comparison of different approaches to document representation in turkish language. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 22, 2 (2018), 569–576.
- [35] ZHAO, J., XIE, X., XU, X., AND SUN, S. Multi-view learning overview: Recent progress and new challenges. *Information Fusion* 38 (2017), 43–54.

Indexación y Administración de Grandes Volúmenes de Datos

Luis Britos, Fernando Kasián, Verónica Ludueña, Franco Merenda, Diego Olivares,
Marcela Printista, Nora Reyes, Patricia Roggero, Pablo Samat

LIDIC, Dpto. de Informática, Fac. de Cs. Físico Matemáticas y Naturales, Universidad Nacional de San Luis

{fkasian, vlud, mprinti, nreyes, proggero}@unsl.edu.ar,

{ldoorz, merenda.franco83, olivarestello.diego, samatpablo}@gmail.com

Karina Figueroa

Universidad Michoacana de San Nicolás de Hidalgo, México

karina@fismat.umich.mx

Claudia Deco

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario

deco@fceia.unr.edu.ar

Resumen

En la actualidad es habitual contar con grandes repositorios de datos no estructurados provenientes de distintas fuentes, los cuales son difíciles de administrar bajo el modelo relacional de base de datos. Este crecimiento sostenido tanto de la cantidad de datos en formato digital disponible, como en la variedad de los mismos, se debe a la rápida evolución de las tecnologías de la información y comunicación. Al considerar tipos de datos tales como texto libre, imágenes, audio, video, secuencias biológicas de ADN o proteínas, entre otros, las consultas no se corresponden a las habituales y forzar una estructuración podría restringir de antemano los diversos tipos de consultas que se puedan responder sobre ellos.

Así, al considerar el procesamiento de grandes conjuntos de datos no estructurados y recuperar a partir de ellos información de interés, se hace evidente la necesidad de proponer nuevos modelos y herramientas para su indexación y administración. Como el objetivo de cualquier sistema de recuperación de información es obtener información valiosa para el usuario, realizando una consulta a una base de datos, para resolverla de manera eficiente es necesario contar con índices apropiados.

Palabras Claves: bases de datos masivas, computación de alto desempeño, recuperación de información.

1. Contexto

Esta línea de investigación “Recuperación de Datos e Información” se encuentra enmarcada dentro del Proyecto Consolidado 3-03-2018 de la Universidad Nacional de San Luis (UNSL - Res. CS 126/18) y en el Programa de Incentivos (Código 22/F834): “Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos”, se desarrolla en el Laboratorio de Investigación y Desarrollo en Inteligencia

Computacional (LIDIC) de la UNSL y finaliza en 2022.

El principal objetivo de esta línea es desarrollar herramientas eficientes para indexar y administrar bases de datos masivas, que almacenan datos no estructurados. El análisis de nuevas técnicas que provean una buena interacción con el usuario, al igual que el desarrollo de nuevas estructuras de datos capaces de manipular eficientemente un gran volumen de datos no estructurados, que aporten a la recuperación de información sobre estos tipos de datos, están orientados en ese sentido. Para ello, se considera el diseño y desarrollo de índices para conjuntos de datos no estructurados masivos (datos multimedia, texto, secuencias de ADN, huellas digitales, etc.), que sean eficientes y escalables, para memorias jerárquicas, aplicando de técnicas de computación de alto desempeño (HPC). Además, se busca su incorporación en un sistema de administración para estas bases de datos, que apoye la recuperación de información sobre estos tipos de datos.

2. Introducción y Motivación

La evolución de las tecnologías de información y comunicación, el uso masivo de internet y la gran disponibilidad de dispositivos electrónicos, ha permitido que continuamente se esté generando un gran volumen de datos y que, por provenir de fuentes muy diversas, los tipos de datos producidos son también muy variados. Este escenario obliga a redefinir las técnicas de procesamiento, análisis y obtención de información útil, y a formular nuevas metodologías más aplicables a las soluciones que se proponen.

En el contexto de problemas de “big data” apare-

cen dos características importantes, que son: el volumen y variedad de los mismos, lo que hace imposible restringir las consultas a búsquedas sobre datos estructurados tradicionales, porque obligaría a representar una visión parcial del problema, dejando fuera información que podría ser relevante para la resolución efectiva del mismo. Las búsquedas por similitud son un tipo de búsqueda más general y aplicable a este contexto. Para lograr resolver dichas búsquedas eficientemente, es necesario considerar *métodos de acceso* o *índices métricos* [5]. Un enfoque útil para sistemas de recuperación por similitud es *la búsqueda basada en contenidos*, donde se usa el dato no estructurado en sí mismo para describir lo que se está buscando.

Entonces, al considerar grandes cantidades de datos no estructurados y necesitar responder consultas de recuperación de información, se pueden utilizar estos índices métricos para lograr eficiencia en la respuesta. Más aún, en aplicaciones reales los índices deben ser eficientes, dinámicos y escalables.

El modelo habitual para las búsquedas por similitud es el de *espacios métricos*. Una de las principales ventajas de este modelo es que, además de brindar un marco formal, es independiente del dominio de la aplicación. Un espacio métrico está compuesto por un *universo* \mathcal{U} de objetos y una *función de distancia* $d : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}^+$, la cual cumple con las propiedades de una métrica. Sobre una *base de datos* $\mathcal{S} \subseteq \mathcal{U}$, se suelen considerar dos tipos básicos de búsqueda por similitud: la *búsqueda por rango* y la *búsqueda de los k vecinos más cercanos*. La función de distancia permite medir la disimilitud entre dos objetos. El cálculo de distancia sobre algunos tipos de datos no estructurados puede ser muy costoso. Por lo tanto, un objetivo importante en el diseño de los índices es ahorrar cálculos de distancia..

Al considerar que $|\mathcal{S}| = n$, cualquier consulta se resuelve de manera trivial examinando los n elementos y calculando así n evaluaciones de distancia. Sin embargo, como la distancia puede ser costosa de calcular (por ej.: comparación de huellas digitales), en la mayoría de las aplicaciones sobre grandes volúmenes de datos no es factible aplicar la solución trivial. Entonces, para resolver consultas minimizando los cálculos de distancia, es necesario preprocesar la base de datos para construir un índice.

En ciertos casos particulares, es probable que la base de datos, el índice, o ambos, no puedan almacenarse en memoria principal y deban hacer uso de niveles más bajos de la jerarquía de memorias, como

la memoria secundaria. Sin embargo, ello acarrea altos costos en las operaciones de E/S. Por lo tanto, para lograr eficiencia, se debe minimizar también el número de operaciones de E/S, considerar el nivel de la jerarquía de memorias sobre la que se trabaja y en algunos casos admitir respuestas no exactas; utilizando, cuando se pueda, técnicas de HPC.

En este contexto, se considera como objetivo principal obtener herramientas de recuperación de información para procesar conjuntos masivos de datos, desarrollando nuevas técnicas y aplicaciones que soporten la interacción con el usuario, diseñando índices que permitan la manipulación eficiente de grandes volúmenes de datos no estructurados y faciliten la realización de diferentes tipos de consultas. De esta manera, se espera contribuir al desarrollo de aplicaciones reales para problemas de big data.

3. Líneas de Investigación

Dado que en esta investigación se pretende contribuir a distintos aspectos de los sistemas de recuperación de información sobre grandes volúmenes de datos no estructurados, se ha considerado el diseño de nuevos índices y la optimización de índices existentes, la resolución de distintas consultas sobre estos tipos de bases de datos y cómo lograr eficiencia y escalabilidad para grandes volúmenes de datos.

Índices

Los índices métricos resultan apropiados para realizar búsquedas sobre bases de datos que contienen datos no estructurados [5]. Éstos aprovechan una propiedad de la función de distancia, la desigualdad triangular, y las distancias almacenadas en el índice, para ahorrar algunos cálculos de distancia y ahorrar tiempo. Si se mantienen algunas distancias precalculadas entre los elementos de la base de datos y objetos particulares o distinguidos, la desigualdad triangular permitirá estimar la distancia entre cualquier objeto de consulta q y los elementos de la base de datos. Los dos enfoques más comunes se diferencian en si esos objetos distinguidos son *pivotes* o *centros*. Si son pivotes se almacenan las distancias de todos los objetos de la base de datos a ellos y si por el contrario son centros se particiona el espacio en zonas denominadas *particiones compactas*, por cercanía a los centros, almacenando en general un radio de cobertura que determina la zona de cada centro.

Los aspectos que se consideran de interés al diseñar índices incluyen: dinamismo, en qué nivel de la jerarquía de memorias deben almacenarse, si pue-

den aplicar técnicas de computación de alto desempeño para mejorar los tiempos de respuesta, si deben proporcionar una respuesta exacta o basta con una respuesta aproximada y la dimensionalidad del espacio métrico considerado.

Como los conjuntos de interés son los de datos no estructurados, los volúmenes de información con los que se debe trabajar (millones de imágenes en la Web) hacen necesario que los índices sean almacenados en memoria secundaria. En este caso, para lograr eficiencia, no sólo se debe considerar que realicen el menor número de cálculos de distancia sino también, que efectúen una menor cantidad de operaciones sobre el disco (E/S), debido a su costo.

Así, esta línea se dedica a diseñar índices adaptados para memoria secundaria, cuyo desempeño en las búsquedas sea adecuado. Éste es el caso de los nuevos índices basados en la *Lista de Clusters(LC)* [5, 4] que son totalmente dinámicos, es decir, admiten inserciones y eliminaciones de objetos y están especialmente diseñados para trabajar sobre grandes volúmenes de datos [13]. La *Lista de Clusters Dinámica (DLC)*, que tiene una buena ocupación de página y operaciones eficientes tanto en cálculos de distancia como en operaciones de E/S, logra un buen desempeño en espacios de alta dimensión. Sin embargo, la necesidad de recorrer toda la lista de centros de los clusters durante las búsquedas, eleva sus costos. Por ello, el *Conjunto Dinámico de Clusters (DSC)*, aunque también mantiene los clusters en memoria secundaria, organiza los centros de clusters en un *DSAT* en memoria principal, para lograr búsquedas con menos cálculos de distancia y accesos páginas/clusters. La información en el *DSAT* permite mejorar los costos en cálculos de distancia, y mantener bajos los costos de acceso a disco.

El *DSC* basa su buen desempeño en la calidad de los clusters generados. Por ello, se está analizando una variante para el *DSC* en la cual, en lugar de insertar los elementos en el índice a medida que van llegando, se demora la incorporación de cada nuevo objeto a un clúster, hasta tener varios elementos y poder determinar un mejor agrupamiento de los mismos, que mejore los costos de búsqueda, al lograr clusters más compactos y que aseguren una total ocupación de la página del disco, achicando el tamaño del archivo y reduciendo los tiempos de acceso. Además, permitiría reducir el costo de construcción por amortizar el costo de la escritura de un clúster en disco, luego de varias inserciones. Esta idea dio lugar a la *Buffered On Line Dynamic List*

of Clusters (BOLDLC), una variante de la *Dynamic List of Clusters (DLC)*, que agrega un espacio en memoria para mantener los objetos que son insertados (bolsa), además de un índice auxiliar (de pivotes) sobre estos elementos. Como pivotes se eligen elementos de la bolsa, y cuando ésta se llena, se selecciona el pivote que necesita el menor radio de cobertura para encerrar la cantidad de elementos que caben en un clúster y todos éstos se sacan de la bolsa y se graban en un clúster completo. La grabación de clusters llenos mejora la cantidad de E/S, pero puede provocar que objetos cercanos, que se insertan en diferentes momentos, queden en clusters diferentes, ocasionando degradación en las búsquedas. Por ello, se está estudiando una variante de la BOLDLC que considera grabar clusters casi llenos, cuya cantidad de objetos depende de un factor de carga ρ . Esto permite un espacio libre en cada clúster, que se considera al insertar un nuevo objeto. Primero se verifica si hay clusters que lo pueden incorporar y de ellos se elige al más cercano a él. Si esto no es posible, se lo incorpora a la bolsa y se actualiza el índice de pivotes. Así, los clusters se adaptan a inserciones y se mejoran los costos en las búsquedas.

Además, se está trabajando en el diseño e implementación de una versión paralela del *Conjunto Dinámico de Clusters (DSC)* [13]; esperando, no sólo bajar la cantidad de cálculos de distancia y operaciones de E/S necesarias para responder a una consulta, sino también la cantidad de consultas simultáneas que puedan resolverse, aprovechando al máximo las operaciones de E/S que se realicen durante las mismas. Para ello, se buscará aplicar y comparar distintas estrategias de paralelización en él.

Como el *DSC* basa su buen desempeño en la calidad de la partición generada por los centros de sus clusters, otro aspecto a considerar es mejorar la poda en estos agrupamientos. Para ello, es posible seleccionar un conjunto de “pivotes” globales que permitan caracterizar sobre cada uno de los clusters las zonas dentro de las cuales existan efectivamente elementos [11]. Así, es posible mantener para cada clúster información sobre la mínima y la máxima distancia a la que cada pivote encuentra elementos de dicho clúster. Esto permite identificar los “huecos” en la zona del clúster para que luego, durante las consultas, se pueda descartar un clúster si la bola de la consulta no posee intersección efectiva con la zona del clúster que realmente contiene elementos de interés. Los clusters en *DSC* almacenan un centro y un radio de cobertura y en las búsquedas se

recupera desde memoria secundaria cada cluster que intersecta la bola de consulta con centro q y radio r . De esta manera, identificando las zonas del clúster que no contienen efectivamente elementos, se puede descartar un cluster, ahorrando cálculos de distancia y lecturas a disco, porque la bola de consulta intersecta al cluster en su zona “hueca”.

Una técnica que ha logrado mucho éxito es la de *Algoritmos Basados en Permutaciones* (PBA). Una manera de indexar estas permutaciones es usando un archivo invertido. En [1] se propone acortar las listas de posteo de cada permutante, considerando un parámetro $m_i < K$, donde K es el número de permutantes considerados. Así, en el archivo invertido, cada permutante p mantiene una lista de pares (u, Φ) , siendo u un elemento y Φ una posición, si la permutación del elemento u tiene al permutante p en la posición Φ y $\Phi < m_i$. Durante las consultas, se introduce un nuevo parámetro m_s , donde $m_s \leq m_i \leq K$ y se consideran las listas de posteo de los permutantes que aparecen en la permutación de la consulta hasta la posición m_s . De esta manera se seleccionan los candidatos y se calcula la distancia entre ellos y la consulta q , para dar la respuesta.

Aunque el archivo invertido es muy útil para determinar la lista de elementos candidatos, en [1] no se utiliza toda la posible información. Así, en [8] se ha propuesto aprovechar toda la información disponible desde el archivo invertido. Además, para ahorrar espacio, se almacena sólo al objeto u en la lista del permutante p , si la permutación de u tiene a p en sus primeras m_i posiciones. Sin embargo, se mantiene además para cada u su permutación acortada a las primeras m_i posiciones. Durante las búsquedas se procede de manera similar, recuperando las listas de los primeros m_s permutantes; pero luego, antes de calcular las distancias entre el objeto de consulta q y los candidatos, se ordenan los candidatos por distancia entre la permutación acortada de q y las de los candidatos; así, los elementos con más probabilidad de aparecer en la respuesta se encuentren primero. Se está evaluando la posibilidad de definir el valor de m_s en función del radio r de la consulta.

Por otra parte, se está estudiando cómo aprovechar los índices sobre conjuntos masivos de datos, como herramienta de apoyo para solucionar un problema de estacionamiento de vehículos.

DBMS para Bases de Datos Multimedia

A pesar de que las operaciones más comunes sobre bases de datos multimedia son las búsquedas por

rango o de k -vecinos más cercanos, la operación de *join* por similitud debería brindarse en un sistema administrador para bases de datos multimedia [15].

Hay distintas variantes del *join* por similitud entre dos bases de datos A y B , con $A, B \subseteq \mathcal{U}$, dependiendo del criterio de similitud Φ utilizado. Si $A = B$, la operación se denomina *auto-join*.

Dadas $A, B \subseteq \mathcal{U}$, se define el *join por similitud* entre A y B ($A \bowtie_{\Phi} B$) como el conjunto de los pares (x, y) , donde $x \in A$ e $y \in B$; es decir, $(x, y) \in A \times B$, tal que $\Phi(x, y)$ es verdadero. 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; entre otras. Al resolver el *join* por similitud es posible que ambas, una o ninguna de las bases de datos posean un índice; o que ambas bases de datos se indexen conjuntamente con un índice diseñado para el *join*. Calcular cualquiera de las variantes del *join* por similitud de manera exacta es muy costoso [14], entonces hay que analizar cómo obtener más rápidamente una respuesta aproximada al *join* y de buena calidad.

PostgreSQL es el primer sistema de base de datos que permite realizar consultas por similitud sobre algunos atributos, particularmente indexa para búsquedas de k -vecinos más cercanos (índices *KNN-GiST*) sobre texto, comparación de ubicación geoespacial, etc. Sin embargo, estos índices proveen plantillas sólo para *árboles balanceados* (*B-tree*, *R-tree*), pero el “balance” no siempre es bueno para los índices para búsquedas por similitud [2]. Por otro lado, no se dispone de este tipo de consultas sobre todo tipo de datos métricos. Así, es importante proveer un DBMS para estos datos y sus operaciones [10].

Como las respuestas a consultas de *join* suelen ser conjuntos muy grandes de pares de objetos y algunos de ellos muy similares entre sí, se planea introducir sobre los *joins* la posibilidad de diversificar las respuestas [16]; asegurando así un conjunto más pequeño y diversificado de respuestas útiles y más rápido de obtener. Además, se ha diseñado un algoritmo simple y eficiente que permite responder consultas de *join* por similitud de k vecinos más cercanos aproximados dentro del mismo conjunto (*auto-join*), con una razonable precisión en la respuesta [6]. Estos desarrollos permitirán tener un DBMS más aplicable en sistemas de información reales.

4. Resultados

Se han publicado en [9, 7] nuevas alternativas para búsquedas aproximadas sobre grandes volúmenes

de datos y en [8] una optimización al archivo invertido de permutaciones. Por otro lado, se está evaluando experimentalmente la versión paralela del índice *DSC*, diseñada para memoria secundaria, que admite inserciones/eliminaciones de objetos y permitirá responder eficientemente lotes de consultas. Además, se encuentran en etapa de implementación las distintas mejoras al *DSC*. Se continúa trabajando en la extensión de *PostgreSQL* para soportar a más tipos de consultas por similitud, sobre distintos tipos de datos, considerar consultas por similitud aproximadas y la diversificación de respuestas para los joins por similitud. Se ha publicado un algoritmo para realizar el auto-join aproximado [6].

5. Formación de Recursos

Se están realizando las tesis de Maestría: (1) - “Estructuras Eficientes sobre Datos Masivos para Búsquedas en Espacios (UNSL)”, (2) - “Sistema Administrador para Bases de Datos Métricas” (UNSL), (3) - “Índices Métricos – Optimización del DSC usando Cortes de Regiones” (UNSJ) y (4) - “Optimización del BOLDLC por la Mejora de la Densidad y Solapamiento de los Clusters”(UNSJ).

Además, está en desarrollo un trabajo final de la Ingeniería en Computación (UNSL).

Referencias

- [1] G. Amato and P. Savino. Approximate similarity search in metric spaces using inverted files. In *Proc. of the 3rd International Conference on Scalable Information Systems, InfoScale '08*.
- [2] E. Chávez, V. Ludueña, and N. Reyes. Revisiting the VP-forest: Unbalance to improve the performance. In *Proc. de las JCC08*, page 26, 2008.
- [3] E. Chávez, V. Ludueña, N. Reyes, and P. Rogero. Faster proximity searching with the distal sat. *Information Systems*, 59:15 – 47, 2016.
- [4] E. Chávez and G. Navarro. A compact space decomposition for effective metric indexing. *Pattern Recognition Letters*, 26(9):1363–1376, 2005.
- [5] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM*, 33(3):273–321, 2001.
- [6] S. Ferrada, B. Bustos, and N. Reyes. An efficient algorithm for approximated self-similarity joins in metric spaces. *Information Systems*, 91:101510, 2020.
- [7] K. Figueroa and N. Reyes. Permutation’s signatures for proximity searching in metric spaces. In *Similarity Search and Applications - 12th International Conference, SISAP 2019*, volume 11807, 151–159, 2019. Springer.
- [8] K. Figueroa, N. Reyes, and A. Camarena-Ibarrola. Candidate list obtained from metric inverted index for similarity searching. In *Advances in Computational Intelligence*, 29–38, 2020. Springer.
- [9] K. Figueroa, N. Reyes, A. Camarena-Ibarrola, and L. Valero-Elizondo. Improving the list of clustered permutation on metric spaces for similarity searching on secondary memory. In *Pattern Recognition*, 82–92, 2018. Springer.
- [10] F. Kasián and N. Reyes. Búsquedas por similitud en PostgreSQL. In *Actas del XVIII Congreso Argentino de Ciencias de la Computación (CACiC)*, 1098–1107, 2012.
- [11] J. Lokoč, J. Moško, P. Čech, and T. Skopal. On indexing metric spaces using cut-regions. *Information Systems*, 43:1–19, 2014.
- [12] G. Navarro. Searching in metric spaces by spatial approximation. *VLDBJ*, 11(1):28–46, 2002.
- [13] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.
- [14] R. Paredes and N. Reyes. Solving similarity joins and range queries in metric spaces with the list of twin clusters. *JDA*, 7:18–35, 2009.
- [15] 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 International Conference on Data Engineering (ICDE)*, 1059–1070, 2017.
- [16] L. Santos, L. Carvalho, W. Oliveira, A. Traina, and C. Jr. Traina. Diversity in similarity joins. In *Similarity Search and Applications*, volume 9371 of *LNCS*, 42–53, 2015. Springer.

INTELIGENCIA Y ANALITICA DE NEGOCIOS PARA LA TOMA DE DECISIONES EN DIFERENTES CONTEXTOS

María Inés Lund^{1,2}, Silvina I. Migani^{1,2}, Cristina Vera¹, Alejandra Orellana Vasallo¹, Adriana María Gómez³, Sonia Eleonora Pinto³, Alejandro Riveros^{1*}, Diego Checarelli³, Mariano Alaniz^{1*}, Leandro Drazic^{1*}, Diego Guevara^{1*}, María Laura Molina^{1*}

¹Departamento de Informática, FCFN, Universidad Nacional de San Juan

^{1*} Alumno Departamento de Informática, FCFN, Universidad Nacional de San Juan

²Instituto de Informática, FCFN, Universidad Nacional de San Juan

³Gobierno de la Provincia de San Juan

mlund@iinfo.unsj.edu.ar, {silvina.migani, aolitaorellana, civerados, a.gomeztapella, pinto.eleo, alejandroriveros09, diegochearelli, mariano.alaniz12, leandrodrazic, guevaradiego08, lauramolinaf86} @gmail.com

RESUMEN

Las organizaciones de hoy enfrentan desafíos cada vez más complejos en términos de gestión y resolución de problemas para alcanzar sus objetivos y metas; están obligados a contar con información y conocimiento que sustenten la toma de decisiones, más allá de la mera intuición. Es justamente la Inteligencia de Negocios (IN), el área de conocimiento que comprende la colección de metodologías, procesos, arquitecturas y tecnologías que hacen posible generar esas necesarias y valiosas evidencias. El elemento esencial es el dato; es necesario entonces aclarar los términos dato, información y conocimiento. Dato es una representación simbólica, no tiene contenido semántico; mientras que la información refiere a un conjunto de datos procesados que tiene un significado; a su vez, la información sintetizada, analizada e interpretada da origen al conocimiento. Estos dos últimos constituyen las piezas claves para una acertada toma de decisiones.

Esta temática originalmente fue aplicada en las grandes empresas. Sin embargo, hoy en día tanto el estado como las micro y pequeñas empresas también tienen la necesidad y la posibilidad de incorporarlas a sus contextos.

Este trabajo expone brevemente algunos resultados del proyecto marco. Ellos incluyen tareas de exploración, investigación y aplicación de diferentes tipos de analítica y herramientas de software propicias para resolver problemáticas provenientes de contextos variados.

Palabras claves: Inteligencia de Negocios, Toma de Decisiones, Analítica de Negocios.

CONTEXTO

Este trabajo forma parte del proyecto aprobado por la UNSJ, E1131-INTELIGENCIA DE NEGOCIOS: TECNOLOGÍAS, APLICACIONES Y TENDENCIAS - Res.N°0591-20-R, para desarrollarse durante los años 2020 y 2021.

El área de conocimiento del proyecto se encuentra estrechamente relacionada a las temáticas abordadas en las asignaturas donde los docentes/investigadores involucrados se desempeñan (dictadas en 3ro, 4to y 5to año de las carreras Licenciatura en Sistemas de Información y Licenciatura en Ciencias de la Computación). Por otro lado, también participan personas externas a la UNSJ que se desempeñan en el gobierno provincial. Esta unión propicia la transmisión de conocimientos en ambos sentidos, y la

orientación de las investigaciones hacia problemas concretos y requeridos por la sociedad.

1. INTRODUCCIÓN

Los procesos de negocio de una organización son todas las actividades interdepartamentales y correlacionadas que juntas logran el funcionamiento lógico de una empresa. Los sistemas de información han contribuido a los negocios innovando las formas de desarrollar los procesos, adoptando nuevas tecnologías en un mundo cada vez más digital, y contribuyendo a una economía más globalizada (Rocha, L. M., citado por [1]).

Hatch, citado en [2], manifestó que: *“Inteligencia de Negocios es la combinación de prácticas, capacidades y tecnologías usadas por las compañías para recopilar e integrar la información, aplicar reglas del negocio y asegurar la visibilidad de la información en función de una mejor comprensión del mismo y, en última instancia, para mejorar el desempeño”*.

La IN resulta fundamental, ya que no sólo permite analizar el estado actual del negocio sino también hacer proyecciones a futuro en base al análisis de grandes cantidades de datos históricos. Por ello, en el mundo actual, donde la evolución tecnológica es vertiginosa, y los datos producidos se caracterizan por ser voluminosos, variados y dispersos, las organizaciones tienen que intensificar sus esfuerzos por incorporarla. (Polo, A. A. M., citado por [1]).

En general, IN brinda apoyo sistemático y estructurado a los procesos y operaciones de una organización, no obstante es especialmente valioso en la toma de decisiones en todos los niveles de la pirámide organizacional (operativo, táctico y estratégico). Los tipos de decisiones y las características de la información que manejan son diferentes, desde decisiones estructuradas

(nivel operativo), donde la información es rutinaria, predefinida, programada, detallada, frecuente, histórica, interna, y enfocada; hasta decisiones no estructuradas (nivel estratégico), donde la información que requieren es específica, por excepción, no programada, resumida, poco frecuente, prospectiva, externa y de amplio alcance [3], [4].

Actualmente, el uso de IN no es exclusivo para empresas de gran tamaño, sino para todas aquellas organizaciones (micro y pequeñas empresas, estado, etc.) que deseen y necesiten tomar mejores decisiones a partir de la información y/o conocimiento generado del análisis de sus datos [5]. Es así como el estado nacional y provincial están involucrados en planes de modernización tendientes a un gobierno inteligente [6].

Los sistemas datawarehouse (DW) constituyen una parte esencial de la IN. Ellos coleccionan, transforman, estructuran y almacenan datos organizacionales en un repositorio que posteriormente es utilizado para realizar tareas de análisis [7]. Algunos autores referentes, tales como Inmon, R. Kimball y M. Ross [8], [9] expresan que un DW (o almacén de datos en español), no es una mera copia de los datos transaccionales en una plataforma diferente; sino que tiene necesidades, objetivos, clientes y ritmos profundamente diferentes a los sistemas operacionales. Otra característica distintiva es que ellos preservan el contexto histórico de los datos para poder analizar el rendimiento de una organización a lo largo del tiempo. Además, están optimizados para consultas de alto rendimiento ya que a menudo requieren que se explore un gran volumen de datos.

Por lo tanto un DW es mucho más que un repositorio o una base de datos organizacional, involucra un conjunto de tecnologías y procesos articulados para lograr el objetivo final, que es asistir a la toma de decisiones. La implementación puede realizarse en diferentes

tipos de gestores de bases de datos, sin embargo la alternativa relacional es la más utilizada; los esquemas típicos son estrella, copo de nieve y galaxia [10] [11]. Sus principales inconvenientes surgen en relación a la rigidez de sus estructuras de datos y al rendimiento en entornos de BIG DATA [12], [13]. Obviamente ellos posibilitan elaborar análisis de gran interés y obtener importantes ventajas. Sin embargo, cuando la cantidad y variedad de datos es excesiva, este abordaje no resulta suficiente y el uso de sistemas de gestión de bases de datos NoSQL constituye una buena alternativa de solución. Concretamente, los gestores NoSQL basados en grafos prometen ser altamente beneficiosos [14], [15].

Todo lo expuesto hace considerar valioso investigar, identificar, analizar y clasificar las diferentes teorías, tecnologías, sistemas, prácticas, tipos de analítica y herramientas de software existentes en el área de IN, para luego proponer un marco que sirva de referencia al encarar un proyecto concreto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto marco de este trabajo busca estudiar, analizar, comparar y experimentar técnicas, métodos, tipos de análisis y herramientas de software de IN propicias fundamentalmente en dos contextos: micro y pequeñas empresas por un lado, y por otro, el estado. Se mencionan a continuación algunas de las líneas de investigación seguidas:

- Investigación y desarrollo de inteligencia y analítica de negocios en la educación, como innovación educativa.
- Investigación y desarrollo de inteligencia y analítica de negocios en la educación, aplicado al contexto de pandemia.
- Investigación y desarrollo de inteligencia y analítica de negocios en la pequeña empresa.

- Investigación y desarrollo de inteligencia y analítica de negocios aplicando un abordaje no relacional.
- Investigación y desarrollo de inteligencia y analítica de negocios aplicado al ámbito gubernamental, para la gestión de proyectos de innovación.

3. RESULTADOS

En el marco del proyecto, y avanzando con las distintas líneas de investigación, se realizó la publicación y exposición, en las 49 Jornadas Argentinas de Informática – JAIIO-SADIO, Octubre de 2020, del trabajo: “Promoviendo el desarrollo de habilidades blandas en entornos distribuidos y colaborativos. Una estrategia de enseñanza-aprendizaje en tiempos de pandemia”. Se expuso el análisis de datos realizado y la información/conocimiento obtenido aplicando analítica tradicional, permitiendo valorar objetivamente, en base a evidencias (encuestas de pares), la aplicación de una estrategia de enseñanza-aprendizaje en entornos distribuidos. Bajo esa misma línea de investigación, se presentó en el IV Congreso Internacional de Ciencias de la Computación y Sistemas de Información 2020 – CICCSI 2020, Noviembre del 2020, el trabajo denominado “REFLEXIONES SOBRE APRENDIZAJE PRESENCIAL VS. REMOTO. Un caso de estudio”. En él se expuso una comparación y análisis de los resultados del aprendizaje y de la medición de satisfacción de alumnos correspondientes al cursado 2019 (con modalidad presencial) y 2020 (con modalidad virtual) considerando diferentes variables que permitieron visualizar las ventajas y desventajas de cada una de las modalidades.

En relación a otro dominio de estudio, en el mismo congreso IV CICCSI 2020, desarrollado en forma virtual en Noviembre del 2020, se presentó y publicó el trabajo “Analítica de Negocio en la Gestión de Vuelos Gubernamentales: Una Revisión Sistemática

de la Literatura”. El objetivo fue realizar una revisión sistemática de la literatura para identificar métodos, técnicas y herramientas propias de Business Intelligence y Business Analytics utilizados en el área gubernamental, con el fin de administrar, operar y mantener eficientemente los recursos aeronáuticos del Estado.

Siguiendo otra línea de investigación, se expuso y publicó el trabajo “Comparando un Sistema de Bases de Datos de Grafos con un Sistema de Bases de Datos Relacional”, también en el IV CICC SI 2020, presentando parte de la investigación referida a la gestión de datos no relacionales, estructurados en grafos concretamente, que supone ventajas prometedoras en el ámbito de la analítica de datos. Dentro de esta misma área, se desarrolló también la conferencia “Neo4j: Un sistema de gestión de Base de Datos orientado a Grafos”, expuesta en la 1º Jornada en Tecnología de la Información y Comunicación 2020 (Septiembre 2020), y organizada por el Colegio de Profesionales de Ciencias Informáticas de la provincia de San Juan. En este caso se presentó una experiencia realizada en Neo4j, con datos simulados relativos a ciudadanos de la provincia de San Juan en relación a la pandemia COVID19.

En el contexto de las micro empresas, se trabajó en un modelo para la generación de datamarts (refiere a DWs pequeños, sobre un objeto de análisis) en microempresas aplicando técnicas tradicionales de analítica de negocios. Requirió adaptar prácticas y propuestas metodológicas pensadas para grandes organizaciones, e investigar herramientas de software libre, ya que constituyen la única alternativa viable en este tipo de entornos. Así, se logró definir un procedimiento simplificado que puede ser replicado de manera sencilla y ordenada. Esto se desarrolló tomando como ámbito de trabajo

una micro empresa sanjuanina dedicada a la venta de productos para el agro.

En relación a la línea de investigación sobre analítica de negocios en la educación, como innovación educativa, la Lic. Alejandra Orellana Vasallo **obtuvo el título de Magíster en Informática (UNSJ)** con el tema de investigación: “Marco de referencia para la implementación de Learning Analytic en educación superior”. Se propuso una taxonomía de métricas para medir el proceso enseñanza-aprendizaje, además de buenas prácticas sobre el uso de herramientas analíticas asociadas a cada métrica, de limpieza, transformación y visualización, obteniendo también un prototipo de “dashboard docente”.

4. FORMACIÓN DE RECURSOS HUMANOS

En términos generales, el proyecto que abarca este trabajo, está focalizado a la profundización y consolidación del conocimiento en el área de la analítica e inteligencia empresarial, por parte de cada uno de sus integrantes docentes, profesionales y alumnos. Sin embargo, dentro del grupo de investigación se distinguen los siguientes:

En relación a tesis de grado, tres alumnos:

- Alejandro Riveros trabajando en el uso de procesos y tecnologías de IN en la Dirección Provincial de Aeronáutica de la provincia de San Juan con el propósito de lograr una mejor gestión de los recursos, manteniendo las prioridades de la repartición.
- Mariano Alaniz en la aplicación de inteligencia de negocios en una pequeña empresa dedicada a la venta de productos informáticos, focalizada a mejorar la gestión de stock y la toma de decisiones en general.
- María Laura Molina desarrollando un modelo que permita relacionar datos y

obtener conocimiento para generar perfiles de posibles estudiantes de carreras tecnológicas/informáticas, aplicando técnicas modernas de analítica de negocios.

En cuanto a tesis de postgrado:

- La Lic. Alejandra Orellana Vassallo presentó y aprobó su tesis de Maestría en Informática, con el trabajo “Marco de referencia para la implementación de Learning Analytic en educación superior”.
- La Lic. Cristina Vera se encuentra en los últimos tramos de su tesis de Maestría en Informática. La temática abordada refiere al uso (transaccional y para el análisis) de gestores de bases de datos orientados a grafos y su comparación con los robustos y ampliamente conocidos relacionales.
- El Lic. Diego Checarella, que se desempeña como profesional en la Dirección Provincial de Informática del gobierno de la provincia de San Juan, se encuentra iniciando su trabajo de Tesis de Maestría en la temática de gestión y seguimiento del estado de avance de los proyectos estatales, utilizando técnicas propias de IN.

Por otra parte, se incorporaron en el segundo semestre del 2020 dos alumnos, Leandro Drazic y Diego Guevara. Ambos bajo la modalidad de adscripción al proyecto, con el fin de involucrarse en los temas y elaborar sus proyectos de tesis de grado.

5. BIBLIOGRAFÍA

- [1] H. Muñoz Hernández, R. C. Osorio Mass, and L. M. Zúñiga Pérez, “Inteligencia de los negocios. Clave del Éxito en la era de la información,” *CLIO América, ISSN-e 2389-7848, ISSN 1909-941X*, vol. 10, no. 20, pp. 194–211, 2016.
- [2] J. M. Rodríguez Parrilla, *Cómo hacer inteligente su negocio: business intelligence a su alcance*. Mexico: Grupo Editorial Patria., 2015.
- [3] J. C. V. Briano, *Sistemas de información gerencial: tecnologías para agregar valor a las organizaciones*, 1ª. Prentice Hall/Pearson Education, 2011.
- [4] K. C. Laudon and J. P. Laudon, *Sistemas de información gerencial*, 14ed ed. Pearson Educación, 2016.
- [5] C. Loreti, “Aplicación de una solución de Business Intelligence que permita a miembros de pequeñas y medianas empresas gestionar la información para conseguir mejoras en las tomas de decisiones organizacionales,” *U.Siglo XXI*, 2016.
- [6] P. Clusellas, E. Martelli, and M. J. Martelo, *Un gobierno inteligente*, 1ra ed. CABA, Argentina: Boletín oficial, 2019.
- [7] J. C. Trujillo, J. Pardillo Vela, and J. N. Mazón López, *Diseño y explotación de almacenes de datos: conceptos básicos de modelado multidimensional*. Editorial Club Universitario, 2011.
- [8] W. H. Inmon, *Building the data warehouse*. Wiley Pub, 2005.
- [9] R. Kimball and M. Ross, *The data warehouse toolkit: the definitive guide to dimensional modeling*. Wiley, 2013.
- [10] D. L. Moody, D. L. Moody, and M. A. R. Kortink, “From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design,” 2000.
- [11] A. Vaisman and E. Zimányi, *Data Warehouse Systems - Design and Implementation*. Springer-Verlag Berlin Heidelberg, 2014.
- [12] J. Bhogal and I. Choksi, “Handling Big Data Using NoSQL,” in *Proceedings - IEEE 29th International Conference on Advanced Information Networking and Applications Workshops, WAINA 2015*, 2015.
- [13] R. Cattell, “Scalable SQL and NoSQL data stores,” *SIGMOD Rec.*, vol. 39, no. 4, 2010.
- [14] R. Angles and C. Gutierrez, “Survey of graph database models,” *ACM Comput. Surv.*, vol. 40, no. 1, 2008.
- [15] R. Angles and C. Gutierrez, “An introduction to graph data management,” *arXiv*. 2017.

Mejoras Algorítmicas para Problemas de Búsquedas en Datos Masivos

Gabriel Tolosa^{1,2}, Tomás Delvechio¹, Pablo Lavallén¹, Andrés Giordano¹, Agustín González¹,
Claudia Reinaudi², Santiago Ricci¹, Tomás Juran^{1,2}, Esteban A. Rissola^{1,3}
{tolosoft, tdevechio, plavallen, agiordano, agonzalez, creinaudi, sricci, tjuran}@unlu.edu.ar
esteban.andres.rissola@usi.ch

¹Departamento de Ciencias Básicas, Universidad Nacional de Luján

²CIDETIC, Universidad Nacional de Luján

³Faculty of Informatics, Università della Svizzera italiana

Resumen

El procesamiento de datos masivos propone a diario nuevos desafíos, debido tanto a cuestiones vinculadas a los datos mismos como a la variedad de aplicaciones y soluciones que requieren los usuarios. En el primero de los casos, los datos crecen a tasas exponenciales pero también existen diversidad de nuevas fuentes a considerar, incluyendo aquellas en las cuales se producen en flujos en tiempo real. Estas características exigen mayores capacidades de hardware a los proveedores de servicios e imponen restricciones a los usuarios en la facilidad de acceso.

En este escenario, los algoritmos que resuelven problemas de búsquedas (en sentido amplio) requieren de mejoras tanto conceptuales como ingenieriles que les permitan escalar con el tamaño del problema. La eficiencia es un requerimiento fundamental para procesar datos masivos, debido al tamaño, la complejidad y la dinámica de las fuentes actuales de información digital.

Este proyecto presenta el abordaje de problemas relacionados con dos escenarios actuales. Por un lado, el procesamiento de colecciones masivas de documentos, para la construcción de motores de búsqueda de escala web. Por otro lado, el procesamiento de grafos en cuanto a las métricas de distancias, para aplicar, por ejemplo, a búsquedas de caminos más cortos entre usuarios de redes sociales. Las líneas de investigación enfatizan el estudio, diseño y evaluación de algoritmos eficientes (y estructuras de datos asociadas) que permitan aumentar las prestaciones de los sistemas de búsqueda haciendo un uso racional de los recursos de hardware.

Palabras clave: algoritmos eficientes, búsquedas web, grafos, datos masivos.

Contexto

Esta presentación se encuentra enmarcada en el proyecto de investigación “Estrategias y Algoritmos para Problemas de Búsquedas a Gran Escala” (Disposición CD-CB N° 350/19) del Departamento de Ciencias Básicas (UNLu).

Introducción

La cantidad, variedad y velocidad a la que se produce información digital desafía día a día a los sistemas de búsquedas. Además, el número de usuarios que interactúa con diversas plataformas online también se incrementa y, en general, se deben ofrecer respuestas a estos usuarios con restricciones de tiempo. Muchas de estas respuestas están basadas en diferentes formas de *búsquedas*, ya sea sobre documentos, bases de datos estructuradas, grafos, flujos en tiempo real, entre otras. Contar con herramientas eficientes que aborden este tipo de problemas es un requerimiento [24].

La Recuperación de Información es una de las áreas de las Ciencias de la Computación que ofrece un ámbito para la investigación y abordaje de estos problemas, en particular siendo aplicada en el contexto de Datos Masivos (*Big Data* [13]). Así, es posible enfocarse en temas de representación, almacenamiento y procesamiento que permitan ofrecer a los usuarios resultados relevante en tiempo y forma [4]. Esto aplica, por ejemplo, a motores de búsqueda de escala web pero, además, muchas de sus técnicas (como indexación y compresión) se usan también junto con algoritmos que manejan grandes estructuras de datos como los grafos masivos, estructura subyacente las redes sociales.

En general, este tipo de problemas presenta características particulares como la masividad de datos (documentos que forman la web o millones de nodos/aristas de un grafo social), los tiempos de respuesta acotados, la necesidad de estructuras de datos específicas y combinaciones de algoritmos sofisticados que permitan el procesamiento eficiente, considerando también parámetros de eficacia.

Para abordar este tipo de problemas, el área de RI se complementa con técnicas de disciplinas relacionadas, lo que, por un lado, ha complejizado las soluciones pero, por el otro, ha abierto nuevos problemas y temas de investigación y transferencia para abordar. Por ejemplo, se han incorporado a la disciplina estrategias basadas en aprendizaje automático para clasificar o rankear documentos [19] y técnicas de estimación y muestreo (no aleatorio) para seleccionar porciones de un grafo masivo que puedan ser procesadas en un tiempo prudencial [2]. En este contexto, aparecen nuevas oportunidades de soluciones que exploran el *tradeoff* entre la eficacia y la eficiencia y que, además, alguna de éstas puedan derivar en soluciones ingenieriles que sean transferibles a problemas concretos.

En el caso de los motores de búsqueda de escala web, el procesamiento de consultas es uno de los desafíos más difíciles de manejar debido al constante crecimiento tanto en datos como en usuarios [15]. Mantener las prestaciones requiere de algoritmos que combinen diversas técnicas tales como poda dinámica [22], caching [20], ranking [3] (como *Learning to Rank*), compresión de las estructuras de datos [17] o selección de recursos [11], entre otras. Dadas las características de estos problemas también se requieren estrategias de distribución de la carga de trabajo optimizadas para cada caso [24].

Los problemas de eficiencia en búsquedas son continuamente identificados como uno de los más importantes en RI [7] y reciben atención permanentemente, tanto de la academia como de la industria [21]. Por lo tanto, esta propuesta considera este tipo de problemas en escenarios de datos masivos. La idea general de aumentar la eficiencia en las búsquedas permite procesar mayor cantidad de datos con menos recursos, impactando positivamente en el mantenimiento de las infraestructuras de hardware (*datacenters*) en los cuales se ejecutan estos sistemas, disminuyendo costos operativos y mejorando el impacto ambiental.

Líneas de I+D

Las líneas de I+D del grupo se enfocan a mejoras algorítmicas y representaciones basadas en colecciones de documentos (índices invertidos) o elementos relacionados (grafos).

a. Búsquedas a Gran Escala

La estructura de datos comúnmente utilizada para soportar la recuperación eficiente es el índice invertido. De forma simple, está compuesto por un vocabulario (V) con todos los términos extraídos de los documentos y, por cada uno de éstos, una lista de los documentos (*posting list*) donde aparece dicho término junto con información usada para el ranking. En el caso de los algoritmos, la eficiencia está dada por analizar la menor cantidad de documentos para satisfacer una consulta, o bien, poder seleccionar adecuadamente un subconjunto de nodos que puedan responderla.

Algoritmos para Top-k: Existen dos estrategias predominantes para recorrer un índice invertido: DAAT (Document-at-a-Time) y TAAT (Term-at-a-Time). Dado un *query* con n términos ($q = \{t_1, t_2 \dots t_n\}$), DAAT recorre las n listas en paralelo intentando determinar en qué momento detener la evaluación sin llegar al final de todas (*dynamic pruning*). En el caso de TAAT, las listas de los términos se procesan una a la vez, siguiendo la misma idea. No obstante, las dos estrategias predominantes actualmente siguen el criterio DAAT (Maxscore [22] y WAND [5]). En ambos casos, la idea subyacente es contar con un valor umbral (*upper bound*) que permita determinar en qué momento finalizar la evaluación. La evolución sobre éstas consiste en combinarlas con una estructura de índice particular basada en bloques fijos [8] o de longitud variable [14].

En esta línea se trabaja en una extensión de MaxScore en la cual se almacenan múltiples valores umbrales en una estructura similar a una *skip list* [6], dotando al algoritmo de más información para mejorar la eficiencia del procesamiento. Resultados preliminares muestran que la evaluación de documentos se puede reducir hasta un 50 % favoreciendo a consultas con términos muy populares.

Por otro lado, se aborda una propuesta que combina una estructura de datos donde la *posting list* completa está dividida en dos secciones: una primera sección con documentos ordenados por la frecuencia del término en el documento y una segunda sec-

ción ordenada por identificador de documento. Cada sección se recorre usando técnicas DAAT y TAAT según corresponda, usando particiones variables en cada caso, de acuerdo a propiedades estadísticas de las listas.

Búsquedas sobre Flujos: Las búsquedas sobre flujos de información en tiempo real (como en redes sociales) desafían las arquitecturas de procesamiento distribuido, los algoritmos y las estructuras de datos empleadas en los motores de búsqueda. El reto radica en que millones de usuarios (por ejemplo, más de 300 millones en Twitter) publican *documentos cortos* desde diferentes tipos de dispositivos (generalmente, móviles). A su vez, estos documentos deben estar disponibles casi de inmediato, lo que implica una dinámica no presente en las búsquedas web clásicas. Así, este problema puede ser abordado particionando la colección en porciones denominadas *shards*, distribuyendo los mismos en los nodos de procesamiento disponibles y, a la hora de la búsqueda, enviando la consulta solo a un número reducido de nodos que sean los más *adecuados* para resolverla. A este enfoque, se lo denomina “búsquedas selectivas” (*selective search*) e implica el diseño, estudio y desarrollo de métodos y estrategias de partición de la colección, selección de los recursos adecuados, estrategias de *caching* y de fusión de resultados.

En esta línea de investigación se intenta mejorar la eficiencia en la recuperación de información de gran escala sobre flujos de documentos en tiempo real mediante el enfoque de búsquedas selectivas. Los problemas a abordar incluyen los criterios de actualización del índice invertido (particionado), implementar estrategias de caché y definir estrategias de selección de recursos para el algoritmo de búsqueda.

Compresión de Índices: El tamaño de un índice invertido es un factor a considerar no solo en cuanto al medio de almacenamiento persistente, sino también a la posibilidad de que resida en memoria principal (parcial o completamente). Aplicar técnicas de compresión específicas inciden en un menor uso de espacio y son requerimiento en sistemas de gran escala actuales. El objetivo es lograr un *tradeoff* adecuado entre espacio final y tiempo de decodificación acorde a la aplicación particular.

Existe un gran cuerpo de literatura en el tema que presenta enfoques para la compresión de *posting lists* con diferentes propiedades. Una clasificación posible es entre aquellos algoritmos (*codecs*) que se consideran *libres de parámetro* (codifican cada ente-

ro de forma individual) [17, 12, 26] y los *adaptativos a lista* (considera porciones de la lista para comprimir) [26, 12, 25, 23, 16].

En esta línea se aborda la compresión de un índice invertido en bloques pero usando múltiples codecs de acuerdo a cada porción particular y sus propiedades. Por ejemplo, se pueden combinar métodos como *Varint* (eficiente en tiempo) e *Interpolative* (eficiente en espacio). Trabajos preliminares del grupo [10] en los que se propone un esquema *multicompresión* muestran que es posible diseñar una solución de compromiso entre el espacio que ocupa un índice invertido y el tiempo de resolución de consultas, contemplando la partición de las listas tanto de forma uniforme como variable.

b. Procesamiento de Grafos

La representación de datos en estructuras de redes o grafos sigue siendo relevante dada su versatilidad para organizar datos de muy variada naturaleza. Una de las métricas más importantes que se calcula sobre grafos es la distancia entre dos nodos. Como los grafos son utilizados para representar, por ejemplo, a redes sociales (siendo los nodos los usuarios, y las aristas las relaciones), el cálculo de la distancia o la búsqueda del camino más corto (*Shortest Path*) sirve, entre otras cosas, para saber qué tan conectada se encuentra esa red. La masividad de las redes sociales actuales hace que ese cálculo sea muy costoso de realizar online y sobre la totalidad de los nodos.

Estimación de Distancias: El caso de las métricas de centralidad o de distancias en grafos, es un ejemplo donde algoritmos clásicos y que aún al día de hoy se consideran estado del arte en grafos pequeños (en relación a la baja cantidad de aristas y nodos presentes), no responden de manera óptima en entornos distribuidos y para grafos masivos. Un enfoque derivado de lo anterior es el de calcular de forma aproximada el valor de una métrica (por ejemplo, mediante métodos basados en *landmarks* [18]). Una de las líneas de trabajo se basa en métodos de cálculo aproximados que permiten estimar el valor de la distancia reduciendo el grado de error asociado [9] y mejorando los tiempos de respuestas.

Corrección de la Estimación: Complementando la línea anterior, cuando se llega a los límites en la performance de un método de estimación de la distancia es posible aplicar estrategias de corrección

de la misma. Esto se puede realizar, por ejemplo, mediante el ajuste de funciones que modelen la distribución de los errores y permitan corregir la estimación inicial. Continuando una línea de trabajo del grupo [9] se busca obtener la o las funciones que mejor ajusten los valores de distancia estimados con los reales, o incluso combinaciones de funciones según intervalos de error. Se proponen nuevas funciones de ajuste y su aplicación en datasets heterogéneos, intentando establecer si alguna característica estructural del grafo influye en el valor del error. Además, se busca comprender cómo utilizar esta información para una mejor corrección.

Estrategias de Partición y Procesamiento:

Una forma eficiente de abordar el procesamiento de grafos masivos es hacerlo mediante procesamiento distribuido. Para esto, una opción es particionar el grafo en porciones que puedan ser derivadas a diferentes nodos de un cluster. El problema de la partición de un grafo no es nuevo, pero el área se mantiene activa debido a los múltiples factores a considerar [1]. Por último, una problemática que resulta de interés es la de calcular estas métricas en grafos que pueden evolucionar en el tiempo (grafos dinámicos). El desafío aquí consiste en lograr métodos de cálculo que puedan reutilizar procesamientos previos y no tener que realizar cálculos completos (dada la escala del grafo). En esta línea se investigan estrategias de partición que permitan realizar el cálculo de métricas de un grafo que evoluciona en el tiempo minimizando el intercambio de datos.

Resultados y objetivos

Los problemas de eficiencia en las búsquedas sobre datos masivos siguen siendo desafiantes y ofrece múltiples oportunidades para desarrollos científico/tecnológicos [7]. En este sentido, estas líneas de investigación proponen estudiar, desarrollar, evaluar y transferir modelos, algoritmos y técnicas específicas de búsquedas y procesamiento de datos masivos en documentos y grafos. En particular, se propone:

- Diseñar y evaluar variantes de los algoritmos para recuperación *top-k* usando umbrales de poda/corte dinámicos para reducir el número de elementos procesados.
- Desarrollar estructuras de datos (y algoritmos de recorrido) que combinen ordenamientos mixtos de las listas basados en frecuencias

e identificadores de documentos, de forma variable.

- Desarrollar esquemas multicompresión de listas usando combinaciones de codecs y parámetros específicos para diferentes problemas.
- Diseñar y evaluar estrategias de indexación distribuida (y políticas de asignación de documentos) y resolución de consultas para flujos en tiempo real. Esto incluye métodos de selección de nodos y técnicas de *caching*.
- Definir y evaluar estructuras y modelos de cómputo distribuido sobre clusters de *hardware commodity* para problemas de escalabilidad en el cálculo exacto de métricas sobre grafos masivos.
- Considerar y analizar el impacto en el rendimiento y escalabilidad de utilizar estrategias de procesamiento para grafos masivos evolutivos. De forma similar, analizar el efecto de utilizar técnicas de particionado en estos entornos distribuidos.
- Proponer y evaluar nuevas estrategias de estimación de distancias entre nodos de un grafo masivo para problemas de búsqueda. Complementariamente, se propone extender el estudio de estrategias de corrección de la estimación de distancias mediante diferentes familias de funciones que ajusten la distribución de los errores.

Formación de Recursos Humanos

En el marco de estas líneas de investigación se están dirigiendo tres tesis de Licenciatura en Sistemas de Información (UNLu). Además, asociados al proyecto de investigación hay una Beca Estímulo a las Vocaciones Científicas (CIN) y dos pasantías internas en la UNLu.

Referencias

- [1] Z. Abbas, V. Kalavri, P. Carbone, and V. Vlasov. Streaming graph partitioning: an experimental study. *VLDB Endowment*, 11(11), 2018.
- [2] N. K. Ahmed, N. Duffield, T. L. Willke, and R. A. Rossi. On sampling from massive graph streams. *Proc. VLDB Endow.*, 10, 2017.

- [3] Q. Ai, T. Yang, H. Wang, and J. Mao. Unbiased learning to rank: Online or offline? *ACM Trans. Inf. Syst.*, 39(2), Feb. 2021.
- [4] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval: The Concepts and Technology Behind Search*. Addison-Wesley Publishing Company, 2nd edition, 2011.
- [5] A. Z. Broder, D. Carmel, M. Herscovici, A. Soffer, and J. Zien. Efficient query evaluation using a two-level retrieval process. In *Proc. of the 12th Int. Conf. on Information and Knowledge Management, CIKM '03*, 2003.
- [6] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms*. MIT press, 2009.
- [7] J. S. Culpepper, F. Diaz, and M. D. Smucker. Research frontiers in information retrieval (swirl 2018). *SIGIR Forum*, 52(1), 2018.
- [8] S. Ding and T. Suel. Faster top-k document retrieval using block-max indexes. In *Proc. of the 34th Int. Conf. on Research and Development in Information Retrieval*. ACM, 2011.
- [9] A. Giordano and G. Tolosa. Improved landmark-based shortest path length estimation in large graphs with distance correction. In *IEEE/WIC/ACM Int. Joint Conf. on Web Intelligence and Intelligent Agent Tech.*, 2020.
- [10] A. González and G. Tolosa. Multicompresión de grandes listas de enteros para sistemas de búsquedas. In *Simposio Argentino de GRAN-des DATos*. JAIIO, 2020.
- [11] Y. Kim, J. Callan, J. S. Culpepper, and A. Moffat. Efficient distributed selective search. *Information Retrieval Journal*, 20(3), June 2017.
- [12] D. Lemire and L. Boytsov. Decoding billions of integers per second through vectorization. *Softw. Pract. Exper.*, 45(1), Jan. 2015.
- [13] S. Madden. From databases to big data. *IEEE Internet Computing*, 16(3), 2012.
- [14] A. Mallia, G. Ottaviano, E. Porciani, N. Tonello, and R. Venturini. Faster blockmax wand with variable-sized blocks. In *Proc. of the 40th Int. Conf. on Research and Development in Information Retrieval*, 2017.
- [15] A. Mallia and E. Porciani. Faster blockmax wand with longer skipping. In L. Azzopardi, B. Stein, N. Fuhr, P. Mayr, C. Hauff, and D. Hiemstra, editors, *Advances in Information Retrieval*. Int. Publishing, 2019.
- [16] G. Ottaviano and R. Venturini. Partitioned elias-fano indexes. In *37th Int. ACM SIGIR Conf. on Research & Development in Information Retrieval, SIGIR '14*, 2014.
- [17] G. E. Pibiri and R. Venturini. Techniques for inverted index compression. 53(6), 2020.
- [18] M. Potamias, F. Bonchi, C. Castillo, and A. Gionis. Fast shortest path distance estimation in large networks. In *18th ACM Conf. on Information and knowledge management*, 2009.
- [19] N. Tax, S. Bockting, and D. Hiemstra. A cross-benchmark comparison of 87 learning to rank methods. *Information Processing & Management*, 51, 2015.
- [20] G. Tolosa, E. Feuerstein, L. Becchetti, and A. Marchetti-Spaccamela. Performance improvements for search systems using an integrated cache of lists + intersections. *Information Retrieval. Journal*, 20(3):172–198, 2017.
- [21] A. Trotman, J. Degenhardt, and S. Kallumadi. The architecture of ebay search. In *SIGIR 2017 Workshop on eCommerce (ECOM '17)*, 2017.
- [22] H. Turtle and J. Flood. Query evaluation: Strategies and optimizations. *Inf. Process. Manage.*, 31(6), Nov. 1995.
- [23] S. Vigna. Quasi-succinct indices. In *Sixth ACM Int. Conf. on Web Search and Data Mining, WSDM '13*, 2013.
- [24] Y. Wang, L. Wu, L. Luo, Y. Zhang, and G. Dong. Short-term internet search using makes people rely on search engines when facing unknown issues. *PLoS One*, 12(4), 2017.
- [25] H. Yan, S. Ding, and T. Suel. Inverted index compression and query processing with optimized document ordering. In *18th Int. Conf. on World Wide Web*, 2009.
- [26] J. Zhang, X. Long, and T. Suel. Performance of compressed inverted list caching in search engines. In *17th Int. Conf. on World Wide Web*, 2008.

Procesamiento inteligente de grandes volúmenes de información y de flujos de datos

W. Hasperué¹, C. Estrebou¹, G. Camele^{1,3}, P. López², P. Jimbo Santana⁴, G. Reyes Zambrano⁵,
L. Lanzarini¹, A. Fernandez Bariviera⁶

¹ Instituto de Investigación en Informática LIDI*, Facultad de Informática, UNLP, La Plata, Argentina

² Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina

³ Becario postgrado UNLP

⁴ Facultad de Ciencias Administrativas, Universidad Central del Ecuador, Quito, Ecuador

⁵ Facultad de Ciencias Físicas y Matemáticas, Universidad de Guayaquil, Guayaquil, Ecuador

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, cesarest, gcamele, pdlopez, laural}@lidi.info.unlp.edu.ar

prjimbo@uce.edu.ec, gary.reyesz@ug.edu.ec, 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 “Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data” perteneciente al Programa de Incentivos (2018-2021).

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de Big Data y 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.

Las investigaciones correspondientes al procesamiento de datos masivos están enfocadas en dos temas: el estudio y desarrollo de técnicas de reducción de características y el diseño de estrategias que faciliten el procesamiento masivo de datos a usuarios no informáticos. En lo referido a reducción de características, dado que se está trabajando con bases de datos genómicas, el foco está puesto en las estrategias de selección de atributos. El análisis a realizar sobre estos datos tiene por objetivo identificar grupos de genes cuyos patrones de expresión

se encuentren asociados a fenotipos específicos. Por otro lado, se está desarrollando una librería con el objetivo de facilitar el manejo de bases de datos en contextos Big Data. Esto tendrá un impacto directo en el trabajo conjunto que se viene desarrollando junto con la Facultad de Ciencias Veterinarias de la UNLP en relación al análisis de datos de progenie de distintas especies animales.

En cuanto a las investigaciones relacionadas con la Minería de Datos se centran en la construcción de modelos que faciliten la interpretación de los patrones obtenidos y la posterior extracción del conocimiento. En particular el énfasis está puesto en la resolución de dos problemas de sumo interés en distintas áreas: las técnicas de agrupamiento aplicables a flujos de datos y la generación de reglas de clasificación.

Palabras clave: Big Data, Minería de Datos, Reducción de características, Flujos de datos, técnicas de optimización, Redes Neuronales.

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 medidos en la solución de

problemas pertenecientes a distintas áreas. A continuación, se detallan las investigaciones realizadas durante el último año.

1.1. BIG DATA

Reducción de características

En el área de la minería de datos y su aplicación con técnicas de machine learning, los algoritmos de reducción de características juegan un papel muy importante. El objetivo de esos algoritmos es el de reducir las entradas a un tamaño apropiado para su procesamiento y análisis. La reducción de características en una base de datos implica la elección de ciertos atributos y/o creación de nuevos atributos en función de los existentes, tal que, con ese subconjunto de atributos, las “propiedades naturales” de los datos pertenecientes a un dataset no sean alteradas (o lo sean con una pequeña pérdida de información).

Cuando el volumen de información a procesar crece, la ejecución de los algoritmos de selección de atributos convencionales incrementa notablemente su tiempo de procesamiento. Si bien puede considerarse la separación o el análisis independiente de cada atributo, muchas veces resulta útil poder analizar correlaciones entre dos o más variables. Por ello contar con algoritmos que puedan realizar este tipo de análisis en grandes volúmenes de datos resulta de mucho interés.

Actualmente, en el III LIDI se están realizando tareas de investigación con bases de datos genómicas. La medicina genómica ayuda a entender de forma más precisa por qué enfermamos y el peso que tiene en una enfermedad la existencia de defectos genómicos frente a factores medioambientales que pueden desencadenar una enfermedad concreta. En esta área se destaca el análisis de perfiles de expresión génica que tienen como objetivo principal la identificación de un grupo de genes, cuyo patrón de expresión se encuentre asociado a un fenotipo en particular: concepto conocido como *gene signature*.

Un objetivo particular de los *gene signatures* es su utilidad como biomarcador diagnóstico, pronóstico o predictivo de una patología en estudio [1]. Los biomarcadores con valor pronóstico permiten una mejor estratificación de pacientes. En la actualidad la tarea del descubrimiento de nuevos *gene signatures* es realizada mayormente de manera manual por expertos. Es por ello que se están desarrollando estrategias de soporte automático que permita seleccionar aquellos genes que resulten más representativos y por ende ser interpretados como un posible *biomarcador con poder pronóstico*. Las estrategias que se están desarrollando están basadas en algoritmos de selección de características.

En relación a las técnicas de selección de atributos, y en entornos de procesamiento de flujos de datos, se están desarrollando estrategias basadas en técnicas de Aprendizaje Automático que permitan la selección de los atributos más relevantes, brindando resultados en tiempos de respuestas cortos los cuales se adaptan de manera dinámica a la llegada de nuevos datos.

Acceso al procesamiento de Big Data

Uno de los frameworks más utilizados en entornos Big Data es Spark. Este framework brinda facilidad en el análisis de grandes bases de datos ya que ofrece una capa de alto nivel para el procesamiento distribuido y paralelo de los datos. El análisis de los datos se puede realizar a través de su propia API la cual trabaja con sus bases de datos distribuidas internas (RDDs). Sobre esta API han aparecido luego más abstracciones que permiten el análisis usando DataFrames e incluso el lenguaje de consultas SQL.

Cuando los datos a analizar pueden ser almacenados en bases de datos equivalentes a tablas relacionales de cualquier motor de bases de datos, entonces no hay mayores problemas para su uso, pero cuando la información está organizada en forma de árbol, el análisis se vuelve un tanto complicado, ya que para cualquier consulta deben realizarse varias operaciones *Join*. En

particular, se requerirá una operación por cada nivel del árbol que se desea explorar. El uso de múltiples *joins* incrementa la complejidad de la consulta a realizar, aún incluso en SQL y dificulta el análisis posterior de los resultados

En relación a esta línea, uno de los desarrollos que se están llevando a cabo actualmente consiste en la implementación de una librería que permita el fácil tratamiento de los datos cuando estos están organizados y relacionados como un árbol, como por ejemplo, los datos de progenie de cualquier especie animal. Esta librería puede ser incluida en cualquier desarrollo realizado en Spark y su objetivo es permitir a los investigadores de distintas áreas realizar análisis entre individuos de distintas generaciones de una manera sencilla y amena [2].

1.2. MINERÍA DE DATOS

Agrupamiento de flujos de datos

El avance tecnológico ha dado lugar a la generación de datos masivos en tiempo real en áreas muy diversas: consultas en la web, video vigilancia, flujos en redes sociales, redes de sensores, análisis de mercado de valores, supervisión de tráfico, etc.

La minería de flujos de datos posee múltiples aplicaciones en la vida real y ha cobrado fuerza debido a su capacidad de extraer patrones ocultos en dichos flujos. Los algoritmos involucrados deben responder a diversos desafíos para satisfacer restricciones tales como: memoria limitada, paso único de los datos, respuesta en tiempo real, adaptación y clasificación de la deriva de concepto (concept drift) y manejo de datos multidimensionales.

Los desarrollos tecnológicos han cambiado la forma en que la gente almacena, comunica y procesa los datos. Para procesarlos, es preciso utilizar algoritmos capaces de generar, de manera incremental, modelos que incorporen la nueva información de los datos más recientes mientras eliminan los efectos de los datos antiguos.

Por lo tanto, el procesamiento de flujos de datos presenta varios retos. El hecho de reentrenar un modelo con nuevos ejemplos es ineficaz e inadecuado en función del volumen y la velocidad con que se generan.

Un área clave de la minería de flujos de datos es el uso de técnicas de agrupamiento. Dada la necesidad de mantener un modelo dinámico, las estrategias partitivas basadas en centroides requieren de una estructura adicional para conformar cada grupo. En cuanto a la representación interna, deben utilizarse estructuras que resuman el flujo preservando el significado de los datos originales sin la necesidad de guardarlos. Esto puede observarse en [3] donde, además de utilizar una estructura de datos particular, emplea un factor de olvido para controlar el dinamismo del modelo y una estructura en forma de árbol para reunir las distintas partes que conforman un mismo grupo.

En esta dirección se han realizado diferentes investigaciones en el framework Spark [4] [5] [6]. Actualmente, se está trabajando en una adaptación del algoritmo de procesamiento de trayectorias definido en [7], incorporándole una estrategia de procesamiento incremental, capaz de operar con flujos de datos, basada en el tiempo de registro de cada ubicación dentro de la trayectoria original.

Extracción de Reglas de Clasificación

Cuando se busca construir un modelo predictivo para resolver un problema de clasificación a partir de datos estructurados, las reglas de clasificación resultan sumamente atractivas por su capacidad explicativa. La literatura muestra distintas alternativas de generación basadas en árboles de clasificación contruidos total o parcialmente. En ambos casos, el antecedente de cada regla se forma a partir de la conjunción de ítems conformados por el valor que aparece en el nodo y el valor indicado en la rama correspondiente del árbol. Por su parte el consecuente estará dado por la clase mayoritaria de ejemplos presentes en la hoja [8] [9]. Otra forma de construcción consiste en generar las reglas a través de un proceso iterativo que, en cada paso, analiza

cuál es el valor de clase a utilizar en el consecuente, construye de manera incremental el antecedente y retira del conjunto de ejemplos de entrada los correctamente cubiertos. Este proceso de construcción incremental da lugar a una lista de reglas de clasificación que debe ser aplicado en el orden en que fueron generadas [10].

A diferencia de estas alternativas, en el III-LIDI se trabaja en la generación de reglas de clasificación utilizando técnicas de optimización por su habilidad para considerar la conformación del antecedente completo en lugar de hacerlo en pasos sucesivos. Dado que se utilizan estrategias poblacionales computacionalmente costosas, se ha buscado reducir el tiempo necesario para realizar el proceso de búsqueda utilizando alguna técnica de clustering. En esta dirección se han efectuado pruebas con redes neuronales competitivas supervisadas y no supervisadas tanto sobre datos de repositorio como reales [11]. Como resultado de estas investigaciones en julio de 2020 se ha defendido una tesis de doctorado en cotutela entre la UNLP y la URV (España) donde se detallan los resultados obtenidos al aplicar estas estrategias sobre datos reales correspondientes a distintas bases de datos de instituciones financieras en el Ecuador [12].

Además, con el objetivo de ayudar al usuario en el momento de ponderar la credibilidad de la regla, se ha incorporado un factor de confianza que le permite ponderar el riesgo de utilizar dicha regla al momento de tener que tomar una decisión. Esto es algo interesante ya que pueden existir condiciones adicionales, ajenas a los atributos considerados durante el proceso de construcción, que justifiquen este tipo de acciones. Los resultados obtenidos con esta nueva variante pueden consultarse en [13].

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Diseño e implementación de un algoritmo de selección de atributos capaz de operar en batch y en streaming.

- Estudio de algoritmos de selección de atributos para la detección de *gene signatures*.
- Estudio e implementación de técnicas inteligentes en el framework Spark Streaming.
- Análisis de bases de datos con información de progenie en entornos Big Data.
- Modelización de trayectorias espacio-temporales con capacidad para establecer características comunes y detectar situaciones anómalas.
- Estudio de técnicas de clustering dinámico basadas en densidad para modelar trayectorias GPS e identificar sectores de posible congestión.
- Estudio de técnicas de optimización poblacionales y redes neuronales artificiales para la obtención de reglas difusas de tipo IF-THEN.

3. RESULTADOS OBTENIDOS

- Desarrollo de una plataforma pública de acceso web para ejecutar análisis de correlación entre grandes bases de datos de genes y moduladores de expresión.
- Desarrollo de una librería que facilita el desarrollo de pequeñas aplicaciones en Spark, para el tratamiento de bases de datos con información de progenie.
- Diseño e implementación de una técnica de agrupamiento dinámico para flujos de datos basada en densidad.
- Diseño e implementación de un nuevo método de agrupamiento de trayectorias GPS aplicable a la predicción de congestiones vehiculares.
- Desarrollo de un método de obtención de reglas de clasificación difusas que incorpora un factor de confianza que ayuda al usuario a ponderar el riesgo de su uso.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 2 profesores doctores con dedicación exclusiva, 3 tesistas de Doctorado en Cs. Informáticas (1 con beca de postgrado de la UNLP), 2 tesistas de grado y 1 profesor extranjero.

Dentro de los temas involucrados en esta línea de investigación, en los últimos 3 años se han finalizado 2 tesis de doctorado, 1 tesis de especialista y 5 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 3 tesis de doctorado, 2 tesis de especialista y 4 tesinas de grado de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Abba, M. C.; Lacunza, E.; Butti, M.; Aldaz, C.M. Breast cancer biomarker discovery in the functional genomic age: a systematic review of 42 gene expression signatures. *Biomarker Insights*; 5:1-16. 2010.
- [2] López, P. D.; Hasperué, W.; Rearte, R.; De La Sota, R. L. Herramienta informática para el análisis de Progenie. *Innovación y Desarrollo Tecnológico y Social*. La Plata: Universidad Nacional de La Plata. Vol.2 n°1. pp 25-54. ISSN 2683-8559. 2020.
- [3] Barbosa, N.; Travé-Massuyès, L.; Grisales-Palacio, V. DyClee: Dynamic clustering for tracking evolving environments, *Pattern Recognition*. Volume 94, Pages 162-186, ISSN 0031-3203. 2019.
- [4] Molina, R.; Hasperué, W. D3CAS: un Algoritmo de Clustering para el Procesamiento de Flujos de Datos en Spark. XXIV CACIC. UNS. Tandil. 2018
- [5] Molina, R.; Hasperué, W.; Villa Monte, A.; D3CAS: Distributed Clustering Algorithm Applied to Short-Text Stream Processing. *Communications in Computer and Information Science*. : Springer. p211 - 220. ISBN 978-3-030-20786-1. 2019
- [6] Reyes-Zambrano, G.; Lanzarini, L.; Hasperué, W.; Fernández-Bariviera, A. GPS trajectory clustering method for decision making on intelligent transportation systems. *Journal of Intelligent & Fuzzy Systems*, vol. Pre-press, pp. 1-6. ISSN 1064-1246. 2020.
- [7] Liu, L. X.; Song, J. T.; Guan, B.; Wu, Z. X.; He, K. J. Tradbscan: a algorithm of clustering trajectories,” in *Applied Mechanics and Materials*, vol. 121. Trans Tech Publ. pp. 4875–4879. 2012.
- [8] Frank, E.; Witten, I. H. Generating accurate rule sets without global optimization. In *Proceedings of the Fifteenth International Conference on Machine Learning*, pp. 144–151, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc. 1998.
- [9] Quinlan, J. R. *C4.5 Programs for Machine Learning*, San Mateo, CA: Morgan Kaufmann. 1992.
- [10] Clark, P; Niblett, T. The CN2 induction algorithm. *Machine learning journal*. 3:4, pp. 261-283. 1989.
- [11] Jimbo, P.; Lanzarini, L.; Fernandez-Bariviera, A. Fuzzy Classification Rules with FRvarPSO Using Various Methods for Obtaining Fuzzy Sets. *Journal of Advances in Information Technology*, 11(4), 233-240. ISSN 1798-2340. 2020.
- [12] Jimbo, P. Obtención de reglas de clasificación difusas utilizando técnicas de optimización. Caso de estudio Riesgo Crediticio. Tesis de doctorado en Ciencias Informáticas realizada en la UNLP en cotutela con la Universitat Rovira i Virgili (URV) (España). <http://sedici.unlp.edu.ar/handle/10915/101163>. 2020.
- [13] Jimbo, P.; Lanzarini, L. Fernandez-Bariviera, A. FRvarPSO as an alternative

to measure credit risk in financial institutions. Serie Springer Advances in Intelligent Systems and Computing. Sep. 2021.

Propuesta para la migración de datos desde SIGEVA-UNSJ hacia Dspace

Prog. Luis Olguin, Lic. Alejandro Ariza

Instituto de Informática – Secretaría de Ciencia y Técnica UNSJ

Facultad de Ciencias Exactas, Físicas y Naturales Universidad Nacional de San Juan

lolguinunsj@gmail.com; aariza@unsj.edu.ar

RESUMEN

Los procesos de ETL (Extract, Transform, Load) son un método que permite la captura de datos desde una base de datos, el reformateo de estos datos, la limpieza de los mismos para finalmente depositarlos en otra base de datos.

La Universidad Nacional de San Juan (UNSJ), desde la Secretaría de Investigación, se encuentra evaluando la puesta en marcha de un Repositorio Institucional (RI-HURU) en el que se pretende depositar y visibilizar la producción de los docentes/investigadores de la UNSJ.

Esta propuesta apunta a evaluar herramientas de software que permitan realizar una limpieza y normalización de los datos que describen los docentes en SIGEVA-UNSJ para que se utilicen en la *ingesta de items* en el RI-HURU

1. CONTEXTO

Para el trabajo con RI-HURU se ha conformado un equipo interdisciplinario de bibliotecarios e informáticos que elaborarán

un plan de ingesta de información en el repositorio.

El grupo de docentes del área informática lo constituyen los integrantes del proyecto “Evaluación De Visualizaciones Eficientes En Ciencia De Datos” enmarcado en el Laboratorio de Sistemas Inteligentes para Extracción de Conocimiento en Datos Masivos del Instituto de Informática de la Facultad de Exactas Físicas y Naturales de la Universidad Nacional de San Juan –FCEF– UNSJ–. En esta ámbito se habrán de analizar, programar y evaluar métricas de similitud entre documentos que permitan la determinación de parecidos entre múltiples versiones de una misma producción o documento cargado en su momento en el auto archivo generado por diferentes investigadores en el sistema SIGEVA-UNSJ, generando así, un único documento a transferir a la nueva base de datos RI-HURU asequible por la totalidad de los investigadores de la UNSJ.

Mediante el análisis detallado de las tablas que conforman la BBDD SIGEVA-UNSJ se propone aplicar una metodología que permita la migración de la información auto-archivada por los docentes/investigadores de la UNSJ hacia el repositorio HURU de la UNSJ.

2. INTRODUCCIÓN

En la actualidad se define como *dato* (DAMA, 2017) a cualquier representación de entidades o hechos que se haga mediante textos, números, gráficos, imágenes, sonido o video.

Wang (citado en Muñoz Reja, 2018) expresa que los datos conforman la materia prima en un proceso de fabricación de producto de datos que llevarán a tener una información cuando se usen en un determinado contexto.

Galway y Hanks (citado en Muñoz Reja, 2018) señala que existen problemas de tipo operacional cuando se transcriben datos en un sistema que generan datos inexistentes, datos imprecisos o no válidos que sin una política de calidad de datos no son detectados.

El Sistema Integral de Gestión y Evaluación (SIGEVA) es un conjunto de aplicaciones informáticas a las que se puede acceder de forma segura a través de una plataforma web (intranet) y por medio de un navegador de Internet. (<https://sigeva.conicet.gov.ar/>).

Con esta herramienta los docentes/investigadores de la UNSJ generan el autoarchivo de sus datos personales, formación, antecedentes, producción científica, tecnológica y artística.

SIGEVA no incluye un control de *calidad de los datos* que ingresa el operador, *confía en lo que carga cada investigador*, por lo que es un sistema débil en este aspecto, propenso a errores (omisión de datos importantes, campos rellenos con información inválida para el contexto, campos nulos, etc.). Un problema importante es la duplicación de datos ya que permite que dos o más operadores ingresen la misma información (por ej. dos autores de un mismo documento). El trabajo que pretendemos abordar busca extraer los datos de la producción de los docentes/investigadores de la UNSJ (autoarchivados por los mismos docentes) para generar una nueva base de datos sobre las que se puedan generar procesos de mejora, normalización de la información obtenida con el objeto de obtener un lote de nuevos ítems que puedan ser incorporados al repositorio institucional de la UNSJ para su *curatoria*⁸ por parte de profesionales de bibliotecología y finalmente ponerlos a disposición del público usuario del RI HURU.

(López, 2009) indica que la existencia de anomalías e impurezas, en estas grandes cantidades de datos, aparece como un fenómeno que distorsiona los resultados que se obtienen de la interpretación y el análisis de los datos, y provocan como consecuencias

la elevación de los costos y la disminución de los beneficios de su procesamiento.

Pretendemos abordar el tratamiento de anomalías sintácticas de los datos para tratar de lograr una normalización; anomalías de datos con valores ausentes y tratamiento de los registros con duplicidad de atributos, es decir datos que representan la misma entrada dos o más veces.

3. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Si se realiza una búsqueda en internet acerca de herramientas para procesos de ETL se descubre que en el mercado existen gran cantidad de herramientas especializadas generalmente asociadas a sistemas CRM.

Ruiz Borja, Jhoan (2018) expone una serie de herramientas de código abierto e indica que estos softwares utilizan técnicas muy diferentes a la hora de manipular los datos, utilizan distintos tipos de elementos de entrada y salida, y al ser varias herramientas de código abierto, algunas fueron desarrolladas por necesidades propias del programador. Por todo ello hay que tener claro, los tipos de datos, de dónde serán extraídos y qué formato deben de tener estos para poder usarse posteriormente.

Nuestra propuesta es diseñar un procedimiento válido que permita analizar y adecuar los datos extraídos de SIGEVA-UNSJ con el fin de cumplir los objetivos propuestos de normalización de entradas de autores; anomalías de datos con valores ausentes y tratamiento de los registros con duplicidad de atributos entre otros.

Se propone evaluar la herramienta Open Refine para el tratamiento (limpieza, normalización, etc.) de los datos extraídos de las BBDD SIGEVA.

Open Refine (openrefine.org) es una herramienta originalmente creada por Google para el manejo de bases de datos. Nos permite limpiar bases de datos, exportarlas en diferentes formatos, y arreglar y manejar las bases para un mejor uso. Los archivos que podemos importar para trabajar pueden tener las extensiones TSV, CSV, XML, JSON, XLS, e incluso Google Spreadsheets, entre otros. También nos permite transformar archivos de cualquiera de estos formatos a otro. Open Refine funciona como ejecutable sobre cualquier navegador web y está disponible para Windows, Mac y Linux.

4. RESULTADOS OBTENIDOS / ESPERADOS

Objetivos.

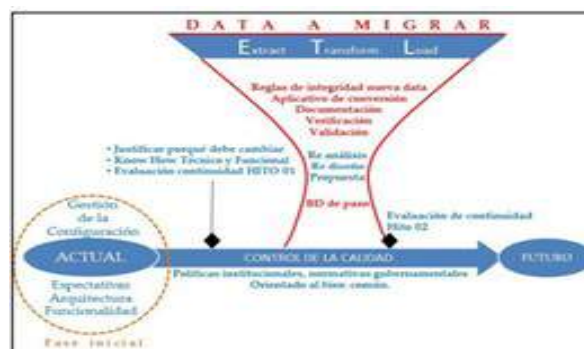
- Recopilar y describir experiencias de migración de datos desde SIGEVA-UNSJ hacia repositorios institucionales.
- Evaluar la herramienta Open refine para el tratamiento de datos (limpieza, detección de duplicados, etc.).
- Diseñar la tabla de correspondencia (crosswalk table) que establezca las relaciones entre los atributos de la BBDD SIGEVA-UNSJ y los metadatos Dublin Core⁵ definidos en RI-HURU

Metodología a aplicar

A fin de ordenar el proceso de migración de datos desde SIGVEVA se adoptará las etapas propuestas por (Pérez Agudo, 2014) comenzando con una Fase Inicial de análisis del aplicativo SIGEVA-UNSJ y el acceso a su BBDD para la obtención de los datos “crudos”.

Luego analizar los datos a migrar (selección de tablas y atributos). Posteriormente modelar el “data stage” donde se almacenarán los datos extraídos para su limpieza y normalización completando el proceso con la

generación del XML compatible con los requerimientos de Dspace para la ingesta de datos.



Fuente (Pérez Agudo, 2014)

5. FORMACIÓN DE RRHH

Olguín, Luis Alberto. Integrante del proyecto. Es actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan, siendo el título de su trabajo de tesis Red de Co-Préstamo en Bibliotecas

Formación de alumnos avanzados de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información, en el uso de herramientas informáticas investigadas.

Formación de docentes de las carreras del Departamento de Informática, integrantes del proyecto, en el tratamiento de datos relativos a la problemática abordada.

Los estudiantes se verán favorecidos por el nuevo conocimiento además de la posibilidad concreta de transferencia al medio. Las temáticas abordadas en el proyecto brindarán a los docentes participantes nuevas líneas de investigación a abordar, así como la

posibilidad de realizar aportes al conocimiento en las áreas relacionadas.

6. REFERENCIAS

Gómez, Emilio (2014). Normalización con OpenRefine. Recuperado de <https://www.altergeosistemas.com/blog/2014/01/12/normalizacion-datos-openrefine/>

López Porrero, B. E. (2009). Limpieza de datos. Santa Clara, Cuba: Editorial Feijóo. Recuperado de <https://elibro.net/es/ereader/bibliotecafranklin/71744?page=6>.

Muñoz-Reja, I. C. Gómez Carretero, A. I. y Gualo Cejudo, F. (2018). Calidad de datos. RA-MA Editorial. Recuperado de <https://elibro.net/es/ereader/bibliotecafranklin/106516?page=28>.

Nash, Jacob, 2016. Desktop Batch Import Workflow for Ingesting Heterogeneous Collections: A Case Study with DSpace 5. D-Lib Magazine. 22. 10.1045/january2016-nash.

Pérez Argudo, Washington Oswaldo (2014). Creación de una metodología que cubra la etapa de análisis del CVDS, a ser aplicada en el proceso de migración de los sistemas informáticos del Grupo Automotriz Quito Motors. Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología.

Universidad de las Fuerzas Armadas ESPE. Extensión Latacunga

Ruiz Borja, Jhoan (2018). Comparación de herramientas ETL de código abierto (Maestría). Universidad Nacional de Colombia, Medellín (Colombia).

Saiful, Amin (2017) .Convert Spreadsheet (Excel or Calc) Data Into DSpace Simple Archive Format Package. Recuperado de <https://www.semanticconsulting.com/blog/2017-12-09-convert-spreadsheet-excel-or-calc-data-dspace-simple-archive-format-package>

Seth van Hooland, (2017). Limpieza de datos con OpenRefine. Recuperado de, <https://doi.org/10.46430/phes0017>

Talent.con, (s.f.).En que consiste un proceso ETL? Recuperado de <https://www.talend.com/es/resources/what-is-etl/>

TutorialAndExample.com, (2009). ETL Tutorial. Recuperado de <https://www.tutorialandexample.com/etl-tutorial/>

Walsh, Maureen. (2013). Batch Loading Collections into DSpace: Using Perl Scripts for Automation and Quality Control. Information Technology and Libraries. 29. 10.6017/ital.v29i3.3137.

Técnicas de Indexación para Bases de Datos Avanzadas

**Norma Herrera, Darío Ruano, Paola Azar, Daniel Welch
Libertad Speranza, María de los Ángeles de la Torre**

Dpto de Informática, Universidad Nacional de San Luis, Argentina

{nherrera, dmruano, epazar, dwelch}@unsl.edu.ar

{libertadsperanza, angelesdlt}@gmail.com

Anabella De Battista, Andrés Pascal

Departamento Ingeniería en Sistemas de Información

FRCU, Universidad Tecnológica Nacional, Entre Ríos, Argentina

{anadebattista, andrespascal22}@gmail.com

Resumen

Las bases de datos han incluido la capacidad de almacenar datos no estructurados tales como imágenes, sonido, texto, video, etc. La problemática de almacenamiento y búsqueda en estos tipos de base de datos difiere de las bases de datos clásicas, dado que no es posible organizarlos en registros y campos, y aun cuando pudiera hacerse, la búsqueda exacta carece de interés. Es en este contexto donde surgen nuevos modelos de bases de datos capaces de cubrir las necesidades de almacenamiento y búsqueda de estas aplicaciones. Nuestro interés se basa en el diseño de índices eficientes para estas nuevas bases de datos.

Contexto

Este trabajo se desarrolla en el ámbito de la línea Técnicas de Indexación para Datos no Estructurados del Proyecto Tecnologías Avanzadas de Bases de Datos (22/F814), cuyo objetivo es realizar investigación básica sobre manejo y recuperación eficiente de información no tradicional.

1 Introducción

Gran parte de la información disponible en formato digital involucra el uso de datos no estructurados tales como imágenes, sonido, texto, video, etc. Debido a que no es posible organizar estos tipos de datos en registros y campos, las tecnologías tradicionales de bases de datos para almacenamiento y búsqueda de información no son adecuadas en este ámbito.

En este contexto surgieron nuevos modelos de bases de datos capaces de cubrir las necesidades de almacenamiento y búsqueda de estas aplicaciones. Nuestro interés se basa en el diseño de técnicas de indexación eficientes para estas nuevas bases de datos. Describimos a continuación los modelos de bases de datos en los que actualmente estamos trabajando.

Bases de Datos de Texto. Una base de datos de texto es un sistema que mantiene una colección grande de texto y que provee acceso rápido y seguro al mismo. Sin pérdida de generalidad, asumiremos que la base de datos de texto es un único texto T que posiblemente se encuentra almacenado en varios archivos.

Una de las búsquedas más comunes en bases de datos de texto es la *búsqueda de un patrón*: el usuario ingresa un string P (*patrón de búsqueda*) y el sistema retorna las ocurrencias

del patrón P en el texto T . Para resolver eficientemente estas búsquedas se hace necesario construir un índice que permita acelerar el proceso de búsqueda.

Mientras que en bases de datos tradicionales los índices ocupan menos espacio que el conjunto de datos indexado, en las bases de datos de texto el índice ocupa más espacio que el texto, pudiendo necesitar de 4 a 20 veces el tamaño del mismo [7, 11]. Una alternativa para reducir el espacio ocupado por el índice es buscar una representación compacta del mismo, manteniendo las facilidades de navegación sobre la estructura [5, 6, 10, 15, 8, 9, 12, 13]. Pero en grandes colecciones de texto, el índice aún comprimido suele ser demasiado grande como para residir en memoria principal. Por esta razón, el estudio de índices comprimidos y en memoria secundaria para texto es un tema de interés

Un *trie de sufijos* [7] es un full-text índice [14]. Básicamente es un *trie* construido sobre el conjunto de todos los sufijos del texto, en el cual cada hoja mantiene el índice del sufijo que esa hoja representa.

Espacios Métricos. El modelo de espacios métricos permite formalizar el concepto de búsqueda por similitud en bases de datos no tradicionales. [2].

Un espacio métrico está formado por un conjunto de objetos \mathcal{X} y una función de distancia d definida entre ellos que mide cuan diferentes son. La base de datos será un subconjunto finito $\mathcal{U} \subseteq \mathcal{X}$.

Una de las consultas más comunes en este modelo de bases de datos es la *búsqueda por rango*. En esta búsqueda dado un elemento $q \in \mathcal{X}$, al que llamaremos *query*, y un radio de tolerancia r , la búsqueda por rango consiste en recuperar los objetos de la base de datos cuya distancia a q no sea mayor que r . Para evitar examinar exhaustivamente la base de datos, se preprocesa la misma por medio de un *algoritmo de indexación* con el objetivo de construir un *índice*, diseñado para ahorrar cálculos en el momento de la búsqueda. En [2] se presenta un desarrollo unificador de las soluciones existentes en la temática.

Bases de Datos Métrico-Temporales. Este modelo permite almacenar objetos no estructurados con tiempos de vigencia asociados y realizar consultas por similitud y por tiempo en forma simultánea. Formalmente un *Espacio Métrico-Temporal* es un par (U, d) , donde $U = O \times N \times N$, y la función d es de la forma $d : O \times O \rightarrow R^+$. Cada elemento $u \in U$ es una tripla (obj, t_i, t_f) , donde obj es un objeto (por ejemplo, una imagen, sonido, cadena, etc) y $[t_i, t_f]$ es el intervalo de vigencia de obj . La función de distancia d , que mide la similitud entre dos objetos, cumple con las propiedades de una métrica (positividad, simetría y desigualdad triangular).

Una consulta métrico-temporales se definen formalmente en símbolos como: $(q, r, t_{iq}, t_{fq})_d = \{o / (o, t_{io}, t_{fo}) \in X \wedge d(q, o) \leq r \wedge (t_{io} \leq t_{fq}) \wedge (t_{iq} \leq t_{fo})\}$ Esta consulta implica buscar todos los objetos o de la base de datos que estén a una distancia a lo más r de q , y cuyo tiempo asociado t se solapa con el tiempo de la consulta.

Varios índices métrico-temporales se han propuesto en este ámbito, todos estos índices fueron desarrollados para ser eficientes en memoria principal. El *Historical-FHQT* (H-FHQT) [4] es un índice métrico-temporal que utiliza tanto la componente métrica como la temporal para resolver eficientemente búsquedas métrico-temporales. Este índice consiste en una lista de instantes válidos donde cada uno contiene un FHQT que indexa a todos los objetos vigentes en dicho instante.

2 Líneas de Investigación

2.1 Espacios Métricos

En este ámbito, avanzamos en el diseño de un prototipo de aplicación de índices métricos al comercio electrónico. El objetivo de este trabajo es agilizar las búsquedas por similitud sobre un conjunto de productos, con el fin de identificar eficiente y eficazmente aquellos productos con una descripción similar a la dada por un usuario.

En este caso, hemos usado como universo de datos un conjunto de productos disponibles en la plataforma Mercado Libre; la función que provee la medida de distancia, es la distancia de edición (o Levenshtein) aplicada al título principal de los productos. Este trabajo es la primera etapa de un proyecto mayor cuyo objetivo final es construir un sistema de recomendación basado en el modelo de espacios métricos.

En esta etapa, hemos usado el enfoque de indexación basado en pivotes para resolver la búsqueda de productos similares. Estudiamos el comportamiento de las búsquedas por rango utilizando las técnicas de selección de pivotes incremental y aleatoria. Los resultados obtenidos hasta el momento se pueden resumir en los siguientes puntos:

- A medida que aumentamos la cantidad de pivotes mejora el comportamiento del índice, pero no en la misma proporción que ocurre en los espacios de prueba generalmente utilizados en esta temática.
- Se pudo observar que la selección incremental no siempre mejora a la selección aleatoria.
- El rango de búsqueda apropiado fue establecido experimentalmente dando como valor adecuado $r = 23$. Este valor difiere de los valores usados con los datasets habitualmente utilizados en esta temática, donde el rango de búsqueda mas apropiado es $r = 1, 2, 3, 4$.
- El universo de datos es bastante más singular, ya que se trata de títulos de productos reales, cuya redacción está a cargo del usuario que publica el producto para su venta y donde la única limitante es el tamaño de ese título (60 caracteres). Esta particularidad tiene como consecuencia un universo de datos variado y heterogéneo. Pudimos establecer experimentalmente que los espacios métricos obtenidos en este ámbito son de alta dimensionalidad.

Como trabajo futuro nos proponemos estudiar el comportamiento de índices basados en parti-

ciones compactas y diseñar además un índice en memoria secundaria, para luego implementar un sistema de recomendación basado en espacios métricos.

2.2 Bases de Datos de Texto

En este ámbito, usando como base el trie de sufijos, nos encontramos abocados al diseño de un índice en memoria secundaria, que además sea eficiente en espacio.

En [16] se propuso y evaluó una representación secuencial del trie de sufijos (que denotaremos con TS_s) que es adecuada para un posterior proceso de paginación. En [17] se presentó una mejora en espacio del trie de sufijos (denotada con TS_{dac}), que consiste en el uso de códigos DAC (Directly Addressable Codes [1]) en algunas componentes de (TS_s). En [3] se propuso una mejora en tiempo del mismo índice (que denotaremos con TS_{dacn}) que implica agregar una componente más a TS_s para mejorar las operaciones que se necesitan realizar durante la búsqueda de un patrón.

De la evaluación experimental de estas tres versiones se concluyó que:

- En tiempo de búsqueda TS_{dacn} es la de mejor desempeño, logrando mejoras en tiempo de un 90% aproximadamente.
- En cuanto al espacio ocupado, TS_{dac} es la mejor opción, usando un 50% menos de espacio que TS_s . Sin embargo TS_{dacn} , con muy poco espacio adicional (0.01%) aproximadamente) logra mejorar substancialmente los tiempos de búsqueda.
- En cuanto a tiempos de construcción TS_s es la de mejor desempeño, pero no es significativa la diferencia de tiempo de construcción con las otras dos versiones.

Por lo expuesto, TS_{dac} es la versión seleccionada para continuar nuestro desarrollo. Nos encontramos trabajando en la implementación de un proceso de paginado de la misma.

2.3 Modelo Métrico-Temporal

En este ámbito nuestro objetivo es el diseño de índices en memoria secundaria. Específicamente estamos trabajando con el índice H-FHQT [4].

En este índice la búsqueda implica primero buscar por igualdad en la lista de instantes de tiempo y luego buscar por similitud sobre el FHQT asociado a cada instante seleccionado. Entonces, para paginar procedemos de la siguiente manera:

- Se pagina la lista de instantes válidos usando un Árbol B^+ .
- Cada FHQT de cada instante de tiempo se pagina de manera separada.

Notar que en el H-FHQT la aridad de los nodos de los FHQT involucrados queda determinada por la cantidad de elementos que hay a cierta distancia de cada uno de los pivotes, por lo que la aridad no se puede modificar para adecuarla a una página de disco, como se hace en un Árbol B^+ . Por esta misma razón el árbol tampoco se puede balancear.

Para paginar cada FHQT procedemos en forma bottom-up tratando de condensar en una única parte un nodo con uno o más de los subárboles que dependen de él. En este proceso de particionado las decisiones se toman en base a la *profundidad* de cada nodo involucrado, donde la profundidad indica la *cantidad de accesos a disco* que deberá realizar el proceso de búsqueda para llegar desde esa parte a una hoja del árbol. Comenzamos asignando cada hoja a una parte con profundidad 1 y luego, en forma bottom-up, procesamos cada uno de los nodos de este árbol r-ario.

Sea x el nodo corriente a procesar, los hijos de x se ordenan de mayor a menor según su profundidad; para aquellos hijos de igual profundidad se ordenan de menor a mayor según su tamaño. Este ordenamiento es solo a los efectos de la paginación, no implica modificar la topología del árbol. Para paginar el subárbol de raíz x , se consideran los siguientes casos:

Caso 1: x y su primer hijo de mayor profundidad d entran en una página de disco:

- Se coloca en una misma parte x y tantos hijos de x como entren en una página, teniendo en cuenta en este proceso el orden ya establecido.
- Se cierran las partes de aquellos hijos que no conforman la nueva parte creada.
- Si todos los hijos de mayor profundidad d se han agregado a la nueva parte creada, se establece que esta nueva parte tiene profundidad d . Si algún hijo de mayor profundidad d es cerrado, la profundidad de la nueva parte se establece en $d + 1$.

Caso 2: x y su primer hijo de mayor profundidad d no entran en una página de disco.

- Se cierran todas las partes hijas y se crea una nueva parte para el nodo corriente.
- La profundidad de la nueva parte creada se establece en $d + 1$, donde d es el máximo de las profundidades de los hijos.

En este proceso, cerrar una parte significa grabarla en disco y reemplazarla en el árbol original por una hoja que contiene el número de página donde esa parte se grabó. Actualmente estamos trabajando en la implementación de este proceso de paginado.

3 Resultados Esperados

Se espera obtener índices eficientes, tanto en espacio como en tiempo, para el procesamiento de consultas en bases de datos textuales, en bases de datos métricas y en bases de datos métrico temporales. Los mismos serán evaluados tanto analíticamente como empíricamente.

4 Recursos Humanos

El trabajo desarrollado en esta línea forma parte del desarrollo de dos Trabajos Finales de Licenciatura, dos Tesis de Maestría y una Tesis de Doctorado, todas ellas en el ámbito de la Universidad Nacional de San Luis.

Bibliografía

- [1] N. Brisaboa, S. Ladra, and G. Navarro. Directly addressable variable-length codes. In *Proc. 16th International Symposium on String Processing and Information Retrieval (SPIRE)*, LNCS 5721, pages 122–130. Springer, 2009.
- [2] E. Chávez, G. Navarro, R. Baeza-Yates, and J.L. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [3] J. Cornejo, D. Ruano, and N. Herrera. Una mejora en tiempo del trie de sufijos. In *Congreso Argentino de Ciencias de la Computación*, Rio Cuarto, Córdoba, Argentina, 2019.
- [4] 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, Argentina, 2007.
- [5] P. Ferragina and G. Manzini. Indexing compressed text. *J. ACM*, 52(4):552–581, 2005.
- [6] T. Gagie and G. Navarro. *Compressed Indexes for Repetitive Textual Datasets*. Springer, 2019.
- [7] G. H. Gonnet, R. Baeza-Yates, and T. Snider. *New indices for text: PAT trees and PAT arrays*. Prentice Hall, New Jersey, 1992.
- [8] R. González and G. Navarro. Compressed text indexes with fast locate. In *Proc. 18th Annual Symposium on Combinatorial Pattern Matching (CPM)*, LNCS 4580, pages 216–227, 2007.
- [9] R. González, G. Navarro, and H. Ferrada. Locally compressed suffix arrays. *ACM Journal of Experimental Algorithmics*, 19(1):article 1, 2014.
- [10] Roberto GROSSI and Jeffrey Scott VITTER. Compressed suffix arrays and suffix trees with applications to text indexing and string matching. *SIAM journal on computing*, 35(2):378–407, 2006.
- [11] U. Manber and G. Myers. Suffix arrays: A new method for on-line string searches. *SIAM Journal of Computing*, 22(5):935–948, 1993.
- [12] E. Moura, G. Navarro, N. Ziviani, and R. Baeza-Yates. Fast and flexible word searching on compressed text. *ACM TrFansactions on Information Systems (TOIS)*, 18(2):113–139, 2000.
- [13] J. I. Munro, G. Navarro, and Y. Nekrich. Fast compressed self-indexes with deterministic linear-time construction. *Algorithmica*, 82(2):316–337, 2020.
- [14] G. Navarro and V. Mäkinen. Compressed full-text indexes. *ACM Computing Surveys*, 39(1):article 2, 2007.
- [15] G. Navarro and K. Sadakane. *Compressed Tree Representations*. Springer, 2nd edition, 2015.
- [16] D. Ruano and N. Herrera. Representación secuencial de un trie de sufijos. In *XX Congreso Argentino de Ciencias de la Computación*, Buenos Aires, Argentina, 2014.
- [17] D. Ruano and N. Herrera. Indexando bases de datos textuales: Una representación compacta del trie de sufijos. In *Congreso Nacional de Ingeniería Informática / Sistemas de Información*, Buenos Aires, Argentina, 2015.

Web Mining y Text Mining: enfoques avanzados para analizar el contenido de grandes cantidades de información

José Federico Medrano¹, Valeria Barriento²

jfmedrano@fi.unju.edu.ar, barriento.valeria@gmail.com

¹VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy (UNJu) - Ítalo Palanca 10, +54 (388) 4221587

² Escuela Superior de Ciencias Jurídicas y Políticas / Universidad Nacional de Jujuy (UNJu)

RESUMEN

A medida que aumenta la cantidad de información contenida y disponible en la web, analizarla, descubrir patrones y conocimiento informativo demanda grandes cantidades de tiempo. Los buscadores y los motores de bases de datos pueden facilitar en parte la tarea de encontrar contenido adecuado, pero en sitios web grandes, donde los resultados de las búsquedas se cuentan por miles o decenas de miles es necesario aplicar enfoques avanzados que permitan relacionar el contenido buscado de algún modo. Este trabajo plantea la aplicación de técnicas de *Web Mining* y *Text Mining* para procesar grandes cantidades de información de sitios web de noticias para ofrecer contenido relevante y relacionado a partir de una búsqueda inicial. Una de las técnicas a emplear será el modelado temático, que permitirá por un lado conocer los distintos temas o tópicos que tratan estas noticias y por otro lado, una vez identificados los conjuntos de temas, hallar las diversas interrelaciones entre ellos. Esto permitirá describir y analizar de un modo objetivo la información ofrecida por este tipo de portales. Del mismo modo, este trabajo también plantea el estudio y

análisis de sitios web de avisos clasificados, de manera de caracterizar por un lado la oferta de inmuebles y por otro la demanda de perfiles para distintos puestos de trabajo.

Palabras clave: *Web Mining; Text Mining; Procesamiento del Lenguaje Natural; Topic Modeling; Recuperación de Información*

CONTEXTO

La línea de investigación aquí presentada se encuadra dentro del Proyecto Bidual 2020-2021 D/B035 denominado “*Agentes Inteligentes para Recuperación de Información y Analítica Visual en Big Data*”, aprobado y financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales de la Universidad Nacional de Jujuy (SeCTER – UNJu).

Continuando con lo expuesto en (Medrano, 2020), este proyecto es llevado a cabo por el grupo de investigación Visualización y Recuperación Avanzada de Información

(VRAIn) de la Facultad de Ingeniería de la UNJu.

1. INTRODUCCIÓN

El *Text Mining* o Minería de Texto, un tipo particular de minería de datos, tiene como objetivo extraer conocimientos útiles como relaciones, patrones y tendencias de datos no estructurados o semiestructurados, por ejemplo, documentos de texto (Feldman & Sanger, 2006). El proceso principal en la minería de textos es transformar el texto en datos numéricos utilizando métodos estadísticos para extraer el contenido textual en una matriz organizada documento-término, que abarca las siguientes dos dimensiones: las palabras (o términos, compuestos por n palabras) y los documentos (Moro, Pires, Rita, & Cortez, 2019). Estas técnicas aportan una gran ventaja al momento de analizar corpus textuales de miles de registros, puesto que automatizan parte del proceso de extracción de nuevo conocimiento, facilitando los análisis y la obtención de conclusiones de manera más sencilla.

La minería de textos extrae información relevante que se encuentra contenida en los diferentes formatos en los que se puede encontrar un conjunto textual como pueden ser: artículos de revistas, páginas web, entre otros (Ariza-Colpas, Oviedo-Carrascal, & De-la-hoz-Franco, 2019). En este sentido ha sido ampliamente utilizada en diversas áreas del conocimiento como la biomedicina (Kim & Delen, 2018), análisis de sentimiento y opiniones (Liu, 2012), recuperación de información (Meystre, Savova, Kipper-Schuler, & Hurdle, 2008), ciencias sociales (Ignatow & Mihalcea, 2016).

El análisis y procesamiento de grandes cantidades de texto no es una tarea sencilla; encontrar relaciones o nuevo conocimiento, a veces oculto entre miles o millones de registros, se torna impracticable sin recurrir a modernas técnicas y algoritmos de Aprendizaje Automático y Procesamiento del Lenguaje Natural. Al respecto, una de las técnicas muy populares y que ha dado buenos resultados al analizar conjuntos enormes de datos es el modelado de temas o *topic modeling*. Un tipo de modelo estadístico para descubrir los “temas” abstractos que ocurren en una colección de documentos. El modelado de temas es una herramienta de minería de texto (*text mining*) de uso frecuente para el descubrimiento de estructuras semánticas ocultas en un cuerpo de texto.

Como lo indica (García-Marco, Figuerola, & Pinto, 2020), la misión del modelado de temas consiste, en identificar el conjunto de temas de la colección documental y en establecer la proporción de cada tema en cada documento. Estas operaciones se basan en la coocurrencia de palabras en los mismos documentos; y permiten establecer conjuntos de palabras definitorias, con mayor o menor peso, de cada tema. La presencia de unas u otras palabras en cada documento permite también estimar el porcentaje o proporción que cada tema juega en el contenido de ese documento.

La minería web es en realidad un área de minería de datos relacionada con la información disponible en internet. Es un concepto de extracción de datos informativos disponibles en páginas web (Kumar & Singh, 2016; Mughal, 2018).

Para la extracción de datos de una página web se emplean diferentes herramientas y algoritmos, la información a extraer puede ser

la estructura (hipervínculos), el uso (páginas visitadas, uso de datos), o el contenido (documento de texto, páginas), siendo este último el más empleado. Este proceso involucra áreas como la Recuperación de Información (Baeza-Yates & Ribeiro-Neto, 1999) y el Procesamiento del Lenguaje Natural (PLN), para poder recuperar los registros necesarios y luego procesarlos de un modo adecuado de acuerdo a la necesidad del problema; por ello como se mencionó previamente, el modelado temático, una técnica de PLN, permitirá la identificación de patrones y relaciones dentro del conjunto textual.

Uno de los objetivos de este trabajo es la descripción avanzada o mejorada del contenido de grandes sitios web. Por ejemplo, para portales de noticias como La Nación¹, resultados de búsquedas como “violencia de género”, “Alberto Fernández”, “dólar” o “vacuna covid”, por citar algunos ejemplos, ofrecen miles, decenas de miles y en algunos casos centenas de miles de resultados. Para poder procesar esta enorme cantidad de información y caracterizar de manera objetiva el contenido textual de las mismas, es necesario recurrir a las técnicas avanzadas que se mencionaron. Puesto que automatizando gran parte de los procesos de recolección, limpieza y procesamiento de datos, se podrá destinar gran parte del tiempo para el análisis finito de las relaciones identificadas.

Así mismo, este trabajo se plantea caracterizar la demanda de empleo y la oferta de inmuebles a partir de la recolección de avisos clasificados. De este modo se podrán comparar o establecer el precio a un inmueble de acuerdo a las características propias y del conjunto de viviendas circundantes, o se

podrá conocer, identificar y monitorear los requerimientos para un perfil de empleo determinado (Karakatsanis, y otros, 2017; Papoutsoglou, Mittas, & Angelis, 2017).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La presente investigación se enfocará en dos aspectos claves:

- Extracción automática de grandes cantidades de información de sitios web.
- Empleo de técnicas de PLN para encontrar información relevante.

Debido al dinamismo de la web, el contenido cambia, se actualiza, se agrega o elimina constantemente, por ello para analizar el contenido de un sitio web sería necesario contar con una instantánea completa que permita tener un vistazo de un momento determinado. Es aquí donde cobra importancia el *web scraping*, una técnica de Recuperación de Información empleada para extraer información de sitios web donde no se cuenta con una API o cuando los datos disponibles son escasos o limitados. De este modo, una vez extraídos los datos relevantes, se conforma un *dataset* para ser procesado y analizado para hallar patrones o tendencias que permitan relacionar distintos conjuntos temáticos o agrupaciones de registros.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera en una primera instancia construir un *crawler* específico para un portal para poder extraer y recolectar toda la información necesaria. Puesto que los sitios web objeto de estudio no disponen de mecanismos que

¹ <https://www.lanacion.com.ar/>

permitan recolectar la información por medio de una API, esta tarea se llevará a cabo mediante *web scraping*.

Por otro lado, se espera caracterizar la información recolectada, empleando un modelado de temas y visualizando los datos con librerías especializadas en la materia.

4. FORMACIÓN DE RECURSOS HUMANOS

El Equipo de Trabajo está conformado por docentes investigadores de la Universidad Nacional de Jujuy. Los mismos llevan adelante esta línea de investigación desde hace años. Cada año se incorporan al proyecto alumnos avanzados de distintas carreras, quienes trabajan en temas relacionados con las temáticas planteadas. Del mismo modo, los integrantes del equipo participan en el dictado de asignaturas/cursos de grado y postgrado de la UNLP, UNJu y UCSEDASS.

5. BIBLIOGRAFÍA

Ariza-Colpas, P., Oviedo-Carrascal, A., & De-la-hoz-Franco, E. (2019). Using K-Means Algorithm for Description Analysis of Text in RSS News Format. *International Conference on Data Mining and Big Data*, (págs. 162-169). Obtenido de <https://academic.microsoft.com/paper/2963086233>

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Obtenido de <https://academic.microsoft.com/paper/1660390307>

Feldman, R., & Sanger, J. (2006). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Obtenido de

<https://academic.microsoft.com/paper/2024228866>

García-Marco, F. J., Figuerola, C., & Pinto, M. (2020). Análisis de la evolución temática de la investigación sobre Información y Documentación en español en la base de datos LISA mediante modelado temático (1978-2019). *Profesional de la información*, 29(4).

Ignatow, G., & Mihalcea, R. (2016). *Text Mining: A Guidebook for the Social Sciences*. Obtenido de <https://academic.microsoft.com/paper/2748167422>

Karakatsanis, I., AlKhader, W., MacCrorry, F., Alibasic, A., Omar, M. A., Aung, Z., & Woon, W. L. (2017). Data mining approach to monitoring the requirements of the job market: A case study. *Information Systems*, 65, 1-6.

Kim, Y.-M., & Delen, D. (2018). Medical informatics research trend analysis: A text mining approach. *Health Informatics Journal*, 24(4), 432-452. Obtenido de <https://academic.microsoft.com/paper/2772572665>

Kumar, A., & Singh, R. K. (2016). Web mining overview, techniques, tools and applications: A survey. *International Research Journal of Engineering and Technology (IRJET)*, 3(12), 1543-1547.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Obtenido de <https://academic.microsoft.com/paper/2108646579>

Medrano, J. F. (2020). Agentes inteligentes para recuperación de información y analítica visual en big data. *XXII*

Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz).

- Meystre, S., Savova, G., Kipper-Schuler, K., & Hurdle, J. (2008). Extracting information from textual documents in the electronic health record: a review of recent research. *Yearb Med Inform, 17*(1), 128-144. Obtenido de <https://academic.microsoft.com/paper/2114388055>
- Moro, S., Pires, G., Rita, P., & Cortez, P. (2019). A text mining and topic modelling perspective of ethnic marketing research. *Journal of Business Research, 103*, 275-285. Obtenido de <https://academic.microsoft.com/paper/2736934374>
- Mughal, M. J. (2018). Data mining: web data mining techniques, tools and algorithms: an overview. *Information Retrieval, 9*(6).
- Papoutsoglou, M., Mittas, N., & Angelis, L. (2017). Mining people analytics from stackoverflow job advertisements. *43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, (págs. 108-115).

Computación Gráfica, Imágenes y Visualización

Análisis de datos biométricos en huellas dactilares y su relación con la DM2 a través de ciencia de datos y procesamiento de imágenes.

María Silvia Vera Laceiras¹ Dana K. Urribarri³ Lucas A. Sosa Valerga²
Marta P. Del Valle² Silvia M. Castro³

¹Facultad de Exactas Químicas y Naturales. Universidad Nacional de Misiones. Posadas, Misiones UNNE UTN.

² Dpto. de Ciencias de la Salud, Universidad Nacional del Sur, Bahía Blanca, Bs. As.

³ VyGLab(UNS-CICPBA) ICIC(UNS-CONICET) Dpto. de Cs. e Ing. de la Computación
Universidad Nacional del Sur, Bahía Blanca, Bs. As.
vlhsilvia@gmail.com, {smc.dku}@cs.uns.edu.ar

Resumen:

La diabetes mellitus tipo 2 (DM2) es una enfermedad metabólica que se asocia con el desarrollo de complicaciones macrovasculares, microvasculares y neuropáticas que impactan seriamente en la expectativa y calidad de vida del paciente.

La finalidad de esta línea de investigación es analizar la existencia de la relación entre simetrías/asimetrías y la DM2 proponer un método alternativo, de bajo costo y no invasivo para detectar posible riesgo de DM2, a través del análisis y clasificación de datos biométricos (huellas dactilares) con técnicas de procesamiento digital de imágenes.

La existencia de una relación entre las huellas digitales y la DM2 agregaría un marcador que colaboraría con la identificación de individuos en riesgo y, por lo tanto, en la intervención temprana para evitar o retrasar el inicio de la enfermedad.

Palabras claves: Procesamiento de imágenes, reconocimiento de patrones, visualización, simetría en dermatoglifos.

Contexto:

El presente trabajo se lleva a cabo entre docente investigador del Instituto de Investigación, diseño e innovación y de la Facultad de ciencias exactas, químicas y naturales de la UNaM y docentes investigadores del Departamento de Ciencias de la Salud de la Universidad Nacional del Sur y docentes investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur; además, se enmarca dentro del proyecto Detección de riesgo de Diabetes Mellitus tipo 2 a través de las huellas dactilares (PGI 24/ZN35) financiado por la Secretaría General de Ciencia y Tecnología de la Universidad Nacional del Sur.

Introducción:

El estudio de los dermatoglifos [1] es el estudio de los patrones que se forman en la palma de la mano, los dedos y la planta del pie se conoce como dermatoglifía.

Esta ciencia se basa en estudios embriológicos que han determinado que el desarrollo de las huellas dactilares se manifiesta durante la gestación temprana (6-17 semanas) [2].

Dado que tanto la diabetes [3] como las huellas dactilares están influenciadas por la genética y el medio ambiental gestacional [1,4,5,6,7,8,9], se han realizado varios estudios de detección temprana de Diabetes [3] y vinculando además diferentes características de las huellas dactilares con el riesgo de padecer DM2.

La finalidad de esta línea de investigación es analizar la existencia de la relación entre las huellas digitales y la DM2 con el objetivo de proponer un método alternativo, de bajo costo y no invasivo para detectar posible riesgo de DM2. Esto agregaría un marcador para la identificación temprana de individuos en riesgo.

El aumento del sobrepeso, la obesidad y la inactividad física, sumados al crecimiento y envejecimiento general de la población, han convertido a la diabetes en una epidemia a nivel mundial. Hay varios estudios que relacionan los dermatoglifos en general, y las huellas dactilares en particular, con diversos trastornos médicos [10,11,12]. La diabetes mellitus tipo 2 (DM2) es una enfermedad metabólica que se asocia con el desarrollo de complicaciones macrovasculares, microvasculares y neuropáticas que impactan seriamente en la expectativa y calidad de vida del paciente. Si bien, el 76.8% de la población del país se ha realizado alguna vez un control de glucemia [13], un tercio de las personas con DM2 tienen complicaciones en el momento del diagnóstico y la duración de la hiperglucemia está directamente

relacionada con la aparición de estas manifestaciones [3,13]. La intervención en el estilo de vida puede retrasar o prevenir la DM2 en aquellos sujetos de alto riesgo de desarrollar la enfermedad [14]. Por lo tanto, detectar cuanto antes a los sujetos en riesgo de desarrollar diabetes podría favorecer la implementación de estrategias preventivas para evitar o retrasar el inicio de la enfermedad y sus complicaciones relacionadas [15,16].

Los modelos actuales para determinar el riesgo de DM2 son útiles en algunas situaciones [17]; sin embargo, en su mayoría, incluyen características fenotípicas (alto índice de masa corporal y circunferencia de la cintura) que a menudo se manifiestan después del desarrollo de hiperglucemia y/o después del momento oportuno para intervenir con el éxito deseado. Los modelos de riesgo para la DM2 que incorporan variables genéticas pueden aportar información sobre el riesgo a desarrollar la enfermedad [18], pero aún los polimorfismos genéticos que se han identificado sólo representan una pequeña proporción de la población diabética (10% de los casos), además de constituir métodos costosos y de difícil aplicación a la población en general. Desde hace varias décadas se está evaluando extensamente la relación entre los dermatoglifos y varios tipos de condiciones médicas. El objetivo es poder utilizar a los dermatoglifos como una herramienta de diagnóstico no invasiva para detectar y predecir tales condiciones, especialmente en centros médicos de poca tecnología.

Se han realizado estudios evaluando la relación entre varias patologías médicas y los dermatoglifos: hipertensión [19, 20], caries [21], maloclusión dental [22] enfermedades de las arterias coronarias [23], enfermedades del riñón [24] y

esquizofrenia [25], entre otras. Estos estudios se basan en la hipótesis de que las huellas dactilares manifiestan características genéticas porque se forma del mismo tejido embrionario y durante el mismo período que las crestas epidérmicas [22].

Por otro lado, basándose en que las huellas dactilares no cambian a lo largo de los años y son únicas para cada persona, se ha realizado mucho trabajo en el reconocimiento biométrico a través de las huellas [26, 27, 28, 29, 30, 31, 32, 33, 34].

Líneas de Investigación y Desarrollo:

Este trabajo se basa en la hipótesis que, dado que la predisposición a diabetes y las huellas dactilares se originan en la misma etapa embrionaria, es posible detectar el riesgo de DM2 de un individuo a través de las características de sus huellas dactilares. En los últimos años, basándose en tal hipótesis, se han analizado diferentes marcadores de la diabetes: la simetría/asimetría entre huellas de diferentes dedos [16], la predominancia de patrones particulares [17, 14] y relaciones entre la cantidad de valle y crestas de diferentes dedos [13, 16]. Si bien estos trabajos han arrojado resultados positivos e indicado cierta relación entre las huellas y la incidencia de la enfermedad, no fueron concluyentes.

Esta línea de investigación tiene como objetivo utilizar diferentes técnicas de procesamiento de imágenes, detección de patrones y visualización por computadora para analizar si existe la relación de simetría/asimetría de las huellas dactilares con la DM2.

Resultados obtenidos/esperados:

La revisión de la literatura acerca de la temática presentada en estos artículos nos determina las bases dejando la posibilidad

de analizar e identificar si existe una relación entre simetrías/asimetrías en huellas dactilares y el riesgo de DM2. Se espera encontrar una forma de determinar la simetría o factor de simetría entre huellas, y un clasificador de huellas de personas con/sin riesgo de diabetes.

Formación de Recursos Humanos:

Dentro de esta línea de investigación se desarrolla una tesis de posgrado, correspondiente a la carrera de Doctorado en Tecnologías de la Información dictada por la Universidad Nacional del Nordeste (UNNE), la Universidad Nacional de Misiones (UNaM) y la Universidad Tecnológica Nacional (UTN).

Bibliografía:

- [1] G Morizon, M Aspillaga - Dermatoglifos. Revista chilena de pediatría, 1977 - scielo.conicyt.cl
- [2] Harris MI, Eastman RC. "Early detection of undiagnosed diabetes mellitus: a US perspective". *Diabetes Metab Res Rev.* 2000;16:230-236.
- [3] Engelgau MM, Narayan KVM, Herman WH. Screening for type 2 diabetes. *Diabetes Care.* 2000;23:1563-1580.
- [4] G. Floris, E. Marini & G. Mulliri. Dermatoglyphic asymmetry and diversity in some pathologies. *International Journal of Anthropology*, 20(1):139, Jan 2005.
- [5] Henry Kahn, Mariaelisa Graff, Aryeh Stein & L.H. Lumey. A fingerprint marker from early gestation associated with diabetes in middle age: The dutch hunger winter families study. *International journal of epidemiology*, 38:101-9, 09 2008.
- [6] Pushpa Burute, S.N. Kazi, Vatsalaswamy & Vasanti Arole. Role

- of Dermatoglyphic Fingertip Patterns in the prediction of Maturity Onset Diabetes Mellitus (Type II). *IOSR Journal of Dental and Medical Sciences*, 8:01–05, 01 2013.
- [7] NS Rakate and BR Zambare. Comparative Study of the Dermatoglyphic Patterns in Type II Diabetes Mellitus Patients with Non Diabetics. *International Journal of Medical Research & Health Sciences*, 2(4), 10-12 2013.
- [8] Molly Morris, Bjoern Ludwar, Evan Swingle, Mahelet Mamo & Jay Shubrook. A New Method to Assess Asymmetry in Fingerprints Could Be Used as an Early Indicator of Type 2 Diabetes Mellitus. *Journal of Diabetes Science and Technology*, 10, 02 2016
- [9] P Manjusha, S Sudha, P M Shameena, Chandni Radhakrishnan, Sujatha Varma & Deepak Pandiar. Analysis of lip print and fingerprint patterns in patients with type II diabetes mellitus. *Journal of Oral and Maxillofacial Pathology*, 21:309, 05 2017.
- [10] I. C. Fuller. Dermatoglyphics: a diagnostic aid? *Journal of Medical Genetics*, 10(2):165–169, Jun 1973
- [11] Eugene Kobylansky, K Yakovenko, M Bejerano, Bat-Miriam & M Katznelson. Relationship between genetic anomalies of different levels and deviation in dermatoglyphic traits. *International Journal of Anthropology*, 20:85–109, 01 1999.
- [12] Prashant Chaware, Shiva Singh, Alkesh Khurana, Abhijit Pakhare, Hemant Arode & Apoorva Tripathi. Study of Fingerprint Patterns to Evaluate the Role of Dermatoglyphics in Early Detection of Bronchial Asthma. *The Journal of Natural Science, Biology and Medicine*, 7:43–46, 01 2016.
- [13] Tabák AG, Jokela M, Akbaraly TN, Brunner EJ, Kivimäki M, Witte DR. Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet*. 2009;373(9682):2215-2221.
- [14] Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. *Diabetes Care*. 002;25(12):2165-2171.
- [15] Laaksonen DE, Lindström J, Lakka TA. Physical activity in the prevention of T2DM: the Finnish diabetes prevention study. *Diabetes*. 2005;54:158-165.
- [16] Sanghera DK, Blackett PR. Type 2 diabetes genetics: beyond GWAS. *J Diabetes Metab*. 2012;3(198):6948.
- [17] Kobylansky E, Bejerano M, Katznelson MBM, Malkin I. Relationship between genetic anomalies of different levels and deviations in dermatoglyphic traits. *Stud Hist Anthropol*. 2006;4:61-121.
- [18] A.W. Drong, Cecilia Lindgren & M.I. McCarthy. The Genetic and Epigenetic Basis of Type 2 Diabetes and Obesity. *Clinical pharmacology and therapeutics*, 92:707–715, 10 2012.
- [19] John T. Ramapuram, Deepak Madi & Ravi Raj Singh Chouhan. A "Handy" tool for hypertension prediction: Dermatoglyphics. *Indian Heart Journal*, 2018
- [20] Ganesh Chakravathy P., Arun Shirali, K. Nithyananda Chowta,
- [21] Shaik Mohammed Asif, Dara Bg Babu & Shaik Naheeda. Utility of

- Dermatoglyphic Pattern in Prediction of Caries in Children of Telangana Region, India. *The Journal of Contemporary Dental Practice*, 18 6:490–496, 2017.
- [22] N. Eslami, A. Jahanbin, A. Ezzati, E. Banihashemi & H. Kianifar. Can Dermatoglyphics Be Used as a Marker for Predicting Future Malocclusions? *Electron Physician*, 8(2):1927–1932, Feb 2016.
- [23] Hong Lu, Wenli Qian, Zhi Geng, Youjing Sheng, Haochen Yu, Zhanbing Ma & Zhenghao Huo. Dermatoglyphs in Coronary Artery Disease Among Ningxia Population of North China. *Journal of clinical and diagnostic research: JCDR*, 9(12):AC01–4, 2015.
- [24] Buddhika T. B. Wijerathne, Robert J. Meier, Sujatha S. Salgado & Suneth B. Agampodi. Dermatoglyphics in kidney diseases: a review. *SpringerPlus*, 5(1):290, Mar 2016
- [25] Aastha, Bina Isaac & Santhosh K. Mathangi. Dermatoglyphic patterns as possible predictor of treatment resistance in schizophrenia. *Journal of the Anatomical Society of India*, 63(2):110 – 116, 2014.
- [26] Kalle Karu and Anil K. Jain. Fingerprint classification. *Pattern Recognition*, 29(3):389 – 404, 1996.
- [27] Alessandra Lumini, Dario Maio & Davide Maltoni. Continuous versus exclusive classification for fingerprint retrieval. *Pattern Recognition Letters*, 18(10):1027 – 1034, 1997.
- [28] R. Cappelli, A. Lumini, D. Maio & D. Maltoni. Fingerprint classification by directional image partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):402–421, May 1999.
- [29] A. K. Jain, S. Prabhakar & Lin Hong. A multichannel approach to fingerprint classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):348–359, April 1999.
- [30] L. C. Ern & G. Sulong. Fingerprint classification approaches: an overview. In *Proceedings of the Sixth International Symposium on Signal Processing and its Applications (Cat.No.01EX467)*, volume 1, pages 347–350, Aug 2001.
- [31] Bir Bhanu & Xuejun Tan. Fingerprint indexing based on novel features of minutiae triplets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):616–622, May 2003.
- [32] Davide Maltoni, Dario Maio, Anil K. Jain & Salil Prabhakar. *Handbook of Fingerprint Recognition*. Springer Professional Computing, 2003.
- [33] Jun Li, Wei-Yun Yau & Han Wang. Continuous Fingerprints Classification by Symmetrical Filters. In *Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security, ASIACCS '06*, pages 357–357, New York, NY, USA, 2006. ACM.
- [34] K. C. Leung and C. H. Leung. Improvement of Fingerprint Retrieval by a Statistical Classifier. *IEEE Transactions on Information Forensics and Security*, 6(1):59–69, March 2011.

ANÁLISIS VISUAL DE DATOS MULTIDIMENSIONALES

Antonella S. Antonini¹, M. Luján Ganuza¹, Florencia Gargiulo², Gabriela Ferracutti², Ernesto A. Bjerg², Silvia M. Castro¹, Krešimir Matković³, Eduard Gröller⁴

⁽¹⁾Laboratorio de I+D en Visualización y Computación Gráfica
Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica,
(UNS-CIC Prov. de Buenos Aires)

antonella.antonini@cs.uns.edu.ar; mlg@cs.uns.edu.ar; smc@cs.uns.edu.ar

⁽²⁾INGEOSUR, Dpto. de Geología, Universidad Nacional del Sur
florenciagargiulo@gmail.com; gferrac@uns.edu.ar; ebjerg@ingeosur-conicet.gob.ar

⁽³⁾VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH,
Donau-City-Strasse 11, 1220 Viena, Austria.

Matkovic@vrvis.at

⁽⁴⁾Technische Universität Wien, Institut für Computergraphik und Algorithmen,
Favoritenstrasse 9-11 / E186, A-1040 Viena, Austria

groeller@cg.tuwien.ac.at

RESUMEN

El objetivo de la visualización es comunicar información de una manera gráfica, interactiva y comprensible. La visualización de datos multivariados, en particular, es un campo de investigación activo con una gran aplicación en diversas áreas.

En el análisis de conjuntos de datos multidimensionales, las cuestiones relacionadas con la detección de eventos, correlaciones, patrones y tendencias, desempeñan un papel cada vez más importante.

A medida que aumenta la dimensionalidad de los datos, la realización de una visualización eficaz se vuelve cada vez más compleja ya que muchas técnicas tradicionales solo pueden generar vistas parciales de los datos [3, 11, 12, 13], imposibilitando la detección de posibles relaciones entre todas las variables involucradas en el proceso.

Por lo tanto, tener un conjunto de metáforas visuales y técnicas de visualización relacionadas, es muy útil para nuestro análisis. Estas representaciones deben poder transmitir de manera efectiva las características del espacio de información e inspirar el descubrimiento.

En este contexto, el objetivo general es contribuir al desarrollo de tecnologías en torno al análisis visual de datos multidimensionales. En esta línea en particular, nos centraremos en la aplicación de estas nuevas tecnologías en el campo de la geología, pero estos métodos son extensibles a diferentes disciplinas.

Palabras Clave: Análisis Visual; Datos Multidimensionales; Visualización de Datos Geológicos

CONTEXTO

Se trabaja en estrecha colaboración con investigadores de centros de investigación de

prestigio nacional e internacional, entre los que se encuentran TuWien (Institute of Computer Graphics and Algorithm, Austria), VRVis (Center for Virtual Reality and Visualization, Austria), INGEOSUR-CONICET (Instituto Geológico del Sur) y el Departamento de Geología de la Universidad Nacional del Sur. Además, contamos con la participación de investigadores del VyGLab (Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica) del Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS).

1. INTRODUCCIÓN

La visualización de información tiene como objetivo analizar datos de forma rápida, interactiva e intuitiva utilizando representaciones visuales. La exploración y el análisis de datos, son pasos cruciales en la investigación científica. Los médicos, físicos, matemáticos y otros científicos examinan, exploran y analizan datos para comprender mejor los problemas.

En la actualidad, los datos aumentan constantemente en término de volumen y diversidad, por lo que las técnicas de visualización se enfrentan al desafío de hacer frente a conjuntos de datos cada vez más grandes en términos de representación, interacción y rendimiento.

La visualización de datos multivariados es una tarea desafiante. El objetivo no es sólo la visualización de múltiples dimensiones de datos, sino ayudar al usuario en la comprensión de los mismos.

Las técnicas tradicionales de visualización, tales como diagramas de dispersión, diagramas de líneas, diagramas de barras, gráficos circulares, etcétera; sólo producen una visión parcial de la información.

Al aumentar la dimensionalidad de los datos, lograr una visualización eficiente de los mismos se convierte en una tarea cada vez más compleja, ya que los espacios multidimensionales son difíciles de entender por su propia naturaleza, debido a la gran cantidad de dimensiones que se deben representar en un espacio muy limitado.

En este contexto, resulta de gran utilidad contar con diferentes herramientas para el análisis visual de datos multidimensionales que permitan inferir conclusiones sobre las posibles relaciones existentes entre ellos. Dada la particularidad de los datos en uso, ninguna de las técnicas de visualización logra satisfacer los requerimientos de análisis y comprensión por sí sola.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Desde los inicios de la investigación, nos centramos particularmente en la visualización de datos geológicos, de características multidimensionales, pero con el objetivo presente de lograr resultados que puedan aplicarse a distintas disciplinas en las que también deba obtenerse información a partir de este tipo de conjuntos de datos.

Se establece como objetivo general el contribuir al desarrollo de tecnologías y soluciones en torno al análisis visual de datos multidimensionales; métodos que puedan aplicarse a distintas disciplinas en las que también deba obtenerse información a partir de conjuntos de datos.

El diseño y desarrollo de las técnicas y herramientas se orienta a la visualización de datos geoquímicos obtenidos a partir del análisis mineral de rocas [2, 9, 10]. Al trabajar en conjunto con investigadores del área de Geología del Instituto Geológico del Sur (UNS-CONICET), podemos establecer el

análisis de minerales del grupo de los espinelos como caso de estudio y validar las técnicas de visualización propuestas mediante la utilización de conjuntos de datos con los que ya cuentan estos profesionales.

3. RESULTADOS OBTENIDOS Y ESPERADOS

De manera sostenida en el tiempo, se ha trabajado en temas relacionados con la Visualización de Datos aplicada a Geociencias en general, y a los minerales del grupo de los espinelos en particular [4,5]. Se desarrollaron nuevas técnicas de visualización para datos geológicos aplicados a la mineralogía y a la prospección de recursos naturales [6, 7, 8].

En el último año, se desarrolló e implementó una herramienta web para la visualización de datos geológicos que integra la mayoría de los diagramas que utilizan los geólogos comúnmente para analizar las características químicas de los minerales e incorpora visualizaciones interactivas en 3D [1]. También proporciona vistas coordinadas e interacciones adecuadas para que los usuarios interactúen con sus conjuntos de datos.

Ya mencionamos que nuestro principal objetivo es desarrollar técnicas y tecnologías que puedan aplicarse a diversas disciplinas, por lo que actualmente estamos trabajando en una segunda herramienta web que integrará un conjunto de metáforas visuales y técnicas de visualización que permita analizar cualquier conjunto de datos.

Por otro lado, en conjunto con investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) se lleva a cabo un relevamiento del estado del arte de las técnicas existentes para datos multidimensionales con el objetivo de, en una próxima instancia, diseñar nuevas técnicas de

visualización para el tratamiento de datos con estas características.

4. FORMACIÓN DE RECURSOS HUMANOS

A continuación, se detallan las tesis finalizadas y en desarrollo, proyectos de investigación vinculados con las líneas de investigación presentadas y las becas adquiridas en centros de investigación científica.

Tesis Finalizada: “Interacciones en Visualización”, tesis de doctorado en Ciencias de la Computación. Alumna: M. Luján Ganuza. Directora: Dra. Silvia Castro.

Proyecto Final Finalizado: “Técnicas 3D para Visualización de Datos Geológicos en la Web”. Proyecto final de la Ingeniería en Sistemas de Computación. Alumno: Gonzalo Picorel. Directora Silvia Castro. Codirectora: M. Luján Ganuza.

Proyecto Finalizado: PGI 24//N037 “Análisis Visual de Grandes Conjuntos de Datos”. Directora: Silvia M. Castro.

Proyecto Vigente: PGI 24//N048 “Análisis Visual de Datos”. Directora: Silvia M. Castro.

Proyecto Vigente: PICT 2017-1246 “Análisis Visual de Datos en Geociencias”. Directora: Silvia M. Castro.

Becaria: Antonella S. Antonini. Título del plan propuesto: “Visualización de Datos Geológicos en la Web”. Beca de entrenamiento 2018. Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Adjudicada a partir de abril de 2018 y por un término de 12 meses.

Becaria: M. Luján Ganuza. Título del plan propuesto: “Análisis Visual de Datos Multidimensionales en Espacios Ad-Hoc”. Beca Posdoctoral 2018 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas de la República Argentina. Adjudicada a partir de abril de 2019 y por un término de 24 meses.

Becaria: Antonella S. Antonini. Título del plan propuesto: “Análisis Visual de Datos Multidimensionales”. Beca doctoral 2019 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas de la República Argentina. Adjudicada a partir de abril de 2019 y por un término de 5 años.

5. BIBLIOGRAFÍA

- [1] 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.
- [2] Barnes, S. J., and Roeder, P. L. The range of spinel compositions in terrestrial mafic and ultramafic rocks. *J. Petrol.* 42, 12 (2001), 2279–2302.
- [3] Card, S. K., Mackinlay, J. D., and Shneiderman, B. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.
- [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 Conference on Visual Analytics Science and Technology, VAST 2015, Chicago, IL, USA, October 25-30, 2015* (2015), 197–198.
- [9] Lindsley, D.H. Oxide minerals: petrologic and magnetic significance. *Review in Mineralogy* 25 (1991).
- [10] Roeder, P. L. Chromite; from the fiery rain of chondrules to the kilauea iki lava lake. *The Canadian Mineralogist* 32, 4 (1994), 729–746.
- [11] Spence, R. *Information Visualization: Design for Interaction*. 2007.
- [12] Ware, C. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2004.
- [13] Ward, M. O., Grinstein, G., Keim, D. *Interactive data visualization: foundations, techniques, and applications*. CRC Press, 2010.

ANÁLISIS VISUAL DE MOVIMIENTOS MICROSACÁDICOS

Leandro Luque^{1,2}, M. Luján Ganuza^{1,2,3}, Silvia M. Castro^{1,2,3} y Osvaldo E. Agamennoni³
{leandro.luque, mlg, smc}@cs.uns.edu.ar, oagamen@uns.edu.ar

¹Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab)
(UNS-CIC Prov. de Buenos Aires)

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur
(DCIC-UNS)

²Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

³Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC)

Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur
(DIEC-UNS)

RESUMEN

El gran volumen de datos que se generan durante los experimentos de eye tracking hace necesario el uso de técnicas que permitan explorar estos conjuntos de manera efectiva y poder así encontrar patrones en los datos.

Los modelos tradicionales empleados en el análisis de los datos oculares resultan muy complejos para un primer análisis de éstos. En particular, muchas veces estos modelos resultan complejos de comprender debido a su naturaleza cuantitativa y no facilitan el entendimiento específico de un comportamiento. En el marco del análisis conductual de cómo un sujeto procesa la información que adquiere visualmente, un tipo de movimiento ocular denominado microsacada ha emergido como potencial candidato para marcador cognitivo que sea invariante a factores externos.

El objetivo de esta Línea de Investigación es el diseño e implementación de técnicas visuales que permitan analizar y evaluar el efecto de los movimientos microsacádicos durante la realización de diferentes tareas cognitivas. Mediante el uso de las mismas será posible determinar la naturaleza de este fenómeno bajo distintas condiciones.

Palabras claves: *Análisis Visual de Datos, Eye Tracking, Visualización de Datos Multidimensionales, Visualización de Datos Espacio-Temporales, Movimientos Microsacádicos*

CONTEXTO

El presente trabajo se realiza en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación (DCIC), de la Universidad Nacional del Sur (UNS) y en colaboración con el Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC, UNS-CIC). Los trabajos realizados en esta línea implican la participación de docentes investigadores y becarios doctorales y posdoctorales.

1. INTRODUCCIÓN

El análisis de datos recolectados mediante un eye tracker (ET) ha cobrado relevancia en los últimos años por su utilidad en diversos campos de aplicación, como el *gaming*, las neurociencias, las interfaces centradas en el usuario, entre otros. En particular, para el caso del campo de las neurociencias, los

movimientos oculares han cobrado especial relevancia por su simplicidad y utilidad para medir el desempeño cognitivo de las personas al momento de realizar tareas de alto orden cognitivo como complemento a las técnicas empleadas tradicionalmente. Para ello, los expertos necesitan analizar estos registros oculares en contexto con el resto de las variables involucradas tanto para obtener una idea general del comportamiento particular de un grupo bajo estudio, como para llevar a cabo el análisis de casos específicos.

Considerando lo expuesto, es clara la necesidad de investigación y desarrollo de nuevas técnicas que permitan tratar con las limitaciones que exponen los modelos de trabajo tradicionales para tratar con este gran volumen de datos. Es así que el empleo de técnicas de visualización adecuadas resulta de especial interés para obtener una primera visión acerca del problema estudiado, que luego servirá como base para la generación de hipótesis de trabajo y su posterior corroboración. Es importante que este nuevo conjunto de técnicas contemplen la naturaleza espacio-temporal de los datos, así como la limitaciones y utilidad de las técnicas preexistentes, además de las limitaciones de los métodos preexistentes definidos para tareas y/o estímulos similares.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las técnicas actuales de visualización de datos provenientes de un ET [1, 15, 11, 2] permiten obtener información relevante acerca de cómo se procesa la información por uno o varios participantes [3, 12]. Sin embargo, sólo se refleja un conjunto limitado de características a visualizar, con lo cual muchas veces no se tienen en cuenta todas las variables que intervienen en el proceso de adquisición visual de la información,

conduciendo a una mala interpretación de los resultados que se obtengan.

Es bien conocido que los marcadores que permiten determinar qué tan eficientemente una persona procesa la información adquirida son el tamaño de la pupila, el tiempo y la cantidad de fijaciones y el número de regresiones, entre otras [9, 6, 7]. Sin embargo, durante los últimos años ha surgido una nueva rama de desarrollo que involucra el uso de un tipo particular de movimiento ocular denominado microsacada como medida de desempeño. En el contexto del eye tracking se pueden distinguir 2 grupos bien definidos de movimientos: las fijaciones y las sacadas [8]. Las fijaciones se refieren a la región del estímulo a la cual está mirando el sujeto y tienen una duración típica del orden de los 100ms a los 700ms. Las sacadas, por otro lado, son los movimientos de transición entre dos fijaciones seguidas y tienen una duración aproximada de 50ms a 150ms. El foco principal de los estudios actuales está en el estudio de las fijaciones, dado que en éstas es donde el cerebro recopila la información visual y la integra con la memoria de trabajo. El valor de este movimiento se da en comprender qué rol juegan una serie de movimientos intra-fijación denominados tremor, *drift* (deriva) y microsacada. Los dos primeros se refieren a movimientos que en su mayoría se dan como consecuencia del cambio entre dos fijaciones, ya que implican la estabilización del ojo luego de una sacada, debido al cambio de velocidad. Por otro lado, las microsacadas son movimientos similares a las sacadas que se dan de manera imperceptible durante una fijación, pero con propiedades totalmente diferentes a su contraparte (por ejemplo, la velocidad y la amplitud). Dada la naturaleza de las microsacadas es posible determinar cómo trabaja el proceso de comprensión en un sujeto y qué regiones del estímulo necesitan

mayor (re) atención para completar una determinada tarea [4]. Las funciones principales de estos movimientos [14] están vinculadas con la exploración de detalles espaciales finos (imágenes, escenas naturales [5] y/o video), las implicaciones en la visión foveal (región de la retina donde la agudeza visual es mayor) y la estrategia ocular empleada (como la aparición de las microsacadas está vinculada al tipo de tarea que se realiza).

Actualmente, los movimientos oculares son analizados mediante el uso de modelos oculares estadísticos [13] que intentan integrar la mayor cantidad de variables posibles para explicar el fenómeno bajo estudio. Sin embargo, la gran cantidad de interrelaciones que puede haber entre las variables y su grado de variabilidad como resultado del estímulo que se esté analizado, resulta en modelos complejos que no son fácilmente entendibles y que no resultan eficientes en término de costo/tiempo para obtener patrones emergentes. Es por ello, que el uso de las técnicas de visualización resulta de especial interés por su utilidad para comprender información multivariada, así como de carácter espacio-temporal.

En la línea de investigación propuesta se investigará el uso de los movimientos microsacádicos (y nuevas propiedades que se puedan inferir) como marcadores del desempeño cognitivo de las personas durante la realización de diversas tareas de distinta complejidad. Por otro lado, se buscará diseñar e implementar las correspondientes técnicas de visualización que permitan mostrar estos movimientos, de manera tal de comprender cómo la información se interrelaciona y cuál es su impacto en el desempeño de una determinada tarea.

3. RESULTADOS OBTENIDOS Y ESPERADOS

El desarrollo de esta línea se orienta al diseño e implementación de herramientas y técnicas de visualización que permitan analizar los microsacádicos y su impacto en el desempeño cognitivo de los sujetos en tipos particulares de estímulos (lectura, *free-viewing*, *point-based*). Para ello se emplearán las técnicas *ad-hoc* desarrolladas por los autores para aprovechar las características inherentes de los estímulos empleados en conjunción con los movimientos oculares registrados. De esta manera se espera que se puedan encontrar *insights* que permitan determinar qué tan útil es este tipo de movimiento como marcador cognitivo en comparación con los tradicionales (dilatación de la pupila, tiempo de fijación, cantidad de regresiones, etc.) presentados en la literatura.

4. FORMACION DE RECURSOS HUMANOS

En cuanto a la formación de recursos humanos se detallan las tesis en desarrollo relacionadas con las líneas de investigación presentadas, así como también los proyectos de investigación y los becarios vinculados.

Tesis Finalizada (febrero 2021): “Desarrollo de modelos del comportamiento ocular”, Juan Andrés Biondi. Tesis de Doctorado en Ciencias de la Computación. Directores: Dra. Silvia Castro, Dr. Osvaldo Agamennoni.

Becario: Leandro Luque. Plan de trabajo: “Análisis Visual de Datos provenientes de Registradores de Movimientos Oculares”. Beca doctoral CONICET 2018, adjudicada a partir de abril de 2019.

Becaria: M. Luján Ganuza. Denominación del plan de trabajo: “Análisis Visual de Datos Multidimensionales en Espacios Ad-Hoc” Beca posdoctoral CONICET 2018, adjudicada a partir de abril de 2019.

Proyecto: PGI 24/N048, “Análisis Visual de Datos” (PGISeCyT-UNS, 2019-2022).
Directora: Dra. Silvia Castro.

Proyecto: PICT-2016-1009, “Modelado de la dinámica ocular para la evaluación del desempeño cognitivo” (ANPCyT, 2017-2019). Categoría: Plan Argentina Innovadora 2020. Director: Dr. Osvaldo Agamennoni.

Proyecto: PGI 24/K085, “Dinámica de los Sistemas Cognitivos” (PGISeCyT-UNS, 2019-2022). Director: Dr. Osvaldo Agamennoni, Co-Directora: Dra. Silvia Castro.

5. BIBLIOGRAFÍA

- [1] Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2014, June). State-of-the-Art of Visualization for Eye Tracking Data. In EuroVis (STARS).
- [2] Blascheck, T., Raschke, M., & Ertl, T. (2013, August). Circular heat map transition diagram. In Proceedings of the 2013 Conference on Eye Tracking South Africa (pp. 58-61).
- [3] Blascheck, T., Schweizer, M., Beck, F., & Ertl, T. (2017, June). Visual comparison of eye movement patterns. In Computer Graphics Forum (Vol. 36, No. 3, pp. 87-97).
- [4] Bowers, N., Rucci, M., & Poletti, M. (2016). Microsaccades during reading. *Journal of Vision*, 16(12), 858-858
- [5] Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, 10(10), 28-28.
- [6] Fernández, G., Buedo, P., Orozco, D., & Agamennoni, O. (2017). Eye Movement Behavior Analyses for Studying Cognitive Performance and Conversion to Pathologies. In *Psychiatry and Neuroscience Update-Vol. II* (pp. 281-292). Springer, Cham.
- [7] Fernández, G., Castro, L. R., Schumacher, M., & Agamennoni, O. E. (2015). Diagnosis of mild Alzheimer disease through the analysis of eye movements during reading. *Journal of Integrative Neuroscience*, 14(01), 121-133
- [8] Holmqvist, K., & Andersson, R. (2017). Eye tracking: A comprehensive guide to methods. Paradigms and Measures.
- [9] Hyönä, J., Lorch Jr, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In *The Mind's Eye* (pp. 313-334). North-Holland.
- [10] Kurzhals, K., Heimerl, F., & Weiskopf, D. (2014, March). ISeeCube: Visual analysis of gaze data for video. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 43-50).
- [11] Kurzhals, K., Hlawatsch, M., Heimerl, F., Burch, M., Ertl, T., & Weiskopf, D. (2015). Gaze stripes: Image-based visualization of eye tracking data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1), 1005-1014.
- [12] Noton, D., & Stark, L. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11(9), 929-938.
- [13] Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the EZ Reader model: Exploring the interface between cognition and

eye-movement control. *Cognitive Psychology*, 52(1), 1-56.

[14] Rucci, M., & Poletti, M. (2015). Control and functions of fixational eye movements. *Annual Review of Vision Science*, 1, 499-518.

[15] Yang, C. K., & Wacharamanotham, C. (2018, April). Alpscarf: Augmenting scarf plots for exploring temporal gaze patterns. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).

Integración del Procesamiento Imágenes e Internet de las Cosas en la estimación temprana del rendimiento de cultivos frutales

Gustavo A. Dejean¹, Federico Balaguer², Alejandra Yommi³, Jorge H. Doorn¹,
María A. David³, Natalia L. Murillo⁴, Ignacio A. García Ravlic¹, Dante H. Mendoza¹

¹Escuela de Informática, Universidad Nacional del Oeste

²Stream S.A., ³INTA, Estación Experimental Balcarce

⁴INTA, Agencia de Extensión Rural Otamendi

dejean2010@uno.edu.ar, federico.balaguer@gmail.com, yommi.alejandra@inta.gov.ar, jdoorn@uno.edu.ar,
{david.maria, murillo.natalia}@inta.gov.ar, gnachoxp@gmail.com, hmendoza@uno.edu.ar

RESUMEN

La actividad frutícola necesita un conjunto de actividades para lograr altos niveles de eficiencia y productividad. Una de esas actividades claves consiste en estimar lo más exactamente posible la producción total, ya que la misma condiciona los aspectos logísticos que deben ser atendidos en el breve lapso en el que transcurre la recolección de frutos. La producción depende esencialmente de dos factores, la cantidad de frutos por unidad de superficie y el tamaño de los frutos. Predecir el tamaño de los frutos requiere realizar un seguimiento continuo de los aspectos climáticos, tanto los de la región, cómo el microclima de la plantación. Por su parte se aspira a estimar la cantidad de frutos en base al procesamiento de imágenes de estados tempranos del cultivo, tales como botones florales, flores o frutos inmaduros. El caso de estudio es público-privado interinstitucional e interdisciplinario y se aboca a predecir la producción de kiwi en una plantación de Miramar, cuyo productor pertenece a la Cámara de Productores de Kiwi y tiene relación desde hace años con INTA Balcarce.

Palabras clave: Procesamiento de Imágenes, Reconocimiento de objetos, Telemedición, Internet de las cosas.

CONTEXTO

La línea de investigación que se presenta es parte del proyecto de investigación

“Estimación temprana de rendimiento del cultivo de kiwi mediante el procesamiento de imágenes” que se desarrolla en forma conjunta por la Universidad Nacional del Oeste, el Instituto Nacional de Tecnología Agropecuaria (INTA) y la empresa Stream S.A. Cada una de estas entidades aborda los aspectos relacionados con su área de especialización.

1. INTRODUCCIÓN

Lograr, en la producción frutícola, un nivel de eficiencia que permita alcanzar resultados económicamente sustentables, requiere la realización de un conjunto de actividades, tales como la poda, atado, riego, fertilización, control de heladas y de plagas, polinización, raleo y cosecha, con alto grado de planificación. Durante la cosecha, se necesita una adecuada gestión del conjunto de recursos que dependen fuertemente de la producción. Es así que, poder estimar tempranamente la producción permite una planificación adecuada. Actualmente, se estima el volumen de cosecha contando los frutos por unidad de superficie en una etapa avanzada de desarrollo de los mismos. Sin embargo, esta estimación tardía es demasiado cercana a la cosecha y muy difícil en plantaciones de medianas a grandes. En ese sentido, una estimación temprana del rendimiento permitiría tomar decisiones operativas y estratégicas. Desde el punto de vista operativo, podría orientar en cuál será la necesidad de recursos permitiendo un uso más racional de los mismos y desde el punto de vista estratégico, posibilitaría

anticipar la negociación de la producción en el mercado externo, además de cuantificar tempranamente el espacio necesario de cámaras de frío para almacenar la fruta destinada a la venta en el mercado interno, así como el material de empaque necesario.

La estimación temprana requiere monitorear la evolución fenológica de la planta, acompañada por información de las condiciones climáticas. El análisis de las condiciones climáticas comienza, inclusive, en la etapa invernal cuando la planta aparentemente no tiene actividad, pero la acumulación de bajas temperaturas es fundamental para romper la dormición de las yemas (muchas de las cuales dan lugar a flores y frutos). También se deben considerar otros factores climáticos que afectan diferentes momentos del proceso productivo: heladas durante la brotación, bajas temperaturas y lluvias durante la polinización, estrés térmico durante el crecimiento de los frutos, condiciones de temperatura y humedad relativa ambiente, que predisponen el desarrollo de enfermedades y plagas, entre otros. El monitoreo de la fenología del cultivo incluye identificar temporalmente la ocurrencia de los distintos estados según escalas internacionalmente reconocidas y luego asociarlos a variables climáticas. La identificación, conteo y seguimiento de las yemas, botones florales o frutos pequeños podría ser utilizado como insumo para la estimación temprana de rendimiento. Se estima que es posible lograrlo con herramientas y tecnologías de cómputo que lleven a cabo la manipulación asistida de imágenes. En resumen, este proyecto aborda un problema que no puede ser resuelto sin el conocimiento y la interacción entre profesionales de diferentes disciplinas: Fruticultura, Fisiología Vegetal, Informática, Procesamiento digital de la información.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El kiwi es una planta perenne introducida en Argentina a fines de los años '80. El área plantada muestra un crecimiento muy marcado y sostenido a partir del año 2003, sumando en la actualidad aproximadamente 1000 has. en

todo el país. Alrededor de la mitad de la superficie se encuentra en el sudeste de la provincia de Buenos Aires. El kiwi es una planta caduca, es decir, que tiene reposo invernal. El número de frutos obtenido por planta varía con el número de yemas, el porcentaje de brotación de las yemas, el porcentaje de brotes reproductivos y el número de inflorescencias por brote reproductivo. El éxito de una plantación de kiwi depende en gran medida de la cantidad de frutos cosechados por hectárea y de la proporción obtenida de frutos medianos a grandes, debido a que estos tamaños suelen tener precios diferenciales, mejorando el retorno económico [Testolin 1992].

Las prácticas agronómicas permiten lograr un determinado número de yemas por hectárea. Cangi y Atalay comprobaron que el peso promedio del fruto de kiwi 'Hayward', de alrededor de 120 g, no varía con 180 a 300 yemas por planta, pero la cantidad de frutos se duplica con la densidad más alta, por lo tanto, también el rendimiento [Cangi 2006]. En Italia, 150.000 a 200.000 yemas/ha (15 a 20 yemas/m²) es la densidad óptima para una carga frutal de kiwi tal que maximiza la cantidad de frutos sin afectar el tamaño y la calidad del fruto, aunque estos valores pueden variar con las condiciones agroclimáticas. En el norte y centro de Italia, un aumento en la densidad reduce la brotación de yemas reproductivas y la calidad del fruto mientras que, en el sur, resulta beneficioso contar con plantaciones de hasta 250.000 yemas/ha [Testolin 2009].

La variedad 'Hayward' requiere entre 700 y 850 horas de frío, por debajo de 7 °C, acumuladas en la etapa invernal para la brotación de las yemas. Además del componente genético, las variables climatológicas, tales como la temperatura media, máxima, mínima, acumulación de horas de frío y heliofanía entre otras condicionan la proporción de yemas reproductivas y el número de flores por yema [Grant 1984] [Warrington 1986] [Snelgar 1988] [Snelgar 1992] [Austin 2002].

Debido a la alta frecuencia de fuertes vientos y a la caída ocasional de granizo en la provincia

de Buenos Aires, el uso de mallas antigranizo y de cortinas, tanto artificiales como forestales, es una práctica recomendada y muy utilizada en las plantaciones de kiwi [David 2020].

Se ha probado que, el uso de techos con mallas monofilamento en producciones frutícolas reduce la velocidad del viento, y dependiendo de la densidad y color de la malla, pueden interceptar de radiación fotosintéticamente activa y la temperatura del aire, creando un microclima [Dussi 2005] [Bosco 2018].

En plantaciones de kiwi se ha reportado que las mallas antigranizo, en función al color, pueden afectar el rendimiento y la calidad de los frutos [Basile 2012]. Se desconoce si las mallas antigranizo que se utilizan en la zona productiva del sudeste de la provincia de Buenos Aires crean un microclima y si así fuera, si el mismo es predecible utilizando la información de los registros de las centrales meteorológicas cercanas.

Sin duda, una estimación temprana sería una herramienta esencial para adecuar el suministro de insumos apuntando a una producción más sustentable, sin resignar cantidad y calidad de la fruta obtenida en cada temporada. Posibilita dimensionar las toneladas a cosechar y negociar con anticipación las ventas en diferentes canales de comercialización. Zhou y Wulfsohn mencionaron que una predicción temprana y precisa del rendimiento tendría grandes beneficios para la fruticultura de precisión [Zhou 2012] [Wulfshohn 2012]. Los productores tendrían un mayor control de la producción, gastos, y podrían adecuar la capacidad de almacenamiento. A su vez, se podrían diferenciar áreas de la plantación con diferente productividad, lo cual conlleva a manejar los recursos adecuadamente para cada situación, diseñando un manejo específico de cada área del lote. La cuantificación temprana podría realizarse por recuento de diferentes órganos vegetativos y reproductivos mediante la interpretación de imágenes.

Algunos avances en este sentido se encuentran publicados. Ciertos reportes internacionales refieren a la estimación de rendimiento de diferentes frutales a partir de imágenes tomadas a los frutos en estados tempranos de

desarrollo [Dunn 2004] [Zhou 2012], encontrando una aceptable correlación con el rendimiento obtenido. La estimación del rendimiento a partir del número de flores ha dado aceptables resultados en manzanos [Aggelopoulou 2010] y poca correlación en almendros [Underwood 2016]. En manzanas, se ha avanzado en modelos de estimación del rendimiento por medio de la digitalización de imágenes [Stajanko 2011]. Recientemente en kiwi, Fu et al. desarrollaron un modelo para detectar los frutos con imágenes, pero en un estado de desarrollo avanzado de los mismos [Fu 208].

Probablemente, el estado de yema brotada en kiwi sea uno de los más apropiados para hacer una predicción temprana mediante la interpretación de imágenes, ya que en estados subsiguientes hay mayor desarrollo del follaje y buena parte de los órganos podrían quedar escondidos entre las hojas.

La cuantificación por imágenes del número de botones florales y de frutos en estados tempranos de desarrollo probablemente mejore la precisión de la estimación del rendimiento, pero implica un número muchísimo mayor de órganos a contar por unidad de superficie respecto a las yemas, y el uso de técnicas para mejorar la capacidad de captura debido al ocultamiento y solapamiento de los mismos. Otro aspecto relevante, asociado con el procesamiento de imágenes, está relacionado con el tipo de cámara a ser utilizada y la iluminación necesaria. Se desconoce si existen diferencias importantes entre las imágenes tomadas con cámaras infrarrojas, de luz visible o luz ultravioleta. También interesa determinar si es útil o necesario utilizar una fuente de iluminación artificial o es suficiente con la radiación natural.

En lo que se refiere al tamaño de los frutos, la estimación del mismo a partir de la información climatológica es una actividad que requiere mayores plazos que los del presente proyecto, por lo que esta parte del proyecto está centrada en determinar las características del microclima en la plantación y comprobar el grado de predictibilidad del mismo a partir de datos de estaciones meteorológicas de la región, así como analizar

las posibilidades de mejorar el mismo en el contexto de los recursos disponibles.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se seleccionarán al azar 50 plantas femeninas, cada una de ellas será georreferenciada y etiquetada. En cada planta a muestrear se colocará un marco de 2,5 m x 1 m de caño plástico pintado con un color contrastante al verde, atándolos en sus cuatro extremos a la planta para delimitar el área de estudio. El lado menor del rectángulo se apoyará sobre los brazos o branquetas de la planta sobre la línea de plantación y el lado mayor, irá en el sentido del largo de los cargadores, llegando aproximadamente a la mitad de la entrefila.

El recuento se realizará durante la primavera 2021 y 2022 en tres estados fenológicos: yema brotada a mediados de septiembre, botón floral a mediados de octubre y frutos en estados tempranos de desarrollo de principio a mediados de diciembre, respectivamente [Salinero 2009] [David 2018]. En caso de ocultamientos se repetirán las imágenes y el conteo se estimará mediante alguna de las técnicas de captura y recaptura, basadas en los modelos de Otis [Otis 1978], para lo cual habrá que evaluar la capacidad de captura de cada imagen respecto de las demás [Briand 2000].

Se instalarán cinco sensores en ubicaciones al azar debajo de la malla antigranizo, cerca de la periferia y en el interior del lote. Se tomarán datos de temperatura del aire, humedad relativa y radiación fotosintéticamente activa cada hora. También se colocarán dos sensores para contar con información climática por encima de la malla antigranizo. Dependiendo de la conectividad disponible, se utilizarán sensores IoT (Internet de las Cosas) o cajas de adquisición (datalogger).

Se calcularán las horas de frío acumuladas durante el reposo invernal, se registrará la ocurrencia de heladas, se evaluarán las condiciones de temperatura durante la polinización, durante el crecimiento y desarrollo del fruto, así como también se registrará el momento y volumen de las precipitaciones. Se evaluará la diferencia entre los datos obtenidos en condiciones protegidas

y a cielo abierto, e incluso se considerarán datos de estaciones meteorológicas cercanas. Se analizará la existencia o no de un microclima en la plantación y se calcularán las regresiones para cada variable climática medida y calculada en las diferentes condiciones.

Con los datos obtenidos a partir del procesamiento de las imágenes y la cuantificación manual del primer año, se definirán modelos para estimar el número de frutos a cosechar.

Todos los frutos de cada uno de los marcos colocados en las plantas en estudio se recolectarán manualmente en mayo del 2022, aunque se planifica realizar un conteo adicional en mayo del 2023 en el marco de un nuevo proyecto. Debido al ciclo del cultivo, el período del proyecto permitirá contar con un segundo año de datos de los dos primeros estados fenológicos. Una extensión del proyecto permitiría obtener la cuantificación del número de frutos mediante imágenes y por conteo manual en diciembre del año 2022, con la verificación del número de frutos realmente cosechados en mayo del siguiente año, contando de esa manera con datos de dos años lo que permitiría una mejor verificación del modelo de predicción propuesto.

Se comparará el número de frutos estimado mediante los modelos predictivos a partir del procesamiento de imágenes y por conteo manual, respecto al número de frutos cosechados. En caso de que el grado de precisión de la estimación por algunos de los métodos probados no fuera el deseado, se establecerá cual es la información más precisa, se replicarán también los conteos manuales y se estudiarán las respectivas dispersiones. Eventualmente, se desarrollarán herramientas que complementen los mismos, de ser necesarios. De esta forma, se estimará el rendimiento en términos de número de frutos/ha en función a la cuantificación temprana de órganos vegetales.

Si bien esta estimación será válida para las condiciones climáticas y de manejo de ese primer año, o de dos años en caso de poder contar con una extensión del proyecto, el modelo se irá validando y ajustando en

sucesivos años para mejorar la predicción.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto participan seis investigadores, y dos alumnos de grado, ambos con una beca de investigación.

5. BIBLIOGRAFÍA

- [Aggelopoulou 2010] Aggelopoulou, K. D., Wulfsohn, D., Fountas, S., Gemtos, T. A., Nanos, G. D. and Blackmore, S. 2010. Spatial variability of yield and quality in a small apple orchard. *Precision Agric.* 11:538–556. DOI: 10.1007/s11119-009-9146-9.
- [Austin 2002] Austin, P. T., Hall, A. J., Snelgar, W. P. and Currie, M. J. 2002. Modelling kiwifruit budbreak as a function of temperature and bud interactions. *Ann. Bot.* 89(6):695–706. DOI: 10.1093/aob/mcf113.
- [Basile 2012] Basile, B., Giaccone, M., Cirillo, C., Ritieni, A., Graziani, G., Shahak, Y. and Forlani, M. 2012. Photo-selective hail nets affect fruit size and quality in Hayward kiwifruit. *Scientia Hort.* 141:91–97. DOI:10.1016/j.scienta.2012.04.022.
- [Bosco 2018] Bosco, L. C., Bergamaschi, H., Cardoso, L. S., Paula, V., Marodin, G.A.B. and Brauner, P.C. 2018. Microclimate alterations caused by agricultural hail net coverage and effects on apple tree yield in subtropical climate of Southern Brazil. *Bragantia* 77(1):181–192. DOI: 10.1590/1678-4499.2016459.
- [Briand 2000] Briand, L., El Emam, K., Freimut, B., and Laitenberger, O. 2000. A Comprehensive Evaluation of Capture-Recapture Models for Estimating software Defects Contents. *IEEE Trans. Software Eng.* 26(6):518–540. DOI: 10.1109/32.852741.
- [Cangi 2006] Cangi, R. and Atalay, D.A. 2006. Effects of different bud loading levels on the yield, leaf and fruit characteristics of Hayward kiwifruit. *Hort. Sci.* 33:23–28. DOI: 10.17221/3736-HORTSCI
- [David 2018] David, M.A., Yommi, A., Sánchez, E. 2018. Fenología del cultivo de kiwi en el sudeste de Buenos Aires. <https://inta.gov.ar/documentos/fenologia-del-cultivo-de-kiwi-en-el-sudeste-de-buenos-aires>.
- [David 2020] 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.
- [Dunn 2004] Dunn, G.M., and Martin, S. R. 2004. Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. *Australian J. Grape and Wine Res.* 10:196–198. DOI: 10.1111/j.1755-0238.2004.tb00022.x.
- [Dussi 2005] Dussi, M., Giardina, G., Sosa, D., González Junyent, R., Zecca, A. and Reeb, P. 205. Shade nets effect on canopy light distribution an quality of fruit and spur leaf on apple cv. Fuji. *Spanish Journal of Agricultural Research* 3(2):253–260. DOI: 10.5424/sjar/2005032-144
- [Fu 2018] Fu, L., Feng, Y., Elkamil, T., Liu, Z., Li, R. and Cui, Y. 2018. Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks. *Trans. Chinese Soc. Agric. Eng.* 34(2):205–211. DOI: 10.11975/j.issn.1002-6819.2018.02.028.
- [Grant 1984] Grant, J.A. and Ryugo, K. 1984. Influence of within-canopy shading on fruit size, shoot growth, and return bloom in kiwifruit. *J. Am. Soc. Hortic. Sci.* 109: 799–802.
- [Otis 1978] Otis, D.L., Burnham, K.P. White, G.C., and Anderson, D.R. 1978. *Statistical inference from Capture on Closed Animal Populations.* Wildlife Monograph 62.
- [Salinero 2009] Salinero, M. C., Vela, P. and Sainz, M. J. 2009. Phenological growth stages of kiwifruit (*Actinidia deliciosa* 'Hayward'). *Scientia Hort.* 121(1):27–31. DOI: 10.1016/j.scienta.2009.01.013.
- [Snelgar 1988] Snelgar, W.P., Bayley, G. and Manson, P.J. 1988. Temperature studies on kiwifruit vines using relocatable greenhouses. *New Zealand J. Exp. Agric. Res.* 16: 329–339.
- [Snelgar 1992] Snelgar, W.P., Manson, P.J., and Martin, P.J. 1992. Influence of time of shading on flowering and yield of kiwifruit vines. *J. Hortic. Sci.* 67: 481–487.
- [Stajanko 2011] Stajanko, D., and Blanke, M. M. 2011. Yield prediction in fruit crops using image analysis. *Acta Hort.* 903:1115–1119. DOI: 10.17660/ActaHortic.2011.903.155.
- [Testolin 1992] Testolin, R., and Costa, G. 1992. Modelling a kiwifruit orchard. *Acta Hort.* 313:99–103.
- [Testolin 2009] Testolin, R. and Ferguson, A. 2009. Kiwifruit (*Actinidia* spp.) production and marketing in Italy. *NZ J. Crop Hort. Sci.* 37(1):1–32. DOI: 10.1080/01140670909510246.
- [Underwood 2016] Underwood, J. P., Whelan, B., Hung, C. and Sukkarieh, S. 2016. Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors. *Computers and Electronics in Agric.* 130:83–96. DOI: 10.1016/j.compag.2016.09.014.
- [Warrington 1986] Warrington, I.J. and Stanley C.J. 1986. The influence of pre- and post-budbreak temperatures on flowering in kiwifruit. *Acta Hort.* 175:103–107.
- [Wulfshohn 2012] Wulfsohn, D., F.A. Zamora, C. Potin Tellez, I. Zamora Lagos and M. García-Fiñana. 2012. Multilevel systematic sampling to estimate total fruit number for yield forecasts. *Precision Agric.* 13:256–275. DOI: 10.1007/s11119-011-9245-2.
- [Zhou 2012] Zhou, R., Damerow, L., Sun, Y. and Blanke, M.M. 2012. Using colour features of cv. 'Gala' apple fruits in an orchard in image processing to predict yield. *Precision Agric.* 13:568–580. DOI: 10.1007/s11119-012-9269-2.

INTERACCIÓN HUMANO-COMPUTADORA: ESTRATEGIAS HACIA LA EMPATÍA COMPUTACIONAL

G. Rodríguez, N. Jofré, Y. Alvarado, J. Fernandez, R. Guerrero
Laboratorio de Computación Gráfica / Dpto. Informática / FCFMyN
Universidad Nacional de San Luis
Ejército de los Andes 950, Tel: 0266 4420823, San Luis, Argentina
{gbrodriguez, npasinetti, ymalvarado, jmfer, rag}@unsl.edu.ar

RESUMEN

Las Nuevas Tecnologías de la Información y la Comunicación han provisto a las personas de nuevas herramientas de socialización.

Entre otros aspectos claves de la comunicación humana, estas deben tener en cuenta a la empatía como un factor crucial para la interacción social. La empatía se describe como la capacidad de comprender el estado emocional de un individuo y, a menudo, expresar una respuesta que resuena con él.

Sin embargo, la capacidad de expresar empatía disminuye a medida que las personas dependen, cada vez más, de los recursos tecnológicos para interactuar.

Este fenómeno demanda hoy, más que nunca, la adecuación de los sistemas informáticos, a través del diseño de estrategias que les permitan comunicarse adecuadamente con los humanos.

Por lo tanto, es necesario maximizar los niveles de empatía, a través de agentes y ambientes virtuales, que permitan optimizar la comunicación de información entre el usuario y un sistema.

Existen diferentes enfoques, técnicas, métodos o mecanismos que permitirían promover la empatía en estas interacciones.

En la presente propuesta de trabajo se propone identificar y diseñar mecanismos para la empatía computacional.

Palabras clave: Interacción Humano Computadora, Empatía computacional, Valle inquietante, Realidad Virtual, Realidad Aumentada.

CONTEXTO

La propuesta de trabajo se lleva a cabo dentro del proyecto “*Realidades Alternativas como lenguaje generativo aplicado a la solución de problemas reales*”. Este proyecto es desarrollado en el ámbito del Laboratorio de Computación Gráfica de la Universidad Nacional de San Luis.

1. INTRODUCCIÓN

Durante las últimas décadas, nuestro entorno social se ha poblado de tecnología que está revolucionando drásticamente nuestra forma de vida: las Nuevas Tecnologías de la Información y la Comunicación (NTIC). Estas se caracterizan por la presencia de artefactos tecnológicos que actúan de forma autónoma en los entornos sociales, en relación con los procesos de la información y de la comunicación, destacándose en los procesos telemáticos y de carácter comunicativo.

Sin embargo, para que estos nuevos sistemas habiten en nuestro entorno social, deben ser capaces de interactuar con humanos. Esta interacción debe ser lo más natural posible y estar inspirada en la forma en que las personas interactúan.

En este sentido, estos sistemas no sólo deben ser reactivos y proactivos sino que también deben tener capacidades sociales, especialmente cuando interactúan con seres humanos [1,2]. Así pues, un sistema con capacidades sociales debe estar dotado con la habilidad de comprender a los demás y sus intenciones, motivaciones y sentimientos. Esta capacidad implica percibir y comprender

los estados afectivos de los demás y actuar en consecuencia, una capacidad que suele denominarse empatía.

La empatía es un fenómeno socioemocional complejo que puede describirse como la capacidad de comprender y reaccionar ante los sentimientos, pensamientos y experiencias de los demás. Esta capacidad nos permite percibir el punto de vista del otro al razonar con sus emociones. La empatía es fundamental en el estudio de las relaciones sociales humanas y, por tanto, uno de los principales elementos de la interacción social humana. La empatía promueve relaciones y participaciones más sólidas [3,4,5].

Dado que la empatía no tiene una definición universalmente aceptada, sus múltiples definiciones pueden subdividirse en tres grandes categorías [6]: (a) la empatía como respuesta afectiva a los estados emocionales de los demás (*empatía afectiva*), (b) la empatía como la comprensión cognitiva de los estados emocionales de los demás (*empatía cognitiva*), y (c) la empatía como un componente tanto afectivo como cognitivo.

Por otro lado, se ha demostrado que los humanos sienten empatía no sólo por otros humanos, sino también por personajes de ficción, personajes de juegos, e incluso robots [7]. Dado el importante papel que desempeña la empatía en la forma en que se establecen la comunicación y las relaciones sociales en los humanos, es evidente que la empatía es un elemento importante en la interacción entre humanos y computadoras.

Uno de los aspectos usualmente considerados al momento de hablar de empatía hace referencia al realismo o apariencia humana de los agentes sintéticos. En el año 1970 Masahiro Mori [8] planteó la hipótesis de que la respuesta emocional de los humanos hacia objetos parecidos a los humanos es directamente proporcional a su grado de semejanza con los humanos. Sin embargo, alcanzado un determinado grado de similitud, las respuestas emocionales se vuelven repentinamente repulsivas.

El correspondiente declive producto de una discontinuidad emocional en la función de respuesta se denomina “valle inquietante”

Considerando la importancia de la empatía en las interacciones sociales, la integración de la capacidad empática en los sistemas computacionales sería de gran utilidad. Ésta podría mejorar los sistemas interactivos, como las aplicaciones educativas, los asistentes médicos, los acompañantes, la psicoterapia, y los juegos, en los que la capacidad social es de gran importancia [9-14]. Además, como amalgama de los procesos afectivos y cognitivos, la empatía computacional puede proporcionar las bases para examinar el vínculo entre las emociones y la cognición.

En consecuencia, la empatía computacional puede ayudar a profundizar en la investigación de las tecnologías empáticas o sistemas empáticos y, en última instancia, a materializarlas en innovaciones significativas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las crecientes capacidades de los sistemas informáticos en cuanto a detección biométrica, han hecho posible el procesamiento de señales sociales y afectivas (voz, gestos faciales, movimientos corporales, etc.), y en consecuencia, la mejora en la interacción natural con las personas.

Se ha demostrado que sistemas interactivos como agentes virtuales contribuyen a diferentes causas, como la gestión del estrés en entrevistas de trabajo [15], el lidiar con situaciones de acoso [16], la construcción de relaciones socioemocionales a largo plazo con personas [17], etc. De la misma manera, existen sistemas capaces de recrear acontecimientos naturales altamente realistas provocando cierto impacto en las personas [18], y consecuentemente, acciones de ayuda y protección, entre otros.

La incorporación de una componente empática en un sistema aportaría grandes beneficios al momento de brindarle al sistema la capacidad de comprender el estado emocional de un individuo y ser capaz de expresar una respuesta acertada, mejorando

notablemente la interacción natural humano-computadora.

Como resultado, la empatía computacional desarrollada en los sistemas permitiría la categorización de los mismos en dos tipos:

- *Sistemas con Agentes Empáticos*: que optimizan los componentes básicos de una comunicación empática a fin de mejorar la conversación con el usuario [19-21]. Según Paiva [22], existen dos formas de generar empatía a través de un agente virtual social: 1- considerando al usuario como el observador (ser empático) y al agente como el que desencadena la empatía en el interlocutor humano, 2- considerando a los agentes como observadores que empatizan (ente empático) con otros agentes o personas.
- *Sistemas con Ambientes Empáticos*: que consideran la relación del usuario con el entorno virtual (Usuario-Ambiente Virtual) [23,24]. Estos pueden generar y simular empatía en ambos sentidos con distintos fines. Por un lado, considerando al usuario como observador del ambiente virtual (ser empático) y el ambiente como un simple ente generador de empatía (por ej., entornos que generan conciencia ambiental para proteger ambientes naturales reales cuyo equilibrio ecológico puede amenazar el futuro del planeta). Por otro lado, considerando al ambiente virtual como observador (ente empático) del usuario, detectando estados anímicos indeseados o dañinos del usuario para sumergirlo en ambientes virtuales particulares cuyo objetivo sea devolver al usuario a estados anímicos saludables.

3. RESULTADOS OBTENIDOS/ESPERADOS

Las actividades realizadas hasta el momento por este grupo se han enmarcado dentro de 1 proyecto de Investigación de la UNSL, 1 proyecto de la Comunidad Europea, 4 proyectos de Desarrollo Tecnológico de la Secretaría de Políticas Universitarias y 3 proyectos de Extensión de Interés Social.

Como consecuencia del trabajo elaborado, se ha logrado desarrollar varios sistemas relacionados a RV con interacción multimodal involucrando aspectos verbales y gestuales, así como también se han elaborado aplicaciones de innovación y desarrollo asociadas a la RV y RA.

Actualmente las acciones se encuentran focalizadas en la incorporación de nuevas estrategias que permitan alcanzar una mejor percepción e interacción, al mismo tiempo que se logra incentivar al usuario tanto cognitiva como físicamente.

Como resultante, se analizará el impacto logrado por el uso de sistemas de RV y RA en la mejora y adquisición de habilidades funcionales en procesos físicos y cognitivos asociados a las áreas de la salud y la educación, entre otras. De esta manera, se pretende detectar y evaluar la evidencia científica resultante para determinar la envergadura de dichas intervenciones.

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 (4 finalizados), tesis de Especialización en Educación Superior (1 finalizada), tesis de Maestría en Cs. de la Computación (2 en ejecución y 1 finalizada) y tesis de Doctorado en Ciencias de la Computación (1 en ejecución).

Adicionalmente se ha obtenido una beca de iniciación a la investigación y una beca de perfeccionamiento en investigaciones otorgadas por la Secretaría de Ciencia y

Técnica de la UNSL; y una beca doctoral de CONICET.

5. BIBLIOGRAFÍA

1. Michael Wooldridge. 2002. *An Introduction to Multiagent Systems*. John Wiley & Sons.
2. Hana Boukricha, Ipke Wachsmuth, Maria Nella Carminati, and Pia Knoeferle. 2013. A computational model of empathy: Empirical evaluation. In *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*. IEEE, Geneva, Switzerland, 1–6.
3. Martin L. Hoffman. 2001. *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge University Press.
- Doori Jo, Jooyun Han, Kyungmi Chung, and Sukhan Lee. 2013. Empathy between human and robot? In *Proceedings of the 8th ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 151–152.
4. Mark H. Davis. 1994. *Empathy: A Social Psychological Approach*. Westview Press.
5. Frans B. M. De Waal. 2010. *The Age of Empathy: Nature's Lessons for a Kinder Society*. Three Rivers Press.
6. Becky Lynn Omdahl. 1995. *Cognitive Appraisal, Emotion, and Empathy*. Lawrence Erlbaum Associates.
7. Tinwell, A. 2014. The uncanny valley in games and animation.
8. Mori, M. 1970/2005. The uncanny valley. (K. F. MacDorman, & T. Minato, Trans.). *Energy*, 7, 33–35.
9. Jofre, N., Alvarado, Y., Fernández, J., Guerrero, R., Rodríguez, G. 2015. A serious game about recycling rules. In *XXI Congreso Argentino de Ciencias de la Computación (Junín, 2015)*.
10. Jofré, N., Rodríguez, G., Alvarado, Y., Fernández, J., & Guerrero, R. 2017. *Natural User Interfaces: A Physical Activity Trainer*. In *Argentine Congress of Computer Science* (pp. 122-131). Springer, Cham.
11. Alvarado, Y., Jofré, N., Rosas, M. V., Guerrero, R. A. 2018. AR4Axo: Aumentando la inteligencia espacial para proyecciones en sistema axonométrico. In *XIII Congreso de Tecnología en Educación y Educación en Tecnología (Posadas, 2018)*.
12. Alvarado, Y., Guerrero, R., Serón, F. 2018. Be civic: An immersive serious game. *EAI Endorsed Transactions on Serious Games*, 4(15).
13. Alvarado, Y., Jofré, N., Rosas, M., Guerrero, R. 2019. *Aplicaciones de Realidad Virtual y Realidad Aumentada como soporte a la enseñanza del Dibujo Técnico. Enseñanza y Aprendizaje de Ingeniería de Computadores*. (9).
14. Zúñiga, M. E., Liendo, C., Rosas, M. V., Rodríguez, G., Jofré Pasinetti, N., Alvarado, Y., Guerrero, R., Fernández, J. 2020. *Tecnologías emergentes para pedagogías emergentes*. In *XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020)(Neuquén, 6 y 7 de julio de 2020)*.
15. H. Prendinger and M. Ishizuka. 2005. The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19:267–285.
16. A. Paiva, J. Dias, D. Sobral, R. Aylett, S. Woods, L. Hall, and C. Zoll. 2005. Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, 19:235–266.
17. T. Bickmore and R. Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327.
18. D. Shin. 2018. Empathy and embodied experience in virtual environment: To what extent can virtual reality

- stimulate empathy and embodied experience?. *Computers in Human Behavior*, pp. 64-73.
19. Ochs, Magalie & Pelachaud, Catherine & Sadek, David. 2018. An empathic virtual dialog agent to improve human-machine interaction. *Autonomous Agent and Multi-Agent Systems (AAMAS)*. 1. 89-96. 10.1145/1402383.1402401.
 20. Yalçın Ö. N. 2018. Modeling Empathy in Embodied Conversational Agents. In *Proceedings of the 2018 on International Conference on Multimodal Interaction* (pp. 546-550). ACM.
 21. Zoll C., Enz S., Schaub H. Empathic Virtual Agents. In: Rist T., Aylett R.S., Ballin D., Rickel J. 2013. (eds) *Intelligent Virtual Agents. IVA 2003. Lecture Notes in Computer Science*, vol 2792. Springer, Berlin, Heidelberg.
 22. Paiva, Ana & Leite, Iolanda & Boukricha, Hana & Wachsmuth, Ipke. 2017. Empathy in Virtual Agents and Robots: A Survey. *ACM Transactions on Interactive Intelligent Systems*. 7. 1-40. 10.1145/2912150.
 23. Louie, A.K., Coverdale, J.H., Balon, R., Beresin, E.V., Brenner, A.M., Guerrero, A.P., & Roberts, L.W. 2018. Enhancing empathy: A role for virtual reality?. *Academic Psychiatry*, 42(6), 747–752..
 24. Jon Rueda, Francisco Lara. 2020. Virtual Reality and Empathy Enhancement: Ethical Aspects. *Frontiers in Robotics and AI*. Vol. 7. pp160.

La Televisión Digital Interactiva para el mejoramiento de los pueblos latinoamericanos

María José Abásolo¹²  Magdalena Rosado³  Telmo Silva⁶  Joaquín Pina⁴  Raisa Socorro⁴
 Raoni Kulesza⁵  Guido Lemos de Souza Filho⁵  Armando De Giusti¹  Marcelo Naiouf¹
 Patricia Pesado¹ 

¹ Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática, Universidad Nacional de La Plata (UNLP), Argentina

{mjabasolo, degiusti, mnaiouf, ppesado}@lidi.info.unlp.edu.ar

² Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

³ Universidad Católica de Santiago de Guayaquil, Ecuador

⁴ Universidad Tecnológica de La Habana José Antonio Echeverría (CUJAE), Cuba

{jpina, raisa}@ceis.cujae.edu.cu

⁵ Universidade Federal de Paraíba (UFPB), João Pessoa, Brasil

{raoni, guido}@lavid.ufpb.br

⁶ Universidad de Aveiro (UA), Aveiro, Portugal

tsilva@ua.pt

Resumen

La línea de investigación y desarrollo presentada consiste en estudiar, desarrollar y evaluar aplicaciones de Televisión Digital Interactiva (TVDi) y tecnologías complementarias para el mejoramiento de los pueblos latinoamericanos. Uno de los principales objetivos es la formación de recursos humanos y fortalecimiento de la investigación mediante el trabajo intergrupar entre diferentes instituciones nacionales y extranjeras.

Palabras Clave: Televisión digital, TVDi, aplicaciones móviles, Interactividad

Contexto

Desde el año 2012 diversas universidades iberoamericanas conforman la Red de Aplicaciones y Usabilidad de la Televisión Digital Interactiva (RedAUTI) coordinada por la UNLP. El objetivo general es apoyar la formación de los recursos humanos, estimular la investigación de forma transdisciplinaria y fortalecer los grupos de investigación de las universidades iberoamericanas en temas relacionados con TVDI, avanzando en el diseño, implementación y despliegue de aplicaciones, servicios y producción de contenidos para TVDI, en sus múltiples plataformas, de código abierto, de interés colectivo, para dar solución a problemas del contexto iberoamericano.

Actualmente se lleva a cabo el proyecto internacional: “PGTF INT/19/K08: TDTi*ABC La Televisión Digital Terrestre aplicada al mejoramiento de los pueblos de los países en vías de desarrollo” financiado por el Fondo Fiduciario Pérez Guerrero de la Oficina de Naciones Unidas para la Cooperación Sur-Sur. Dicho proyecto es coordinado por la UNLP y se lleva a cabo con la CUJAE (Cuba) y la UFPB (Brasil).

Introducción

La televisión digital interactiva (TVDi) proporciona impulso a la generación de contenidos y servicios digitales, permitiendo crear entornos favorables para el desarrollo de la Sociedad de la Información por medio de un dispositivo presente en todos los hogares. En particular el grupo de investigación III-LIDI ha presentado en ediciones anteriores de este workshop las líneas de investigación relacionadas con la TVDi [1-5]. En [6] Rosado et al. presenta una revisión de proyectos que se han desarrollado en el Ecuador en vinculación con la TVDi, señalando el tipo de servicio, los enfoques de aplicación (educativo, salud, información, entretenimiento), y el tipo de interactividad y tecnologías empleadas.

A continuación se presentan experiencias llevadas a cabo en países latinoamericanos que pueden servir de base para que otros países las promuevan para el mejoramiento de la población.

Adultos mayores y TVDi

La Organización Mundial de la Salud señala que el sujeto que pasa los 65 años puede presentar en su gran mayoría dificultades en su andar lo que puede provocar accidentes por desequilibrios por efectos de gravedad afectando en su condición física, es así que el uso de las Tecnologías de la Información y la Comunicación (TIC) por profesionales -médicos y no médicos-, han contribuido a que existan nuevos programas para ayudar a superar las afectaciones en la condición física. En [7] [8] Rosado et al. presentan una revisión sistemática, desde 2012 hasta principios de 2019, de experiencias realizadas con personas mayores de 65 años, utilizando las TIC con el objetivo de mejorar su autocuidado y empoderamiento asistencial para el envejecimiento activo. La mayoría de los estudios analizados se centran en el seguimiento y asistencia a los adultos mayores, mientras que hay una minoría de estudios centrados en la rehabilitación. Las TIC más utilizadas para el cuidado de la salud son las aplicaciones móviles y los servicios web. También prevalece el uso de sensores específicos para monitorear o controlar a los adultos mayores. En relación con la promoción de ejercicios, el objetivo principal es la prevención de caídas de los adultos mayores. Pocos proyectos utilizaron TVDi, los cuales se enfocan en establecer recordatorios para atención médica y utilizar programas de entretenimiento para tener bienestar físico, mental y social. En particular, se encontraron muy pocos proyectos relacionados con la promoción de la actividad física.

Debido a la familiaridad de uso de la televisión por parte del adulto mayor, en [9] Rosado et al. plantean desarrollar una aplicación con contenidos interactivos para TVDi que tenga por objetivo reducir las alteraciones de la marcha. Abordan consideraciones para la evaluación de la condición funcional a través de un test, el desarrollo de planes de ejercicio localizados, el proceso a considerar para el diseño funcional y la planificación de la evaluación del estudio. Para la evaluación se plantean dos grupos de usuarios: terapeutas de la Universidad Católica Santiago de Guayaquil y adultos mayores que asisten al centro

gerontológico Municipal “Dr. Arsenio de la Torre Marcillo” de la ciudad de Guayaquil.

Desarrollo y evaluación de aplicaciones para TVDi

El consumo de contenidos audiovisuales se ha convertido en una actividad cada vez más compleja a medida que se popularizan nuevas interfaces para la interacción, el intercambio, el aumento de la potencia informática y los diferentes modos de disfrute. Al observar las nuevas herramientas de distribución de contenido como Netflix, Amazon Prime, Apple TV plus, Disney Plus, emisoras como BBC, ABC, CBS, CNN, Rede Globo y fabricantes como Sony, Samsung y LG, es evidente que estas empresas se han desarrollado en los últimos años, productos y servicios que han ampliado la experiencia de ver contenidos de vídeo en la televisión. Estos nuevos productos y servicios son ejemplos de cómo la producción, distribución y especialmente la recepción de contenidos audiovisuales es ahora una actividad más compleja. Es necesario considerar que parte del conocimiento para desarrollar este tipo de productos es el resultado de acciones iterativas, costosas y confidenciales en estas empresas. Este conocimiento técnico o metodológico generalmente no es de dominio público, ni se publica en revistas científicas, ni siquiera es común a todos los profesionales del área de software y audiovisuales. Como consecuencia, se vuelve un trabajo arduo, a veces un desafío, organizar los conocimientos y habilidades para integrar contenidos e interfaces en sistemas audiovisuales de complejidad similar a los actores del mercado. Este contexto motivó al grupo de la Universidad Federal de Paraíba a trabajar en proyectos de investigación en el área de evaluación de aplicaciones para TVDi.

La propuesta para la herramienta de evaluación de interfaces de interacción de sistemas audiovisuales presentada en [10] se llevó a cabo sobre la base de un mapeo de patrones de interfaz de usuario y se diseñó una aplicación con un enfoque en dos usuarios principales: administradores y probadores. El sistema fue desarrollado y probado internamente para su operación en herramientas automatizadas y mediante una evaluación de dos productos de una

empresa en el mercado brasileño. Esta evaluación, aunque preliminar, permitió los siguientes aprendizajes: (a) La organización de la jerarquía a través de las estrategias propuestas se presentó como un enfoque claro y viable para la implementación final del servicio en los sectores de control de calidad y producto; (b) El desafío inicial de establecer una inspección de calidad con estándares de mercado se ha abordado satisfactoriamente; (c) La lógica de la herramienta y los criterios implementados se pueden replicar en elementos como el funcionamiento básico (el dispositivo se enciende, reconoce la señal) y la accesibilidad.

Canal Social

La aplicación “Canal Social” surgió en 2015 a través de un proyecto creado por el Ministerio de Desarrollo Social y Lucha contra el Hambre del Gobierno Federal de Brasil con el objetivo de establecer una comunicación directa de las políticas y programas sociales del gobierno con la población beneficiaria de los servicios gubernamentales que buscan facilitar el acceso a la información y estimular la generación de competencias y el desarrollo social, a partir de contenidos audiovisuales con información, entretenimiento y prestación de servicios.

Desde un principio, la solución fue diseñada para funcionar en una plataforma de TVDi basada en el middleware Ginga y los lenguajes NCL y Lua. Actualmente la solución se encuentra en su sexta versión con funcionalidad para acceder a información sobre diversos programas sociales: acceso al cronograma de pagos de beneficios provistos por el gobierno, ubicación de puntos de servicios sociales, juegos educativos para niños, disponibilidad de videos de campañas publicitarias gubernamentales, películas y acceso a una colección de recetas culinarias para una alimentación saludable y de bajo costo. La solución también incluye opciones de accesibilidad como el uso guiado por voz y el uso de las recomendaciones del W3C.

Entre diciembre de 2019 y febrero de 2020 se realizó una evaluación inicial mediante pruebas de campo del uso de la aplicación por parte de los usuarios con el fin de mapear los perfiles que constituyen una posible audiencia para la

aplicación y comprender mejor la experiencia de estos usuarios. Para ello, se utilizó una muestra de 131 usuarios en una ciudad capital de la región Nordeste de Brasil con 404,614 inscritos de un total de 96,022 familias. Esta audiencia está compuesta en gran parte por personas con bajos niveles de educación y la descripción de audio ayuda a facilitar la navegación y la comprensión de los recursos y la información que se ofrecen. Se puede encontrar más información sobre este estudio en [11].

Soberanía en la TVDi

La televisión digital interactiva (TVDi) es un escenario emergente moderno que permite transmitir contenidos informativos utilizando un medio de gran penetración en la mayoría de los pueblos del mundo. Sin embargo, actualmente la TVDi se utiliza principalmente para transmitir contenido multimedia que está controlado por empresas y corporaciones que no responden a los intereses de los pueblos. Por otro lado, el contenido útil está disperso en Internet, lo que dificulta que la mayoría de las personas se apropien de él fácilmente. En [12] se presenta una nueva solución de software que demuestra las posibilidades de la TVDi en un escenario real. La solución llamada TVC+ recoge información útil disponible en Internet y la integra con los servicios de TVDi. Algunas de sus funcionalidades ya se han desplegado en Cuba, demostrando su utilidad en algunas áreas de los objetivos de Desarrollo Sostenible 2030 de la ONU: Educación, Salud, Alimentación y Patrimonio.

Otro aspecto clave en la soberanía de la TVDi son los procesos tecnológicos para lograr la conformación y transmisión al aire del contenido siguiendo el estándar definido por cada país. En [13] se muestra que para el empaquetado y entrega del contenido de los servicios interactivos de la TDT se requiere de un proceso de alta complejidad, dependencia tecnológica y dificultad de adaptación a nuevos requerimientos, generando gastos económicos al país para mantener esta infraestructura. En dicho trabajo se presenta una nueva solución que obtiene los datos de interactividad, los empaqueta siguiendo el estándar definido y entrega los TS tanto por

IP/UDP como mediante un dispositivo de modulación. El proceso se probó a nivel de laboratorio y se verificó su correcto funcionamiento en varios STB disponibles y en el análisis de los TS formados siguiendo la norma. El sistema PaqTVC+ es multiplataforma y requiere pocos recursos para su funcionamiento, por lo que puede ser utilizado como una herramienta práctica en la TVDi cubana y abre nuevas posibilidades para su desarrollo.

Del lado del usuario final, en [14] se muestra como los avances en las tecnologías de la información (TI) están logrando la convergencia tecnológica en la mayoría de los escenarios de la vida cotidiana. Los sistemas informáticos desde la nube hasta los teléfonos móviles y los dispositivos portátiles se conectan entre sí para proporcionar a las personas contenido interactivo y una amplia variedad de servicios que se presentan en diferentes formatos de representación. Sin embargo, estas ventajas no se aprovechan plenamente y cada escenario de aplicación, como la televisión digital y las salas inteligentes, utilizan diferentes dispositivos informáticos que aumentan la dependencia de proveedores extranjeros, aumentan los costos de producción o comercialización y no pueden reutilizarse en otros escenarios similares, ni para ampliar sus posibilidades y alargar su período de obsolescencia. Esta situación se agrava aún más en el contexto de Cuba como país bloqueado y asediado por los Estados Unidos de América. Las industrias de equipos médicos y turismo son ejemplos actuales de estas limitaciones al no poder comprar suministros electrónicos, ver limitada su comercialización o no poder ofrecer una mayor comodidad de los servicios de TI a sus clientes. A partir de esta situación, este trabajo se planteó como objetivo identificar posibles módulos informáticos que puedan adaptarse a diferentes escenarios mediante acoplamiento modular de interfaces de expansión para lograr la conexión con diferentes periféricos (tales como: USB, entrada RF y salida RCA, salida HDMI, Interfaz Ethernet, WiFi y otros). Los diferentes escenarios propuestos al principio fueron: monitor de parámetros médicos vitales, decodificador de televisión digital terrestre (dos modelos: uno minimalista con funciones muy básicas y otro

mejorado con funcionalidades de minicomputadora), domótica, control de acceso, agricultura de precisión y otros.

En los trabajos previamente referenciados [12][13][14] las tecnologías utilizadas se ajustan a la filosofía del software libre, lo que permite adaptarlo a otros escenarios tecnológicos existentes y abaratar los costos y, lo más importante, ganar en la soberanía del contenido y las tecnologías asociadas a su producción, transmisión y recepción de la TVDi.

Líneas de investigación y desarrollo

Las líneas de investigación y desarrollo principales son:

- Contenidos y Aplicaciones Interactivas para TVDi, con especial interés en asistencia a adultos mayores
- Herramientas informáticas para el desarrollo y evaluación de aplicaciones para TVDi
- Experiencias de usuario con la TVDi: usabilidad y accesibilidad

Resultados

- Se organiza anualmente, desde 2012 a la actualidad, un evento científico que permite el intercambio de las investigaciones en el ámbito iberoamericano, denominado Jornadas de difusión y capacitación de Aplicaciones y Usabilidad de la TVDi [15-29]. Durante 2020 se realizó la 9a Conferencia Iberoamericana de Aplicaciones y Usabilidad de la TVDi jAUTI 2020, Aveiro, Portugal, 18 dic 2020.
- Se están desarrollando contenidos y aplicaciones de TVDi para asistir el mejoramiento de las alteraciones de la marcha en personas mayores, con el objetivo de probar los desarrollos realizados en Ecuador
- Se planifica la realización de cursos de doctorado conjunto con profesores de la UNLP y las instituciones extranjeras colaboradoras

Formación de recursos humanos

Participan en el proyecto PGTF un investigador formado de la UNLP en conjunto con dos investigadores formados de LAVID Universidad de Paraíba (Brasil) y dos investigadores formados de CUJAE (Cuba).

En la actualidad se está desarrollando una tesis de doctorado en el marco de esta línea de investigación dirigida conjuntamente con un investigador de la Universidad de Aveiro:

- Magdalena Rosado “Televisión Digital Interactiva (TVDi) para reducir las alteraciones de la marcha en adultos mayores”. Directores: M.J.Abásolo y T. Silva (tesis de Doctorado en Ciencias UNLP en curso)

Referencias

- [1] Abásolo M.J., De Giusti A., Naiouf M., Pesado P., Rosado Alvarez M., Silva T., Pina J., Kulesza R.(2019) *Aplicaciones de la Televisión Interactiva y aplicaciones móviles para el mejoramiento de los pueblos latinoamericanos*. XXII Workshop de Investigadores en Ciencias de la Computación WICC 2020, RedUNCI, ISBN 978-987-3714-82-5
- [2] Abásolo M.J., De Giusti A., Naiouf M., Pesado P., Rosado Alvarez M., Pina J., Kulesza R., Silva T. (2019) *Aplicaciones de la Televisión Interactiva y tecnologías afines para el mejoramiento de los pueblos latinoamericanos*. XXI Workshop de Investigadores en Ciencias de la Computación WICC 2019, Universidad Nacional de San Juan, RedUNCI, ISBN 978-987-3619-27-4
- [3] Abásolo M.J., De Giusti A., Sanz C., Pesado P., Martorelli S., Artola V., Naiouf M., Zangara A., Santos G., Casas S. (2016) *RedAUTI Red de Aplicaciones y Usabilidad de la Televisión Digital Interactiva (TVDi)* XVIII Workshop de Investigadores en Ciencias de la Computación WICC 2016, Entre Ríos, Argentina
- [4] Abásolo M.J., De Giusti A., Naiouf M., Sanz C., Zangara A., Perales Lopez F., Santos G., Casas S., González N., Paz M. (2015) *RedAUTI Red de Aplicaciones y Usabilidad de la Televisión digital Interactiva*. XVII Workshop de Investigadores en Ciencias de la Computación WICC 2015, Salta, Argentina
- [5] Abásolo, M.; Mitaritonna, A.; Giacomantone, J.; De Giusti, A.; Naiouf, M.; Perales, F.; Manresa, C.; Vénere, M.; García Bauza, C. *Realidad Virtual, Realidad Aumentada y TVDI*. WICC 2014 XVI Workshop De Investigadores en Ciencias de la

Computación, Ushuaia, Argentina

- [6] Rosado M., Abásolo M.J., Silva T. (2021). *Revisión de experiencias de aplicaciones interactivas para la televisión digital ecuatoriana*. En Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: jAUTI 2020. Universidad de Aveiro (en prensa)
- [7] Rosado, M., Abásolo, M. J., & Silva, T. (2019). *ICT Oriented to the Elderly and Their Active Aging: A Systematic Review*. In Iberoamerican Conference on Applications and Usability of Interactive TV (pp. 134-155). Springer, Cham.
- [8] Rosado M., Abásolo M.J., Silva T. (2020) *Revisión de TIC Orientadas al Adulto Mayor y su Envejecimiento Activo*. En Abásolo M.J., Kulesza R. y Pina-Amargós J. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: jAUTI 2019. Universidad Nacional de La Plata.
- [9] Rosado M., Abásolo M.J., Silva T. (2021). *Contenidos interactivos para TVDI destinados a reducir las alteraciones de la marcha en adultos mayores*. En Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: jAUTI 2020. Universidad de Aveiro (en prensa).
- [10] Kulesza, R.; Toscano, R.; Alves, K.; Costa, R.; Donin Noleto, J.; Aires Moreira, G.; Medeiros, F.; Dias, C. (2021) *Uma proposta de ferramenta de avaliação de interfaces de interação para TV Digital*. En Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: jAUTI 2020, Aveiro, Portugal, 18 dic. 2020. Universidad de Aveiro (en prensa)
- [11] Kulesza, R.; Toscano, R.; Alves, K.; Costa, R.; Nobrega, R.; Feitosa, M.; Souza Filho, G. (2021) *Canal Social: desenvolvimento e avaliação de uma aplicação de TV Interativa voltada para serviços sociais no contexto de governo eletrônico brasileiro*. En Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: jAUTI 2020, Aveiro, Portugal, 18 dic. 2020. Universidad de Aveiro (en prensa)
- [12] Pina-Amargós, J.D., Socorro-Llanes, R. y col.: *Incorporation of immediacy, dynamics and interactivity to digital terrestrial television services in Cuba through TVC+*. In: *Iberoamerican Conference on Applications and Usability of Interactive TV*. pp. 3–15. Springer (2019)
- [13] González-Fernández, J.C., Pina-Amargós, J.D, y col. (2019) *Nueva solución para el empaquetado y entrega de los servicios interactivos de la TDT en Cuba*. En Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: jAUTI 2020, Aveiro, Portugal, 18 dic. 2020. Universidad de Aveiro (en prensa)

- [14] Pina-Amargós, J.D, Valdés-Zaldívar, E.E. (2019) *Selección de módulos de cómputo configurables para escenarios emergentes*. En Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: *jAUTI 2020*, Aveiro, Portugal, 18 dic. 2020. Universidad de Aveiro (en prensa)
- [15] Abreu, Abásolo, Almeida, Silva. (eds) Libro de aplicaciones y usabilidad de la televisión digital interactiva: *jAUTI 2020*, Aveiro, Portugal, 18 dic. 2020. Universidad de Aveiro (en prensa)
- [16] Abásolo, M. J., Kulesza, R., & Amargós, J. D. P. (Eds.). (2020). *Applications and Usability of Interactive TV: 8th Iberoamerican Conference, JAUTI 2019, Rio de Janeiro, Brazil, October 29-November 1, 2019, Revised Selected Papers* (Vol. 1202). Springer Nature.
- [17] Abásolo M.J., Kulesza R. y Pina-Amargós J. (eds) *Libro de aplicaciones y usabilidad de la televisión digital interactiva : AUTI 2019 8va Conferencia de Aplicaciones y Usabilidad de la TVDi, Rio de Janeiro, Brasil, 30 oct-1 nov 2019*. Universidad Nacional de La Plata, ISBN 978-950-34-1861-1
- [18] Abásolo M.J., Silva T. and González D. (eds.) (2019) *Applications and Usability of Interactive Television 7th Iberoamerican Conference, jAUTI 2018, Bernal, Argentina, October 16–18, 2018, Revised Selected Papers*. Communications in Computer and Information Science 1004, Springer ISBN 978-3-030-23862-9
- [19] Abásolo M.J., González D. (eds.) (2019) *Anales 7ma Conferencia Iberoamericana JAUTI 2018: Televisión Digital, estudios del audiovisual y nuevas plataformas y VIII Jornadas Transversales de TV Digital*. Universidad Nacional de Quilmes
- [20] Abásolo, María José; Ferraz de Abreu, Jorge; Almeida, Pedro; Silva, Telmo (eds.) (2018) *Applications and Usability of Interactive Television 6th Iberoamerican Conference, jAUTI 2017, Aveiro, Portugal, October 12-13, 2017, Revised Selected Papers*. Communications in Computer and Information Science 813, Springer ISBN 978-3-319-90170-1
- [21] Ferraz de Abreu, Jorge; Abásolo Guerrero, María José; Almeida, Pedro; Silva, Telmo (Eds.) (2017) *Anales jAUTI'17. 6ª Jornadas de Aplicaciones y Usabilidad de la Televisión Digital Interativa*. UA Editora, ISBN 978-972-789-521-2
- [22] Abásolo, María José; Pedro Almeida; Joaquín Pina Amargós (eds.) (2017) *Applications and Usability of Interactive TV 5th Iberoamerican Conference, jAUTI 2016 La Habana, Cuba November 21-15, 2016 Revised Selected Papers*. Communications in Computer and Information Science 605, Springer ISBN 978-3-319-63321-3, julio 2017
- [23] Abásolo Guerrero, María José; Pina Amargós, Joaquín (2017) *Libro de aplicaciones y usabilidad de la televisión digital interactiva. V Jornadas Iberoamericanas sobre Aplicaciones y Usabilidad de la Televisión Digital Interactiva, jAUTI2016. Artículos seleccionados*. UNLP, ISBN: 978-950-34-1514-6
- [24] Abásolo, María José; Perales López, Francisco; Bibiloni Coll, Antoni (eds.) (2016) *Applications and Usability of Interactive TV 4th Iberoamerican Conference, jAUTI 2015 and 6th Congress on Interactive Digital TV, CTVDI 2015 Palma de Mallorca, Spain, October 15–16, 2015 Revised Selected Papers*. Communications in Computer and Information Science 605, Springer ISBN 978-3-319-38906-6, mayo 2016
- [25] A.Bibiloni, F. Perales, M.J.Abásolo (Eds.) (2016) *“Proceedings del VI Congreso de Televisión Digital Interactiva - jAUTI 2015 IV Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva”* celebrado en Palma (España) del 14 al 16 de octubre 2015. Universitat de les Illes Balears (España). ISBN: 978-84-608-5567-5
- [26] Abásolo, M.J., Kulesza, R. (Eds.) (2015) *Applications and Usability of Interactive TV Third Iberoamerican Conference, jAUTI 2014, and Third Workshop on Interactive Digital TV, Held as Part of Webmedia 2014, João Pessoa, PB, Brazil, November 18-21, 2014. Revised Selected Papers* ISBN: 978-3-319-22656-9, Communications in Computer and Information Science Springer-Verlag,
- [27] Abásolo, M.J., Kulesza, R. (Eds.) (2015) *“Proceedings de jAUTI 2014 III Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva | Workshop WTVI Webmedia 2014”* celebrado en João Pessoa, Paraíba (Brasil), Noviembre 2014, ISBN: 978-950-34-1188-9, Universidad Nacional de La Plata (Argentina), 2015
- [28] Abásolo, M.J., Castro, C. de (Eds.) (2014) *“Anales jAUTI 2013 II Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva”* celebrado en Córdoba (España) en septiembre 2013, ISBN: 978-84-697-0302-1. Universidad de Córdoba (España), 2014.
- [29] Abásolo, M.J., De Giusti A. (Eds.) (2013) *“Anales jAUTI 2012 I Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva”* celebrado en La Plata (Argentina) en octubre 2013. ISBN: 978-950-34-0945-9. Universidad Nacional de La Plata (Argentina), 2013.

Métricas de Inmersión para Sistemas de Realidad Virtual mediante Modelos de Regresión

Matías N. Selzer, Martín L. Larrea y Silvia M. Castro

Departamento de Ciencias e Ingeniería de la Computación,
Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica, (UNS-CIC Prov. de Buenos Aires)
{matias.selzer, mll, smc}@cs.uns.edu.ar

Resumen

La Realidad Virtual es una tecnología que intenta sumergir a los usuarios en un mundo virtual generado por computadora. Los avances tecnológicos ayudan a generar sistemas cada vez más inmersivos, pero aún no está claro qué factores afectan la inmersión del sistema. Esta línea de investigación tiene como objetivo relevar dichos factores y analizar su impacto y relación con distintas técnicas computacionales y estadísticas. Se busca desarrollar métricas que predigan el nivel de inmersión percibido en cualquier sistema de Realidad Virtual dado. Hasta el momento, realizamos diversos experimentos con usuarios para poder obtener un conjunto de datos que se pueda analizar de manera estadística y a partir de los cuales se puedan generar métricas de inmersión mediante modelos de regresión. Estas métricas y la importancia de cada factor brindará una guía para poder seleccionar el sistema y técnicas de Realidad Virtual que mejor se adapten a cada tipo de sistema y aplicación.

Palabras clave: *Realidad Virtual, Interacción Humano Computadora en Realidad Virtual, Inmersión*

Contexto

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. Introducción

La Realidad Virtual (RV) es una tecnología inmersiva que se ha vuelto muy popular en los últimos años. Mediante el uso de esta tecnología, y más específicamente mediante el uso de dispositivos de hardware como visores especiales, los usuarios pueden adentrarse, explorar e interactuar en mundos 3D generados por computadora. Aunque esta tecnología ha sido mayormente explorada en el área de videojuegos, también se ha utilizado en gran medida en áreas como la medicina [1, 2, 3], la arqueología [4, 5], el entrenamiento militar [6], o diferentes tipos de simulaciones [7].

Para decidir qué tan “bueno” es un sistema de RV, la literatura presenta distintas métricas, como *presencia*, *inmersión* y *realismo*. En RV, la presencia se define como “la sensación de estar ahí”, es decir, la sensación de que uno de verdad pertenece al mundo virtual que actualmente está experimentando [8]. La inmersión, por el otro lado, está más relacionada al hardware utilizado y se define como “qué tanto de nuestros sentidos está comprendido por el sistema de RV?” [9]. Por este motivo, la presencia es una medida subjetiva que depende de la percepción de cada usuario; mientras que la inmersión es una medida objetiva que depende del sistema de RV que se esté utilizando. Se suele pensar que en un sistema de RV es fundamental maximizar el nivel de realismo gráfico para obtener la mayor sensación de presencia. Sin embargo, está demostrado que esto no es necesario [10].

Además de las características de hardware, que mejoran constantemente, existen muchas otras características de los sistemas de RV que afectan la experiencia del usuario. Variables como la resolución de pantalla, el sistema de audio utilizado, la forma de caminar por el ambiente virtual, entre otras, son algunas de las variables que, de acuerdo a la literatura, influyen en decidir cuándo un sistema de RV es mejor que otro [11, 12, 13]. La decisión acerca de qué variables considerar en un sistema de RV y los valores que éstas deben tomar es un problema desafiante. Además, existe mucha discusión entre la verdadera influencia de dichos factores e incluso la relación entre ellos.

Por este motivo, es de vital importancia realizar un estudio exhaustivo de todos aquellos factores que pueden influir en la sensación de inmersión producida por cualquier sistema de RV. Dicho análisis revelará los factores más relevantes que deben tenerse en cuenta a la hora de crear un nuevo sistema o a la hora de seleccionar uno de los sistemas existentes. Este análisis proveerá guías para la generación de nuevos sistemas de RV, tanto en el área de hardware como de software.

2. Líneas de Investigación, Desarrollo e Innovación

Actualmente, en esta línea de investigación, se han realizado estudios respecto al análisis y relevamiento de todos aquellos factores que influyen en la sensación de inmersión en un sistema de RV. Dichas variables han sido categorizadas según el tipo de inmersión que proveen en *inmersión visual*, *inmersión auditiva* e *inmersión táctil*. Como hasta el momento no existe una forma de cuantificar la inmersión de un sistema de manera automática, se deben utilizar distintos test con usuarios, en los que éstos evalúan el grado de inmersión del sistema de RV dado.

Para esto se han desarrollado experimentos en

los que los usuarios deben realizar una tarea mientras experimentan un sistema de RV. Este sistema está caracterizado por todas las variables analizadas cuyos valores cambian cada prueba. De esta manera, se obtienen valores de inmersión asociados a los valores de las variables utilizadas. Cada vez que se realiza un test, se obtiene una evaluación de inmersión de un conjunto de variables y esto genera una muestra en un conjunto de datos que se va construyendo.

A partir de este conjunto de datos se realizan análisis estadísticos para evaluar la relación entre las variables y cómo inciden éstas en la inmersión. Además, dentro de esta línea de investigación, se pretende hallar una métrica que relacione aquellos factores que influyen en la inmersión. Esta métrica permite predecir entonces el nivel de inmersión que brindaría cualquier sistema de RV dado. Para esto se están investigando distintas estrategias y técnicas, pero mayormente se están utilizando análisis de regresión.

En síntesis, en esta línea de investigación se trabaja en:

- **Selección y Clasificación de Variables:** luego de haber realizado los correspondientes experimentos, se vio que algunas de las variables que se consideraban relevantes resultaron no serlo, en tanto que algunas que se consideraban no relevantes resultaron serlo. Dentro de esta línea se analiza y mejora la clasificación de variables que efectivamente resultan influyentes en la sensación de inmersión de un sistema de RV. El análisis de estas variables es muy importante ya que de esta manera se puede avanzar en la optimización de la o las métricas de inmersión.
- **Experimentos:** a medida que se realizan experimentos surgen nuevas ideas y estrategias para mejorar su diseño y poder recolectar así una mayor cantidad de datos. Esta línea analiza los experimentos realizados, así como otros experimentos de

la literatura para poder obtener mejores resultados. Además, se planifican nuevos experimentos para analizar distintas relaciones entre las variables.

- **Análisis Estadísticos:** a partir de los datos generados en los experimentos se pueden comparar los factores entre sí para determinar si existe alguna relación estadísticamente significativa entre los mismos. De esta forma se puede calcular la interacción entre ellos y la influencia que tienen en el valor final de inmersión.
- **Predictor de Inmersión:** dado un determinado sistema de RV, resulta importante predecir el grado de inmersión que éste produce. En esta línea se trabaja en análisis de regresión, reducción de variables y verificación de modelos para la generación de dicho predictor a partir de un determinado *dataset* ya construido. Se está profundizando en técnicas de Machine Learning para lograr obtener distintas métricas en función de los mismos datos.
- **Optimización de las Métricas:** resulta muy interesante poder hallar qué variables y qué valores de dichas variables maximizan la inmersión. Como la cantidad de factores que influyen en el nivel de inmersión de los sistemas de RV es muy grande (al menos 30), resulta difícil optimizar manualmente cada uno de ellos cuando se desea desarrollar un nuevo sistema. Por este motivo, es importante identificar aquellos factores que más influyen en el nivel de inmersión final. Al reducir considerablemente la cantidad de factores, el diseñador de un nuevo sistema puede enfocarse en la optimización de los mismos, mejorando así el grado de inmersión del sistema. Además, se están estudiando distintas estrategias de optimización de funciones para poder maximizar los modelos de inmersión obtenidos.

3. Resultados y Objetivos

Sobre los ejes presentados se han obtenido resultados concluyentes. Actualmente se ha realizado una clasificación de todas aquellas variables que influyen en la inmersión y se ha desarrollado un sistema de RV que genera aleatoriamente diferentes escenarios en función de cada uno de los factores identificados. Por ejemplo, un escenario tendrá determinada resolución de pantalla, nivel gráfico, tipo de audio, tipo de tracking, etc., en tanto que otro escenario generado tendrá factores completamente distintos. Este sistema permite a los usuarios experimentar un escenario generado durante un tiempo determinado para que finalmente dichos usuarios califiquen el escenario según el nivel de inmersión experimentado.

Se han realizado diversos experimentos utilizando este sistema. Éstos han permitido obtener un conjunto de datos que relaciona las variables del sistema de RV con la inmersión percibida por el usuario. A partir de este conjunto de datos se han realizado análisis estadísticos que nos han brindado información relevante sobre qué variables influyen en mayor o menor medida en la inmersión. Este análisis de variables nos ha servido para seleccionar aquellas variables a considerar en los modelos de regresión.

Finalmente se han realizado análisis de regresión para obtener modelos de inmersión en función de las variables del sistema. Estos modelos varían según la eficiencia de predicción, la cantidad de variables que utilizan y la cantidad de coeficientes.

Actualmente se están analizando estos modelos obtenidos, particularmente se están testeando en sistemas de RV comerciales y se están evaluando diversas técnicas para maximizar la inmersión.

4. Formación de Recursos Humanos

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo y

concluidas relaciones con las líneas de investigación presentadas:

Tesis Concluidas

“Interacción Humano Computadora en Ambientes Virtuales”, tesis de Magister en Ciencias de la Computación. Alumno: Matías Selzer. Director: Martín Larrea.

“La Realidad Virtual como Herramienta para el Desarrollo Arquitectónico”, tesis de grado de Ingeniería en Sistemas de Información. Alumnos: Facundo Reissing, Sebastián Vicente. Director: Martín L. Larrea. Colaborador: Matías N. Selzer.

“Métricas de Inmersión para sistemas de Realidad Virtual”, tesis de Doctorado en Ciencias de la Computación. Alumno: Matías Selzer. Director: Silvia M. Castro. Co-Director: Martín L. Larrea.

5. Bibliografía

1. Freedman, S. A., Dayan, E., Kimelman, Y. B., Weissman, H., & Eitan, R. (2015). Early intervention for preventing posttraumatic stress disorder: an Internet-based virtual reality treatment. *European journal of psychotraumatology*, 6.
2. Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B. & Ressler, K. J. (2014). A randomized, double-blind evaluation of D-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in Iraq and Afghanistan War veterans. *American Journal of Psychiatry*.
3. Gorini, A., & Riva, G. (2014). Virtual reality in anxiety disorders: the past and the future. *Expert Review of Neurotherapeutics*.
4. Lynch, J., & Corrado, G. (2014). Arqueología virtual aplicada al sitio Villavil, Catamarca, Argentina Virtual Archaeology applied to the site Villavil, Catamarca, Argentina.
5. Gagne, R., Gouranton, V., Dumont, G., Chauffaut, A., & Arnaldi, B. (2014). Immersia, an open immersive infrastructure: doing archaeology in virtual reality. *Archeologia e Calcolatori, supplemento 5*, 1-10.
6. Carroll, J. M. (Ed.). (2003). *HCI models, theories, and frameworks: Toward a multidisciplinary science*. Morgan Kaufmann.
7. Schreuder, H. W., Persson, J. E., Wolswijk, R. G., Ihse, I., Schijven, M. P., & Verheijen, R. H. (2014). Validation of a novel virtual reality simulator for robotic surgery. *The Scientific World Journal*, 2014.
8. Bob G Witmer and Michael J Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.
9. Mel Slater. Measuring presence: A response to the witmer and singer presence questionnaire. *Presence*, 8(5):560–565, 1999.
10. Rheingold, H. (1991). *Virtual reality: exploring the brave new technologies*. Simon & Schuster Adult Publishing Group.
11. Khanna, P., Yu, I., Mortensen, J., and Slater, M. (2006). Presence in response to dynamic visual realism: A preliminary report of an experiment study. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST '06*, pages 364–367, New York, NY, USA. ACM.
12. Jsselsteijn, W., Ridder, H. d., Freeman, J., Avons, S. E., and Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments*, 10(3):298–311.

13. Skarbez, R., Brooks, Jr., F. P., and Whitton, M. C. (2017a). A survey of presence and related concepts. *ACM Comput. Surv.*, 50(6):96:1–96:39.

Realidad Extendida y Aplicaciones Móviles aplicadas a la Educación y a brindar Información al Ciudadano

María José Abásolo^{1,2}, Wilma Gavilanes³, Evelyn Del Pezo Izaguirre⁴, Tomas Allisiardi⁵, María José Bouciguez⁶, Graciela Santos⁶, Cesar A. Collazos⁷, Armando De Giusti¹, Marcelo Naiouf¹, Patricia Pesado¹, Cecilia Sanz¹

¹ III-LIDI Instituto de Investigación en Informática,
Facultad de Informática – Universidad Nacional de La Plata (UNLP), Argentina
{mjabasolo, degiusti, mnaiouf, ppesado, csanz}@lidi.info.unlp.edu.ar

² Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA), Argentina

³ Universidad Técnica de Ambato, Ecuador

⁴ Universidad del Río, Ecuador

⁵ Facultad de Ciencias Humanas, Universidad Nacional del Centro de la Pcia. de Bs.As., Argentina

⁶ Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Pcia. de Bs.As., Argentina

⁷ Departamento de Sistemas, Universidad del Cauca, Colombia

Resumen

La línea de investigación y desarrollo presentada en este artículo realiza actividades de análisis, desarrollo y evaluación de aplicaciones de Realidad Extendida, siendo este un término para englobar en este contexto aplicaciones de Realidad Virtual, Realidad Aumentada, Interfaces Multimodales, Visión por computador. Uno de los principales objetivos es la formación de recursos humanos y el fortalecimiento de la investigación mediante el trabajo intergrupar entre diferentes instituciones nacionales y extranjeras.

Palabras Clave: Realidad Extendida, Realidad Aumentada, Realidad Virtual, Interfaces multimodales, Visión por Computador

Contexto

Las diferentes líneas de investigación están enmarcadas en diversos proyectos:

- Proyecto IDEAS-CICPBA “AMAPAS Aplicaciones Móviles para la Medición de Agentes Peligrosos para el Ambiente y la Salud” (2019-2020) dirigido por M.J.Abásolo.

- Proyecto "Computación de Alto Desempeño: Arquitecturas, Algoritmos,

Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real" (2018-2021) acreditado UNLP Programa de Incentivos, dirigido por M. Naiouf.

- Proyecto "Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso" (2018-2021) acreditado UNLP Programa de Incentivos dirigido por P. Pesado

Introducción

Desde hace varios años el grupo de investigación III-LIDI se dedica al desarrollo de aplicaciones en las áreas de Visión por Computador, Realidad Aumentada y Realidad Virtual, con resultados publicados en ediciones anteriores de WICC [1-9]. En este artículo se presentan los principales avances alcanzados en esas diferentes líneas de investigación fruto de un trabajo conjunto con investigadores de diferentes instituciones del país en colaboración con el extranjero.

Realidad Aumentada en Educación Universitaria

El uso de Realidad Aumentada (RA) en educación ha tenido un gran efecto sobre todo por las posibilidades que presenta este recurso

en las aulas de clase. La motivación y el interés del alumnado han sido algunas de las potencialidades que se evidencian dentro del quehacer educativo, según señala Gavilanes et. al. [10] en su resumen de revisiones sobre RA en la educación.

Como una experiencia innovadora en la Universidad Técnica de Ambato en Ecuador se viene trabajando desde el 2018 en el diseño de materiales educativos con RA. Se trabaja con los estudiantes de la carrera de Docencia en Informática quienes son los diseñadores de los recursos. Para ello se implementó la metodología OARA desarrollada por la doctoranda Gavilanes [11][12], compuesta por 4 fases: Estructura de Información, Diseño Multimedia, Ensamblaje-Publicación y Validación. Cada una de las fases determina las acciones necesarias para desarrollar el material educativo, y una vez verificada su funcionalidad se lo publica en un sitio Web.

Los materiales educativos desarrollados son evaluados por los estudiantes consumidores quienes lo revisan y manipulan utilizando sus dispositivos móviles. Se les presenta una encuesta de aceptación tecnológica utilizando el modelo TAM y otra encuesta de validación del diseño. Los resultados alcanzados demuestran que el uso de los materiales educativos con RA desarrollados para la enseñanza universitaria tiene un alto grado de aceptación y motivación. Se determinó una mejora en el rendimiento académico y se analizaron ventajas y desventajas de las herramientas utilizadas para la producción.

Aplicaciones móviles y realidad extendida para niños sordos

La evolución de las tecnologías como el internet, la Realidad Extendida, el uso de los dispositivos móviles, entre otros, se han convertido en elementos claves para el desarrollo de las tareas cotidianas de las personas en el trabajo, la escuela, el hogar, etc.. Sin embargo, hay sectores específicos de la población como los sordos que necesitan que el enfoque de su uso satisfaga en alguna medida sus necesidades de comunicación con

el entorno, más aún cuando se promueve desde los gobiernos no solo mejorar sus condiciones de vida sino también su inclusión en la sociedad.

En una publicación de la doctoranda Del Pezo [13] se evidencia que en relación a la comunidad sorda se utiliza la tecnología móvil, pero muy poca utiliza la realidad extendida (específicamente la RA). Las aplicaciones se presentan más a manera de prototipos y se orientan a satisfacer en primera línea como intérprete la comunicación entre sordos y oyentes antes que para apoyar procesos educativos. Esto, sumado a que en muchos países se promueve la enseñanza de la primera lengua, en sordos la lengua de señas, se encontró que ésta lidera como sistema de comunicación en las aplicaciones encontradas. Los estudios actuales han priorizado la enseñanza a niños en áreas específicas como lenguaje, vocabulario, matemáticas, la misma lengua de señas, entre otras, para de esta manera contribuir a su desarrollo intelectual en el ámbito académico. Estos hallazgos ratifican la importancia de enseñar técnicas complementarias que fortalezcan el proceso de comunicación entre sordos y oyentes, como la lectura labial, la que expertos recomiendan que se fortalezca en edades tempranas. Por esta razón, el interés de próximos estudios está orientado a identificar las metodologías que se aplican en la enseñanza de esta técnica para poder sugerir nuevas opciones de apoyo al proceso educativo utilizando métodos colaborativos y lúdicos a través del uso de la tecnología móvil y extendida.

Videojuegos para la enseñanza de la Física

Las simulaciones, las herramientas de modelado, los mundos virtuales, los laboratorios virtuales y más recientemente los videojuegos, son algunas de las herramientas mediadoras utilizadas en la enseñanza de la Física. Bouciguez et al describen en [14] algunos videojuegos educativos para enseñar

Física y se caracterizan en función de la tecnología que involucra, los aportes y limitaciones al aprendizaje desde una perspectiva constructivista social.

El proceso de desarrollo de un videojuego educativo implica integrar componentes lúdicos, tecnológicos y educativos, lo cual requiere un trabajo articulado y continuo en equipos interdisciplinarios. La doctorando Bouciguez [15] trabajó el desarrollo del videojuego basado en misiones "SpaceEscape: The F.E.M." utilizado para una secuencia didáctica para enseñar la inducción electromagnética.

La jugabilidad educativa se define como el conjunto de propiedades que describen la experiencia del jugador en el entorno de juego, cuyo objetivo principal es proporcionar diversión a la vez de aprendizaje. La evaluación cuantitativa de la jugabilidad se puede realizar gracias a métricas que se basan en el estándar ISO 9126-4:2004 y son procedurales ya que están enfocadas al uso del videojuego y, por lo tanto, a medir la experiencia del jugador durante el proceso de juego con el producto. Al ser las métricas demasiado genéricas, González Sánchez [16] propone como complemento una evaluación cualitativa basada en inspecciones heurísticas, criterios de evaluación y cuestionarios. Las preguntas de los cuestionarios interrogan al jugador sobre su experiencia a la hora de jugar. Las mismas además de ofrecer información sobre qué propiedades de la jugabilidad educativa tienen más relevancia en la experiencia global del juego, permiten identificar qué elementos de un videojuego influyen en mayor o menor medida en ellas, permitiendo mejorar el diseño de dichos elementos.

Considerando lo antes mencionado, este tipo de evaluación adoptada requiere fundamentalmente prueba del videojuego con los jugadores, observación de dichos usuarios y la realización de cuestionarios basados en la jugabilidad educativa. Atendiendo al marco metodológico antes descrito para evaluar la jugabilidad educativa se recogió información que permitiera realizar el análisis cuantitativo

y cualitativo, en tres momentos claves: previo a la prueba con el videojuego se obtiene información sobre el perfil de jugador que tienen los estudiantes; durante la prueba se registran las jugadas de los jugadores-estudiantes mientras utilizan el videojuego; posterior a la prueba se obtiene información brindada por los participantes después de jugar sobre su experiencia con el videojuego.

Los elementos de recolección de datos principales fueron los registros en video y audio, para cuyo análisis se establecieron métricas, y cuestionarios o listas de cuestiones (listas de heurísticas), buscando investigar el grado de cada atributo de jugabilidad para obtener una medida de la experiencia del jugador.

Aplicaciones móviles para información de agentes peligrosos para la salud y el ambiente

Actualmente la sociedad está fuertemente informatizada y emplea dispositivos móviles tanto para lo laboral como para la vida cotidiana, resultando indispensables para el desarrollo de diferentes tareas sin distinción de edades en detrimento de quedar aislados si no hay acceso a estos. El desarrollo de los dispositivos móviles dio lugar al surgimiento de los Sistemas de Información Geográfica (SIG) móviles, lo cual permitió el crecimiento del mercado de los SIG y los conocimientos específicos del ámbito geográfico y cartográfico, ampliando la posibilidad de creación de datos e información geográfica directamente en campo, para luego ser aplicada y analizada directamente en un SIG de escritorio. Los dispositivos móviles permiten recoger con más precisión cualquier tipo de dato espacial sobre el terreno, al mismo tiempo que facilitan la creación de dicho dato espacial o la edición de uno ya existente en función de lo observado. Se unen en este punto la capacidad del dispositivo para conocer las coordenadas de su localización y las capacidades de las aplicaciones SIG para edición de datos.

Se tiene por objeto brindarles a los ciudadanos información georreferenciada referida a diferentes mediciones ambientales desde una aplicación móvil, de manera que puedan acceder a dicha información de acuerdo a su posicionamiento global. La finalidad última consiste en visualizar los datos tomados mediante una aplicación móvil basada en realidad aumentada, de modo tal que cualquier ciudadano/a pueda acceder a dicha información al enfocar con su cámara su entorno.

En particular, nos hemos centrado en mediciones de radiaciones de alta frecuencia emitidas por las antenas de telecomunicación. Se realizó una comparativa de diferentes aplicaciones móviles disponibles que permitan la entrada de datos georeferenciados. Se seleccionó la aplicación QField¹ como aquella que mejor cumplía con los requerimientos, considerando sus principales características técnicas y el gran alcance de sus herramientas al combinarlas con el software libre QGIS². Se estableció un protocolo de medición de radiaciones electromagnéticas, y se llevó a cabo la toma de datos en diferentes puntos de la ciudad de Tandil a partir de un equipo medidor de radiación electromagnética. Las mediciones fueron georreferenciadas en un SIG móvil con el objetivo de obtener un mapa que refleje dichas mediciones espacialmente. De este mismo modo, se creó una capa que contiene la ubicación de las antenas próximas a los núcleos urbanos.

Líneas de investigación y desarrollo

- Aplicaciones móviles para brindar información y servicios al ciudadano
- Aplicaciones móviles para enseñanza a niños sordos
- Realidad Aumentada aplicada a la educación universitaria

- VideoJuegos en la enseñanza de las ciencias

Resultados y Objetivos

- Con el objetivo de formar recursos humanos desde el año 2012 se dicta la carrera de postgrado “Especialización en Computación Gráfica, Imágenes y Visión por Computadora”, Facultad de Informática de la UNLP. (Nº 11.162/12).
- Se realizó un mapa de la ciudad de Tandil donde se ubicaron antenas y mediciones de radiaciones electromagnéticas de alta frecuencia. Por el momento el mapa puede ser visualizado en Google Maps³. Se trabaja actualmente en el desarrollo de una aplicación visualizadora de los datos georeferenciados basada en RA.
- Se han desarrollado diversos materiales educativos basados en RA siguiendo la metodología OARA, y se continúa con su evaluación en el ámbito universitario de la Universidad Técnica de Ambato (Ecuador), como parte de una tesis doctoral.
- Se está realizando un relevamiento en centros educativos para niños sordos, en la ciudad de Guayaquil (Ecuador) en base al cual se espera poder definir requerimientos de una aplicación móvil lúdica para la enseñanza de lectura labial, a desarrollarse como parte de una tesis doctoral.
- Se utilizó el videojuego “SpaceEscape: The F.E.M.” en una clase universitaria de la materia de Física II, en la temática inducción electromagnética, y se evaluó la experiencia de 27 alumnos como parte de una tesis doctoral.

Formación de recursos humanos

El equipo cuenta con un investigador formado trabajando en colaboración, para llevar a cabo la dirección conjunta de tesis doctorales, con

¹ <https://qfield.org/>

² <https://www.qgis.org/es/site/>

³

<https://www.google.com/maps/d/u/0/viewer?mid=12TAnlxBkzDtwqZ5pMGTrxmr72z-MSZEK>

tres investigadores de universidades nacionales, de España y de Colombia.

Las siguientes diferentes tesis de grado y postgrado se desarrollan en el marco de esta línea de investigación:

- Evelyn Del Pezo I. “Modelo colaborativo y lúdico basado en Aplicaciones Móviles para apoyar la enseñanza de la lectura labial a niños sordos” Director: Abásolo M.J. y Codirector Collazos C.A. (tesis de Doctorado en Ciencias en curso).

- Wilma Gavilanes “Metodología para la evaluación del impacto de experiencias con Realidad aumentada en educación superior” Director: Abásolo M.J. y Codirector J. Cabero (tesis de Doctorado en Ciencias en curso)

- María José Bouciguez “Ambientes virtuales altamente interactivos basados en videojuegos y simulaciones para la educación en ciencias” Directores: Santos, G. y Abásolo, M.J. (tesis de Doctorado en Ciencias, Fac. Inf. de la UNLP, finalizada pendiente de entrega)

- Florencia Puppo “Reconocimiento óptico de caracteres de la etiqueta nutricional de productos alimenticios”. Directores: Abásolo M.J., Ronchetti F. (tesina de grado de Licenciatura en Sistemas en curso)

- Tomás Allisiardi. “Mapa de antenas y mediciones de radiaciones electromagnéticas de alta frecuencia en Tandil” Tutora: Abásolo M.J. (Proyecto de prácticas profesionales de la Tecnicatura en Sistemas de Información Geográfica, UNICEN).

Además, se colabora en la formación de recursos humanos de otras universidades, entre los cuales se enumeran el becario doctoral de Conicet :

- Lucas Benjamin Cicerchia “Detección de enfermedades y falta de nutrientes en cultivos utilizando algoritmos de Active Learning aplicados al sensado remoto” Directores: Claudia Russo (UNNOBA), María José Abásolo (tesis de Doctorado en Ciencias UNLP en curso)

Referencias

[1] Abásolo M.J., De Giusti A., Naiouf M., Pesado P., Barbieri S., Gavilanes W.,

Mitaritonna A., Vincenzi, M., Bria, O.; Ronchetti, F.; Montero F., Perales López F.; Springer, V. *Aplicaciones de realidad extendida y Aplicaciones Móviles*. XXII Workshop de Investigadores en Ciencias de la Computación WICC 2020, RedUNCI, ISBN 978-987-3714-82-5

[2] Abásolo M.J., De Giusti A., Naiouf M., Pesado P., Sanz C., Barbieri S., Boza R., Gavilanes W., Mitaritonna A., Prinsich N., Vincenzi, M., Montero F., Perales López F. *Aplicaciones de realidad virtual, realidad aumentada e interfaces multimodales*. XXI Workshop de Investigadores en Ciencias de la Computación WICC 2019, Universidad Nacional de San Juan, RedUNCI, ISBN 978-987-3619-27-4

[3] Abásolo M.J., Mitaritonna A., Castañeda, S., Sanz C., Boza R., Prinsich N., Silva, T., Rosado, M.; Naiouf, M.; Pesado, P.; De Giusti, A. (2018) *Aplicaciones de visión por computador, realidad aumentada y TVDi*. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). p. 389-395, RedUNCI, ISBN: 978-987-3619-27-4

[4] Abásolo, M.J.; Sanz, C.; Naiouf, M.; De Giusti, A.; Santos, G.; Castro, M.; Bouciguez, M.J. G. (2017) *Realidad Aumentada, Realidad Virtual e Interacción Tangible para la Educación*. XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017), pp. 1312-1316, RedUNCI, ISBN 978-987-42-5143-5

[5] Abásolo, M.; Mitaritonna, A.; Bouciguez, M.; Encina, N.; Vincenzi, M.; De Giusti, A.; Naiouf, M.; Giacomantone, J.; Manresa Yee, C. *Realidad Aumentada, Realidad Virtual, Interfaces Avanzadas y Juegos Educativos*. XVIII Workshop de Investigadores en Ciencia de la Computación (WICC 2016); Concordia, Entre Ríos, RedUNCI, ISBN 978-950-698-377-2

[6] Abásolo, M.; Mitaritonna A.; Encina N.; Vincenzi M.; Borelli L.; De Giusti A.; Naiouf

- M.; Giacomantone J. *Realidad Aumentada y Realidad Virtual XVII Workshop de Investigadores en Ciencia de la Computación (WICC 2015)*, RedUNCI, ISBN 978-987-633-134-0
- [7] Abásolo, M.; Mitaritonna, A.; Giacomantone, J.; De Giusti, A.; Naiouf, M.; Perales, F.; Manresa, C.; Vénere, M.; García Bauza, C. *Realidad Virtual, Realidad Aumentada y TVDI*. WICC 2014 XVI Workshop De Investigadores en Ciencias de la Computación, RedUNCI, ISBN 978-950-34-1084-4
- [8] Manresa-Yee, C.; Abásolo, M.J.; Mas Sansó, R.; Vénere, M.. (2011) *Realidad Virtual, Realidad Aumentada e Interfaces Basadas en Visión*. XV Escuela Internacional de Informática, XVII Congreso Argentino de Ciencia de la Computación CACIC 2011. Editorial EDULP, ISBN 978-950-34-0765-3
- [9] Mitaritonna, A., Abásolo, M.; Montero Simaro, F. (2020) *Resumen de tesis: Empoderamiento de la Conciencia Situacional en Operaciones Militares utilizando Realidad Aumentada*. XXII Workshop de Investigadores en Ciencias de la Computación WICC 2020, RedUNCI, ISBN 978-987-3714-82-5
- [10] Gavilanes, W., Abásolo, M. J., Cuji, B. (2018) “*Realidad Aumentada en la Educación: una Revisión desde la Perspectiva Pedagógica*”, Revista ESPACIOS. ISSN 0798 1015
- [11] Gavilanes, W., Cuji, B., Toalombo, O., & Fiallos, J. C. (2021). Augmented Reality as an Academic Training Experience in Higher Education. *Advances in Intelligent Systems and Computing*, 1277, 106–116. https://doi.org/10.1007/978-3-030-60467-7_9
- [12] Gavilanes López W.L., Cuji B.R., Salazar Mera J.V., Abásolo M.J. (2020) *Methodology for the Production of Learning Objects Enriched with Augmented Reality by University Students*. In: Auer M., Hortsch H., Sethakul P. (eds) *The Impact of the 4th Industrial Revolution on Engineering Education*. ICL 2019. *Advances in Intelligent Systems and Computing*, vol 1134. Springer, Cham.
- [13] Del Pezo Izaguirre, E., Abásolo, M. J., & Collazos, C. A. (2020). *Uso de tecnologías móviles y realidad extendida para personas sordas: Una revisión sistemática de la literatura de acceso abierto*. XV Conferencia Latinoamericana de Tecnologías de Aprendizaje.
- [14] Bouciguez, M. J.; Santos, G. ;Abásolo, M. J. (2013) *Potencialidad de los videojuegos en el aprendizaje de Física*. Actas de WEFA 2013 I Workshop de Enseñanza de Física en Argentina, Universidad Nacional del Centro de la Pcia. de Bs. As., ISBN 978-950-658-342-2.
- [15] Bouciguez, M.J., Braunmüller, M., Bravo, B., Santos, G.y Abásolo, M. J. (2019) *Desarrollo del videojuego “SpaceEscape: The F.E.M.” para una secuencia didáctica de inducción electromagnética*. X Congreso Iberoamericano de Educación Científica Enseñanza y Aprendizaje de las Ciencias en Debate. Montevideo, Uruguay, 25, 26-28 de Marzo del 2019
- [16] González Sánchez, J. L. (2010). *Jugabilidad. Caracterización de la experiencia del jugador en videojuegos*. Tesis Doctoral, Universidad de Granada.

VERIFICACIÓN Y VALIDACIÓN DE REPRESENTACIONES VISUALES Y SUS INTERACCIONES

Martín Larrea, Dana K. Urribarri, Matías Selzer,

Clara Casalini, M. Luján Ganuza, Silvia M. Castro

Laboratorio de I+D en Visualización y Computación Gráfica

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)

Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

Laboratorio de I+D en Visualización y Computación Gráfica,

(UNS-CIC Prov. de Buenos Aires)

{mll, dku, matias.selzer, mcca, mlg, smc}@cs.uns.edu.ar

RESUMEN

Tanto en el ámbito profesional como personal, cada día más decisiones son tomadas en base al análisis visual mediante el uso de representaciones visuales de grandes conjuntos de datos. 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 este contexto resulta fundamental desarrollar nuevas y mejores metodologías y herramientas que permitan asegurarnos el correcto funcionamiento de las representaciones visuales y sus interacciones.

Palabras Clave: *Verificación y Validación, Visualización de Datos, Análisis Visual de Datos*

CONTEXTO

El proyecto de investigación se desarrollará dentro del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur. Actualmente el VyGLab cuenta con investigadores, becarios y estudiantes de postgrado trabajando intensamente en las áreas de Visualización, Computación Gráfica e Interacción Humano Computadora. El lugar de trabajo cuenta además con una biblioteca especializada, equipo propio de cómputo, acceso a revistas especializadas por Internet y varios

dispositivos con los cuales se podrán implementar y experimentar los resultados teóricos producidos. El proyecto propuesto está estrechamente relacionado con el proyecto PICT-2017-1246 “Análisis Visual de Datos en Geociencias” bajo la dirección de la Dr. Silva Castro. El VyGLab se encuentra en continuo contacto con el Institute of Computer Graphics and Algorithms of the Vienna University of Technology, un centro de investigación de prestigio internacional. Se han realizado pasantías en el mismo y se están buscando nuevas formas de colaboración y trabajo en conjunto. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La toma de decisiones se ha transformado en una herramienta clave en toda organización e individuo, pasando de ser un proceso basado en la experiencia y la intuición a uno cada vez más establecido en el análisis de datos. Los procedimientos usados para tomar decisiones, en especial en base a grandes conjuntos de datos, son más sofisticados y requieren software especializado. Un tipo de software especializado es el de representaciones visuales. Las representaciones visuales son de suma importancia en la actualidad ya que permiten la exploración efectiva de un

conjunto de datos y facilitan la tarea de identificar patrones y extraer conclusiones. Los sistemas de representaciones visuales no sólo ofrecen una vista de los datos, sino también proveen interacciones mediante las cuales el usuario puede manifestar sus necesidades y a partir de las cuales obtener la perspectiva deseada del conjunto de datos. Cada día más decisiones son tomadas en base al análisis visual mediante el uso de representaciones visuales. No sólo hay un aumento cuantitativo sino también cualitativo, las decisiones son más críticas y con mayor impacto en la sociedad, el medio ambiente y los individuos. En este contexto resulta fundamental desarrollar nuevas y mejores metodologías y herramientas que permitan asegurarnos el correcto funcionamiento de las representaciones visuales y sus interacciones ([1]).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Basándonos en el estado del arte de las representaciones visuales de grandes conjuntos de datos ([2]), y teniendo en cuenta sus aplicaciones ([3]), los estados intermedios de los datos ([4]) y las posibles interacciones ([5]), es posible definir e implementar metodologías y herramientas que permitan detectar errores y/o problemas en los software de representaciones visuales de grandes conjuntos de datos. La principal línea de investigación y desarrollo consiste en el estudio y desarrollo de metodologías para la validación y verificación de representaciones visuales de grandes conjuntos de datos. De esta línea principal se desprenden sublíneas relativas al diseño y desarrollo de herramientas que asistan en el desarrollo de representaciones visuales de grandes conjuntos de datos para asegurar la calidad del producto final, así también como herramientas de testeo para software de

propósito general ([13, 14]). Estas herramientas buscarán eliminar o reducir la incidencia de errores en los sistemas de representación visual. Esta línea de investigación expande nuestro trabajo sobre el área de Visualización ([6, 7, 8, 9, 10]).

3. RESULTADOS OBTENIDOS Y ESPERADOS

Sobre los ejes presentados se han obtenido resultados parciales. Desde el punto de vista de la teoría de base se ha comenzado a realizar un relevamiento de las diferentes metodologías disponibles en el ámbito de la Verificación y Validación del Software que sean aplicables a un dominio tan particular como el de la Visualización de Información. En lo referido al desarrollo de nuevas metodologías y herramientas se ha comenzado a publicar los resultados obtenidos hasta el momento. Tal es el caso de “Increasing Confidence in Industry 4.0 through new Software Verification and Validation Techniques” ([13]) y “TAPIR: An Object-Oriented Programming Testing Framework based on Message Sequence Specification with Aspect-Oriented Programming” ([14]). Durante el año 2020 se avanzó en nuevos desarrollos que se esperan publicar en el transcurso del 2021.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo y concluidas relaciones con las líneas de investigación presentadas, así también como un proyecto de grupo de investigación.

Tesis en Desarrollo: “Testing de Accesibilidad en Formularios Web. SIU-Kolla como Caso de Estudio.”, tesis de grado para la Ingeniería en Sistemas de Información. Alumna: Constanza Giorgetti.

Director: Martín Larrea, Colaborador: Matías Selzer.

Tesis en Desarrollo: “Metodologías de Análisis Estático en el Contexto de la Enseñanza de la Programación.”, proyecto final de carrera para la Ingeniería en Computación. Alumno: Tomas Miguel. Director: Martín Larrea.

Tesis concluida: “Testing de Visualizaciones Mediante Expresiones Regulares”, proyecto final de carrera para la Ingeniería en Computación. Alumno: Martín Schiaffino. Director: Martín Larrea.

Tesis concluida: “Testing de Software mediante Análisis Dinámico usando Aspectos en C#”, proyecto final de carrera para la Ingeniería en Computación. Alumno: Martín Hredil. Director: Martín Larrea, Co directora: Dana K. Urribarri.

Proyecto: PGI 2019 “VERIFICACIÓN Y VALIDACIÓN DE REPRESENTACIONES VISUALES Y SUS INTERACCIONES”. Director: Martín Larrea, Miembros: Matías Selzer, Dana K. Urribarri, M. Clara Casalini. Proyecto acreditado para incentivos.

5. BIBLIOGRAFÍA

[1] Kirby, Robert M., and Cláudio T. Silva. “The need for verifiable visualization”. *Computer Graphics and Applications*, IEEE 28.5, 78-83, 2008.

[2] Rees, D., and R. S. Laramée. “A Survey of Information Visualization Books.” *Computer Graphics Forum*. 2019.

[3] Börner, Katy, Andreas Bueckle, and Michael Ginda. “Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments.” *Proceedings of*

the National Academy of Sciences 116.6 (2019): 1857-1864.

[4] Hagen, Hans. “DFU, Volume 2, Scientific Visualization: Interactions, Features, Metaphors, Complete Volume.” *Dagstuhl Follow-Ups*. Vol. 2. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012.

[5] Ganuza, M. L. “Thesis Overview: Interactions in Visualization.” *Journal of Computer Science & Technology* 18.2 (2018): 178-179.

[6] Castro, Silvia Mabel, et al. “Métricas, técnicas y semántica para la visualización de datos.” *XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste)*. 2018.

[7] Urribarri, Dana K., and Silvia M. Castro. “Prediction of data visibility in two-dimensional scatterplots.” *Information Visualization* 16.2: 113-125. 2017

[8] Ganuza, María Luján, et al. “Visualización de Datos aplicada a resolver problemas en Ciencias Geológicas.” *XVI Workshop de Investigadores en Ciencias de la Computación*. 2014.

[9] Larrea, Martín Leonardo, et al. “Ontologías y Semántica en el Proceso de Visualización.” *XVI Workshop de Investigadores en Ciencias de la Computación*. 2014.

[10] Selzer, Matías Nicolás, et al. “Modelos de interacción y aplicaciones en realidad virtual mediante dispositivos móviles.” *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*. 2017.

[11] Larrea, Martín Leonardo. "Black-Box Testing Technique for Information Visualization. Sequencing Constraints with Low-Level Interactions." *Journal of Computer Science & Technology* 17. 2017.

[12] Silva, J.I.R., Larrea, M.: White-box testing framework for object-oriented programming based on message sequence specification. In: *XXIV Congreso Argentino de Ciencias de la Computación* (Tandil, 2018). 2018.

[13] Larrea, M. & Urribarri, D.: Increasing Confidence in Industry 4.0 through new Software Verification and Validation Techniques. *International Conference on Production Research ICPR Americas*, 10. In press. 2020.

[14] Larrea, M. & Urribarri, D.: TAPIR: An Object-Oriented Programming Testing Framework based on Message Sequence Specification with Aspect-Oriented Programming. *XXVI Congreso Argentino de Ciencias de la Computación*. In press. 2020.

VISUALIZANDO LA INFORMACIÓN EN CIENCIA DE DATOS

Mag. María Alejandra Malberti, Mag. Graciela Elida Beguerí, Mag. Raúl Oscar Klenzi, Lic. Manuel Ortega, Prog. Luis Olguín, Lic. Fabrizio Amaya, Joaquín Cortez

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: {amalberti, grabeda, rauloscarklenzi, manuel.ortega, lolguinunsj, fabrizio.amaya88, joaquinortez19}@gmail.com

RESUMEN

El presente trabajo plantea los avances del proyecto “Evaluación de visualizaciones eficientes en Ciencia de Datos” que tiene, entre otros, los siguientes objetivos: Análisis de distintos aspectos que atañen a una representación visual, búsqueda de datos abiertos y otras fuentes de datos provenientes de actividades de cooperación, análisis de herramientas de software libre en cuanto a sus potencialidades de visualización y análisis comparativo de los lenguajes Python y JavaScript como soporte de visualizaciones. Para tal fin, se están considerando diversas métricas tales como escala, longitud, área, color, entre otros. Así como visualizaciones de distintos tipos de datos e información con softwares libres y lenguajes de código abierto.

Palabras clave: Visualización, Ciencia de Datos, Lenguajes de código abierto, Software Libre.

CONTEXTO

El Proyecto articula líneas de investigación de un grupo de investigadores y docentes de las carreras Licenciatura en Sistemas de Información y Licenciatura en Ciencias de la Computación del Departamento de Informática de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan (FCEFN-UNSJ), y se encuentra contenido en el Laboratorio de

Sistemas Inteligentes para Extracción de Conocimiento en Datos Masivos del Instituto de Informática de la misma facultad. Asimismo, el proyecto se encuentra aprobado y subsidiado por Consejo de Investigaciones Científicas y Técnicas y de Creación Artística CICITCA- UNSJ.

La propuesta cuenta con antecedentes logrados en el tema conforme a sucesivos proyectos aprobados y subsidiados por el ente mencionado en los que el grupo viene trabajando desde el año 2003, siendo los desarrollados desde el año 2014:

- “Extracción de Conocimiento en Datos Masivos” 21/E-951, período 2014-2015
- “La Ciencia de Datos en grandes colecciones de datos” 21/E1014, período 2016-2017
- “Visualización y Deep Learning en Ciencia de Datos” 21/E1071, período 2018-2019

1. INTRODUCCIÓN

El constante aumento de datos y la complejidad de los mismos, ha traído como consecuencia problemas y más desafíos a la hora de su visualización.

Entre los problemas se pueden citar espacios de alta dimensión y relaciones complejas y como gran desafío la cantidad de datos que se generan permanentemente.

Es por ello que el proyecto citado “Evaluación de visualizaciones eficientes en

ciencia de datos” tiene como objetivo general proponer criterios para la evaluación de visualizaciones eficientes para Ciencia de Datos.

Behrisch, en el paper “Quality Metrics for Information Visualization” expone el pipeline de visualización –figura 1– con el cual da el caso en que el encargado de diseñar las visualización se enfrenta con el dilema de elegir una entre una infinidad de posibilidades de procesamiento de datos y una cantidad, aún mayor, de opciones de visualizaciones potenciales.

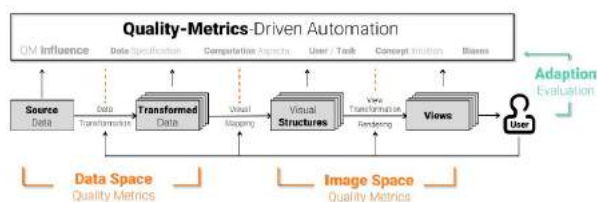


Figura 1. Analítica Visual basada en métricas de calidad con la capa adicional de automatización (Behrisch, et al., 2018)

https://bib.dbvis.de/uploadedFiles/QMSTAR_QualityMetricsForInformationVisualization_FINAL.pdf

En la publicación, ¿Cuántos datos se producen en un minuto?, del Grupo Bit se afirma que “*Un buen uso y análisis de los datos le puede dar a una marca o a una empresa la posibilidad de conocer características o Insights de sus consumidores que antes eran simplemente una hipótesis o eran desconocidos*” (Grupo Bit, 2020)

La Ciencia de Datos es el campo interdisciplinario tendiente a extraer conocimiento o un mejor entendimiento de datos en sus diferentes formas, ya sean estructurados o no estructurados.

Dada la gran cantidad de herramientas posibles de aplicar en cada proceso de visualización, el determinar cuáles son los objetivos perseguidos al visualizar es de relevancia. Una caracterización inicial propone dos grandes grupos: Visualizaciones para el análisis vs. Visualizaciones para la comunicación. En el primero de los grupos, lo que se busca es la exploración visual de los datos para que *ellos mismos hablen de su*

estructura y patrones (NIST/SEMATECH, 2012). Herramientas como Rstudio y Python poseen numerosas opciones para visualización.

El visualizar para comunicar implica conocer el usuario hacia quien va dirigido el mensaje de la visualización, por lo que se hace necesario *traducir la complejidad* hacia una forma más empática con el observador. Herramientas como Flourish, Data Studio, Tableau, RawGraphics aportan soluciones en este sentido. (Vega, R., 2019)

Jankun-Kelly et al. (2006) propone que para comprender la Visualización, es recomendable analizarla a partir de responder a tres preguntas claves: cómo se crea la visualización, qué sucede durante la visualización, y qué beneficio recibió el usuario o qué lo motivó para trabajar con la visualización. Estas preguntas sugieren tres tipos de modelos básicos:

- Modelo de transformación para describir el método de la visualización
- Modelo de exploración para describir el uso de la visualización, y
- Modelo de diseño para predecir o medir el éxito de un método de visualización. (Ponjuán, D., 2010) (Kelly, et al., 2006)

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las tareas que se están llevando a cabo comprenden varias líneas de investigación relacionadas con:

- Ciencia de Datos, principalmente lo relativo a Visualización de Información.
- Deep Learning a través del análisis y diferencias entre Boosted Decision Tree Regression y Redes Neuronales del tipo LSTM (Long Short Term Memory) en entorno de visualización open source Knowledge.
- Herramientas de software libre para arquitecturas secuenciales, paralelas y

distribuidas particularmente KNIME Analytics cuyo enorme potencial didáctico permite visualizar mediante workflows las diferentes etapas que constituyen el pipeline del Data Science.

- Lenguajes de programación de código abierto tales como Python, JavaScript y R.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Se está realizando, sobre diversos conjuntos de datos abiertos, una comparativa entre las bibliotecas de visualización disponibles en el lenguaje Python. Hasta ahora las bibliotecas exploradas son: Matplotlib, Seaborn, Plotly, Bokeh, Altair, y Folium. Se encuentra en desarrollo un documento con los siguientes criterios de evaluación de las bibliotecas citadas:

- Código e interfaz: qué características o particularidades tiene el código necesario para producir las visualizaciones en estas bibliotecas. Cuál es el nivel de su interfaz.
- Entorno de ejecución: sobre qué plataformas se puede ejecutar, otros programas que sean necesarios para el funcionamiento de la biblioteca, donde se visualizan los gráficos.
- Tipo de visualizaciones: qué tipo de visualizaciones va a estar en condiciones de generar la biblioteca en cuestión.
- Interactiva: Si la biblioteca genera gráficos con los que se pueda interactuar, o, por el contrario, son estáticos.
- Aportes y documentación: en este criterio se analiza qué tan completa es la documentación y cuántos aportes existen para esta biblioteca.
- Dificultad de uso: qué tan sencillo o complejo se puede llegar a tornar trabajar con estas bibliotecas.

También, se han creado varios Jupyter notebooks, con diversos conjuntos de datos libres, con los que se ejemplifican el uso de las bibliotecas.

En el caso particular del espacio vectorial de las palabras, área del procesamiento del lenguaje natural, se aplicaron técnicas de aprendizaje automático basadas en redes neuronales y se empleó el método Word2vec mediante la librería Gensim de Python, para su visualización.

Para el caso de datos temporales, se han revisado varias publicaciones y vale la pena mencionar, en particular, uno de los artículos científicos: “Integration of temporal data visualization techniques in a KDD-based DSS Application in the medical field” (Mohamed, et al. 2014) dado que en éste, se desarrollan temas tales como análisis de datos temporales y la visualización de estos; además de proporcionar una integración de técnicas de visualización (para representaciones interactivas). Ver figura 2

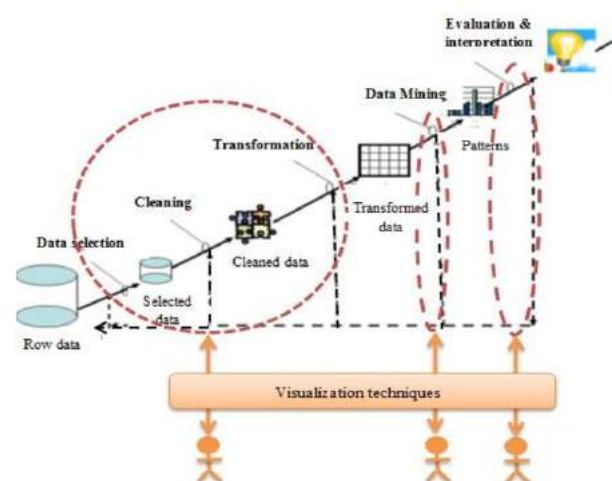


Figura 2. Integración de técnicas de visualización. Recuperado de “Integration of temporal data visualization techniques in a KDD-based DSS Application in the medical field”, de Mohamed, E, (2014).

<https://www.semanticscholar.org/paper/Integration-of-temporal-data-visualization-in-a-DSS-Mohamed-Ltifi/0acdc5c6758c34769ae5898a57781e4bd0580be4>

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto de investigación se encuentra conformado por ocho docentes-investigadores de distintas áreas de conocimientos que atañen a la Ciencia de Datos, tales como Estadística, Inteligencia Artificial, Lenguajes de Programación y Estructuras de Datos, y cuatro alumnos.

Los investigadores trabajan desde hace varios años en forma conjunta lo que ha permitido generar y asesorar, de modo integral, varios trabajos finales de grado, tesis de posgrado y Beca CIN - Consejo Interuniversitario Nacional- periodo 2020-2021.

Pertencen a las líneas de investigación los siguientes trabajos finales de la carrera Licenciatura en Ciencias de la Computación.

- “Herramienta de apoyo al aprendizaje de Metaheurísticas” (Autor: Olivares Juan Ignacio. Defendido en 2020)
- “Descripción de los procesos de recolección de datos y extracción de información en Redes Sociales en Ambientes Paralelos y Distribuidos” (Autor: Gouric Guillermo. Defendido 2020)
- Herramienta tecnológica de apoyo al aprendizaje: Problema del Viajante de Comercio, caso asimétrico (Autor: Cocinero, Pablo. En ejecución)

Con docentes investigadores del Instituto de Automática de la Facultad de Ingeniería INAUT-FI y personal del Instituto Nacional de Tecnología Agropecuaria INTA-San Juan, se lleva adelante la tesis de Maestría en Informática "Análisis de Fenómenos en Estaciones Agrometeorológicas mediante Ciencia de Datos" a cargo de un integrante del equipo de trabajo.

Es de destacar que en todos los trabajos citados, está planteado un apartado específico sobre visualización.

La Beca CIN “Abordaje de la Analítica Visual desde un lenguaje de programación-Python como caso de estudio” ha sido

otorgada al estudiante Joaquín Cortez, alumno de quinto año de la carrera Licenciatura en Ciencias de la Computación.

Como aporte a la sociedad se ha planificado un curso destinado a las pequeñas y medianas empresas de la provincia de San Juan, a partir de una iniciativa del gobierno y con el aval éste. Dicho curso tiene como finalidad introducir los conceptos básicos de Ciencia de Datos y mostrar las virtudes o ventajas de extraer conocimiento de los datos.

5. BIBLIOGRAFÍA

- Abela, A. (2008). *Advanced presentations by design: Creating communication that drives action*. John Wiley & Sons.
- Barcellos, R., Viterbo, J., Bernardini, F., & Trevisan, D. (2018, July). An Instrument for Evaluating the Quality of Data Visualizations. In 2018 22nd International Conference Information Visualisation (IV) (pp. 169-174). IEEE.
- Behrisch, M., Blumenschein, M., Kim, N. W., Shao, L., El-Assady, M., Fuchs, J., ... & Keim, D. A. (2018, June). Quality metrics for information visualization. In *Computer Graphics Forum* (Vol. 37, No. 3, pp. 625-662). https://bib.dbvis.de/uploadedFiles/QM_STAR_QualityMetricsForInformation_Visualization_FINAL.pdf
- Benoit, G. (2019). *Introduction to Information Visualization: Transforming Data Into Meaningful Information*. Rowman & Littlefield.
- Cady, F. (2017). *The data science handbook*. John Wiley & Sons.
- Chen, M., Feixas, M., Viola, I., Bardera, A., Shen, H. W., & Sbert, M. (2017). *Information theory tools for visualization*. AK Peters/CRC Press
- Cielen, D., Meysman, A., & Ali, M. (2016). *Introducing data science: big data, machine learning, and more*,

- using Python tools. Manning Publications Co..
- Erraissi, A., Mouad, B., & Belangour, A. (2019, April). A Big Data visualization layer meta-model proposition. In 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO) (pp. 1-5). IEEE.
 - Grupo Bit (2020) ¿Cuántos datos se producen en un minuto? <https://business-intelligence.grupobit.net/blog/cuantos-datos-se-producen-en-un-minuto>
 - Grus, J. (2019). Data science from scratch: first principles with python. O'Reilly Media.
 - How to choose a visualization (2019). <https://www.kdnuggets.com/2019/06/how-choose-visualization.html>
 - Klenzi, R. O., Malberti, A., & Beguerí, G. (2020). Evaluación de visualizaciones eficientes en ciencia de datos. In *XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz)*.
 - Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals. John Wiley & Sons.
 - Laura Igual Muñoz, & Santi Seguí Mesquida. (2017). Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications. Springer.
 - Lo, L. Y. H., Ming, Y., & Qu, H. (2019). Learning Vis Tools: Teaching Data Visualization Tutorials. arXiv preprint arXiv:1907.08796.
 - Mohamed, E.B., Ltifi, H., & Ayed, M.B. (2014). Integration of temporal data visualization techniques in a KDD-based DSS Application in the medical field. <https://www.semanticscholar.org/paper/Integration-of-temporal-data-visualization-in-a-DSS-Mohamed-Ltifi/>
 - [0acdc5c6758c34769ae5898a57781e4bd0580be4](https://doi.org/10.1007/978-1-4939-9878-8_10)
 - Murray, S. (2017). Interactive data visualization for the web: an introduction to designing with D3" O'Reilly Media, Inc."
 - Sosulski, K. (2018). Data Visualization Made Simple: Insights Into Becoming Visual. Routledge.
 - VanderPlas, J. (2016). Python data science handbook: essential tools for working with data. " O'Reilly Media, Inc."
 - Wang, J., Hazarika, S., Li, C., & Shen, H. W. (2018). Visualization and visual analysis of ensemble data: A survey. IEEE transactions on visualization and computer graphics.
 - Wang, C., & Shen, H. (2011). Information Theory in Scientific Visualization. Entropy, 13, 254-273.
 - Ware, C. (2012). Information visualization: perception for design. Elsevier
 - Wenqiang Cui (2019) Visual Analytics: A Comprehensive Overview. IEEE Xplore Digital Library. Volume 7: 2019 <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8740868>
 - Why Data Visualization Is The Most Important Skill in a Data Analyst Arsenal (2019). <https://www.kdnuggets.com/2019/08/simpliv-data-visualization-data-analyst.html>

Innovación en Educación Informática

Adecuación de estrategias y métodos para fomentar el pensamiento computacional

Gladys N. Dapozo, Cristina L. Greiner, Raquel H. Petris, Yanina Medina, Ana M. Company, María C. Espíndola

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste. 9 de Julio N° 1449. Corrientes. Argentina.
{gndapozo, cgreiner, rpetris, yanina, acompany, mcespindola}@exa.unne.edu.ar

RESUMEN

Este proyecto de investigación se orienta al estudio y aplicación experimental de estrategias educativas que incorporen métodos y herramientas innovadores que contribuyan a promover el pensamiento computacional a través de la enseñanza de la programación. En la universidad para mitigar indicadores de deserción y desgranamiento en estudiantes de Informática, y en formación docente de los niveles educativos preuniversitarios para apoyar las políticas públicas orientadas a incorporar las Ciencias de la Computación en las escuelas, iniciativas que buscan incidir en un mayor desarrollo de las áreas STEM (*Science, Technology, Engineering and Mathematics*). Se muestra un resumen de las actividades realizadas, los resultados obtenidos y las actividades previstas para continuar en estas líneas de investigación.

Palabras clave: Enseñanza de la programación. Didácticas específicas. Pensamiento computacional.

CONTEXTO

Las líneas de I/D corresponden al proyecto 16F018 “Promoción del pensamiento computacional para favorecer la formación en STEM”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE), iniciado en el año 2017.

1. INTRODUCCIÓN

Existe consenso a nivel global sobre las competencias requeridas para el siglo XXI, así como de la importancia de la formación en las áreas STEM (*Science, Technology, Engineering and Mathematics*). En particular, en América Latina y el Caribe el desempeño de los jóvenes en ciencias, y la inversión en innovación y desarrollo es menor que en otras regiones del mundo [1].

También existe una creciente demanda de desarrolladores de software mientras disminuye la cantidad de graduados en carreras de informática. Los países a nivel mundial desarrollan políticas para atender esta realidad.

Por otra parte, si bien es marcada la tendencia mundial en educación de fomentar el pensamiento computacional y la enseñanza de la programación desde niveles iniciales, una dificultad importante está vinculada con la formación de los docentes. Muchos de ellos no cuentan con las habilidades necesarias para enfrentar este desafío y, en muchos casos, tampoco están preparados con los nuevos enfoques de la didáctica de la programación [2].

En el marco de esta problemática, desde este grupo de investigación se estudian métodos y herramientas que se incorporan a la enseñanza para promover el pensamiento computacional. Este conocimiento se vuelca en acciones concretas que abarcan: la incorporación de nuevos

métodos para mejorar el proceso de enseñanza-aprendizaje de la programación en la universidad, favoreciendo la retención y permanencia de los alumnos, y la formación de los docentes de los niveles preuniversitarios con las competencias tecnológicas y pedagógicas necesarias para abordar el desafío de la enseñanza de la programación en las escuelas.

Contexto de las acciones

Conscientes de la necesidad de llevar adelante la formación de formadores capaces de lograr el objetivo de incorporar la Computación en las escuelas, la Fundación Sadosky convocó a las universidades nacionales para contribuir con los equipos docentes y la infraestructura para el dictado del curso “La Programación y su Didáctica”, destinado a los docentes de niveles preuniversitarios, de cualquier área de conocimiento, priorizando a docentes del área TIC y de Matemáticas.

En Corrientes, la UNNE (Universidad Nacional del Nordeste) participó en la formación de docentes de niveles preuniversitarios, dictando el curso “La Programación y su Didáctica” para 4 cohortes, desde el 2015 al 2018.

La experiencia adquirida en los cuatro años de capacitación a docentes y el contacto directo con los docentes y sus entornos de trabajo, permitió realizar un relevamiento de las necesidades fundamentales y detectar la motivación de los docentes y maestros de los diferentes niveles educacionales. Esto motivó a un equipo de docentes de la UNNE a elaborar una propuesta para la incorporación de la programación y la robótica en los niveles educacionales no universitarios, con el propósito de contribuir a la alfabetización y a la cultura digital, promovidas por las políticas públicas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación de este proyecto están enfocadas en:

- a) Estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática.
- b) Evaluación de métodos y herramientas para la enseñanza de programación y robótica para docentes de los niveles educativos no universitarios.
- c) Adecuación de la enseñanza de la programación a un modelo híbrido que aproveche las ventajas y los recursos de la modalidad de educación a distancia.

3. RESULTADOS OBTENIDOS/ESPERADOS

En la línea de estrategias educativas para la enseñanza de programación en carreras de Informática, en [3] se describe una estrategia para mejorar la comprensión de los conceptos básicos de la programación en cursos introductorios, utilizando un conjunto de actividades lúdicas, basadas en la programación por bloques. A través de estas actividades se introdujo un método de resolución de problemas que fue aplicado posteriormente en el desarrollo de los contenidos propios del curso. Para evaluar la estrategia se realizó una encuesta a los estudiantes y se analizaron los resultados parciales de su desempeño académico. Se pudo comprobar que los estudiantes consideraron de utilidad la propuesta pedagógica y se observó una leve mejoría en su desempeño académico. Adicionalmente, la estrategia resultó útil para mejorar la comprensión de los conceptos, especialmente algunos temas que en la enseñanza tradicional presentan dificultades, y se logró la

aplicación del concepto de legibilidad en las soluciones.

Por otra parte, el inicio del año lectivo 2020 fue fuertemente impactado por la pandemia por la Covid19. Enseñar programación en la modalidad remota de emergencia en una carrera de Informática es el desafío que abordó el equipo docente de la asignatura Algoritmos y Estructuras de Datos I, de primer año de la carrera Licenciatura en Sistemas de Información de la Universidad Nacional del Nordeste. Con el foco puesto en el aprendizaje, se elaboró una metodología para el dictado remoto que consideró la evaluación formativa como eje de la propuesta y el diseño de actividades para obtener evidencias de los aprendizajes, en el marco de las restricciones impuestas por el ASPO. En [4] se describe la metodología y los resultados obtenidos en el dictado realizado en el primer cuatrimestre del ciclo lectivo 2020. Se destaca como positivo haber revalorizado la evaluación formativa como facilitadora para el seguimiento del aprendizaje y como elemento fundamental para la evaluación sumaria que acredita los aprendizajes.

En la línea de formación docente, se evalúan métodos y estrategias que forman parte de la oferta académica denominada “Diplomatura en Programación y Robótica Educativa”, implementada desde el año 2019. Esta propuesta educativa tiene como principal objetivo acompañar las políticas públicas orientadas a incorporar las ciencias de la computación en las escuelas, mediante la formación de los docentes del nivel primario, secundario y terciario en el manejo de tecnologías de programación y robótica, adecuadas a los niveles preuniversitarios.

Se destaca que esta capacitación enfatiza la adquisición de habilidades digitales a través de una intensa práctica de programación y robótica, teniendo en cuenta permanentemente el enfoque didáctico, buscando no solo que el docente aprenda los temas disciplinares sino también, y muy importante, que los sepa transmitir adecuadamente.

Por ello, las actividades de evaluación, además de las cuestiones específicas de cada módulo, evaluarán la adquisición de las competencias digitales propuestas por PLANIED: Creatividad e innovación, Comunicación y colaboración, Pensamiento crítico y Uso autónomo de las TIC.

En esta línea, en [5] y [6] se muestran los resultados de la implementación de la Diplomatura en Programación y Robótica Educativa. Se destaca el interés que suscita esta formación por la participación de docentes de localidades alejadas de la sede de la unidad académica dictante, provenientes de distintas áreas de conocimiento y con una motivación en consonancia con los objetivos de esta propuesta educativa. Así también se comprueba que los docentes han incorporado los conocimientos y las técnicas propias de la programación y han consolidado competencias para la cultura digital, tales como la colaboración y la creatividad, necesarias para lograr en los alumnos un aprendizaje más efectivo y motivador.

Cabe destacar que esta actividad formativa también tuvo que adaptar sus estrategias de enseñanza debido a la suspensión de las clases presenciales por las restricciones impuestas por la pandemia 2020. Los temas específicos de robótica fueron los que presentaron el mayor desafío dado que en las clases presenciales se

contaba con los kits de robótica provistos por la institución. Además, la modalidad de enseñanza consistía en el desarrollo de actividades prácticas con supervisión constante y personalizada en la manipulación de los kits.

Para lograr los objetivos de aprendizaje de los módulos Programación de placas Arduino y Programación de microcontroladores, se propuso que:

- a) Los cursantes adquirieran un kit de robótica para realizar la práctica. Los docentes responsables propusieron un modelo básico a modo de sugerencia.
- b) Utilizar la plataforma Moodle de la UNNE Virtual para el dictado de las clases, incluyendo videos explicativos sobre las actividades prácticas a realizar.

Los módulos de robótica se desarrollaron en la modalidad de clases teórico- prácticas, enfatizando en la metodología de aprendizaje de “aprender haciendo”. Se armaron y programaron diferentes circuitos de complejidad incremental.

La dinámica para los estudiantes consistía en participar del encuentro sincrónico programado, leer el material puesto a disposición por los docentes, observar los videos explicativos y realizar los ejercicios prácticos. Los cursantes presentaban las propuestas de solución a los ejercicios planteados en un espacio de la clase, a modo de revisión de las tareas y puesta en común de soluciones. También se propuso una prueba de evaluación publicada en la plataforma.

Con respecto al trabajo integrador, este fue definido de carácter grupal, cada equipo debía diseñar un proyecto de computación física a desarrollar en

un hipotético curso de un nivel que se especificaba en el proyecto.

Es importante destacar, que uno de los requisitos para la elaboración del proyecto, era la utilización de algún componente de reutilización (cartón, residuo plástico, madera, residuos tecnológicos, etc.). El trabajo integrador constó de tres etapas: la presentación de la idea del proyecto, la entrega de un informe de avance y la presentación de un informe final. Esta última, realizada de forma sincrónica, consistió en una presentación y defensa del prototipo en un máximo de 20 minutos por grupo. La evaluación se realizó utilizando rúbricas diseñadas especialmente.

Cabe destacar que esta implementación se llevó a cabo en cuestión de semanas y que la respuesta al cambio resultó satisfactoria. El 62% de los cursantes pudo cumplir las condiciones para aprobar la diplomatura.

En cuanto a Resultados Esperados, el proyecto continuará con las siguientes actividades:

- a) Análisis del grado de incorporación del método de resolución de problemas por parte de los estudiantes de la asignatura Algoritmos y Estructuras de Datos I de la carrera Licenciatura en Sistemas de Información de la UNNE. Este método, aplicado en actividades lúdicas que se resuelven con programación en bloque, incorpora herramientas conceptuales y herramientas del lenguaje para propiciar una forma de pensar que facilite el proceso de abstracción que la programación requiere. Se evaluará con una rúbrica especialmente diseñada la resolución de un desafío concreto que pondrá en juego sus

habilidades y estrategias adquiridas.

- b) Dado que desde este grupo de investigación se ha realizado capacitación en Didáctica de la Programación a los docentes de los niveles educativos preuniversitarios desde el año 2015, interesa saber cuántos estudiantes sobre el total de ingresantes a la carrera tienen la programación en bloque como conocimiento previo, y en cuánto contribuye con un buen rendimiento en la carrera de Informática.

En la línea vinculada con estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática, se proponen las siguientes actividades:

- Relevar y evaluar herramientas software utilizadas en la enseñanza inicial de la programación, y realizar una clasificación según sus características
- Definir criterios de evaluación de las herramientas, en función de los distintos aspectos que surgirán de la etapa de relevamiento y evaluación de las mismas.
- Diseñar y validar un marco de referencia que oficie de guía en la selección de la herramienta más apropiada para el logro de los objetivos de aprendizaje

Como consecuencia de la obligada virtualización de la Diplomatura en Programación y Robótica, se contempla el rediseño de la misma para ser ofrecida en la modalidad A distancia (EaD). Esta modalidad permitiría que docentes que residan en lugares más alejados de la universidad puedan acceder a esta capacitación. En este contexto, se abordará el estudio de recursos de la

EaD que faciliten el logro de los objetivos de formación en temas de robótica educativa y programación.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto están involucradas seis docentes investigadoras, entre ellas dos tesis de posgrado que desarrollan su Trabajo Final de la Maestría en Tecnologías de la Información de la UNNE.

5. REFERENCIAS

[1] Ochoa, L. A., Valenzuela, A., Estela, D., Márquez, F. (2018). "La indagación como estrategia para la educación STEAM". Organización de Estados Americanos. Disponible en: <https://tinyurl.com/y9ptbgs1>

[2] The Royal Society. (12 de 01 de 2012). "Shut down or restart? The way forward for computing in UK schools". Disponible en: <https://royalsociety.org/-/media/education/computing-in-schools/2012-01-12-computing-in-schools.pdf>

[3] G. N. Dapozo, C. L. Greiner, R. H. Petris, M. V. Godoy and M. C. Espíndola. "Enseñanza de programación en la universidad. Estrategia basada en programación por bloques". Libro "Innovation and Practice in Education". Páginas 89-99. Editorial CIATA.org Ciudad Real, España. I SBN 978-84-09-09792-0. Junio 2019.

[4] G. N. Dapozo, C.L., Greiner, R. H., Petris, M.F. Piragine, A.M. Company, M. C., Espíndola. "Estrategias de evaluación formativa en la enseñanza de programación en modalidad remota de emergencia". XXVI Congreso Argentino de Ciencias de la Computación - CACIC 2020. Libro de actas. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/114458>

[5] G. N. Dapozo, Y. Medina, R. H. Petris, M. C. Espíndola, A. M. Company “Oferta educativa en programación y robótica para docentes de los niveles preuniversitarios”. I Simposio Argentino de Educación en Informática (SAEI 2019) - JAIIO 48 (Salta). ISSN: 2683-8958. Disponible en:
<http://sedici.unlp.edu.ar/handle/10915/88932>

[6] G. N. Dapozo, C.L., Greiner, R. H., Petris, Y. Medina, M. C., Espíndola, A.M. Company. “Motivación y logros en la formación de docentes para introducir la programación y la robótica en los niveles educativos no universitarios”. Presentado en IEEE ARGENCON 2020, realizado del 1 al 4 de diciembre de 2020. UTN Facultad Regional Resistencia. Chaco (en etapa de publicación).

Análisis del impacto de la clase invertida a través del uso de herramientas de la Web 3.0 en la adquisición de competencias del perfil profesional del Ingeniero en Sistemas de Información

Maurel, María del Carmen, Sandobal Verón, Valeria; Barrios, Teresita H.; Marín, Bianca
Grupo de Investigación Educativa Sobre Ingeniería – GIESIN. Facultad Regional Resistencia
Universidad Tecnológica Nacional
mmaurel_38@yahoo.com.ar, @vsandobal@frre.utn.edu.ar, barriosth@gmail.com,
mbiancamarin@gmail.com

RESUMEN

En este trabajo se describe el proyecto inscripto en la línea de investigación sobre TIC aplicada a educación.

El proyecto se propone indagar si con el modelo de clase invertida, utilizando herramientas de la Web 3.0 se generan competencias para el perfil de profesionales de Ingeniería en Sistemas de Información que figuran en el Libro Rojo del CONFEDI. En particular nos enfocaremos en las clasificadas como sociales, políticas y actitudinales: aprender en forma continua y autónoma, desempeñarse de manera efectiva en equipos de trabajo. Asimismo, se consideró relevante de las competencias tecnológicas considerar: identificar, formular y resolver problemas de ingeniería y utilizar de manera efectiva las técnicas y herramientas de aplicación en la ingeniería. La aplicación de metodologías activas y centradas en los alumnos trae aparejados cambios de roles tanto de profesores como de alumnos. En este sentido, se consideró pertinente recabar información sobre cómo impactan en los docentes estos cambios, teniendo en cuenta que su posición en el aula es de guía, tutor y mediador en el aprendizaje.

El alumno ocupa un papel central por ello es necesario determinar si el uso de estas herramientas permite el desarrollo de habilidades para acceder a fuentes de información soportadas por la tecnología y desarrollar competencias tecnológicas para adquirir, utilizar, transferir y producir más información.

Palabras clave: clase invertida, web 3.0, competencias profesionales, educación superior, ingeniería

CONTEXTO

La línea de I/D presentada en este trabajo forma parte de las actividades definidas en el marco del proyecto “Análisis del impacto de la clase invertida a través del uso de herramientas de la web 3.0 en la adquisición de competencias del perfil profesional del Ingeniero en Sistemas de Información”, homologado por la Secretaría de Ciencia y tecnología de la Universidad tecnológica Nacional (UTN).

El proyecto tiene como objetivos:

- Indagar sobre nuevas herramientas web 3.0 disponibles en la web, que por sus características sean factibles de implementarse en el modelo de clase invertida en las cátedras de ISI de la UTN FRRe.
- Caracterizar al modelo de clase invertida que adoptan las cátedras de la carrera de Ingeniería en Sistemas de Información de la UTN FRRe; con la incorporación de las herramientas de la Web 3.0.
- Analizar si las competencias definidas en el proyecto fueron adquiridas por los alumnos, de las cátedras seleccionadas en la muestra; al utilizar el modelo de clase invertida a través de las herramientas de la web 3.0.
- Determinar el grado de aceptación, tanto de docentes como de alumnos, de la utilización de las herramientas de la web 3.0 como forma de transformar el modelo de clase tradicional; hacia la inversión de la misma.

1. INTRODUCCIÓN

En la búsqueda de mejorar el rendimiento académico de los alumnos, mantener la motivación en las actividades propuestas y, recientemente, la generación de competencias se implementan diferentes metodologías de enseñanza-aprendizaje. En especial nos centraremos en aquellas que incorporan el uso de las TIC como complemento a las clases tradicionales; y en particular nos interesa la aplicación del modelo “Aula invertida” o “Clase invertida” o “Flipped Classroom”. Este modelo considera un cambio de rol tanto del alumno como del docente. Donde el alumno pasa a tener una participación más activa y el docente actúa de tutor más que de expositor.

El concepto de “aula invertida” tiene su origen en la clase de dos profesores de Química, Bergmann y Sams, preocupados por los alumnos que se ausentaban de las clases no pudieran seguir el ritmo de los temas desarrollados, subsanaron esta situación, grabando sus clases y la presentación de Power Point utilizada. Este material era distribuido a todos los alumnos. Con el tiempo notaron que, no solo los alumnos que no asistían a clases utilizaban esos videos sino los que concurrían habitualmente también. Por lo cual, optaron por distribuir los videos antes de las clases, para luego aprovechar la presencialidad para la realización de prácticos y proyectos. Es así como “la Flipped Classroom es un enfoque pedagógico en el que la instrucción directa se mueve desde el espacio de aprendizaje colectivo hacia el espacio de aprendizaje individual, y el espacio resultante se transforma en un ambiente de aprendizaje dinámico e interactivo en el que el educador guía a los estudiantes a medida que se aplican los conceptos y puede participar creativamente en la materia” [1].

Actualmente, se está trabajando en la generación de competencias en los alumnos de ingeniería. Como resultado de este trabajo en el año 2018 se generó lo que se

denomina “Libro rojo” [2]. Los nuevos estándares de acreditación aquí propuestos se aplicarán a todas las carreras de Ingeniería, entre las cuales se encuentra la terminal Ingeniería en sistemas de información. Estos estándares están definidos en términos de formación por competencias. En este sentido es necesario aclarar qué entendemos por competencia: es la capacidad de articular eficazmente un conjunto de esquemas (estructuras mentales) y valores, permitiendo movilizar (poner a disposición) distintos saberes, en un determinado contexto con el fin de resolver situaciones profesionales.

Ahora bien, si tenemos en cuenta que buscamos formar a los futuros ingenieros por competencias, que contamos con las TIC que están incorporadas al aula, y que la aplicación del aula invertida tiene entre sus objetivos el trabajo autónomo, colaborativo, buscar una actitud analítica y crítica en el alumno; entonces es menester un trabajo que pretenda analizar si el modelo de clase invertida propicia la generación de competencias en los alumnos de Ingeniería en Sistemas de Información.

En un trabajo realizado por [3], en su tesis “Flipped Classroom y el efecto en las competencias transversales de los alumnos del curso de Electricidad y Electrónica industrial en una universidad pública de Lima”. En este trabajo se llega a la conclusión de que Flipped Classroom es un modelo pedagógico que mejora el proceso de enseñanza-aprendizaje, que fomenta la adquisición de competencias transversales, teniendo en cuenta que se incorpora herramientas TIC para elaborar la sesión de clases. En esta misma tesis realizan algunas recomendaciones al aplicar este modelo, tales como: aplicar el modelo flipped classroom teniendo en cuenta los diferentes estilos de aprendizaje de los alumnos, mejorar el uso de herramientas como video que combinen los fundamentos teóricos con preguntas que incentiven a los alumnos; incorporar el uso de herramientas como Google drive, youtube, redes sociales entre

otros, para lo cual será necesario que el docente adquiera competencias digitales.

También el artículo “Enseñanza de la ingeniería de software basada en competencias fundamentales y aula invertida” de [4] es una experiencia realizada en un curso de pregrado de los cursos de Ingeniería de Software de la Universidad de Caldas. Las fases generales de la propuesta son: Identificar competencias, clasificar competencias y definir estrategias de enseñanza. Como conclusión, los autores mencionan que ha permitido revisar y actualizar el plan de los cursos, desarrollar material adecuado y orientado a las nuevas generaciones, y la realización de actividades más prácticas y colaborativas lo que permite la apropiación de las competencias deseadas.

Hay artículos que hacen especial énfasis en la competencia de aprendizaje de forma continua y autónoma, tales como el de [5] que enfatiza en sus conclusiones sobre la responsabilidad asumida por los alumnos en su proceso de aprendizaje, entendiendo y reflexionando sobre el conocimiento que va adquiriendo. Asimismo, en [6] los autores ven reflejado esa responsabilidad en la obtención de mejores resultados durante el desarrollo de la materia.

Una de las competencias que se consideran relevantes y que es contemplada en este proyecto es la de “desempeñarse de manera efectiva en equipos de trabajo”. Esta competencia se puede identificar explícitamente en los trabajos realizados por [7] y [8]. Donde el primero incorpora dentro de sus conclusiones que el trabajo en equipo y colaborativo es un factor importante al momento de realizar proyectos. En el segundo caso, el trabajo en equipo se considera trascendente teniendo en cuenta que se utiliza la metodología de aula invertida combinando la práctica ágil Scrum. Donde este último busca además equipos que tengan como principales características los equipos que sean autogestionados y multifuncionales.

Una de las características importantes que se mencionan en la aplicación del modelo

de aula invertida es la motivación por parte de los alumnos. Ya que a partir de la implementación de este modelo pasa a tener un papel principal en el proceso de aprendizaje. En el artículo de [9] concluyen en su trabajo que la motivación que han tenido los alumnos durante el proceso de implementación del aula invertida ha traído aparejado un notable mejoramiento en el rendimiento académico de los alumnos. De igual manera, en [10] se menciona que, para poder llevar a cabo la experiencia de aplicar aula invertida, combinada con blended-learning, aprendizaje colaborativo y aprendizaje basado en proyectos es primordial la participación activa de los alumnos y la motivación de ellos para el desarrollo de las actividades.

A nivel nacional, se ha consultado la experiencia desarrollada por [11], donde se lleva a cabo la implementación de la clase invertida en el curso de Ambientación a la vida universitaria (módulo de matemática) destinado a ingresantes a la Facultad de Ciencias Agropecuarias de la Universidad Nacional de Entre Ríos. Para el desarrollo de la experiencia se realizó una “guía de estudio”. Cada semana contaba con los contenidos teóricos que deberían desarrollarse, agregando a este material recursos audiovisuales y multimediales que son subidos al Aula virtual de la asignatura. Luego, en las clases presenciales los alumnos trabajaban en sesiones colaborativas de resolución de problemas y actividades de transferencia a situaciones cercanas a las Ciencias Agropecuarias. Los autores realizan un análisis comparativo de resultados obtenidos con cortes anteriores, donde concluyen que encuentran varias ventajas en la aplicación de este modelo, entre las que se mencionan: la posibilidad de lograr un aprendizaje más profundo, la adquisición de competencias transversales y el aumento de la motivación del estudiante.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación que aborda el proyecto están vinculadas con:

- El desarrollo tecnológico en el ámbito de la tecnología aplicada a educación
- Adaptación del modelo de aula invertida al contexto de la generación de competencias en los alumnos de la carrera de Ingeniería en sistemas de Información

3. RESULTADOS ESPERADOS

Los resultados directos esperados del proyecto son:

- Definición de una lista de herramientas de la Web 3.0 que puedan ser implementadas con el modelo de aula invertida.
- Determinación de las características utilizadas en las cátedras de la carrera de Ingeniería en sistemas de información con la incorporación de las herramientas de la Web 3.0.
- Determinación de las competencias que fueron adquiridas por los alumnos con la incorporación del modelo de clase invertida y la aplicación de las herramientas de la Web 3.0.
- Establecer el grado de aceptación de docentes y alumnos a la utilización de herramientas de la Web 3.0 para invertir el aula.

4. FORMACION DE RECURSOS HUMANOS

En cuanto a la formación de recursos humanos, este proyecto se constituirá en un nuevo ámbito para el desarrollo de Tesinas de la Licenciatura en Tecnología Educativa, como así también en la Maestría en Educación en Entornos Virtuales dictada por la Universidad Nacional de la Patagonia Austral (UNPA) para los docentes que se encuentren cursando dicha maestría. En este sentido la Ing. Claudia García radica su tesis de la Maestría en educación en entornos virtuales de la UNPA.

Esta iniciativa permitirá también incorporar becarios alumnos o recientes egresados en las becas de investigación que para tal fin propone la Secretaría de Ciencia y Tecnología de la UTN.

En particular, sus conclusiones -como todas las generadas por los proyectos del grupo GIESIN-, se transfieren directamente a la formación de recursos humanos a través de la Dirección de Planeamiento y Apoyo Académico de la FRRe.

5. BIBLIOGRAFÍA

- [1] Berenguer Albaladejo C. (2016). Acerca de la utilidad del aula invertida o flipped classroom. XIV Jornadas de Redes de Investigación en Docencia Universitaria. Universidad de Alicante
- [2] CONFEDI - Consejo Federal de Decanos de Ingeniería. Obtenido de <https://confedi.org.ar/librorojo/>
- [3] Benites y Arleque, J. V. (2018). Flipped Classroom y el efecto en las competencias transversales de los alumnos del curso de Electricidad y Electrónica Industrial en una universidad pública de Lima. Lima, Perú: Universidad Peruana Cayetano Heredia - Escuela de Posgrado Victor Alzamora Castro.
- [4] Bedoya, Ó. H. F., & Hurtado, S. (2019). Enseñanza de la ingeniería de software basada en competencias fundamentales y aula invertida. *Encuentro Internacional de Educación en Ingeniería*.
- [5] Matzumura-Kasano, J. P., Gutiérrez-Crespo, H., Zamudio-Eslava, L. A., & Zavala-Gonzales, J. C. (2018). Aprendizaje invertido para la mejora y logro de metas de aprendizaje en el Curso de Metodología de la Investigación en estudiantes de universidad. *Revista Electrónica Educare*, 22(3), 177-197.
- [6] Lopes, A. P. & Soares, F. (2018). Perception and performance in a flipped Financial Mathematics classroom. *The International Journal of Management Education*, 16(1), 105-113.
- [7] Nwokeji, J. C., Stachel, R., & Holmes, T. (2019, October). Effect of Instructional Methods on Student Performance in Flipped Classroom. In 2019 IEEE Frontiers in Education Conference (FIE) (pp. 1-9). IEEE.

- [8] Cornide-Reyes, H. C., & Villarroel, R. H. (2019). Método para Promover el Aprendizaje Colaborativo en Ingeniería de Software. *Formación universitaria*, 12(4), 3-12.
- [9] Álvarez, D. M. L., Aguilar, G. F. C., Conforme, N. C. R., & Alcívar, I. A. M. (2020). Implementación de flipped classroom enfocado a los estudiantes de Ingeniería de software: caso universidad ecuatoriana. *Revista científica ecociencia*, 7(3), 1-18.
- [10] Tong, Y., & Song, P. (2018, January). Research on blended teaching model in the teaching of 'micro-controller principles and applications'. In *Proceedings of the 6th International Conference on Information and Education Technology* (pp. 38-41).
- [11] Ponce, S. L., Marichal, A., Martínez, G., Soldini, M., & Ponce, R. D. (2017). Implementación de la clase invertida en el aula universitaria: posibilidades para la obtención de aprendizajes no superficiales. In *XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET, La Matanza 2017)*.

Aplicación de la herramienta Quizizz como estrategia de Gamificación en la Educación Superior

Lazarte, Ivanna M.^{1,2}; Gómez, Sofía G.¹

¹Departamento de Formación Básica/Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca

²Departamento de Informática/ Facultad de Tecnología y Ciencias Aplicadas/ Universidad Nacional de Catamarca

Maximio Victoria 55, San Fernando del Valle de Catamarca, 0383-4435112
ilazarte@tecono.unca.edu.ar, sofiagom@tecono.unca.edu.ar

Resumen

En los últimos años, la gamificación ha despertado gran interés en el ámbito educativo como consecuencia del auge de las Tecnologías de la Información y la Comunicación (TIC), los videojuegos y las nuevas tendencias educativas que buscan garantizar el aprendizaje competencial y activo del alumno desde la innovación y la eficacia.

La gamificación es la inclusión de elementos de los juegos en contextos que no son juegos. Diversos autores señalan que, en el ámbito de la educación, la gamificación se presenta como una oportunidad para motivar y/o mejorar las dinámicas de grupo, la atención, la crítica reflexiva y el aprendizaje significativo de los estudiantes, potenciando el proceso de enseñanza y aprendizaje en el aula.

Mediante este proyecto se pretende incorporar la gamificación en el proceso de enseñanza y aprendizaje de la asignatura Probabilidad y Estadística que se dicta en el 2° año de las carreras de Ingeniería de la Facultad de Tecnología y Ciencias Aplicadas (FTyCA) de la Universidad Nacional de Catamarca (UNCA). El propósito de este proyecto es implementar y promover ambientes de

aprendizaje dinámicos, incentivando la participación de los alumnos en espacios donde se sientan motivados y se diviertan mientras aprenden y descubren la importancia de la Estadística en la toma de decisiones.

Palabras clave: *Gamificación, Enseñanza, Aprendizaje, Probabilidad y Estadística*

Contexto

Este trabajo de investigación se lleva a cabo en el ámbito de la FTyCA de la UNCA y se encuadra dentro del proyecto de I/D "Gamificación en la Educación Superior: aplicación en la asignatura Probabilidad y Estadística". Bajo la dirección de la Dra. Ivanna Lazarte y el asesoramiento de la Esp. Sofía Gómez, cuenta con aval académico-institucional otorgado mediante Resolución Rectoral N° 171/2020. La línea de investigación está orientada a la investigación aplicada de tecnologías de la información y la comunicación a procesos educativos para utilizarse en el aula de ingeniería tanto presencial como bajo la modalidad a distancia y sus posibles combinaciones como b-learning (aprendizaje mixto) y u-learning (aprendizaje ubicuo) entre otros.

Introducción

En los últimos años, la gamificación ha despertado gran interés en el ámbito educativo como consecuencia del auge de las Tecnologías de la Información y la Comunicación (TIC), los videojuegos y las nuevas tendencias educativas que buscan garantizar el aprendizaje competencial y activo del alumno desde la innovación y la eficacia (Ardila-Muñoz, 2019; Jaber et al, 2016; Ortégón Yañez, 2016).

El término gamificación se define como el uso de las mecánicas de juego en entornos ajenos al juego, con el fin de conducir el comportamiento de los alumnos mediante la participación, la interacción, la adicción o la competición hacia la consecución de un determinado objetivo de aprendizaje (Jaber et al, 2016; Ortégón Yañez, 2016).

Cada vez más instituciones educativas de todos los niveles optan por este modelo como alternativa a las estrategias tradicionales del aula, basándose en algunos de sus beneficios más destacables para estudiantes y docentes:

- Mejora la actitud de los alumnos para con su aprendizaje.
- Elimina el temor al fracaso transmitiendo que perder siempre es una posibilidad.
- Enseña a trabajar en equipo.
- Transmite la importancia de la competencia sana.

La gamificación en el aula combina las clásicas demandas de todo docente con las preferencias de los estudiantes. Requiere crear un ambiente propio de juego en el que los participantes desarrollan habilidades, logran los objetivos propuestos y aprenden mientras se divierten jugando. Asimismo, se convierten en el centro del juego, se sienten involucrados, toman sus propias decisiones, perciben que progresan,

asumen nuevos retos, participan en un entorno social y son reconocidos por sus logros. De este modo, los alumnos mejoran tanto su aprendizaje como su predisposición para aprender, y además reciben retroalimentación en su proceso de aprendizaje (Jaber et al, 2016; Ardila-Muñoz, 2019; Oliva, 2016).

Diversos autores señalan a la gamificación como una estrategia metodológica eficaz para motivar los estudiantes y es este punto, la motivación, el desafío al que los docentes se enfrentan en sus clases. Motivar a los estudiantes del siglo XXI es una tarea difícil, sobre todo en el área de las llamadas ciencias duras, entre las que se cuentan las Matemáticas, cuyo aprendizaje resulta muy complejo para la mayoría de los estudiantes (Farias & Pérez, 2010; Macías Espinales, 2017; Ortégón Yañez, 2016).

Como señala Macías Espinales (2017), en pleno siglo XXI, la matemática se sigue enseñando desde un enfoque tradicional, caracterizado por clases magistrales, poco flexibles, de conocimiento abstracto, rígidas, mecánicas, memorísticas y con poca implicación de las TIC. Como resultado, los estudiantes se vuelven memorizadores de procedimientos matemáticos, se conforman con recibir la explicación del docente, no se esfuerzan por lograr autonomía en su aprendizaje y como consecuencia, tienen pocas habilidades para dar solución a problemas de su contexto.

Por lo expresado, la gamificación puede resultar una estrategia ventajosa en la enseñanza de la asignatura Probabilidad y Estadística para involucrar al alumno en su proceso de aprendizaje y mostrar que no tiene por qué ser un espacio curricular mecánico, rígido y aburrido.

Mediante este proyecto se incorpora la gamificación en el proceso de enseñanza

y aprendizaje de la asignatura Probabilidad y Estadística que se dicta en el 2° año de las carreras de Ingeniería de la FTyCA de la UNCA, con el propósito de implementar y promover ambientes de aprendizaje dinámicos, incentivando la participación de los alumnos en espacios donde se sientan motivados y se diviertan mientras aprenden. Al mismo tiempo, se apunta a que descubran la importancia de la Probabilidad y la Estadística como herramienta para la toma de decisiones ante situaciones de incertidumbre.

Líneas de Investigación, Desarrollo e Innovación

Las innovaciones en educación en ingeniería, - generadas principalmente por la inclusión de las TIC en las prácticas de enseñanza y de aprendizaje y por las competencias que deben desarrollar los graduados de ingeniería en Argentina de acuerdo a los estándares propuestos por CONFEDI - proponen un cambio paradigmático en la formación de ingenieros, en tanto ponen su foco en el estudiante a través de metodologías como el Aprendizaje centrado en el estudiante (ACE) y el Aprendizaje basado en problemas (ABP).

En este trabajo se investiga el impacto que tiene la aplicación de la gamificación en la Educación Superior, particularmente en la asignatura Probabilidad y Estadística. La hipótesis de trabajo planteada es que la gamificación crea un ambiente de aprendizaje significativo e interactivo, que favorece la motivación y la participación de los estudiantes.

Mediante esta investigación se busca obtener datos y conceptualizaciones para la elaboración de propuestas que generen una cultura del buen uso de las TIC como herramientas de apoyo al trabajo académico. La implementación de nuevas

estrategias y métodos de enseñanza llevará a que los docentes conozcan las potencialidades que brinda la gamificación para optimizar el aprendizaje y la adquisición de competencias en contextos universitarios.

Resultados y Objetivos

El objetivo principal del proyecto es aplicar herramientas dinámicas que favorezcan el aprendizaje significativo de la asignatura Probabilidad y Estadística, mediante la incorporación de la gamificación en el aula, apoyada por TIC y otros recursos didácticos.

Esto se fundamenta en el hecho que, los estudiantes de esta Facultad por la naturaleza tecnológica de las carreras que aquí se dictan, han incorporado el uso de la computadora e Internet a sus hábitos de comunicación y de estudio. Asimismo, en su mayoría demuestran destrezas en el uso de herramientas y software específicos. También manifiestan un marcado dominio experiencial en el uso de herramientas de comunicación y redes sociales. En ese sentido, los avances de las TIC han generado cambios en las formas tradicionales de acceso al conocimiento, formas que se mantuvieron vigentes durante años y que ahora conllevan a la adopción de nuevos soportes.

Sin embargo, sigue observándose cierta reticencia por parte de algunos docentes en implementar las TIC, ya sea como fuente de acceso al conocimiento —prevaleciendo el uso de los libros de texto en soporte escrito— o como herramienta para su construcción.

En ese sentido, es dable destacar que la Facultad propicia capacitaciones gratuitas a docentes sobre el uso básico e intermedio de la Plataforma Moodle y talleres varios sobre otras herramientas de la Red aplicables al aula de ingeniería,

hecho que se ha masivizado este último año a partir de la cuarentena impuesta por la crisis sanitaria.

Es por ello que, en el contexto del Aislamiento Social, Preventivo y Obligatorio durante el ciclo lectivo 2020, se decidió aplicar una experiencia de gamificación como estrategia de evaluación¹. Mediante esta estrategia, la evaluación es formativa ya que se realiza durante todo el proceso de enseñanza y aprendizaje, busca comprobar el nivel de comprensión que tienen los alumnos de los temas vistos en clase, sirve como insumo para planificar acciones de aprendizaje pertinentes a las necesidades y sirve de guía tanto del docente como del estudiante para alcanzar las metas de aprendizaje. La gran ventaja de esta estrategia es que quita la presión que sienten los estudiantes al ser evaluados.

Las evaluaciones se planificaron para realizarse al finalizar cada unidad temática de la asignatura para las carreras Ingeniería Electrónica y Minas (Grupo 1) durante el primer cuatrimestre, donde participaron 31 alumnos y las carreras de Ingeniería en Informática y Agrimensura (Grupo 2) en el segundo cuatrimestre, donde participaron 35 alumnos. Las mismas se realizaron utilizando la herramienta Quizizz², la cual permite evaluar a los alumnos mientras se divierten, a través de cuestionarios online. Se seleccionó esta herramienta porque es la más completa (comparada con Kahoot³ y Socrative⁴) permitiendo al docente realizar actividades más dinámicas (Heredia-Sánchez, 2020). Además,

¹ “Conjunto de métodos, técnicas y recursos que utiliza el docente para valorar el aprendizaje del alumno” Díaz Barriga (2016).

² <https://quizizz.com/>

³ <https://kahoot.com/>

⁴ <https://www.socrative.com/>

Quizizz permite incluir imágenes o fórmulas a las preguntas y también a las posibles respuestas, elegir si se muestran o no las respuestas correctas después de un fallo, enviar al alumno (o a su tutor) un archivo PDF con todos los detalles de su evaluación y obtener informes completos de las evaluaciones, con estadísticas muy útiles como, por ejemplo, las preguntas con mayor fallo o acierto.

Para evaluar la estrategia de evaluación propuesta, al finalizar la cursada se realizó una encuesta anónima a los alumnos, mediante la herramienta QuestionPro⁵.

Las preguntas y respuestas obtenidas de ambos grupos se detallan en la Tabla 1.

Tabla 1: Preguntas realizadas a los alumnos para valorizar la estrategia de evaluación gamificada

Nº	Pregunta	Grupo 1	Grupo 2
1	¿Te resulta divertida la evaluación mediante el juego?	SI 93,75% NO 6,25%	SI 100% NO 0%
2	¿La evaluación mediante el juego te requiere más concentración para realizarla?	SI 87,50% NO 12,50%	SI 83,33% NO 16,67%
3	¿El juego te permite aprender los conceptos de la asignatura?	★ 0% ★★ 0% ★★★ 25% ★★★★ 18,75% ★★★★★ 56,25% % (*)	Nada 0% Poco 5,56% Mucho 94,44%
4	¿Premiar con puntos te motiva a esforzarte para superar a tus compañeros y de paso aprender?	★ 0% ★★ 6,25% ★★★ 18,75% ★★★★ 12,5% ★★★★★ 62,5% (*)	Nada 0% Poco 27,78% Mucho 72,22%

(*) En el Grupo 1 los alumnos debían puntuar con estrellas. Para clarificar las opciones de respuesta, en el Grupo 2 se decidió optar por Nada, Poco, Mucho.

Como se observa en la Tabla 1, al 93,75% de los alumnos del Grupo 1 y al 100% del Grupo 2 les resultó divertida la evaluación gamificada. El 87,50% del Grupo 1 y el 83,33% del Grupo 2

⁵ <https://www.questionpro.com/es/>

consideraron que esta estrategia les requiere mayor concentración. En cuanto a la pregunta sobre si el juego les permite aprender los conceptos de la asignatura, el 75% del Grupo 1 y el 72,22% del Grupo 2 respondieron que aprenden mucho. Con respecto al hecho que la premiación con puntos, los motiva a esforzarse para superar a sus compañeros, el 75% del Grupo 1 y el 72,22 del Grupo 2 consignaron que los motiva mucho.

En la encuesta, los alumnos tenían la posibilidad de agregar comentarios o sugerencias y todas fueron positivas. Algunos de estos comentarios son “*Las clases son super útiles. La evaluación mediante el juego es una propuesta muy buena, insta a aprender de manera divertida. No tengo quejas respecto a la cátedra, la Dra. Lazarte enseña muy bien.*” “*Más bien como comentario, estoy conforme con el cursado virtual de la materia, además que me encanta el modo de evaluación y presentaciones de trabajos.*”

También se realizó una evaluación externa a la cátedra donde los alumnos también manifestaron su conformidad con la propuesta.

Estos resultados apoyan la hipótesis de que la gamificación crea un ambiente de aprendizaje significativo, interactivo y divertido, que favorece la motivación y la participación de los estudiantes.

Formación de Recursos Humanos

Los integrantes del proyecto cuentan con conocimientos y experiencia en docencia universitaria de disciplinas tecnológicas y/o el uso de TIC en educación. La directora del proyecto está finalizando la Especialización en Docencia Universitaria en Disciplinas Tecnológicas, mientras que la asesora está finalizando su Tesis de Maestría en

Procesos Educativos mediados por Tecnologías.

Cabe destacar que los docentes están categorizados en el Programa de Incentivos de la Secretaría de Políticas Universitarias (SPU).

Los alumnos que integran el proyecto investigan y prueban diferentes herramientas de gamificación que pueden aplicarse en la Educación Superior. También participaron en jornadas exponiendo los avances de su investigación.

Referencias

- Ardila-Muñoz, J. Y. (2019). Supuestos teóricos para la gamificación de la educación superior. *Revista Internacional de Investigación en Educación*, 12 (24), 71-84.
- Heredia-Sánchez, B. D. C., Pérez-Cruz, D., Cocón-Juárez, J. F., & Zavaleta-Carrillo, P. (2020). La Gamificación como Herramienta Tecnológica para el Aprendizaje en la Educación Superior. *Revista Tecnológica-Educativa Docentes 2.0*, 9(2), 49-58.
- Farias, D., & Pérez, J. (2010). Motivación en la Enseñanza de las Matemáticas y la Administración. *Formación Universitaria*, 3(6), 33-40.
- Jaber, J. R. et. al. (2016). Empleo de Kahoot como herramienta de gamificación en la docencia universitaria. *III Jornadas Iberoamericanas de Innovación Educativa en el ámbito de las TIC*. Las Palmas de Gran Canaria.
- Macías Espinales, A. V. (2017). *La Gamificación como estrategia para el desarrollo de la competencia matemática: plantear y resolver problemas*. Tesis de Maestría: Tecnología e Innovación Educativa. Universidad Casa Grande. Guayaquil, Ecuador.
- Oliva, H. A. (2016). La gamificación como estrategia metodológica en el contexto educativo universitario. *Revista Realidad y Reflexión*, N° 44. ISSN 1992-6510
- Ortegón Yañez, M. E. (2016). *Gamificación de las matemáticas en la enseñanza del valor posicional de cantidades*. Tesis de Maestría: elearning y Redes Sociales. Universidad Internacional de La Rioja. Cali, Colombia.

Avances en la formación de docentes y alumnos como Investigadores Científicos Iniciales en Informática y Ciencias de la Computación

Bianchini Germán¹, Caymes-Scutari Paola^{1,2}, Ontiveros Patricia³, Rotella Carina³, Salinas Sergio⁴, Tagarelli Sandra⁴, Chirino Pamela¹, Galdamez Mariela¹, Díaz Karvin¹, Ponce de Leon Alejo¹

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido,
Departamento de Ingeniería en Sistemas de Información. Facultad Regional
Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Laboratorio de Gobierno Electrónico, Departamento de Ingeniería en Sistemas de
Información. Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza

⁴Laboratorio de Analítica de Datos, Departamento de Ingeniería en Sistemas de
Información. Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza

gbianchini@frm.utn.edu.ar, pcaymesscutari@frm.utn.edu.ar, pontiveros@frm.utn.edu.ar,
carinarotella@gmail.com, s4salinas@gmail.com, stagarelli@gmail.com,
pamelaachirino@gmail.com, mariela.galdamez.16@gmail.com, karvindiaz@gmail.com,
alejo.poncedleon@gmail.com

RESUMEN

El ciclo 2020 y lo que va del actual 2021 han sido muy complejos en virtud del aislamiento y distanciamiento vividos por la situación de pandemia por COVID-19; en particular para el ámbito académico en todos sus niveles. Dentro del área universitaria, el desarrollo de las actividades de investigación tampoco ha sido el normal, pero a pesar de ello se ha logrado avanzar en los proyectos. Todo el esquema de trabajo debió realizarse en modo no presencial, e incluso la participación en eventos (congresos, cursos, etc.) también ha debido adecuarse al formato virtual de videoconferencias. Como tantas otras actividades, el estudio y la investigación no pueden pausarse

indefinidamente, y la investigación e innovación resultan elementos claves e imprescindibles para hacer frente a los retos con que los estudiantes se encontrarán en el mundo laboral. Es responsabilidad de las instituciones abrir un nuevo panorama a los ojos de los estudiantes y docentes universitarios en general, de modo tal de ayudarlos a desarrollar sus capacidades de evaluación, crítica e inventiva.

Un obstáculo histórico hallado es que las actividades de investigación normalmente se encuentran desvinculadas de las actividades académicas de grado, y por tanto esta disociación mantiene alejados a los alumnos y a muchos docentes. Por tal motivo, en el presente trabajo se comenta el proyecto en el que se ha buscado

formalizar y poner en práctica el proceso de transferencia y formación de Investigadores Científicos Iniciales a través de su incorporación en actividades en el marco de diversas áreas informáticas, y la adecuación que se llevó a cabo a medida que la situación de pandemia impuso diversas restricciones.

Palabras clave: Investigación, ASPO, DISPO, Formación, Alumnos, Proyectos.

CONTEXTO

El presente proyecto se realiza en colaboración entre los grupos ADA-Lab (Laboratorio de Analítica de Datos), GE-Lab (Laboratorio de Gobierno Electrónico) y LICPaD (Laboratorio de Investigación en Cómputo Paralelo /Distribuido), en donde se enmarca el mismo. Los tres pertenecen al Departamento de Ingeniería en Sistemas de Información de la UTN-FRM. Cuenta con la acreditación, homologación y financiamiento de la Universidad Tecnológica Nacional a través del proyecto PID TEUTIME0007658TC. Mediante este trabajo conjunto, se busca combinar en el mismo proceso la formación teórica y la práctica de investigadores iniciales para que éstos recorran todas las etapas necesarias que involucran el proceso científico en un ámbito de investigación. Para esto se planteó utilizar un trabajo de investigación específico como facilitador del proceso de aprendizaje, enfocado en las competencias específicas que deben adquirir los graduados de Ingeniería en Sistemas de Información, ya que al fomentarlas se contribuye también a incrementar el número y calidad de las investigaciones realizadas en la institución, además de impulsar el pensamiento crítico de docentes y alumnos frente a nuevos retos.

1. INTRODUCCIÓN

El objetivo general del proyecto que aquí se describe ha sido ofrecer una guía y brindar los elementos y acciones necesarias, como así también el acompañamiento, que permitan estimular la formación integral en el campo de la investigación científica en un nivel y temática acorde a la capacidad, la preferencia, y formación previa de cada docente o alumno, teniendo en cuenta además para estos últimos el estadio en la carrera.

En la UTN-FRM, por ejemplo, se efectúan anualmente ferias de ciencia y encuentros en los que se exponen los avances en materia de investigación que se llevan a cabo en los distintos centros y grupos de los diversos departamentos. Si bien esta actividad constituye una iniciativa positiva, la mayoría de los estudiantes normalmente están ajenos a estas actividades, las cuales requerirían una mejor y mayor difusión, con la orientación adecuada para el alumnado. Además, deberían realizarse en lugares y fechas que propicien y faciliten al alumno de grado el acercarse para informarse sobre qué otras actividades se desarrollan en la institución en la que estudia y modela su perfil profesional. Por otro lado, cabe considerar el aspecto económico relacionado con la realización de actividades de investigación. Indudablemente, es necesario que las instituciones cuenten con un importante apoyo y disponibilidad de presupuesto y recursos, tanto para remunerar el trabajo realizado y así incentivar la dedicación a investigación, como para brindar el entorno de trabajo equipado y adecuado a la temática que se aborde. Respecto de esto, es importante aclarar que encarar un proyecto de investigación científica normalmente es costoso, sobre todo considerando que los resultados suelen

conseguirse a largo plazo. Esto convierte a la investigación en una actividad poco atractiva para quienes buscan respaldar solamente aquellos proyectos cuya finalidad sea aportar a intereses inmediatos. No obstante, es claro que, como inversión a largo plazo, resulta de gran importancia para el desarrollo de las instituciones y la sociedad.

La modalidad usual de trabajo ha sido la de efectuar reuniones semanales, en las que se tratan los temas de interés, se aclaran dudas, se plantean nuevos interrogantes, y se realiza un seguimiento para definir el siguiente paso a dar, el cual se evaluará en las subsiguientes reuniones. Esto, a su vez, se complementa con material de lectura que se brinda (bibliografía, textos electrónicos, artículos científicos, etc.) y con la documentación continua que se solicita mantengan al día los investigadores, la cual es un excelente punto de partida a la hora de redactar artículos. A su vez, normalmente se plantean seminarios temáticos en donde cada alumno, becario o docente, transfiere sus avances y conocimiento al resto del equipo, con el propósito de intercambiar conocimiento y ampliar la capacidad de análisis y la incorporación de posturas y puntos de vista. Tanto la redacción de textos de carácter científico como las competencias vinculadas a la expresión oral, necesaria para expresar y comunicar correctamente los trabajos, se ejercitan a su vez en la asignatura electiva Computación Paralela, brindada por los directores del LICPaD, en donde se brindan lineamientos y recursos para que los alumnos desarrollen estos aspectos de la comunicación.

En marzo de 2020, poco después del inicio de actividades, la situación sanitaria del país y el mundo cambió todo, y la modalidad de trabajo seguida en el laboratorio de investigación no fue la excepción. Hubo que adecuar la forma de

trabajo hacia un esquema que nos permitiese seguir avanzando en las tareas, aunque posiblemente a un ritmo más lento que el esperado.

Las reuniones presenciales se vieron reemplazadas por un esquema de videoconferencias mediante plataformas de videollamadas. A su vez, el desarrollo de temas y explicaciones de conceptos por parte de los directores, anteriormente efectuado durante charlas en reuniones y seminarios, se reemplazó por videos explicativos subidos a la plataforma Youtube [1] que los alumnos pudiesen consultar todas las veces que fuese necesario, y las dudas consultarlas a través de distintos medios como por ejemplo en el contexto de reuniones por videoconferencias, mediante foros creados para dicho propósito en el aula del Campus Virtual de una de las cátedras intervinientes [2], a través de mails o por mensajes de whatsapp [3].

En lo referido a la redacción, revisión y corrección de textos para la participación en eventos científicos, de transferencia y divulgación, se trabajó mediante iteraciones por mail, revisando los borradores y haciendo devoluciones con correcciones y sugerencias sobre cómo tratar los temas y desarrollar los contenidos.

En lo atinente al acceso al clúster de computadoras de la Facultad, el mismo se ha efectuado mediante acceso remoto, de modo que los alumnos becarios pudiesen hacer uso de dicho recurso sin la necesidad de estar físicamente en la institución.

Finalmente, la asistencia a eventos se realizó de acuerdo a lo estipulado por los comités organizadores de cada uno, en general, coincidiendo la mayoría en la realización por modalidad virtual mediante distintas plataformas de videoconferencias como Zoom [4], Google Meet [5] o Webex Meeting [6].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Inicialmente se plantearon líneas amplias de investigación, de las cuales se seleccionaron algunos temas en particular sobre los que se ha ido avanzando y se han publicado resultados preliminares en diversos congresos. Dentro del área del Cómputo Paralelo, las líneas elegidas por los alumnos-becarios fueron: Paralelismo en algoritmos de aprendizaje para redes neuronales [7], Métodos de Inteligencia artificial para la reducción de incertidumbre en modelos de predicción de incendios, Redes neuronales paralelas aplicadas a la visión computacional, Predicción del avance del dengue en zonas endémicas [8] y, finalmente, Multiplicación de Matrices.

Por su parte, los docentes intervinientes de la rama Analítica de Datos orientaron sus temas de investigación hacia Indicadores de procesos y medición, dado que ambos (Mg. Sandra Tagarelli y Dr. Sergio Salinas) culminaron sus tesis de Maestría durante el período 2020 vinculadas a dicha temática.

Finalmente, los docentes de la rama Gobierno Electrónico se centraron en la temática de aplicación de Blockchain Federal (BFA) sobre la gestión de documentos judiciales [9].

3. RESULTADOS OBTENIDOS/ESPERADOS

Si bien la contribución principal que se ha buscado es la transferencia de formación por parte de docentes-investigadores hacia docentes y alumnos de grado y al medio educativo, también se ha impulsado que los alumnos aprendan y experimenten el trabajo en colaboración y por metas, dado que, en general, el trabajo de investigación no es algo que pueda

realizarse de forma individual, puesto que normalmente requiere de espacios inter y transdisciplinarios.

A su vez, dentro de los proyectos individuales llevados a cabo dentro de este proyecto general, se han ido logrando avances que fueron presentados en el marco de otros eventos de ciencia y educación. Los primeros lineamientos de este trabajo y los temas vinculados fueron expuestos en el WICC 2020 mediante cinco artículos [10-14], un artículo en el XXVI CACIC (en prensa), un artículo premiado como 3er mejor trabajo estudiantil en el 8° CONAIISI (en prensa) y otro artículo en E-ICES 15 (en prensa). Por su parte, los docentes investigadores, desde el comienzo del ciclo, han participado en diversos eventos y cursos para facilitar y mejorar la adaptación de la forma de trabajo y contar con más herramientas y experiencias que impulsen la concreción del objetivo global.

4. FORMACIÓN DE RECURSOS HUMANOS

En principio, el proyecto está pensado con alumnos-becarios, lo cual les permitirá dedicar cierta cantidad de horas semanales a la tarea de investigar de modo tal de llegar a cumplir los objetivos formales establecidos. Actualmente, el proyecto cuenta con cuatro becarios de grado de la carrera Ingeniería en Sistemas de Información (1 módulo de 6 horas cada uno). A su vez, se cuenta con seis docentes. Tres son investigadores formados con titulación doctoral, dos de los cuales cumplen el rol de guía del proyecto en general. Por su parte, como se ha mencionado, dos docentes han concluido sus tesis de Maestría en Inteligencia de Negocios durante el pasado 2020. Las actividades que todos ellos realicen en el marco del proyecto

contribuirán a continuar con su formación profesional en investigación.

5. BIBLIOGRAFÍA

- [1] Youtube. <https://youtu.be.com> (fecha de consulta: febrero de 2021)
- [2] Moodle - Open-source learning platform. <https://moodle.org> (fecha de consulta: febrero de 2021)
- [3] Whatsapp - Mensajería confiable. Simple. Segura. <https://www.whatsapp.com> (fecha de consulta: febrero de 2021)
- [4] Videoconferencias, Conferencias Web, Seminarios. <https://zoom.us/> (fecha de consulta: febrero de 2021)
- [5] G Suite essentials: la manera más simple de trabajar en equipo. <https://gsuite.google.com/> (fecha de consulta: febrero de 2021)
- [6] Reuniones en línea, videoconferencias gratuitas. <https://www.webex.com/es/video-conferencing.html> (fecha de consulta: febrero de 2021)
- [7] Chirino P., Galdamez M., Bianchini G., Caymes-Scutari P.: Propuesta de paralelización en redes neuronales. In: Libro de Actas 7mo CONNAIISI 2019. pp.1620--1630. Edit. UNLaM, San Justo (2020)
- [8] Rigau-Pérez J.G., Clark Gary G.: Cómo responder a una epidemia de dengue: visión global y experiencia en Puerto Rico. In: Rev Panam Salud Publica;17(4) pp.282--292. Edit. PAHO (2005)
- [9] Blockchain Federal Argentina. <https://bfa.ar> (fecha de consulta: febrero de 2021)
- [10] Bianchini G., Caymes-Scutari P., Ontiveros P., Rotella C., Salinas S., Chirino P., Galdamez M.: Formación de docentes y alumnos de grado como Investigadores Científicos Iniciales en las áreas de Informática y Ciencias de la Computación. In: Libro de Actas del XXII WICC 2020. ISBN 978-987-3714-82-5. pp.590-594 (2020)
- [11] Rotella C., Ontiveros P., Bianchini G., Caymes-Scutari P., Tagarelli S., Salinas S., Chirino P., Galdamez M.: Uso de la tecnología Blockchain Federal (BFA) para dejar pistas de auditoría y trazabilidad a sentencias y acordadas de la Suprema Corte de la provincia de Mendoza. In: Libro de Actas del XXII WICC 2020. ISBN 978-987-3714-82-5. pp.683-687 (2020)
- [12] Caymes-Scutari P., Bianchini G., Tardivo L., Méndez-Garabetti M.: Sintonización de Aplicaciones científico/ingenieriles: un proceso de medición-mejora para incrementar la eficiencia. In: Libro de Actas del XXII WICC 2020. ISBN 978-987-3714-82-5. pp.748-752 (2020)
- [13] Salinas S., Tagarelli S., Caymes-Scutari P., Bianchini G., Ontiveros P., Rotella C., Chirino P., Galdamez M.: Diseño de un sistema de medición de desempeño para Moodle en Educación Superior. In: Libro de Actas del XXII WICC 2020. ISBN 978-987-3714-82-5. pp.683-687 (2020)
- [14] Tagarelli S., Salinas S., Caymes-Scutari P., Bianchini G., Ontiveros P., Rotella C., Chirino P., Galdamez M.: Visualización de datos en un Tablero de Comando aplicado a plataformas de educación a distancia en el nivel

Superior. In: Libro de Actas del
XXII WICC 2020. ISBN 978-
987-3714-82-5. pp.955-959
(2020)

DISEÑO DE ESTRATEGIAS DIDÁCTICAS PARA LA ENSEÑANZA DE INGENIERÍA DE REQUERIMIENTOS.

Moyano Ezequiel, Jerez Marcela, Romano Lucas, Villarreal Martín, Moncho Matías, Villanueva Tomas, Rigoni Brian, Andrade Daniel, Iost Daniel
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: ++54-2901-432403

{[emoyano](mailto:emoyano@untdf.edu.ar), [mjerez](mailto:mjerez@untdf.edu.ar), [lromano](mailto:lromano@untdf.edu.ar), [mvillarreal](mailto:mvillarreal@untdf.edu.ar), [mmoncho](mailto:mmoncho@untdf.edu.ar)}@untdf.edu.ar, tomasino@hotmail.com, brian.rigonil@gmail.com, eldani17@gmail.com, danieliost@yahoo.com.ar

Resumen

Los procesos de enseñanza-aprendizaje se caracterizan por el papel determinante del docente en identificar, planificar e instrumentar estrategias (a corto, mediano y largo plazo) en la formación de los estudiantes.

La didáctica exige la utilización de estrategias y métodos centrados en el sujeto que aprende, enfocar la enseñanza como un proceso de orientación, para no sólo apropiarse de conocimientos sino desarrollar habilidades.

La enseñanza ingeniería de requerimientos se desarrolla sobre la base de un modelo conceptual de representaciones, supuestos y restricciones, para especificar las necesidades, deseos y expectativas de los stakeholders.

La dificultad de la obtención de los requerimientos radica en la incertidumbre derivada del proceso de transferencia de información de varias fuentes no siempre consistentes, ésta se incrementa cuando se considera el problema de la enseñanza de las técnicas de elicitación de requerimientos.

Las estrategias tradicionales (sin “interacción con un usuario”) no son suficientes y resulta fundamental incorporar nuevas estrategias didácticas (con clientes reales) al proceso de enseñanza para abordar su complejidad, y adecuarlas a las particularidades y dificultades planteadas, que permitan mejorar las prácticas y los aprendizajes de los estudiantes en los procesos de ingeniería de requerimientos.

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 en educación en informática como en tecnología informática aplicada a la educación.

El proyecto actual 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], donde el estudiante de forma consciente elige, planifica, coordina, y aplica métodos y procedimientos para conseguir un objetivo o aprendizaje.

El término estrategia se utiliza cada vez con mayor frecuencia y presupone la planificación de acciones a corto, mediano y largo plazo; son susceptibles al cambio, a la 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, en la actualidad se reconoce la necesidad de una didáctica 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 Software, los futuros profesionales deben tener habilidades y aptitudes determinadas (como abstracción, trabajo colaborativo y cooperativo, 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; los objetivos generales del aprendizaje distinguen tres elementos[5]:

- Conceptual: Relacionado con el incremento del conocimiento teórico del saber de un área.

- Procedimental: Se basa en la ampliación de conocimiento práctico del saber de un área.

- Integrador: Relacionado con el crecimiento de conocimiento sobre las destrezas, aptitudes y actitudes propias del ejercicio de una profesión.

Enseñar 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 aplicación[6].

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.

Para esto se necesitan técnicas que se incorporen en el proceso de enseñanza en el área de ingeniería de software, más concretamente en el área de la ingeniería de requerimientos, con el propósito de crear estrategias de enseñanza que permitan abordar la complejidad inherente, la cual es una actividad de transferencia difícil 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 Ingeniería de Requerimientos 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.

La dificultad de la obtención de los requerimientos radica fundamentalmente en la incertidumbre derivada del proceso de

transferencia de información desde una o varias fuentes, no siempre consistentes, hacia el ingeniero de requerimientos. Esta dificultad inherente a la elicitación se ve incrementada cuando se considera el problema de la enseñanza en un curso universitario de las técnicas de elicitación de requerimientos (grupos focales, JAD, prototipos, entrevistas, escenarios, lluvia de ideas, etc.).

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 son meramente descritas y, en el mejor de los casos, practicadas 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 ingeniería de requerimientos[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

Como parte de la línea de investigación y desarrollo, el equipo de investigadores lleva adelante trabajos relacionados con las temáticas vinculadas a la innovación en educación en

informática como en tecnología informática aplicada a la educación; se plantean los siguientes objetivos a alcanzar en el presente proyecto:

Objetivo General

Diseñar estrategias que permitan promover nuevas propuestas didácticas en la enseñanza de las técnicas de elicitación, especificación y validación de la ingeniería de requerimientos.

Objetivos Específicos

- 1.- Sintetizar el estado del arte respecto a la situación actual de la enseñanza de ingeniería de requerimientos en otras universidades.
- 2.- Diseñar diversas estrategias didácticas en el marco y contexto de la asignatura de Ingeniería de software de la carrera de Licenciatura en Sistemas de la UNTDF.
- 3.- Definir la/s estrategias que mejor se adecúan a las posibilidades áulicas y los tiempos de la asignatura.
- 4.- Implementar la/s estrategias seleccionadas en el curso de Ingeniería de Requerimientos en dos ciclos lectivos (2020 y 2021)
- 5.- Comparar las experiencias y producciones obtenidas respecto a los años previos.

Hipótesis

Como hipótesis general de la investigación se plantea que las estrategias didácticas tradicionales de enseñanza en la ingeniería de requerimientos (simples narrativas descriptivas de un contexto particular), dificultan al alumno comprender cabalmente los problemas que presentan las interacciones con los usuarios reales.

Poder implementar estrategias adecuadas con prácticas más cercanas a la realidad, permitirá mejorar la enseñanza de los docentes, y sobre todo, el aprendizaje de los alumnos en los procesos involucrados en la ingeniería de requerimientos.

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 la primera actividad del proyecto que consiste en el relevamiento de estrategias actuales a partir de la bibliografía existente y la experimentación con aquellos que estén disponibles y que sean considerados de interés para el proyecto

Para lo cual se espera lograr:

- Disponer de nuevas estrategias didácticas para la enseñanza de la ingeniería de requerimientos, que permitan a los estudiantes una mejor comprensión del trabajo de especificar requerimientos, a través de interacciones con usuarios, lo más cercano a la realidad.
- Enriquecer las experiencias de enseñanza y aprendizaje y ofrecer clases más dinámicas
- Reflexionar y resignificar la práctica.
- Establecer un conjunto de criterios que pueden utilizarse para satisfacer distintos espacios curriculares de similares.
- Conocer y comprender el aprendizaje de los estudiantes y aumentar el contacto con las experiencias reales y lúdicas como como filmaciones, juegos[10], entrevistas en vivo, desarrollo a clientes por convenios, etc.
- Implementar nuevas herramientas tecnológicas en el aula que permitan trabajar sobre estas competencias de manera más explícita.

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 docente Martín Villarreal se encuentra cursando el Magister en Ingeniería de Software de la UNLP.

Los docentes Lucas Romano y Marcela Jerez se encuentran cursando el Magister de Tecnología Informática Aplicada en Educación en la Facultad de Informática de la UNLP, y tienen previsto realizar sus tesis de Maestría en temas vinculados con el 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.

Los alumnos Daniel Andrade, Rigoni y Lopez se encuentran cursando el último año de la carrera, los alumnos Iost y Villanueva se encuentran desarrollando el trabajo final para el título de Analista Universitario de Sistemas.

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).

A su vez los docentes podrán nutrirse de la experiencia de quienes manejan diariamente los problemas y conflictos vinculados a la enseñanza de Sistemas.

Referencias

1. Burón, J. *Aprender a aprender: Introducción a la metacognición*. Bilbao: Ed. Mensajero; 1994, p.94.
2. Montes de Oca, N. *Estrategias docentes y métodos de enseñanza-aprendizaje en la Educación Superior*. Rev Hum Med [online]. 2011, vol.11, n.3, pp.475-488. ISSN 1727-8120

3. Fernández, A.; Ginoris, F. *Didáctica y optimización del proceso de enseñanza-aprendizaje*. La Habana: Instituto Pedagógico Latinoamericano y Caribeño (IPLAC); 1998.
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 Enseñanza Univ. Informática, pp.191–198, 2006.
6. Loucopoulos P., Karakostas V., *Systems Requirements Engineering*. McGraw-Hill, 1995.
7. Oliveros, A.; *Aprendizaje de la elicitación y especificación de requerimientos* Conference Paper, October 2014.
8. Montoya Suarez L. M., Echeverri Arias J. A.; *Estrategias didácticas para enseñar Técnicas de Elicitación de Requisitos*. Universidad Católica Luis Amigó, Colombia, 2013
9. Montoya Suárez L. M., Monsalve Gómez j. C., Gaviria Hincapié J. M., *Aplicación de un Caso de estudio para la enseñanza de Ingeniería de Requisitos*. Universidad Autónoma Latinoamericana Medellín, Colombia, 2014.
10. Blanco Palencia S. M., López F. E., Viejo A.; *Juego educativo enfocado al desarrollo de habilidades para aprender a modelar requerimientos con artefactos UML*. 2015; Universidad Abierta de Cataluña; (Proyecto de Investigación). Recuperado de http://journal.seriousgamessociety.org/index.php?journal=IJSG&page=article&op=view&path%5B%5D=41&path%5B%5D=pdf_7.

Enseñanza introductoria de la programación: un estudio curricular por competencias

Ana M. Piccin, Diana Cicinelli

Teoría de la Computación /Facultad de Ingeniería y Tecnología de la Información/

Universidad de Belgrano

ana.piccin; diana.cicinelli @comunidad.ub.edu.ar

RESUMEN

La programación puede aparecer referida en textos académicos tanto como fácil y divertida como difícil y frustrante. Los trabajos analizados revelan la complejidad inherente al estudio de este problema. Una de las variables del problema es la ubicación curricular de la enseñanza de la programación que, por razones históricas y de motivación para los alumnos, se incluye generalmente de manera temprana en los planes de estudios. Ubicar la enseñanza de la programación en los primeros módulos curriculares lleva a una visión eminentemente práctica de la programación, distanciada de los desarrollos teóricos. En la literatura específica se ha referido a esta separación entre teoría y práctica como una posible causa de que los alumnos cometan errores clásicos difícilmente superables sin conocimiento teórico. Este proyecto se propone realizar un estudio curricular basado en competencias que permita organizar las prácticas relacionadas con la enseñanza y la evaluación de la programación cuando dicha enseñanza está ubicada tempranamente en el plan de estudios. Se producirá un modelo curricular por competencias, se obtendrán requisitos funcionales para el diseño de aplicaciones didácticas que induzcan a relacionar la experiencia en la práctica con la teoría de la disciplina. En calidad de prueba de concepto, se desarrollará por lo menos una aplicación didáctica.

Palabras clave: enseñanza introductoria de la programación; diseño curricular por competencias; currículum manifiesto, currículum real.

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 de la universidad, en la Facultad de Ingeniería y Tecnología Informática, se encuentra en ejecución una línea específica para el desarrollo de herramientas para la enseñanza de la programación en los cursos introductorios, “*Framework* para la construcción de estrategias didácticas para la enseñanza introductoria de la programación” (FEDEP), en el que se encuadra el proyecto 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. Han participado alumnos locales y de universidades extranjeras. En todos los casos los proyectos

culminan en el desarrollo de aplicaciones didácticas.

1. INTRODUCCIÓN

Las competencias desarrolladas en los cursos introductorios de programación son consideradas “un fundamento esencial para las competencias informáticas más avanzadas en todas las áreas del conocimiento curricular, y un objetivo central en nuestro currículum” [1] y deberían proveer competencias básicas requeridas para cursos posteriores dentro de los planes de estudios de las carreras de informática [11]. Sin embargo, contrariamente a lo deseado, se verifica que los estudiantes “reprueban o abandonan sus cursos en tasas mayores que lo que es aceptado por las instituciones” [9].

La teoría del diseño curricular señala la diferencia entre el “currículum real” y el “manifiesto” y la necesidad de estudiar la relación entre ambos [2]. El proyecto que se presenta tiene por hipótesis que el diseño de un currículum por competencias para los cursos introductorios de programación de las carreras de grado de informática permitiría revelar el currículum real [3], describirlo y establecer su relación con el currículum manifiesto. Una vez completado el modelo curricular por competencias, se podrán producir los requisitos funcionales para el diseño de aplicaciones didácticas orientadas a facilitar que los alumnos puedan construir una relación conceptual entre práctica y teoría.

Los ejes sobre los que se formuló este proyecto corresponden a dos de las múltiples variables que intervienen en la problemática de la enseñanza de la programación [7]. Una, el escaso o nulo conocimiento de semántica del lenguaje de programación utilizado en las

prácticas [7]; la otra, dudas formuladas por los docentes e investigadores con respecto a calidad de las evaluaciones que ellos mismos diseñan [8] [7].

La programación, aun considerando exclusivamente su dimensión práctica, es una actividad de orden intelectual, de aplicación intensiva de la lógica y la abstracción [12], que si bien puede ser referida como “divertida” [12], fácil para unos o difícil para otros, es indudablemente compleja [13] [5] [2]. El contraste entre el atractivo de su práctica, a la que se accedería con pocos conocimientos teóricos, y sus conceptos fundamentales sorprende al practicante novato. Se lee en [2]: “Parece ser que la programación es una actividad intelectual con características únicas (...) el programador competente debe moverse entre varios niveles semánticos (...). [Se requiere] una agilidad desconcertante para quienes no están acostumbrados a ella”.

En [10] se expresa: “(...) una de las cuestiones más debatidas en la educación en ciencias de la computación es el rol de la programación en el currículum introductorio”, y la decisión de incluirla tempranamente en los currículos “... surge de una cantidad de razones prácticas e históricas”, en contraposición a estar basada en razones epistemológicas, psicológicas o cognitivas, por ejemplo, según las cuales debería ubicarse en ciclos posteriores, cuando el alumno ya ha cursado materias de ciencias básicas. [10]. Esta circunstancia favorecería una disociación conceptual entre los contenidos asociados a la teoría que, por abstractos, serían inabordables por los alumnos en instancias tempranas del plan de estudios, y el objetivo de que éstos desarrollen competencias básicas de programación.

Por su ubicación curricular, los fundamentos teóricos que gobiernan la semántica y por lo tanto la sintaxis y la posibilidad de operar efectivamente en la construcción de soluciones y en el desarrollo de las competencias necesarias en cursos posteriores, son postergados a otras instancias de la carrera facilitando que la idea de práctica desplace a la de teoría.

La inserción curricular temprana de la programación colocaría en un segundo plano sus fundamentos teóricos haciendo prevalecer el sentido de artesanía sobre el de disciplina [13] [4]. Esta condición favorecería la concepción docente por la que la enseñanza es eminentemente organizada sobre la práctica [6] [8].

Un diseño curricular por competencias permitirá analizar la relación entre enseñanza y evaluación y también articularlas. El estudio curricular que se propone prevé la articulación con otras asignaturas del plan de estudios para las que se requieran conocimientos específicos de programación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para completar el proyecto se deberán alcanzar los siguientes objetivos:

Objetivo principal: Construir una aplicación didáctica orientada al desarrollo de competencias específicas de la enseñanza introductoria de la programación, que permita que los alumnos establezcan relaciones entre la práctica que realizan y fundamentos teóricos de la disciplina. Dicha aplicación estará basada en un modelo curricular por competencias,

diseñado acorde a planes de estudio de carreras de grado de informática encuadrados en resolución 786/09 del Ministerio de Educación de la Nación.

Objetivos específicos:

1 Construir un modelo conceptual de competencias incluidas en la enseñanza introductoria de la programación en carreras de grado de informática.

2 Identificar la teoría de la programación relacionada con las competencias o conjuntos de competencias representados en el modelo conceptual referido en el objetivo 1.

3 Desarrollar una herramienta que permita evaluar la coherencia de competencias a lo largo de una asignatura, incluyendo prácticas de la enseñanza y de la evaluación.

4 Determinar requisitos funcionales para el diseño de aplicaciones digitales basadas en un modelo curricular por competencias y que permitan vincular los objetivos pedagógicos de rectores de la práctica con los fundamentos teóricos de la programación.

5 Desarrollar por lo menos una aplicación didáctica.

El logro de dichos objetivos permitirá corroborar la siguiente hipótesis:

El desarrollo de un modelo curricular por competencias de la enseñanza introductoria de la programación permitirá: a) Establecer relaciones entre las competencias prácticas y sus fundamentos teóricos y b) diseñar requisitos funcionales y desarrollar aplicaciones didácticas orientadas a facilitar que los alumnos relacionen las competencias que están desarrollando en la práctica con los fundamentos teóricos sobre los que se fundan.

Metodología: El proyecto está organizado en dos módulos. El primer módulo está caracterizado por la construcción de un mapa conceptual de las competencias implícitas o explícitas en los objetivos pedagógicos del curso introductorio de programación. El núcleo del segundo módulo está constituido por la elicitación de requisitos funcionales para el diseño y desarrollo de aplicaciones didácticas que faciliten que el alumno relacione los aspectos de la práctica con sus fundamentos.

3. RESULTADOS ESPERADOS

Los datos obtenidos en el primer módulo serán interpretados a la luz de la teoría del diseño curricular por competencias [3] y se utilizarán teorías y técnicas de grafos para la organización de un mapa conceptual de competencias. Una vez construido dicho mapa, se apelará a la teoría de la programación para la identificación de núcleos teóricos que fundamenten las competencias incluidas en el mapa conceptual. Los resultados hasta acá obtenidos colaborarán, junto con un análisis desde la teoría general de la didáctica [3], con la definición de requisitos funcionales para el diseño de aplicaciones didácticas. Una vez identificados y descriptos dichos requisitos funcionales, se procederá al desarrollo de una aplicación.

Los conocimientos y herramientas desarrollados durante el estudio curricular podrán ser adaptados a otros planes de estudios de carreras de informática con independencia del nombre o espacio curricular que se otorgue a la enseñanza introductoria de la programación.

Los resultados teóricos y prácticos del proyecto se aplicarán al diseño curricular por competencias de las asignaturas que componen el eje de enseñanza de la programación en las carreras de grado de informática de la Universidad de Belgrano; se diseñarán articulaciones entre ellas. Posteriormente las herramientas desarrolladas a lo largo del proyecto serán aplicadas a asignaturas de otras áreas del plan de estudios que requieran conocimientos previos de programación; por ejemplo, las orientadas a diseño y desarrollo de software o a los sistemas operativos avanzados.

Se han planificado cursos de extensión sobre diseño curricular por competencias dirigidos a docentes de nivel medio y superior.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está integrado por los siguientes investigadores:

- Mg. Ana M. Piccin, Prof. Titular de la asignatura Sistemas Inteligentes, coordinadora del área “Teoría de la computación y lenguajes” de la Facultad de Ingeniería y Tecnología Informática. Directora del equipo de investigación.
- CC Diana Cicinelli, docente investigadora Prof. Adjunta I, a cargo de los cursos introductorios de programación, integrante de equipo, colaboradora de investigación.
- Colaboradores: El proyecto fue diseñado para facilitar la incorporación de alumnos de las carreras de grado de informática de la Universidad de Belgrano o pasantes externos.

Tesis de maestría:

Título: “Enseñanza de la programación: las concepciones de los docentes de los cursos introductorios de las carreras de grado en informática”.

Tesista: Ana M. Piccin; *Directora:* Graciela D. S. Hadad.

Título alcanzado: Magíster en Gestión de Proyectos Educativos.

Universidad CAECE, Ciudad Autónoma de Buenos Aires, Argentina.

5. BIBLIOGRAFÍA

- [1] Bills, Dianne P. y Biles, John A. (2005) “The Role of Programming in IT”, en *SIGITE’05*, October 20-22, 2005, Newark, New Jersey, USA. Copyright 2005 ACM, 43-49. Extraído de <http://www.acm.org> el 10 de noviembre de 2008, p. 43.
- [2] E. Dijkstra, “Programming as a discipline of mathematical nature”, *The American Mathematical Monthly*, Vol. 81, No. 6 pp. 608-612, Jun.-Jul., 1974.
- [3] Gimeno Sacristán, J. y Pérez Gómez, A. I. (2002) *Comprender y transformar la enseñanza*. Ediciones Morata, S. L. Madrid.
- [4] C.A.R., Hoare, “An Axiomatic Basis for Computer Programming”, *Communications of the ACM*, Vol. 1, Num. 10, Oct., 1969, p.576.
- [5] C.A.R., Hoare, “Notes on data structuring”, Conferencias en la Nato Summer School, Marktoberdorf.
- [6] S. M. Merchán Rubiano, O. López-Cruz y E. Gómez Soto, "Teaching computer programming: Practices, difficulties and opportunities," en *IEEE Frontiers in Education Conference (FIE)*, El Paso, TX, 2015, pp. 1-9.
- [7] A. M. Piccin, “Programming Teaching Issues: Presence and Evolution throughout Specialized Literacy”, *2017 XLIII Latin American Computer Conference (CLEI)*, Córdoba, Argentina, 4 a 8 de septiembre de 2017- IEEE Xplore, URL: <https://ieeexplore.ieee.org/document/8226451>.
- [8] A. M. Piccin, “Enseñanza de la programación: las concepciones de los docentes de los cursos introductorios de las carreras de grado en informática”, Departamento de Educación, Maestría en Gestión de Proyectos Educativos, Universidad CAECE, Ciudad Autónoma de Buenos Aires, Argentina, 2020.
- [9] Simon et al (2016) “Benchmarking Introductory Programming Exams: Some Preliminary Results”, en *Proceedings of the 2016 ACM Conference on International Computing Education Research* [103-111].
- [10] The Joint Task Force on Computing Curricula, “Computing Curricula 2001 Computer Science — Final Report”, November 2001 y la Junta de Gobierno de la IEEE-CS, December 2001, pp.22-24.
- [11] The Joint Task Force for Computing Curricula, “Computer Science Curricula 2013: Final Report”, Computing Association for Computing Machinery (ACM); The Association for Information Systems (AIS); The Computer Society (IEEE-CS), 2013.

- [12] D. Watt, *Programming Language Design Concepts*; John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2004.
- [13] N. Wirth, *Algorithms + Data Structures = Programs*, Prentice-Hall Series in Automatic Computation, Prentice-Hall, inc. Englewood Cliffs, New Jersey, 1976.

Enseñanza de la programación: las concepciones didácticas de los docentes de los cursos introductorios

Autora: Ana M. Piccin; ana.piccin@comunidad.ub.edu.ar

Directora: Graciela D. S. Hadad; ghadad@uno.edu.ar

La tesis que se presenta corresponde a la Maestría de Gestión de Proyectos Educativos de la Universidad CAECE. Fue defendida el 26 de noviembre de 2020, obteniendo su autora el título de Magíster.

El tema de esta investigación surgió de la necesidad de producir herramientas conceptuales para la gestión educativa aplicada a los cursos introductorios de programación de las carreras de grado de informática. En la literatura especializada, no solo se informa reiteradamente sobre los fracasos de los alumnos en el aprendizaje de las competencias básicas de programación, asociadas a su vez con los altos niveles de deserción en los primeros años de las carreras de informática, sino que también se suele dar cuenta de la escasez de resultados ofrecidos desde la educación. Así surgió la pregunta que motivó y guio este estudio: ¿cómo creen, piensan, conciben, los docentes de los cursos introductorios de programación, que se debe enseñar?

Esta pregunta dio lugar al planteo del objetivo principal de este estudio: describir las distintas maneras en que los docentes de los cursos introductorios de programación abordan la enseñanza, el que a su vez fue descompuesto de la siguiente manera: 1) Describir concepciones de los docentes con respecto a la enseñanza de la programación; 2) Describir enfoques con los que los docentes abordan la enseñanza; 3) Describir relaciones entre enfoques, concepciones y dimensiones de la enseñanza. El objetivo de establecer relaciones entre enfoques de la enseñanza y concepciones didácticas se alcanzó mediante el proceso de dar respuesta a las siguientes preguntas de investigación: ¿de qué diferentes maneras los docentes enfocan la enseñanza introductoria?; ¿cómo conciben los docentes al objeto de enseñanza?; ¿cómo conciben los docentes a la disciplina? y ¿qué concepciones tienen los docentes con respecto a sus alumnos?

Se consideraron docentes de los cursos introductorios de programación de las carreras de grado de informática, acreditadas por CONEAU, vigentes en el Ámbito Metropolitano de Buenos Aires.

El trabajo se sustentó sobre la teoría de las concepciones didácticas, en cuanto a que las decisiones en la planificación de la enseñanza, así como en su práctica, están sujetas a un conjunto determinado de estructuras conceptuales. Es por ello que se aplicó el método fenomenográfico. Este método cualitativo-interpretativo, basado en *Grounded Theory*, está diseñado específicamente para el estudio de concepciones y está orientado a describir las variaciones con las que distintos individuos experimentan un fenómeno, para este caso, las distintas maneras en que los docentes de los cursos introductorios experimentan la enseñanza de la programación.

Con el objeto de que la muestra resultante estuviera caracterizada por la variedad, se partió de encuestas respondidas por docentes de cursos introductorios de programación que se desempeñan en universidades públicas y privadas de la región estudiada, lo que permitió la selección de una primera muestra.

La recolección de datos se realizó mediante entrevistas estructuradas con un protocolo diseñado según los requisitos metodológicos, el que fue validado mediante una prueba piloto.

El estudio se organizó en dos grandes etapas, cada una con su propia estrategia. En la primera, se obtuvieron categorías de las concepciones didácticas y se las describió. Para ello se aplicó el método fenomenográfico, con la ejecución recurrente de la comparación constante. Sucesivas abstracciones sobre los datos permitieron la organización de las concepciones en categorías, las que se describieron en textos ilustrados con expresiones extraídas de las entrevistas.

En la segunda etapa, para la categorización de enfoques de la enseñanza, se utilizaron dimensiones según criterios propios de la teoría general de la didáctica, descartando las dimensiones emergidas de los datos según la ejecución de la primera etapa. Sucesivos análisis cuantitativos y cualitativos de los datos permitieron organizar y describir los enfoques en cuatro categorías.

1. RESULTADOS OBTENIDOS

Con respecto al objeto de enseñanza. Los docentes comparten la siguiente concepción: “Analizar, entender y resolver problemas, aplicando algoritmos y estructuras de datos, mediante el uso de un lenguaje de programación, independientemente del lenguaje que se utilice, observando las buenas prácticas”.

Con respecto a la disciplina: Los docentes perciben a la programación como una disciplina de orden práctico que requiere de niveles de capacidad de abstracción superiores a las de otras disciplinas; es necesaria en otros cursos de las carreras de informática; requiere y permite desarrollar competencias para la resolución de problemas; es compleja, su complejidad estaría relacionada con la representación del mundo real en un mundo virtual.

Con respecto a los alumnos: La percepción colectiva de los docentes con respecto a los alumnos de los cursos introductorios refleja un individuo que llega a la universidad carente de la preparación intelectual adecuada y de competencias metacognitivas suficientes; sus actitudes hacia el aprendizaje no estarían adecuadamente desarrolladas.

2. CONCLUSIONES

Se han logrado describir cuatro categorías de concepciones didácticas; “Operativa”, “Académica”, “Vincular” y “Contextual”, y cuatro de enfoques de la enseñanza “Asociacionista”, “Cognitivo”, “Personal” y “Social”. Se ha comprobado la coherencia entre concepciones didácticas y enfoques de enseñanza.

Mediante el análisis de los planes de estudios y programas analíticos, se corroboró la correspondencia entre la concepción de los docentes y la documentación institucional. Sin embargo, se llegó a la conclusión de que el término “resolver problemas”, presente en la concepción compartida del objeto de enseñanza, presentaría ambigüedades. De las expresiones de los docentes surge que se debería tener en cuenta, de manera expresa, el concepto “buenas prácticas” en el diseño curricular.

El lenguaje de programación con que se realiza la enseñanza es concebido por los docentes como una herramienta. No se pudieron obtener conclusiones sobre la

importancia que los docentes otorgan al paradigma de programación en la enseñanza introductoria.

Como futuros trabajos se propone: determinar qué tipo de disciplina es la programación como punto de partida para el diseño de una didáctica específica; profundizar en el conocimiento de la relación didáctica entre resolución de problemas, paradigmas de programación y lenguaje que se utilice en la enseñanza; desarrollar modelos de evaluación que satisfagan las necesidades de la disciplina.

Dado el método riguroso seguido, es esperable la replicación de este estudio en distintas regiones y con muestras más extensas para corroborar o refutar los resultados obtenidos.

Los resultados y conclusiones de esta investigación podrán ser aplicados al diseño curricular, a la gestión de contenidos y a la capacitación docente.

EscuelasTIC: estrategias para trabajar el pensamiento computacional en la escuela argentina

Claudia Queiruga, Claudia Banchoff Tzancoff, Paula Venosa, Sofía Martín, Vanessa Aybar Rosales, Soledad Gomez, Isabel Kimura

Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI). Facultad de Informática. Universidad Nacional de La Plata
50 y 120. La Plata

{claudiaq, cbanchoff, pvenosa, vaybar, sgomez}@info.unlp.edu.ar, {smartin, ikimura}@linti.unlp.edu.ar

Resumen

En los últimos años las experiencias de incorporación de tecnologías digitales en las aulas de las escuelas de nuestro país se ha desarrollado de forma diversa, adoptando diferentes enfoques, entre ellos la “informática de usuario”, sin abordar las especificidades de la Informática como disciplina, y en otras experiencias aparece el enfoque del “pensamiento computacional” como nuevo paradigma de enseñanza de la informática en la escuela, ampliando las miradas instrumentales sobre los saberes de la Informática. Este artículo presenta los avances de la línea de investigación, denominada “EscuelasTIC” que problematiza el uso de tecnologías digitales a la vez que propone el desarrollo de nuevos materiales que contribuyan a orientar la incorporación crítica de saberes fuertemente vinculados a dinámicas de inclusión/exclusión social.

Palabras clave: Informática, enseñanza de programación, robótica educativa, pensamiento computacional, formación docente, TIC.

Contexto

La línea de investigación “EscuelasTIC” forma parte del proyecto de investigación

“De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región” del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI) de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). Esta línea de trabajo, que viene desarrollándose desde hace varios años, articula con actividades de extensión que promueven la inclusión de distintos conceptos relacionados a la Informática en las aulas de la escuela. Se focaliza principalmente en las áreas de programación, ciberseguridad y ciencia de datos, con distintas estrategias de aplicación. En este artículo se presenta la evolución de la investigación ya presentada en eventos anteriores [1].

Introducción

La sociedad actual, denominada por algunos autores Sociedad de la Información o Sociedad del Conocimiento se caracteriza por estar mediada por tecnologías digitales, las cuales forman parte de la mayoría de los procesos productivos. Lo podemos observar en nuestras actividades cotidianas y se puso claramente en evidencia con la irrupción de la pandemia

provocada por el COVID-19 en donde la mayoría de las actividades se realizaron en modalidad virtual, desde las clases en las escuelas, las universidades, institutos, la adopción de los trabajos en modalidad home-office, las consultas médicas por videollamadas, las compras on-line, las reuniones sociales, el acelerado crecimiento de las apps para múltiples servicios gubernamentales, son algunos ejemplos de los usos de las tecnologías digitales en nuestra cotidianeidad, a la vez que los problematizan.

Esta sociedad *digital* nos plantea desafíos en múltiples planos, uno de ellos está vinculado a la **formación ciudadana** y en cómo pensamos la relación de los ciudadanos con las tecnologías digitales. Específicamente en la educación obligatoria se plantea la formación de ciudadanos que puedan comprender los lenguajes digitales, cómo funcionan los artefactos digitales, ubicándolos como sujetos críticos y creadores de innovaciones con tecnologías digitales, por sobre la pasividad y el mero consumo tecnológico.

Las TIC se han ido incorporando ampliamente a las prácticas educativas en los distintos niveles de la escolaridad obligatoria de nuestro país, primando el enfoque de “informática de usuario” que hace hincapié en el conocimiento y la capacidad de utilizar las computadoras, y la tecnología relacionada con ellas de manera eficiente, con un enfoque “utilitario” en algunos casos e “integrador” en otros [2]. En los últimos años se ha ido modificando este enfoque y en los diseños curriculares de algunas provincias comienza a aparecer el enfoque del **pensamiento computacional** como nuevo paradigma de enseñanza de la informática que hace referencia a técnicas y metodologías de resolución de problemas, donde principalmente intervienen saberes que provienen del

campo de la Informática [3]. En ese sentido la CAS [4] señala como los cinco elementos claves del pensamiento computacional a los siguientes: a) el **pensamiento algorítmico** en referencia a la capacidad de expresar soluciones a problemas a partir de una serie de pasos que un autómata realiza; b) la **descomposición**, referido a la capacidad de dividir e identificar las partes que componen un problema para facilitar su tratamiento y análisis; c) la **generalización**, entendida como la capacidad para descubrir patrones en los problemas o en las soluciones que son aplicables a ellos; d) la **abstracción**, como la capacidad de elegir las representaciones que destacan las características relevantes a un contexto y ocultan los detalles innecesarios; y e) la **evaluación**, entendida como la capacidad de analizar críticamente las soluciones creadas para detectar y corregir posibles errores, así también como la búsqueda de soluciones que aprovechen mejor los recursos. En su primer artículo, Jeannette Wing señala “ (...) *el pensamiento computacional es una habilidad fundamental para todas las personas, no sólo para los informáticos*” [3].

En cuanto a políticas educativas de nuestro país en relación con este nuevo enfoque, se pueden señalar los Núcleos de Aprendizaje Prioritarios (NAP) definidos por el Consejo Federal de Educación [5] que incorporan saberes sobre programación y robótica en los diferentes niveles de la educación obligatoria. Diferentes provincias han modificado sus diseños curriculares atendiendo a estos NAP. Sin embargo, a pesar de estos avances, aún no se ha logrado permear en las escuelas argentinas los diseños curriculares ni conformar redes de capacitación docente que favorezcan la incorporación de estos saberes. En este sentido podríamos afirmar que el

pensamiento computacional y los conceptos de Informática han comenzado a debatirse en espacios educativos que promueven su inclusión en los nuevos currículos escolares reemplazando las ideas vinculadas al uso eficiente, sin embargo aún se trata de experiencias jurisdiccionales.

Líneas de Investigación, Desarrollo e Innovación

Esta línea de investigación se centra en los siguientes temas:

- Análisis de cómo los conceptos del pensamiento computacional permean la enseñanza de Informática en la escuela.
- Relación entre competencias digitales y pensamiento computacional.
- Formación docente: desarrollo y puesta en acción de propuestas de capacitación y actualización para los niveles primario y secundario que contemplen ideas de aprendizaje situado [6] y de design thinking[7].
- Diseño y elaboración de materiales didácticos para la enseñanza de Informática para los distintos niveles educativos.
- Diseño, ejecución y evaluación de actividades específicas con estudiantes y docentes de los niveles primario y secundario.

La presencia del equipo de investigación en las escuelas de la región de La Plata, Berisso, Ensenada y Magdalena se ha consolidado a lo largo de los años a través de proyectos de extensión, que proveen evidencias con base empírica para continuar indagando sobre las formas de incorporar esta disciplina en los sistemas de educación formales y obligatorios. Si bien la situación producida por la pandemia de COVID-19 ha puesto en suspenso algunas de las actividades, esta misma situación produjo un enorme desafío en la adecuación de las mismas

tanto para el equipo de investigación como para los equipos docentes de las escuelas.

La participación del equipo en el diseño curricular de la carrera “Especialización docente en didáctica de las Ciencias de la Computación” de la provincia de Buenos Aires y su implementación, y la elaboración del “Primer manual de didáctica de las Ciencias de la Computación”¹ ha generado experiencia en cuanto a poder intervenir en formación docente de nuestra provincia [8][9].

Los conceptos trabajados están relacionados con la enseñanza de la programación, de ciberseguridad y de ciencia de datos. Respecto al primero, se han utilizado distintas estrategias basadas en el uso de lenguajes de programación visuales basados en bloques y la manipulación de objetos físicos relacionados a la robótica educativa e “Internet de las cosas”. Respecto al abordaje de temas relacionados a la ciberseguridad se han llevado a cabo instancias de competencias de “captura la bandera” o CTF (Capture The Flag) utilizando una plataforma desarrollada específicamente para este fin. De esta manera se introduce una nueva forma de abordar conceptos de áreas del campo de la Informática en la escuela, con resultados alentadores.

Resultados y Objetivos

El objetivo principal de esta línea de trabajo es fortalecer el aprendizaje crítico y significativo de la Informática en la escuela. Se pone especial atención a la adopción de algunas características del enfoque del pensamiento computacional,

¹El manual está disponible en https://program.ar/descargas/cc_para_el_aula-2do_ciclo_secundaria.pdf. Último acceso: marzo de 2021

el aprendizaje situado, el design thinking y el acercamiento de saberes de programación, ciberseguridad y ciencia de datos orientado a la promoción de competencias vinculadas a diseñar, crear e innovar con tecnologías digitales.

Para alcanzar este objetivo, se proponen los siguientes actividades específicas:

- Desarrollar propuestas pedagógicas que involucren el trabajo de y con Informática para trabajar en la escuela.
- Desarrollar herramientas didácticas digitales que puedan ser utilizadas en distintas propuestas pedagógicas buscando un abordaje holístico para la resolución de problemas. Se propone la aplicación de conceptos de gamification y juegos serios.
- Diseñar, implementar y evaluar las intervenciones con docentes y estudiantes de los diferentes ciclos escolares.
- Difundir las producciones.

En el marco de esta línea de trabajo se han desarrollado y se encuentran en proceso, las siguientes herramientas resultantes de tesinas de grado: una plataforma de competencias CTF; Blokino, una herramienta centrada en la construcción y programación de objetos físicos basados en Arduino; Studium, un sistema de gestión de aprendizaje que integra al EVEA Chamilo² la herramienta “Juegos de Blockly”³ y ENREDADOS, un juego serio para la enseñanza de redes. Los avances de la tesis doctoral sobre aprendizaje automático en la enseñanza secundaria, de la tesis de maestría en Ciencia, Tecnología y Sociedad sobre competencias informáticas en la educación secundaria técnica y la finalización de la tesis de maestría⁴ en

² <https://chamilo.org/es/>

³ <https://blockly.games/>

⁴ Tesis de maestría disponible en <http://sedici.unlp.edu.ar/handle/10915/111306>

Tecnologías Informáticas Aplicadas a Educación, sobre competencias digitales y pensamiento computacional, permiten contar con resultados con base empírica de esta línea de investigación.

Formación de Recursos Humanos

El equipo de trabajo está integrado por docentes-investigadores del LINTI y estudiantes de Informática e Ingeniería en Computación. Se han formulado varias tesinas de grado, tesis de postgrado, proyectos de extensión y actividades de cátedras relacionadas con los temas de esta línea de investigación. Actualmente, se encuentran en desarrollo varias tesinas de grado y tesis de postgrado que contribuirán a avanzar en esta línea de trabajo.

Referencias

[1] Queiruga C., Banchoff Tzancoff C., Venosa P., Martin S., Aybar Rosales V., Gomez S., Kimura I. (2019). Escuelas TIC. El pensamiento computacional en la escuela. En WICC 2019, Universidad Nacional de San Juan. Libro de Actas XX Workshop de Investigadores en Ciencias de la Computación. ISBN: 978-987-3619-27-4.

[2] Levis D. (2007). “Enseñar y aprender con informática/ enseñar y aprender informática. Medios informáticos en la escuela argentina”. En Cabello, R. y Levis D., edits. (2007) Tecnologías informáticas en la educación a principios del siglo XXI Buenos Aires: Prometeo.

[3] Wing, J. M. (2006). Computational Thinking. Communications of the ACM, vol. 49, 33-35.

[4] CAS (2015) Pensamiento Computacional. Guía para profesores. Computing At School. Disponible en

<https://community.computingschool.org.uk/resources/2324/single>. Último acceso: marzo de 2021.

[5] Resolución del Consejo Federal de Educación N° 343/18. Disponible en https://www.argentina.gob.ar/sites/default/files/res_cfe_343_18_0.pdf. Último acceso: marzo de 2021.

[6] Díaz Barriga F. (2003). Cognición situada y estrategias para el aprendizaje significativo. Revista. Electrónica de Investigación Educativa, 5(2). Disponible en: <https://bit.ly/1JSJZpG>. Último acceso marzo de 2021.

[7] Brown T. (2009). Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation. HarperBusiness.

[8] Queiruga C., Banchoff Tzancoff C., Venosa P, Gómez S., Morandi G. (2018). Computer Science and Schools: A Specific Didactics?. In: Pesado P., Aciti C. (eds) Computer Science – CACIC 2018. Communications in Computer and Information Science, vol 995. Springer, Cham. Pag 343-351.

[9] Queiruga C., Banchoff Tzancoff C., Gómez S., Venosa P. (2020). Enseñar a Enseñar Ciencias de la Computación. Una experiencia sobre políticas educativas y contenidos de Ciencias de la Computación. En Actas II Jornadas Argentinas de Didáctica de la Programación / Alejandro Iglesias... [et al.] ; editado por Araceli Acosta... [et al.].- 1a ed.- Córdoba: Universidad Nacional de Córdoba. Facultad de Filosofía y Humanidades, 2020. ISBN 978-950-33-1600-9.

Experiencia de Enseñanza y Aprendizaje del Pensamiento Computacional en las carreras de Ingeniería

Pamela Viale^{1,2}, Claudia Deco^{1,2}, Cristina Bender^{1,2}, Sebastián Velazquez¹

¹ Facultad de Química e Ingeniería del Rosario,
Universidad Católica Argentina, Campus Rosario, Argentina.
{pamelaviale, cdeco, cbender, angelvelazquez}@uca.edu.ar

² Facultad de Ciencias Exactas, Ingeniería y Agrimensura,
Universidad Nacional de Rosario, Rosario, Argentina.
{pamela, deco, bender}@fceia.unr.edu.ar

RESUMEN

En este artículo se presenta una línea de investigación que se está llevando a cabo en la Universidad Católica Argentina, en su sede Rosario en la Facultad de Química e Ingeniería, que está orientada al desarrollo de estrategias motivacionales destinada a estudiantes de las carreras de ingeniería para el estudio y aprendizaje del Pensamiento Computacional y de la Programación. Se desarrollaron algunas primeras experiencias con este nuevo enfoque en los cursos iniciales de dichas carreras y los resultados son promisorios.

Palabras clave: Pensamiento Computacional, Programación, Carreras de Ingenierías.

CONTEXTO

Esta línea de I+D se está llevando a cabo a través de un proyecto del Departamento de Investigación Institucional de la Facultad de Química e Ingeniería del Rosario de la Universidad Católica Argentina.

El proyecto involucrado es un PID UCA (período 2019-2022) titulado “Pensamiento Computacional Aplicado a Educación”, cuya directora es la Dra Claudia Deco.

1. INTRODUCCIÓN

En la actualidad se considera que la enseñanza del Pensamiento Computacional resulta beneficiosa para todos los profesionales, independientemente de la especialidad donde se desarrolle. Se usa el término Pensamiento Computacional (PC) para referirse a un conjunto de aptitudes y/o competencias útiles para la formulación y resolución de problemas de manera que puedan ser resueltos por un agente de procesamiento de información [1, 2, 3].

El PC promueve la utilización de cuatro estrategias principales para la solución de problemas:

- (1) Descomposición de un problema en subproblemas: involucra el análisis de problemas complejos y su descomposición en problemas más pequeños, más fáciles de analizar.
- (2) Reconocimiento de patrones: cada uno de estos problemas más pequeños puede ser analizado en profundidad para identificar problemas similares ya resueltos.
- (3) Abstracción: para buscar las soluciones a los problemas encontrados es necesario focalizarse en los detalles importantes e ignorar información no relevante.
- (4) Pensamiento algorítmico: pueden proponerse una serie de pasos o reglas a seguir para crear una solución para cada uno de los subproblemas encontrados.

Una solución elaborada utilizando las estrategias del PC, puede ser fácilmente implementada en un sistema computacional, construyendo así una solución eficiente a un problema inicial complejo.

El estudio del Pensamiento Computacional ha tenido influencia en las investigaciones relacionadas al entendimiento y desarrollo de los procesos de enseñanza y aprendizaje, provocando que se aborde el tema desde el punto de vista de la formación en los distintos niveles educativos [4, 5, 6].

Se ha decidido entonces focalizar este proyecto en mejorar la enseñanza de las competencias mencionadas anteriormente a los estudiantes de las carreras de Ingeniería de la Facultad de Química e Ingeniería del Rosario (UCA). Para esto, se está trabajando en la elaboración y adaptación de los materiales didácticos para las materias “Informática” e “Informática General” del primer año de las carreras de nuestra facultad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO E INNOVACIÓN

El objetivo de este proyecto es mejorar los procesos de enseñanza y aprendizaje de las competencias involucradas en el Pensamiento Computacional, para que los futuros ingenieros estén más preparados para asumir los retos que demanda la sociedad contemporánea.

Se propone desarrollar una didáctica para la enseñanza del pensamiento computacional y la programación para estudiantes de las carreras de ingeniería y a su vez proponer metodologías activas y lúdicas para la utilización y entrega de los materiales en experiencias educativas presenciales y en línea.

Se trabaja en el desarrollo del pensamiento computacional, como una habilidad importante para la formación de toda persona en esta era digital.

Se comenzó con las primeras experiencias en las materias “Informática” e “Informática

General” que se dictan en el primer año de las carreras de Ingeniería Industrial, Ingeniería Ambiental, Licenciatura en Química y Licenciatura en Tecnología de los Alimentos. Se desarrolló material para estas asignaturas centrados en aquellos conocimientos principales del Pensamiento Computacional, como ser, la abstracción, la descomposición de problemas en subproblemas, el reconocimiento de patrones y algoritmos.

En las prácticas se utiliza la estrategia de aprendizaje basado en proyectos. Se busca motivar a los estudiantes de las diferentes carreras mediante el desarrollo de prototipos que les sean de utilidad en su profesión. La idea es que los estudiantes comprendan, a través del desarrollo de estos proyectos, la importancia que estas habilidades tienen para su vida profesional y personal futura. Se busca incentivarlos mediante el desarrollo de proyectos de su elección.

Los temas de los mismos son libres, pero se les pide a los alumnos que reflexionen sobre prototipos y/o herramientas que les puedan ser de utilidad en su labor futura como profesionales, y con eso en mente propongan un proyecto a desarrollar. La dificultad del mismo es evaluada por el docente, quien, teniendo en cuenta la propuesta de los estudiantes, brinda el acuerdo o adecúa la dificultad teniendo en mente los alcances de la materia.

Con este plan de innovación se busca que nuestros egresados puedan hacer uso de las tecnologías para resolver los problemas que se les presenten en su actividad profesional, para dar soluciones a una sociedad cada vez más exigente, que demanda respuestas cada vez más rápidas y eficientes.

3. RESULTADOS

En vista de estos objetivos planteados, se elaboró material didáctico utilizando este enfoque para las asignaturas “Informática” e “Informática General”. Durante el segundo

semestre del año 2020 se dictaron en forma virtual estas materias con esta perspectiva usando este nuevo material. Los resultados iniciales son promisorios [7].

A modo de ejemplo, podemos citar algunos de los proyectos desarrollados por los alumnos del primer año de la carrera de la Licenciatura en Química, durante el segundo semestre del año 2020. Nuestros estudiantes han desarrollado prototipos para: informar sobre las peligrosidades de reactivos en laboratorios, aceptar o rechazar medios de cultivos coloreados en función de su absorbancia, calcular pesos moleculares, realizar cálculos esquiométricos, y calcular energía en celdas electroquímicas, con el objetivo de optimizar la generación de energía en baterías.

Durante estas primeras experiencias se ha notado una mayor participación de los estudiantes en el proceso de aprendizaje así como también una mayor motivación por la programación. Creemos que esto se debe fundamentalmente a la propuesta de trabajo basado en proyectos de su interés profesional.

Podemos compartir algunos comentarios hechos por nuestros alumnos al finalizar el cursado, que dan cuenta de esto:

- “Nos encontramos, en primera instancia, frente al desafío de pensar en algún tipo de programa que nos ayude a llevar a cabo con menor dificultad alguna tarea relacionada con nuestra carrera. Una vez que decidimos cuál sería la tarea que el programa debería realizar, nos encontramos frente a otra dificultad: pensar en la estructura del mismo y armarlo. Creemos que estos contenidos nos servirán para tener una idea general, sobre qué es el pensamiento computacional o la algoritmia. Esto, nos será de mucha utilidad en un futuro, ya que cada vez, se van introduciendo más y más estos conceptos en el mundo del trabajo.”;
- “Creemos completamente que los contenidos dados sobre pensamiento

computacional y algoritmia nos van a ser útil al finalizar la carrera ya que nos plantea nuevas formas de ver los problemas y encontrarle soluciones más sencillas apoyándonos en la tecnología”;

- “Creemos y esperamos que nos ayude ya que es una nueva manera de pensar y ver las cosas en este mundo tan tecnológico, que va a seguir creciendo en cuanto a la tecnología y por lo tanto cada vez más vamos a tener que apoyarnos en este tipo de contenidos.”

Los trabajos realizados por los estudiantes han sido de calidad y mostraron la adquisición de los conceptos de pensamiento computacional.

Consideramos que estas primeras experiencias son muy positivas y seguiremos planteando nuevas estrategias para guiar y motivar a nuestros alumnos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por las doctoras Claudia Deco y Pamela Viale, la magister Cristina Bender y el Licenciado Sebastián Velázquez, todos investigadores de la Universidad Católica Argentina.

Dentro del marco de esta línea de I+D, el Licenciado Velázquez cursa actualmente una maestría en Educación.

Los investigadores de este grupo realizan constantemente esfuerzos para volcar los conocimientos adquiridos en las carreras que se dictan en nuestra universidad a través de las actividades docentes.

La viabilidad del proyecto está sustentada en los avances obtenidos mediante trabajos previos de los integrantes en el área y su interacción con otros grupos de investigadores latinoamericanos [8].

5. REFERENCIAS

- [1] J. Wing (2014). Computational Thinking Benefits Society. Social Issues in Computing.
- [2] K. Brennan, M. Resnick (2012). New frameworks for studying and assessing the development of computational thinking. In Proceedings of the 2012 annual meeting of the american educational research association (Vancouver, Canada).
- [3] M. Román-González, J. C. Perez-González, C. Jiménez-Fernández (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior*, 72: 678-691.
- [4] M. García, C. Deco, C. Bender, C. Collazos (2017). Invited paper: Robotics Based Strategies to Support Computational Thinking: The Case of the Pascual Bravo Industrial Technical Institute. *Journal of Computer Science and Technology (JCS&T)* Vol. 17 No. 1. April 2017. pp 59-67.
- [5] M. García, C. Deco, C. Bender, C. Collazos (2017). Herramientas de Diseño para el Desarrollo de Competencias en Educación Básica, Media y Tecnológica: Experiencia en el Instituto Técnico Industrial Pascual Bravo de Colombia. *Revista TE&ET (Argentina)*. Mayo 2017.
- [6] A. Casali, C. Deco, P. Viale, C. Bender, D. Zanarini, N. Monjelat (2020). Enseñanza y aprendizaje del pensamiento computacional y la programación en los distintos niveles educativos. Mayo 2020 - XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz). pp 595-599.
- [7] P. Viale, C. Deco, C. Bender (2020). Introduciendo conocimientos sobre el Pensamiento Computacional en los primeros años de las carreras de Ingeniería. In Proceedings del XIII Congreso Argentino Ingeniería Industrial (COINI). Organizado por Asociación Argentina de Carreras de Ingeniería Industrial y Afines (AACINI). Virtual. 19 al 24 Octubre 2020.
- [8] Y. Rodríguez del Rey, I. Cawanga Cambinda, C. Deco, C. Bender, R. Avello-Martínez, K. Villalba-Condori (2020). Developing Computational Thinking with a Module of Solved Problems. In *Computer Applications in Engineering Education*. March 2020.

HERRAMIENTAS Y TECNOLOGÍAS PARA EL DESARROLLO WEB DESDE EL FRONTEND AL BACKEND

Susana Graciela Pérez Ibarra, José Rolando Quispe, Felipe Fernando Mullicundo,
Daniel Alberto Lamas

Facultad de Ingeniería - Universidad Nacional de Jujuy
gperezunju@gmail.com
josexeq@gmail.com

RESUMEN

El desarrollo de aplicaciones web se conforma de la capa de presentación (FrontEnd) y una capa de acceso a datos (BackEnd).

El FrontEnd trabaja la interfaz visual, y hace que el usuario pueda interactuar con nuestro sitio o sistema. Está orientado al lenguaje de marcas y al lenguaje de programación web de ejecución en equipos clientes.

El Backend se encarga de la manipulación de los datos, un BackEnd no sirve de mucho si no existe un FrontEnd de por medio, el desarrollador BackEnd debe de conocer de bases de datos, frameworks y aspectos de seguridad. Él debe encargarse de que la información que llega desde el FrontEnd, sea almacenada a una base de datos. Así mismo se encarga de crear API's para que sus datos puedan consumirse de manera cómoda y pueda mejorar la experiencia del usuario [1].

El avance tecnológico obliga al profesional de la informática, dedicado al desarrollo web, a conocer estos términos y las tecnologías relacionadas.

El presente trabajo pretende centrarse en el estudio y análisis de estas tecnologías, con lo cual se pretende determinar cuáles son las más usadas y relevantes en el ámbito laboral, y en razón de ello, fomentar o motivar la incorporación de sus conocimientos en la formación de los alumnos que están cursando los últimos años de las carreras de informática.

Palabras clave: Tecnologías web, FrontEnd, BackEnd, cliente, educación

CONTEXTO

La presente investigación corresponde a la “Tercera Convocatoria – Proyectos de Investigación orientados a la Investigación Básica y a la Investigación Aplicada en el área de Informática de la Facultad de Ingeniería” el cual fue aprobado por la Facultad de Ingeniería de la Universidad Nacional de Jujuy.

1. INTRODUCCIÓN

En el ámbito del desarrollo web, el FrontEnd está conformado por todas aquellas tecnologías que corren del lado del cliente, es decir, todas aquellas tecnologías que corren del lado del navegador web, generalizándose más que nada en tres lenguajes, Html, CSS y JavaScript. La persona encargada del FrontEnd, trabaja con estas tres tecnologías, aunque esto no significa que desconozca como trabaja el BackEnd (lado servidor), ya que es necesario para que pueda consumir datos y pueda estructurar correctamente un maquetado en HTML y CSS para su mejor comodidad.

¿Qué es el FrontEnd?

FrontEnd se encarga de estilizar la página de tal manera que la misma pueda presentar la

información de forma agradable para el usuario [2]. El responsable del FrontEnd, debe de conocer las técnicas de experiencia de usuario para brindar una mejor interacción entre la persona y la página que visita, así mismo debe tener conocimientos de diseño de Interacción para colocar los elementos de tal manera que el usuario las pueda ubicar de forma rápida y cómoda.

Existen muchas tecnologías relacionadas a estos tres lenguajes que deben ser conocidas por el desarrollador FrontEnd. Por ejemplo, para JavaScript existen Angular y BackboneJS, los cuales se pueden apoyar en librerías como AnimateCSS, JQuery y otras. También existen lenguajes de transferencia de información como XML, JSON y Ajax para hacer solicitudes al servidor sin necesidad de refrescar la página completa.

¿Qué es el BackEnd?

Se denomina BackEnd a la capa de acceso a los datos de un software que no es accesible para el usuario final. Además, esta capa contiene toda la lógica de la aplicación que maneja los datos [3]. Cabe destacar que los datos de una aplicación se encuentran almacenados en una base de datos dentro de un servidor.

El responsable del BackEnd es la persona que trabaja del lado del servidor y debe formarse como desarrollador de aplicaciones web o como desarrollador de aplicaciones multiplataforma. Debe estudiar los diferentes lenguajes de programación que son necesarios para desarrollar su trabajo y según la empresa en la que trabaje serán necesarios unos u otros. Además, necesita conocer las interacciones con diferentes bases de datos, saber las diferencias entre estas y cualidades de las más utilizadas [4]. Esto no significa que un BackEnd deba desconocer por completo el trabajo realizado por un FrontEnd sino que debe poseer los conocimientos necesarios para trabajar en equipo ya que ambos se complementan.

La persona encargada del BackEnd debe tener conocimientos, según el lugar donde trabaje, de los lenguajes del lado del servidor,

como ser: Java, C#, PHP, Node.JS, entre otros. Además, de aquellos que interactúan con la base de datos, como ser: MySQL, PostgreSQL, SQLServer, MongoDB, entre otras [5].

2. JUSTIFICACION DEL PROYECTO

El interés en realizar la investigación propuesta surgió al observar que los recientes graduados universitarios no poseían algunos de los conocimientos que las empresas de desarrollo de software esperan de ellos hoy en día.

Debido a los grandes avances en la informática, muchas de las tecnologías usadas en la actualidad hicieron su aparición o tuvieron mayor demanda en los últimos 2 o 3 años. Por este motivo, muchos de los graduados se encuentran con obstáculos al momento de iniciarse en el ámbito laboral debido a que carecen de algunos conocimientos que buscan las empresas de software en un profesional en la actualidad.

Con la realización de este proyecto se pretende que los futuros graduados conozcan cuales son las tecnologías vigentes en el ámbito laboral y se motive en ellos la experimentación, aplicación y uso de tales tecnologías mientras transitan el cursado de sus carreras. También se pretende que los conocimientos adquiridos en este proyecto puedan ser incorporados en las materias en las que se desempeñan los docentes.

3. OBJETIVOS GENERALES Y PARTICULARES

Objetivos Generales

- Analizar y estudiar las tecnologías usadas en la actualidad tanto en el FrontEnd como en el BackEnd de un desarrollo web y que las empresas de software requieren como parte de la formación de un profesional de la informática.

- Elaborar un documento informativo sobre las tecnologías investigadas, que entenderíamos como necesarias en la formación del futuro profesional, para los directores de las carreras de informática de la Facultad de Ingeniería.
- Promover la actualización de los contenidos curriculares de las asignaturas de las carreras de informática de la Facultad de Ingeniería.

Objetivos Particulares

- Relevante información acerca de las tecnologías actuales referidas al FrontEnd y BackEnd de un proyecto de desarrollo web.
- Fortalecer los procesos de enseñanza y aprendizaje de las materias de los últimos años y/o en aquellas en las que los docentes de este proyecto se desempeñan, siendo en ambos casos, materias de las carreras de informática de la Facultad de Ingeniería.
- Promover la incorporación de los conocimientos de las tecnologías investigadas en las curriculas de las asignaturas pertinentes.
- Fomentar la realización de talleres o jornadas de actualización para alumnos y egresados sobre el uso de las tecnologías investigadas.
- Socializar los resultados de este proyecto en congresos de informática.
- Establecer vínculos con las empresas de software factory de la provincia a fin de favorecer la inserción laboral de los graduados de la Facultad de Ingeniería en tales empresas.

4. RESULTADOS OBTENIDOS Y ESPERADOS

Documento informativo sobre las tecnologías BackEnd y FrontEnd actuales que se entiende como necesarias en la formación del futuro profesional para los directores de las carreras de informática de la Facultad de Ingeniería.

Transferencia de los conocimientos adquiridos a los alumnos y docentes de las carreras de Licenciatura en Sistemas e Ingeniería Informática de la Facultad de Ingeniería de la Universidad Nacional de Jujuy.

5. FORMACIÓN DE RECURSOS HUMANOS

El proyecto está siendo desarrollado por un equipo conformado por docentes investigadores del Grupo de Investigación y Desarrollo en Ingeniería de Software (GIDIS) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy. La estructura del equipo de investigación es la siguiente:

- Directora: Ing. Susana Graciela Pérez Ibarra.
- Codirector: Ing. José Rolando Quispe.

Investigadores:

- Lic. Felipe Fernando Mullicundo. Categoría de Investigación V.

Participa del proyecto un alumno avanzado de la carrera de Ingeniería Informática:

- Daniel Alberto Lamas

Con la realización de este proyecto de investigación se espera la consolidación de los miembros del grupo en especial del alumno como joven investigador. Además, el proyecto brindará un marco propicio para la iniciación de trabajos finales de grado de la carrera Ingeniería Informática.

6. BIBLIOGRAFÍA

- [1] Campus MVP. “Desarrollador web: Front-end, back-end y full stack. ¿Quién es quién?”. Disponible en <https://www.campusmv.es/recursos/post/Desarrollador-web-Front-end-back-end-y-full-stack-Quien-es-quien.aspx>
- [2] SerProgramadores. “¿Que es FrontEnd Y Backend en la programación web?”. Disponible en <https://serprogramador.es/que-es-frontend-y-backend-en-la-programacion-web/>

[3] Platzi. “Qué es Frontend y Backend”. Disponible en <https://platzi.com/blog/que-es-frontend-y-backend/>

[4] Fundació Jesuïtes Educació. “¿Qué es un Backend Developer?”. UOC Universitat Oberta de Catalunya. Julio de 2018. Disponible en: <http://fp.uoc.edu/blog/que-es-un-backend-developer/>

[5] Freelancer. “¿Qué es un programador backend?”. Octubre 2016. Disponible en: <https://www.freelancer.com.ar/community/articulos/que-backend>

INNOVACIÓN EDUCATIVA CON TECNOLOGÍAS EMERGENTES EN EL CONTEXTO DE LAS PRÁCTICAS EDUCATIVAS ABIERTAS

Chiarani Marcela C - Gacia Berta E.- Noriega, Jaquelina E. – Allendes Paola A,– Daza Monica M.– Sosa, Alejandra B.- Aguirre, Jesús F. – Torres, Silvia V. – Gómez, Verónica I. – Viano, Hugo J. – Zangla Urteaga, María S. – Gomez Cintia L.

Departamento de Informática -Facultad de Ciencias Físico Matemáticas y Naturales-UNSL

mcchi@unsl.edu.ar , bertae.garcia@gmail.com , jenoriega@unsl.edu.ar , allendes.paola@gmail.com , sosaab@email.unsl.edu.ar , aguirre.jesus.francisco@gmail.com , svtorres@unvime.edu.ar , vigomez@email.unsl.edu.ar , hviano@gmail.com , soledad.zangla@gmail.com , cintialorenagomez@gmail.com

RESUMEN

Las Tecnologías de la Información y la Comunicación (TIC) modifican las prácticas sociales e influyen en diferentes ámbitos de la vida cotidiana. En la actual Sociedad del Conocimiento, éste se entiende subjetivo, dinámico y producido colaborativamente. La conectividad se constituye en un aspecto clave: permite a los individuos crear y distribuir sus propios materiales, incorporar nuevas ideas, tecnologías y tendencias de uso, permitiendo a los docentes adoptar modelos y metodologías innovadoras. La integración de la tecnología puede promover cambios en el sistema educativo al impactar en las prácticas. Los Recursos Educativos Abiertos (REA), surgen como iniciativa para compartir materiales digitales disponibles en la web de forma libre y gratuita, utilizables en el ámbito académico y de investigación de todo el mundo. A partir de los REA surgen las Prácticas Educativas Abiertas (PEA), las que apoyan la producción, utilización y reutilización de REA de alta calidad a través de políticas institucionales y modelos pedagógicos innovadores. Con la adopción de las nuevas tecnologías se requiere el desarrollo de nuevas teorías, métodos, enfoques de enseñanza, evaluación y organización. Las relaciones entre las pedagogías emergentes y las tecnologías emergentes producen un doble vínculo: algunas tecnologías desarrolladas sin una finalidad educativa específica se usan con fines

educativos y las prácticas pedagógicas se modifican transformando las prestaciones tecnológicas. Las tecnologías emergentes resultan mediadoras de los procesos de cambio y la innovación educativa es el camino para avanzar hacia la calidad. Este proyecto tiene por objetivo: Analizar, diseñar e implementar Modelos Pedagógicos y Estrategias Didácticas para el desarrollo de PEA, aplicado a prácticas educativas innovadoras con tecnologías emergentes, mediante la metodología propia de la Investigación-Acción y con el objetivo amplio de explorar estrategias de mejora para el sistema educativo y social. EL grupo de trabajo se plantea un doble propósito: el de investigación, para generar conocimiento y comprensión y el de acción, para potenciar la reflexión de las prácticas educativas. Se espera que docentes de distintos niveles educativos experimenten las PEA como diseños instruccionales y como objetos mediadores en la apropiación del conocimiento, para impactar en los métodos y estrategias de enseñanza con vistas a que esas prácticas puedan ser replicadas.

Palabras clave: Innovación Educativa - Tecnologías Emergentes - Prácticas Educativas Abiertas

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Innovación educativa con Tecnologías Emergentes en el contexto de las

Prácticas Educativas Abiertas, del Departamento de Informática de la Facultad de Cs. Físico Matemáticas y Naturales; Universidad Nacional de San Luis, PROICO N° 3-0420. Dicho proyecto es la continuación de diferentes proyectos de investigación de esta institución educativa a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e instituciones educativas del nivel medio. Además, cuenta con el reconocimiento del programa de incentivos.

1. INTRODUCCIÓN

Las TIC han modificado y transformado los espacios públicos y las prácticas sociales en la educación. Diferentes ámbitos de la vida cotidiana están influidos por esta transformación.

Sin dudas el conocimiento ocupa un lugar preponderante en la sociedad actual, llamada Sociedad del Conocimiento (SC). Begoña Gros [1] señala que una de las consecuencias más importantes de la Sociedad del Conocimiento es la transformación de los espacios y lugares para el aprendizaje. Cambia, además, la expectativa de relevancia y validez del conocimiento. Lo que en el siglo XX se medía en ciclos de años y décadas, se ha reducido a otros de meses y años [2]. Por otra parte, Begoña Gros afirma que ha cambiado su forma de producción: “hemos pasado de pensar en el conocimiento como algo objetivo, estable, producido por expertos y que se puede transmitir, a algo subjetivo, dinámico y producido de forma colaborativa”. “Por ello, está marcado por la disposición de poner en cuestión las percepciones, las expectativas tradicionales y socialmente aceptadas. Las reglas y evidencias de nuestra sociedad están cada vez más sometidas a procesos de reflexión [1].

El movimiento de los Recursos Educativos Abiertos [3], y [4] se entiende como la iniciativa de compartir materiales digitales disponibles en la web de forma libre y gratuita, para ser utilizados tanto en el ámbito académico como de investigación de todo el mundo.

El aporte significativo de los REA es su potencial en la reutilización, dado que aumenta la productividad de los sistemas de gestión de aprendizaje y disminuye los costos. Cabe destacar que la construcción de estos recursos digitales es compleja y su calidad puede medirse según nos centramos en el producto o en el proceso, en lo técnico o pedagógico. La calidad está relacionada con la reusabilidad y es posible evaluarla a partir de la aplicación de métricas de software. Establecer si están elaborados adecuadamente nos permitirá mejorar los procesos de creación de los mismos.

Es a partir de los REA que emergen las PEA. Diego Leal Fonseca [5] entiende que son dos las definiciones que se encuentran acerca de las PEA. En resumen, podríamos expresar que son prácticas educativas que apoyan la producción, utilización y reutilización de REA de alta calidad a través de políticas institucionales, que promueven modelos pedagógicos innovadores y el respeto y la autonomía de los alumnos, como co-productores en su camino hacia el aprendizaje permanente.

Comprender qué son las PEA contribuye a alternativas que fortalecen la educación y brindan la posibilidad de expandir los conocimientos de una manera accesible. En un modelo educativo interconectado las PEA tienen el potencial de ofrecer oportunidades y experiencias de aprendizaje dentro de una amplia variedad de contextos tan diversos como sea necesario. Se han convertido en una

tendencia educativa creciente basada en el uso de TIC [6].

En este sentido, las tecnologías emergentes resultan mediadoras de estos procesos de cambio, al permitir espacios de interacción y reflexión desde el acceso, la producción y el intercambio de contenidos. Esta visión entiende que el docente, además de acceder al conocimiento disponible, se constituye en autor y productor desde su propio contexto de acción. Así, la innovación educativa es el camino para avanzar hacia la calidad, y las tecnologías digitales son la herramienta para movilizar acciones que lo hagan posible partiendo de la acción y la reflexión en su propia práctica, desde el trabajo con otros.

En función de lo precedente este proyecto tiene por objetivo analizar, diseñar e implementar Modelos Pedagógicos y Estrategias Didácticas para el desarrollo de PEA, aplicado a prácticas educativas innovadoras con tecnologías emergentes.

2. RESULTADOS OBTENIDOS/ESPERADOS

En cuanto a los resultados que se pretende lograr y la contribución del proyecto, este grupo de investigación se plantea un doble propósito: el de acción, para potenciar la reflexión de las prácticas educativas; y el de investigación, para generar conocimiento y comprensión.

Dado que, el diseño de investigación elegido responde a la metodología propia de la Investigación-Acción, se propone intensificar la promoción de la apropiación de las PEA. A partir de los talleres de sensibilización-acción se trabajará con las instituciones educativas y

los docentes interesados en participar de la investigación.

Esta metodología requiere un nivel elevado de participación de los docentes, de manera que lejos de ser meros informantes, éstos se benefician de la construcción conjunta del conocimiento, así como de la incorporación de técnicas de recolección y análisis de datos en su propia labor de investigación, como co-investigadores en el aula.

3. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Las principales líneas de trabajo abordadas son:

Modelos Pedagógicos y Estrategias Didácticas para el desarrollo de PEA.

Las prácticas educativas innovadoras con tecnologías Emergentes a partir de la reflexión sobre la propia práctica.

La difusión de las PEA a fin de promover el intercambio y discusión acerca de las prácticas educativas innovadoras con tecnologías emergentes diseñadas y desarrolladas a partir de la investigación-acción.

Herramientas de software libre que permitan desarrollar REA y dar soporte a las PEA.

Las Plataformas de Software Libre que dan soporte tecnológico a las innovaciones adoptadas en los diferentes ciclos del proyecto de investigación.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del proyecto de investigación, se están llevando a cabo, por un lado, pasantías de investigación con docentes de instituciones educativas de nivel medio y universitario. Por el otro, varias tesis de posgrado,

correspondientes a Maestrías que se dictan en la UNPA, UNC, UNSL y en especial en la Maestría en Enseñanza en escenarios digitales, interinstitucional con modalidad a distancia. Ofrecida en el marco de la Asociación de Universidades Sur Andina.

5. REFERENCIAS

- [1] GROS, B. (2015). La caída de los muros del conocimiento en la sociedad digital y las pedagogías emergentes. *Education In The Knowledge Society (EKS)*, 16(1), 58-68. doi:10.14201/eks20151615868
- [2] Siemens, G. (2006). *Connectivism: Learning theory or pastime of the self-amused*. Manitoba, Canada: Learning Technologies Centre.
- [3] OECD (2010) El conocimiento libre y los recursos educativos abiertos. Serie Sociedad de la Información. JUNTA DE EXTREMADURA. I.S.B.N.-13:978-84-691- 8082-2.
- [4] Ramírez Montoya Maria Soledad, Burgos Aguilar José Vladimir (2010) Coordinadores. *Recursos Educativos Abiertos en Ambientes Enriquecidos con Tecnología: Innovación en la Práctica Educativa*. México, -Primera edición Edición impresa: ISBN 978-607-501-018-2 Edición electrónica: e-ISBN 978-607-501-022-9
- [5] Diego Leal Fonseca. *Prácticas educativas abiertas* <https://reaprender.org/openep/practicas-educativas-abiertas/>
- [6] CHIAPPE, Andrés ; ADAME, Silvia Irene. (2018) *Prácticas Educativas Abiertas: una forma de aprender más allá del conocimiento de acceso libre*. Ensayo:aval.pol.públ.Educ. [online]. vol.26, n.98, pp.213-230. Epub Dec 18, 2017. ISSN 0104-4036.
- [7] Bailón, M; Rabajoli, G. (2014) El desafío de las prácticas educativas abiertas (PEA). Congreso Iberoamericano de Ciencia, Tecnología Educación e Innovación. Buenos Aires, Argentina, noviembre de 2014.
- [8] Castillo, P. y De Benito, B. (2008). *Cambios, novedades y procesos de innovación*.
- [9] En Salinas, J. (Ed.), *Innovación Revista Virtual Universidad Católica del Norte*, Año 2015. Número 45, Mayo-Agosto 51 Educativa y uso de las TIC (pp.31-41). España: Universidad Internacional de Andalucía. Recuperado de http://dspace.unia.es/bitstream/handle/10334/130/004tic_cabero.pdf?sequence=1
- [10] Chiarani, M. (2014). Recursos educativos abiertos: oportunidades y retos en el ámbito universitario. *Argonautas*, 4(4), 112-123. Recuperado de <http://www.argonautas.unsl.edu.ar/files/09%20CHIARANI%20MARCELA.pdf>
- [11] Chiappe Andrés, (2012) *Prácticas Educativas Abiertas como factor de Innovación Educativa con TIC*. Boletín Virtual REDIPE No 818 Noviembre 12 de 2012 – ISSN 2256-1536
- [12] Chiappe Laverde, Andrés; Herrera Arias, Mónica (2015) *Producción abierta de un contenido educativo digital: Un estudio de caso Opción*, vol. 31, núm. 4, 2015, pp. 312-327 Universidad del Zulia Maracaibo, Venezuela ISSN: 1012-1587

La articulación integrada entre academia, investigación y extensión como una experiencia de innovación en educación en Informática para los Cursos de Ingreso de la UNSJ en el marco de SIED UNSJ

Mg. Carmona Emilse, Esp. Ana Dominguez, Mg. Pósito Rosa. Ing. Pickelny, Dr. Mario Gutierrez, Mg. Buteler Luciana, Esp. Canto Isabel, Lic. Balmaceda Silvina, Ing. Sebastián Godoy

Área de Educación a Distancia - SIED UNSJ – RED EDUSyT Educación Subjetividades y Tecnología de la Facultad de Filosofía, Humanidades y Artes – UNSJ – IDICE- FFHA – Referentes en SIED UNSJ de las facultades FCEFyN, FACSO, FFHA, FI

Av. Ignacio de la Roza 590 (O). Complejo Universitario "Islas Malvinas".

Rivadavia, San Juan, Teléfonos: 4234129, Fax 0264-4234980

emilsecoaching@gmail.com, anapato.domin@gmail.com, mpickelny@yahoo.com,
marioluisgutierrez@gmail.com, butelerandalue@gmail.com, isabelcantolopez@gmail.com,
balmaceda.silvina@gmail.com, godoysebastianalejandro@gmail.com

Resumen

En el marco de SIED UNSJ, coordinado por Mg. Rosa Pósito, surgen dos capacitaciones: Curso Taller de Estudiantes Tutores /Tutores Guías/ Ayudantes y Curso Taller Docentes de Ingreso para toda la UNSJ, dictadas secuencialmente a niveles de tiempos, donde el resultado obtenido en la primera capacitación es usado como insumo para la segunda capacitación, quienes evaluaron lo realizado en la capacitación anterior.

Estas actividades de capacitación académica tienen su innovación y sustento teórico en la articulación de distintos proyectos de investigación que se proponen colaborar en especial con el momento en que los estudiantes aspirantes a ingresar a la UNSJ están decidiendo su carrera por medio de los

Cursos de Ingreso que brindan las distintas Facultades y Escuela de Ciencias de la Salud. El diseño innovador se presenta en el enfoque de meta-modelo pedagógico elegido de aprendizaje inverso/invertido/aula inversa on-line en articulación con el diseño de aula tradicional en un entorno Moodle, de SIED UNSJ mostrando claramente la complementariedad de ambos modelos y en el equipo interdisciplinario de Profesores e Investigadores de UNSJ, de las distintas unidades académicas.

Palabras clave: Educación a distancia. Sistema Institucional de Educación a Distancia. SIED UNSJ. Docentes Ingreso. Tutores Ingreso. Aprendizaje inverso/Aula inversa on-line. Interdisciplinariedad.

Contexto

Este trabajo se enmarca en la experiencia de autoría, diseño e implementación de dos capacitaciones: Curso Taller de Estudiantes Tutores /Tutores Guías/ Ayudantes y Curso Taller Docentes de Ingreso para toda la UNSJ, donde intervienen con sustento teórico los proyectos de investigación financiados por la Universidad Nacional de San Juan. Focalizadas ambas en la temática Uso de la herramienta Moodle. Los proyectos de investigación intervinientes son: Proyecto Marco de Referencia para la formulación, gestión e implementación de carreras de pregrado, grado y posgrados en la opción pedagógica a distancia según el SIED UNSJ de la FCEfyN-UNSJ Directora Mg. Rosa Pósito, co director Esp. Ing. Leiva Alfredo. Participa coordinando y gestionando estas propuestas de capacitación la directora Mg. Rosa Pósito y coordinadora general de SIED UNSJ. En articulación con los proyectos: El compromiso de la universidad con la promoción de la innovación en la formación de profesionales universitarios de la FI UNSJ. Directora: Mg. Martínez Dávila, Liliana, codirectora Mg. Carestia, Nirva Ana. Participa en esta propuesta como integrante del proyecto Prof. Esp. Ana Dominguez. Junto al Proyecto Propuestas de enseñanza innovadora con aprendizaje invertido, proyectos inclusivos, retos, gamificación, ubicuidad, en Web 4.0: Fortaleciendo la red EDUSyT (Educación, Subjetividades y

Tecnología) articulando con contextos educacionales de Secundaria Rural FFHA-UNSJ, aplicados a estudiantes, directora Mg. Emilse Carmona, co directora Mg. Silvia Dipp. Participa en esta propuesta la directora Mg. Emilse Carmona y una integrante del equipo Mg. Luciana Buteler. También Proyecto Tecnologías de la Información y comunicación y resiliencia educacional . Nuevas didácticas para nuevos escenarios. Directora Tamara Abigail Bitar. Co director: Mario Gutiérrez. Participa en esta propuesta el codirector Dr. Mario Gutierrez. Por último articulando con distintos referentes de SIED UNSJ y sus componentes Area Tecnológica Esp. Ana Dominguez; area Pedagógica Comunicacional Mg. Emilse Carmona, Esp. Isabel Canto, Mg. Luciana Buteler, con referentes de FACSO Ing. Marcelo Pickelny, y Dr. Mario Gutierrez, area Diseño de Materiales Ing. Sebastián Godoy, Lic. Silvina Balmaceda y coordinadora general de SIED UNSJ Mg. Rosa Pósito.

Introducción

El diseño innovador se presenta por el equipo interdisciplinario que participa en el mismo, e implícitamente en el enfoque del meta- modelo pedagógico elegido de aprendizaje inverso/invertido/aula inversa on-line en articulación con el diseño de aula tradicional en un entorno Moodle, mostrando claramente la complementariedad de ambos modelos para el logro de los objetivos “aprender a crear, diseñar y evaluar en un aula virtual creada en la plataforma Moodle de SIED UNSJ; diseño innovador pedagógico que permite el aprendizaje ubicuo,

colaborativo e inclusivo, e implica un proceso mental de anticipación, ya que se trabaja sobre la representación de aquello que se busca promover que suceda.

La innovación es a niveles pedagógicos en esta experiencia podemos caracterizarla de la siguiente manera:

1- Se interrelacionan dos enfoques a nivel de planificación de la estructura del aula, conformando una planificación híbrida: Siguiendo diseño tradicional con el enfoque al interno de Aprendizaje inverso/aula inversa on-line (Curso Taller Estudiantes Tutores/Tutores guías/ Ayudantes) y el enfoque de Aprendizaje inverso/aula inversa on line (Curso Taller Docentes de ingreso UNSJ) según planificación de la red Flipped Learning Global. El modelo híbrido se produce por la conjunción de dos modelos pedagógicos que si bien la primera tiene un enfoque tradicional por temática al abordar en su interior las mismas desde el enfoque de meta-modelo Aprendizaje inverso/Aula inversa on-line, entregando con anticipación los materiales tutoriales, y destacando la necesidad del espacio individual del estudiante, como el espacio grupal colaborativo. En la segunda capacitación se estructura directamente desde el enfoque del meta-modelo de Aprendizaje Inverso/Aula inversa on-line proporcionado por la estructura basada en la red Flipped Learning Global, a través del conocimiento impartido por Dr. Raúl Santiago, que junto a Jon Bergmann y a otros expertos de la red internacional, generosamente, ante la Pandemia mundial, han querido donar parte de su

tiempo para organizar un mini curso introductorio que permita pensar el Aprendizaje inverso/Invertido/Aula inversa en e-learning. En febrero/marzo 2021 los Tutores, Tutores guías y Docentes están dictando y acompañando a los estudiantes de Ingreso de la UNSJ de todas las facultades.

2- Un aula virtual donde se desarrollan dos cursos taller: El aula virtual se estructuró en dos grandes bloques: Taller de Estudiantes Tutores/Tutores Guías/Ayudantes y bloque de Curso Taller Docentes Ingresantes. El primer bloque, con planificación de enfoque tradicional por temática disciplinar a abordar combinado con Aprendizaje inverso/Aula inversa on-line conformado por cuatro pestañas: Presentación/ Foro/ Tarea/ Cuestionario. Al interno de Presentación se realiza un espacio de Bienvenida y se configura la etiqueta Consultas y Comunicación, con Sala de videoconferencias Big blue botton, chat, foro de consultas, hoja de ruta con la presentación del taller. En las Pestañas Foro/Tarea/Cuestionario se estructuraron las temáticas planteadas por etiquetas y estrategias de aprendizaje aplicando el enfoque del meta-modelo Aprendizaje Inverso/Aula Inversa on-line, precisando los dos espacios que el estudiante debe vivenciar llamando a las etiquetas Actividad Individual / Actividad Grupal Colaborativa.

En el otro bloque Curso Taller para Docentes se estructuró con planificación de enfoque meta-modelo Aprendizaje Inverso/Aula inversa on-line,

estableciendo las pestañas Antes (Antes de la clase en videoconferencia, totalmente asincrónico) / Durante (espacio de presencia en la no presencialidad en sincronía combinado con interacciones asincrónicas/ Después (después de sincronía totalmente asincrónico) / Evaluación (con tres tipos de evaluación: autoevaluación, co-evaluación y siempre evaluación formativa)/ Conclusión En este espacio se coloca un pensamiento que permita darle cierre y a la vez plantear interrogantes que generen el deseo de seguir profundizando en la temática planteada como puentes a nuevos conocimientos.

3. Las actividades planificadas para la capacitación se articulan con las teorías del aprender hacer y en la innovación de enseñar a diferentes áreas disciplinares.

“Se trata entonces de aprender a hacer, y bien, lo que nunca se hizo a gran escala en nuestro continente: una escuela cuya finalidad central sea el aprendizaje de todos; una pedagogía capaz de inspirar y reunir a los profesores en torno a un proyecto coherente con objetivos claros; una didáctica inspirada en concepciones de aprendizaje más contemporáneas en el cual el objeto de conocimiento y el objeto de enseñanza finalmente se reconcilien” (Namo de Mello, 2005: 26, en GUTIÉRREZ, M.; PIEGAIA, L. , 2018: 323).

3. Esta capacitación brinda la posibilidad de dar a conocer el uso de Moodle a 171 destinatarios en un mismo aula virtual,

(entre Estudiantes Tutores/Tutores guías/ Ayudantes y Docentes de ingreso de toda la Universidad Nacional de San Juan).

Se presenta el rol del Tutor Coach en aulas virtuales, como una nueva manera de acompañar a los estudiantes en el ingreso a la UNSJ, fortaleciendo no solo lo disciplinar sino también la emocionalidad en épocas de Pandemia, cuando fuera necesario, se les enseña la importancia de la tutoría coach individual y grupal o en equipos.

4. Articulación entre capacitaciones y proyectos de investigación: El resultado obtenido por la capacitación Estudiantes Tutores/Tutores guías/Ayudantes es utilizado por la capacitación de Docentes de Ingreso para realizar las prácticas de evaluación de los mismos, utilizando como insumo lo trabajado en la primer capacitación. Se produce una articulación que permite la sinergia entre ambas capacitaciones desde una planificación de aula virtual que llamamos híbrida y también la articulación de distintas líneas de investigación para impactar en la comunidad sanjuanina a través de la extensión por medio de Cursos de Ingreso de las distintas facultades y EUCS para los estudiantes que desean ingresar a UNSJ:.

5. Destacar el equipo interdisciplinario de investigadores y profesores de distintas facultades que han intervenido como formadores en ambas capacitaciones articulando SIED UNSJ, a través de su coordinadora general de SIED UNSJ Mg. Rosa Pósito, con representantes de la red

EDUSyT (Educación, Subjetividades y Tecnología) de la facultad FFHA como coordinadora Mg. Emilse Carmona e integrante Mg. Luciana Buteler fortaleciendo dicha articulación al Curso Taller Estudiantes Tutores, tutores guías, ayudantes, coordinado por Prof. Esp. Ana Dominguez, en forma conjunta con la capacitación del Curso Taller Docentes de Ingreso de UNSJ coordinado por Mg. Emilse Carmona. Ambas capacitaciones se pudieron llevar a cabo junto al equipo de docentes interdisciplinario conformado por Representantes del equipo base de SIED UNSJ Ing. Ana Dominguez, Esp. Isa Canto, Lic. Silvina Balmaceda, Representantes del equipo base de FACSO en SIED UNSJ Dr. Mario Gutierrez, Ing. Marcelo Pickelny. Representantes del equipo base de FFHA en SIED UNSJ Mg. Emilse Carmona, Mg. Luciana Buteler, Representante del equipo base de FI en SIED UNSJ Ing. Sebastián Godoy.

Los objetivos teóricos y experimentales

Podemos destacar de cada uno de los proyectos de investigación los siguientes objetivos teóricos experimentales:

Proyecto a) Generar mecanismos operativos que faciliten la articulación de las Unidades Académicas con el área EAD de la Universidad.

Proyecto b) Fortalecer las acciones de seguimiento y tutoría a los ingresantes a través de la capacitación a los tutores guías (estudiantes de los años superiores de cada carrera)

Proyecto c) Fortalecer la red EDUSyT (Educación, subjetividades y Tecnología) en el ámbito de la FFHA UNSJ. Gestar

comunidades académicas y de investigación entre profesores de la FFHA y UNSJ que intercambien experiencias relacionadas con la temática del proyecto.

Proyecto d) Analizar algunos de los tópicos fundamentales del Principio Aprender a Hacer y de la Resiliencia Educativa rescatando la perspectiva de los/as participantes (estudiantes, docentes e investigadores/as).

La contribución en la formación de Recursos Humanos

El estudio de los informes de la UNESCO permite conjugar varios componentes del sistema educativo ya que son instancias indicadoras del proceso de enseñanza-aprendizaje como de las diferentes modalidades de actuación de los profesores y estudiantes. Éstos se han transformado en un ámbito de interés porque se constituyen como actores relevantes y protagónicos en los trayectos desplegados en los sistemas de educación universitarios. Desde el Sistema Institucional de Educación a Distancia (SIED) se brindan espacios de formación para adquirir destrezas y habilidades en el uso de las herramientas que provee el sistema Moodle (PICKELNY M. 2021).

Aprendizaje inverso/Aula inversa evoluciona con la pandemia a Aprendizaje inverso/Aula inversa on-line. Es durante este proceso de planificación híbrida, entonces, que las docentes coordinadoras se encuentran junto con sus equipos interdisciplinarios, frente a la necesidad de tomar decisiones sobre los diferentes aspectos involucrados en la enseñanza de los contenidos de las capacitaciones al pensar la estructura

según un enfoque elegido: Hibridación en el diseño del aula virtual, articulación entre distintos modelos pedagógicos Tradicional y Aprendizaje Inverso/Aula Inversa on line centrado en el estudiante o docente participante; metodologías activas que acompañaron este proceso, aplicando el aprender a Hacer con mirada inclusiva

Resultados

Se capacitaron 171 participantes entre las dos capacitaciones entre Alumnos Tutores y Tutores guía y Docentes de Ingreso de la UNSJ.

Acciones de transferencia en la UNSJ

En el mes de febrero los destinatarios han dictando y acompañando los cursillos de ingreso a UNSJ aproximadamente a 730 ingresantes en Facultad de Ingeniería (FI), a 1900 ingresantes en Escuela Universitaria de Ciencias de la Salud (EUCS), a 3280 en Facultad de Ciencias Sociales (FACSO), a 1020 ingresantes en Facultad de Ciencias Exactas, Físicas y Naturales, a 2002 estudiantes en Facultad de Filosofía, Humanidades y Artes (FFHyA) y 433 de Facultad de Arquitectura, Urbanismo y Diseño (FAUD)

La definición y aprobación del SIED por un Consejo Superior da cuenta de un ejercicio muy interesante en el que se lograron nuevos consensos y la confirmación de propuestas basadas en construcciones previas. (González A y Roig H (2018, pág.5). La impronta de la comunicación eficaz en espacios virtuales de aprendizaje permitirá potenciar no solo

la presencia social, docente, y cognitiva, sino fortalecer los lazos comunicacionales con nuestros estudiantes, potenciando el redescubrir su ingenio creativo y genialidad dormida por el exceso de información (CARMONA, E. 2016, pág. 35).

Referencias

GONZÁLEZ, A. Y ROIG, H. (2018). Normativa de educación a distancia para la universidad argentina: avances y desafíos pendientes. *Virtualidad, Educación y Ciencia*, 16 (9), pp. 152-157

- Sedoff Miguel – Bonetti Sandra, (2018) *Flipped Learning Una guía para darle vuelta a tu clase*. 1ra Editorial Logos Rosario Argentina, pág 167.

- CARMONA, Emilse (2016) *Desafíos innovadores en Educación: PNL3 y Quantum Coaching en aulas virtuales*. Ed. UNSJ. San Juan Argentina.

Maggio, M. (2012) *Enriquecer la enseñanza*. Buenos Aires: S.XXI

Goleman, D. *El cerebro y la inteligencia emocional*. Título original: *The Brain and Emotional Intelligence: New Insights* Traducción: Carlos Mayor 1.ª edición: abril, 2013 ISBN DIGITAL: 978-84-9019-434-8

Pickelny, M (2021) "Aprendiendo a Hacer" *Nuevas Estrategias de Mediación Pedagógica: un intento de sistematización de las actividades desplegadas desde el SIED UNSJ-FACSO*, *Revista Contribuciones Científicas*, ISSN DIGITAL 2684-0596, en <https://www.adicus.org.ar/ccpu2-ahnemp/>

Carmona, 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*, *Revista Contribuciones Científicas*, ISSN DIGITAL 2684-0596, en <https://www.adicus.org.ar/ccpu2-trdde/>

GUTIÉRREZ, M.; PIEGAIA, L. (2018) *Mentes que Hacen: Dos propuestas teóricas que ponen en reflexión las MENTES DEL FUTURO y profundizan el APRENDER A HACER*. 1º Edición. San Juan. Editorial UNSJ.

UNESCO (2004) *Alianzas e innovaciones en proyectos de desarrollo educativo local*. International Institute for Educational Planning 7-9 rue Eugène-Delacroix 75116, Paris, Francia IPE - UNESCO - Sede Regional Buenos Aires Agüero 2071 (C1425EHS) B

LAS REDES NEURONALES ARTIFICIALES Y LA DESERCIÓN EN EL NIVEL UNNIVERSITARIO

Marcelo Omar Sosa¹, Carlos Iván Chesñevar², Eugenia Cecilia Sosa Bruchmann¹

¹Departamento Computación/Facultad de Ciencias Exactas y Naturales/Universidad Nacional de Catamarca

Av. Belgrano N° 300 - Planta alta - C.P: 4700 - San Fernando del Valle de Catamarca
Teléfono: 0383- 4425610 /4420900

²Instituto de Ciencias e Ingeniería de la Computación (Conicet –U.N.S.)/Facultad de Ciencias e Ingeniería /Universidad Nacional del Sur

Av. Alem 1253 -B8000CPB Bahía Blanca, Argentina
Teléfono: 0291- 4595135 ext. 2610

sosamod1@hotmail.com, cic@cs.uns.edu.ar, sosab_ec@hotmail.com

RESUMEN

La deserción como problema en el nivel universitario representa una preocupación constante, se incrementa año a año y toma mayor importancia en el 2.020 con la situación de pandemia. Aunque se viene analizando numerosas soluciones y en la búsqueda constante de alternativas que permitan disminuirla. Con la integración generalizada de las plataformas virtuales de aprendizaje en el proceso educativo el volumen de datos recogidos creció exponencialmente. Permitió registrar cada acción realizada por los estudiantes lo que posibilitará la materialización de perfiles y el análisis de factores que influyen directamente en el rendimiento académico. Para el procesamiento de estos datos educativos se propone la utilización de redes neuronales artificiales (RNA) y el desarrollo de experimentaciones para que los resultados permitan la predicción de los posibles desempeños académicos de los estudiantes y la detección de aquellos que se encuentran en riesgo de abandono. El objetivo principal que se plantea la investigación es la disminución de la deserción basado en los resultados obtenidos de la implementación de las RNA. Se espera que esta predicción permita al docente actuar adecuadamente en temas puntuales e influir positivamente en el desempeño académico de los estudiantes.

Palabras clave: Deserción, redes neuronales artificiales.

CONTEXTO

El presente trabajo de investigación se enmarca dentro de las actividades de la tesis de maestría que se desarrolla conjuntamente con tareas dentro de los centros de investigación de informática y estadística aplicada de la Facultad de Ciencias Exactas y Naturales (FACEN) - Universidad Nacional de Catamarca (UNCA), en el marco del sistema de investigación, desarrollo e innovación. Además, los datos a procesar se obtuvieron de las actividades virtuales realizadas por los estudiantes que cursan las asignaturas seleccionadas en diferentes sedes (San Fernando del valle de Catamarca, Departamento Ancasti y San Miguel de Tucumán). Los docentes del equipo de investigación realizan sus actividades en los campus universitarios de origen y están a cargo de asignaturas de las carreras del área informática de la FACEN de la UNCA, y de la Facultad de Ciencias e Ingeniería de la Universidad Nacional del Sur (UNS).

1. INTRODUCCIÓN

Sin duda el año 2.020 afectó profundamente la vida de las personas a nivel global, produciendo cambios que parecían imposibles e impensados hasta el momento [8]. Algunos aspectos tan importantes como la educación se vieron afectados profundizando y acelerando procesos tales como la virtualización haciendo que las instituciones educativas se aferren a él como si fuera la panacea [9]. Como una consecuencia lógica de este proceso se produjo el inesperado crecimiento en la generación y almacenamiento de datos que representan las actividades virtuales (únicas posibles) en entornos de aprendizaje mientras el estudiante utiliza las diversas plataformas elegidas por cada institución educativa [10]. Además, otra consecuencia no deseada la representa la deserción, aunque si bien es un problema compartido por todos los niveles educativos, en particular en el universitario se vio incrementado de forma preocupante durante el proceso pandémico sufrido [11]. Se debe tener en cuenta que además de los factores que afectaron en años anteriores ahora se les suman otros nuevos como la imposibilidad de acceder a una conexión de internet, no contar con los dispositivos adecuados para el desarrollo de las actividades por nombrar algunos [9]. Esta situación viene siendo analizada desde años anteriores y desarrollando diferentes acciones en pos de lograr una disminución efectiva de la misma [11]. Hoy en día se presenta la oportunidad, como nunca antes, de analizar cada acción realizada por los estudiantes mientras se desenvuelven en las plataformas virtuales de aprendizaje. El volumen de datos generado es tan detallado y voluminoso que sólo puede ser procesado utilizando herramientas de Data mining y Big data [2].

Es de esperar que las instituciones educativas deseen abordar la deserción antes de que ésta ocurra concretamente. Para lo cual es necesario realizar un proceso de predicción de cómo será el desempeño de los estudiantes basado en datos iniciales e intermedios que permitan detectar a aquellos que se encuentran en riesgo de abandono. Esta predicción se

considera un aviso que permitirá al docente actuar adecuadamente apoyando a estos estudiantes en temas puntuales que le permitan influir positivamente en su desempeño.

El data mining está compuesto por numerosas técnicas que tienen como objetivo el buscar patrones e información relevante y oculta dentro de los datos almacenados en grandes bases de datos [6]. Actualmente los diferentes programas disponibles para la aplicación de las técnicas de data mining facilitan la tarea del procesamiento de grandes volúmenes de datos. En particular los datos educativos representan un campo especial para su procesamiento y la búsqueda de patrones tanto en las actividades como en el análisis de los resultados intermedios obtenidos por los estudiantes. Dentro de los datos disponibles se encuentran detalles tan precisos como la cantidad de accesos a los materiales virtualizados, tiempo de permanencia en foros, participaciones, creación de temas nuevos, aportes a las tareas en los grupos, cantidades de consultas y respuestas de preguntas realizadas por otros estudiantes, entre muchos otros. Antes del procesamiento de estos datos se debe realizar un proceso de adecuación, como por ejemplo: borrado de datos incompletos, erróneos o innecesarios y la selección de aquellos atributos que se consideren más representativos del perfil del estudiante. Del procesamiento de los datos puede extraerse información que correctamente interpretada seguramente favorecerá el proceso educativo y servirá como guía al docente en la toma de decisiones. Esta rama del data mining que procesa datos educativos se denomina Educational data mining (EDM)[1] que se aplicará a datos relacionados con el desempeño académico generados en plataformas Learning Management System (LMS).

El análisis de datos educativos puede realizarse desde distintos enfoques de acuerdo a las técnicas que se apliquen, estos son: descripción [3], predicción [4], segmentación [5].

Para este caso particular, las técnicas que se agrupan en la predicción tienen como objetivo el establecer un modelo que represente a los

datos y que permita estimar los valores que pueden tomar a futuro las variables analizadas, en el caso de su aplicación a datos educativos los atributos iniciales de los estudiantes son analizados por estas técnicas obteniéndose resultados que le predicen al docente como puede llegar a ser el desempeño académico durante las actividades a realizar. Estas predicciones se basan en el análisis de los datos históricos, iniciales y resultados intermedios que van obteniendo los estudiantes durante el desarrollo del contenido de la asignatura [4].

Unas de las técnicas más utilizadas para la predicción son las redes neuronales artificiales (RNA), constituyen algoritmos computacionales inspirados en las redes neuronales biológicas, comportándose de forma similar a las mismas en sus funciones más comunes. Las RNA son entrenadas para aprender, generalizando resultados y extrayendo patrones de un conjunto de datos [7].

Los investigadores vienen trabajando en el desarrollo de nuevos prototipos de redes neuronales basados en el tipo de problema a resolver. Para los problemas de predicción las arquitecturas más utilizadas son las denominadas perceptrón de múltiples capas, en ellas se produce el aprendizaje mediante un proceso denominado “retro propagación del error”, que se basa en obtener el mínimo de una función de error calculada a través de la salida del sistema y el modelo obtenido por el entrenamiento de la red [7].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo presenta las expectativas en cuanto a los resultados que se obtendrán de las experimentaciones previstas para su realización sobre la base de datos de los estudiantes de carreras de informática de primer año. Esta base se encuentra en la etapa de preprocesamiento y determinación de los variables y atributos más representativos de la performance desarrollada por los estudiantes en las asignaturas con contenidos básicos de programación. Representa parte de las

actividades a realizar en la tesis de maestría en Ciencias de la Computación, en donde se investiga los aportes del data mining como disciplina, al proceso educativo. Particularmente esta herramienta se orienta a prevenir la deserción utilizando su característica predictiva.

Si bien los resultados obtenidos serán presentados dentro de la tesis de maestría, posteriormente las experiencias se aplicaran en las actividades de asignaturas que aportaron los datos necesarios al proceso. En general el tema del procesamiento de los datos se viene analizando en distintos trabajos y presentaciones en congresos y reuniones científicas que van demostrando los avances y aportes encontrados mientras se desarrollan las actividades previstas en el plan de tesis.

Los objetivos principales son:

- Demostrar la capacidad predictiva de las redes neuronales artificiales en el procesamiento de datos educativos.
- Promover la utilización de esta herramienta como apoyo a la disminución de la deserción en las carreras de informática de la FACEN-UNCA.
- Comprender los procesos mentales que permiten a los estudiantes la construcción de conocimiento en programación.
- Experimentar con la selección de atributos que representen fehacientemente el perfil del estudiante de carreras de informática.

3. METODOLOGÍA

Con esta investigación se busca proponer a las redes neuronales artificiales como una herramienta que permitirá realizar predicciones sobre el rendimiento futuro de los estudiantes basados en datos históricos, iniciales e intermedios. Además de dar una base cierta para la toma de decisiones por parte de los docentes en cuanto a varios aspectos importantes del proceso educativo de modo que permitan una disminución concreta de la deserción de estudiantes en carreras de informática de la FACEN-UNCA.

El estudio de los resultados que se obtengan de la aplicación de las RNA sobre datos educativos se estima que permitirá intervenir precisamente sobre los estudiantes detectados para mejorar su desempeño y evitar la deserción de los mismos. Particularmente para en este estudio se utilizarán datos de asignaturas con contenidos básicos de programación de las diferentes carreras y subsedes de la FACEN-UNCA. Como corolario de esta investigación se estima el poder aportar una herramienta más para abordar la deserción desde una mirada científica que pueda ser aplicada por docentes para confirmar su validez.

4. FORMACIÓN DE RECURSOS HUMANOS

Algunos de los autores del trabajo se encuentran en la etapa de desarrollo de sus tesis de posgrado en carreras relacionadas con el tema de investigación, como la Maestría en Ciencias de la Computación en donde el Mgter. Sosa Marcelo Omar Diógenes desarrolla actualmente la tesis denominada: ***“Aportes de data mining a la mejora del proceso educativo con blended learning: formalización y experimentaciones”*** que se realiza con la dirección del Dr. Chesñevar Carlos Iván perteneciente a la Universidad Nacional del Sur (UNS). Además el investigador se encuentra en la etapa de planificación de su tesis doctoral en el área de minería de datos en el Doctorado en Ciencias dictado en la (FACEN) en convenio con la (UNS).

La Docente Investigadora Esp. Sosa Bruchmann Eugenia Cecilia desarrolla su tesis en la carrera Maestría en Ingeniería en Software de la Universidad Nacional de San Luis denominada ***“La experiencia del usuario desde un nuevo enfoque para el desarrollo de productos interactivos: el comportamiento emocional del usuario y la importancia de los atributos estéticos”*** dirigida por el Dr. Germán Montejano. Los docentes investigadores desarrollan actividades de dirección de tesis de la carrera de Licenciatura en Tecnología Educativa de los siguientes estudiantes: Varela

Marino del Valle cuya tesis se denomina ***“Análisis del impacto de un aula virtual en el proceso de enseñanza y aprendizaje en la escuela de educación técnica N° 7 José Alsina Alcobert”***, Tolaba Rodrigo Alejandro desarrolla la tesis denominada: ***“Aplicación del modelo Van Hiele a la enseñanza de triángulos con geogebra”***. y Santillán Mario de la carrera Especialización en docencia universitaria en disciplinas tecnológicas donde desarrolla la tesis denominada : ***“Diseño de simulador para el desarrollo de destrezas de programación”***

Además desarrollan las siguientes actividades:

- Dirección de proyectos de investigación denominado: “Digitalización de materiales didácticos para la enseñanza de la programación”.
- Investigadores pertenecientes a los centros de investigación e innovación de la FACEN-UNCA.
- Producción de artículos científicos para su presentación en congresos locales, nacionales e internacionales.
- Dictado y participación en cursos de actualización sobre los diferentes temas de investigación de área de cada docente.
- Integrantes de la revista de ciencias de la Facultad de Ciencias Agrarias de la UNCA.
- Actualización y capacitación permanente de los investigadores en talleres o workshop relacionadas con el tema del trabajo.
- Participación de los investigadores como consultores en proyecto afines que se desarrollan en la FACEN en distintas áreas.
- Examinadores de trabajos finales en las diferentes carreras que se dictan en la FACEN- UNCA.
- Dirección de tesis y tesinas finales de las carreras de computación, informática y Licenciatura en tecnología educativa.
- Planificación de seminarios para docentes en temas relacionados con la investigación.

- Participación en convenios con la Facultad de Tecnología y Ciencias Aplicadas para el desarrollo de estudios del área de datamining.

5. CONCLUSIÓN

Se espera que los resultados obtenidos de estas experimentaciones presenten a las herramientas de data mining como una alternativa al procesamiento de datos educativos. En cuanto a las RNA, se espera que los resultados que se obtengan guíen las actividades de los docentes enfocándose en los temas y estudiantes que según los resultados presentan problemas. Además que faciliten la disminución de manera concreta la deserción de los estudiantes de las carreras de informática de la FACEN-UNCA una vez implementadas en forma concreta en las asignaturas establecidas.

• 6. REFERENCIAS

1. Romero C. and Ventura S., "Educational data mining: A Survey From 1995 to 2005", *Expert System with Applications*, vol. 33, pp. 135-146, 2007.
2. Lavrac, N., Kavsec, B., Flach, P. and Todorovski, L., "Subgroup discovery with CN2-SD". *JoR.N.A.I of machine learning research*. 2004.
3. Jain A.K. and Dubes R.C. "Algorithm for clustering data. 1998. Englewood Cliffs. N.J. Prentice Hall.
4. Fayyad, U, Piatetsky-Shapiro, G. and Smyth, P., "The KDD process for extracting useful knowledge from volumes of data". *Communication of ADM 1996*.
5. Agrawal, R., Imielinski, T. and Swami, A.N. "Mining Association Rules between set of item in large databases". In *International conference on management of data*. 1993. Washington D.C. ACM Press.
6. Molina Félix L., "Data mining: torturando los datos hasta que

confiesen". *Accedido en línea 2 marzo de 2016*.

7. Del Brío B. Martín, Sanz Molina A., "Redes neuronales y sistemas borrosos"; pról. de Lotfi A. Zadeh. January 2006. Publisher: RA-MA (Madrid, España), ISBN: 978-84-7897-743-7
8. Ordorika, I. (2020). *Pandemia y educación superior*. *Revista de la educación superior*, 49(194), 1-8.
9. Peñuelas, S. A. P., Pierra, L. I. C., González, Ó. U. R., & Nogales, O. I. G. (2020). *Enseñanza remota de emergencia ante la pandemia Covid-19 en Educación Media Superior y Educación Superior*. *Propósitos y Representaciones*, 8(SPE3), 589.
10. Morado, M. F., & Ocampo Hernández, S. (2019). *Una experiencia de acompañamiento tecno-pedagógico para la construcción de entornos virtuales de aprendizaje en educación superior*. *Revista Educación*, 43(1), 43-61.
11. Martínez, C. R. V., & Pérez, M. C. R. (2007). *La deserción estudiantil en educación superior a distancia: perspectiva teórica y factores de incidencia*. *Revista Latinoamericana de Estudios Educativos (México)*, 37(3-4), 107-122.

7. BIBLIOGRAFÍA Y TRABAJOS CONSULTADOS

- Brijesh Kumar Baradwaj and Saurabh Pal, "Mining Educational Data to Analyze Students' Performance," in *International Journal of Advanced Computer Science and Applications, India*, 2011, pp. Vol. 2, No. 6.
- Fischer Angulo Erwin Sergio. "Modelo para la automatización del proceso de determinación de riesgo de deserción en estudiantes universitarios". Santiago de Chile 2012.
- González, J. C., Ramos, S., & Hernández, S. (2017). *Modelo Difuso*

- del Rendimiento Académico Bi-explicado. Revista de Sistemas y Gestión Educativa*, 55–64.
- Grossi María Delia. "Reglas de predicción aplicables al diseño de un curso de computación". Marzo 2008.
 - Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 133–145.
 - Solarte Martinez, Guillermo Roberto, Ocampo S., Carlos Alberto. *Técnicas de clasificación y análisis de representacion del conocimiento para problemas de diagnósticos. Cientia Et Technica [En Linea] 2009, Xv (Agosto-Sin Mes) : [Fecha De Consulta: 2 de marzo de 2016]*.
 - Usman, O., & Adenubi, A. (2013). Artificial Neural Network (ANN) Model for Predicting Students' Academic Performance. *Journal of Science and Information Technology*, 23–37.
 - Ventura Cristoban Romero, Sebastian, Nykola Pechenizkiy and Rayan Beker. "Handbook of educational data mining". Chapman & Hall CRC press. 2011.

Microcomputadoras para la enseñanza y el aprendizaje

Martha Semken, Mariano Vargas

Instituto de Ciencias, Universidad Nacional de General Sarmiento
General Sarmiento, Buenos Aires, Argentina
{mseken, avargas} @campus.ungs.edu.ar

RESUMEN

Un aprendizaje es significativo cuando los contenidos: Son relacionados de modo no arbitrario y sustancial (no al pie de la letra) con lo que el alumno ya sabe. Por relación sustancial y no arbitraria se debe entender que las ideas se relacionan con algún aspecto existente específicamente relevante de la estructura cognoscitiva del alumno, como una imagen, un símbolo ya significativo, un concepto o una proposición [1].

En las materias del área Sistemas, se presenta mayor dificultad para trabajar sobre dispositivos reales. Se suele trabajar con emuladores para poder simular situaciones o entornos necesarios para llevar a cabo las prácticas de los contenidos vistos. Si bien tienen una primera aproximación, no son entornos de trabajo conocidos, ni utilizados en la actualidad.

El área de sistemas de la carrera Licenciatura en Informática, ha comenzado un proceso de innovación, intentando llevar nuevas herramientas a los alumnos, con tecnologías que se utilizan en el mercado laboral.

Incorporar el uso de microcomputadoras Raspberry Pi. Con la intención de brindar a los alumnos, dándoles acceso a dispositivos físicos logrando una experiencia real. Luego de una primera etapa de acercamiento con dicha herramienta, desarrollar una etapa de investigación basada en este dispositivo.

CONTEXTO

Este proyecto de I/D en esta primera etapa, involucra el área Sistemas y las materias Organización del Computador y Sistemas Operativos y Redes, de la carrera Licenciatura en Informática del Instituto de Ciencias, Universidad Nacional de General Sarmiento. Con la intención de ampliar las áreas de investigación, abarcando al área programación.

Ya iniciado el proyecto, surge la situación actual de distanciamiento social, lo que da la posibilidad de llevar al máximo la capacidad de la plataforma seleccionada. Con un Raspberry Pi online brindar acceso remoto a los alumnos para que realicen sus prácticas.

INTRODUCCIÓN

La materia Organización del Computador tiene como objetivo conocer la organización interna y funcionamiento de una computadora, su interacción y operación. En un sistema de software es esencial comprender la ejecución como así también la utilización de los recursos de los que se dispone.

Para lograr estos objetivos se trabaja con un procesador en particular. En nuestro caso la arquitectura elegida anteriormente, era Intel 8086 (Fig. 1). Dicha arquitectura ya es obsoleta, pero dada su simplicidad, permitía introducir a los alumnos en los temas descritos. Los conceptos presentes en este procesador a pesar de su obsolescencia, como recurso de aprendizaje son válidos, pero para los alumnos es una tecnología que no está ni estará presente en su futuro profesional, ni

como herramienta de trabajo en su trayecto académico, salvo por la materia Organización del Computador.

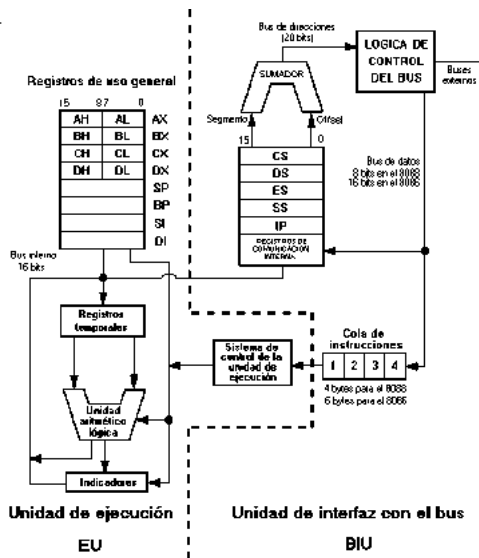


Figura 1 - Arquitectura 8086

En cuanto a la parte práctica, en esta arquitectura, solo se puede trabajar sobre emuladores (Fig. 2). En las prácticas de laboratorio, los alumnos trabajan con Assembler, para lograr comprender como es el proceso para generar un software ejecutable, editar, compilar, vincular, ejecutar. Con frecuencia en los lenguajes de alto nivel, trabajan en entornos de desarrollo (IDE). Lo que hace que el proceso sea prácticamente automático, no llegan a visualizar estos pasos involucrados en el desarrollo de un programa ejecutable. Por esto se incluyen prácticas donde deben realizar estos pasos por separado [2].

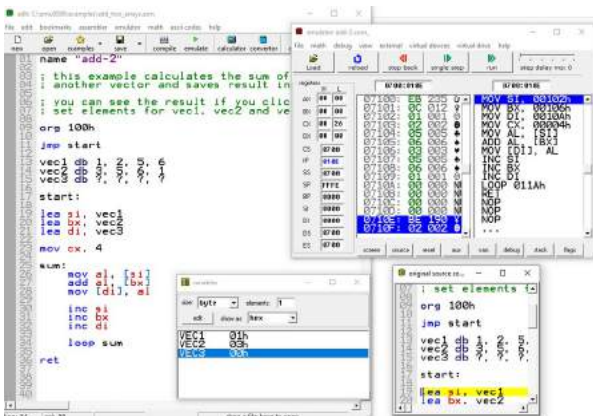


Figura 2- Emu8086, emulador de 8086

Como alternativa a este procesador, el planteo es trabajar sobre ARM (Fig. 3). Este procesador tiene una arquitectura RISC. Dicha arquitectura posee un conjunto reducido de instrucciones [3].

Actualmente es utilizado en celulares y recientemente se conoció la noticia de que los procesadores de Apple serán ARM [4]. Lo que la hace algo actual y tangible.

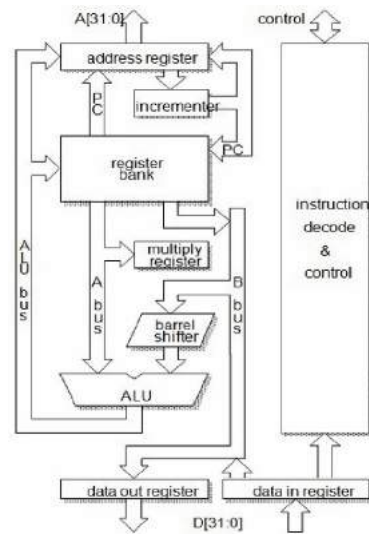


Figura 3 - Arquitectura ARM

Utilizando dispositivos Raspberry Pi, se les puede brindar a los estudiantes una imagen real del hardware sobre el cual trabajar y no un hipotético hardware que nunca verán en su futuro profesional. Estos dispositivos además tienen un muy bajo costo. Brindando un entorno de trabajo basado en el sistema operativo GNU/Linux. Un entorno de trabajo común para desarrollar sistemas en tiempo real.

En la materia Organización del Computador, las prácticas de laboratorio donde trabajan en Assembler de ARM [5], se llevan a cabo sobre un dispositivo al que se accede de forma remota (Fig. 4), lo ideal sería contar con varios dispositivos para distribuir la carga de trabajo.

```

root@raspb01:~/home/oc3q/guap0#
04 /usr/bin/pasta leer el momento/
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Figura 4 – Debug GDB

Estas características permitirán ampliar el uso de la plataforma Raspberry Pi a la materia Sistemas Operativos y Redes y en un futuro brinda un entorno de desarrollo, con múltiples posibilidades. Lograr que cada alumno pueda acceder a la interfaz gráfica, que hasta ahora no es posible por la cantidad de alumnos que se conectan concurrentemente. Instalar diversos servicios de red como por ejemplo Apache Web Server. Realizar análisis de tráfico de red entre varios Raspberry Pi.

En este proyecto la propuesta es aprovechar el interés que los alumnos ya demuestran por la plataforma Raspberry Pi, para ponerlo a trabajar en pro del siguiente objetivo docente: facilitar el estudio de conceptos y técnicas impartidas en varias asignaturas de la carrera y en un futuro desarrollar otros proyectos de investigación. Además de proporcionarles una computadora remota a aquellos que no la poseen, para poder realizar sus prácticas.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de I/D que se proponen abordar mediante el uso del dispositivo Raspberry Pi como herramienta, son múltiples

- Herramienta didáctica: Introducción a los comandos básicos del sistema operativo GNU/Linux, no como concepto teórico sino con una aplicación concreta, para poder realizar prácticas.

Realización de prácticas de laboratorio para la materia Organización del computador, brindando la posibilidad de realizarlas sobre un dispositivo real. Prácticas de laboratorio para la materia Sistemas Operativos y Redes, realizando análisis de tráfico y programación de multithreads.

- Proyectos de investigación: Desarrollo de un Cluster de Raspberry Pi. Aplicaciones de computo de alto rendimiento, basadas en el cluster[6].
- Realización de talleres: Talleres específicos de computación y actividades extracurriculares relacionadas con la programación de servidores, IoT.

RESULTADOS OBTENIDOS/ESPERADOS

En una primera etapa el objetivo de este proyecto fue brindar a los estudiantes una herramienta de trabajo más actual, mediante el acceso a nuevas tecnologías comunes en el mercado laboral, modernizando de esta manera los contenidos de las materias. Esto fue logrado con éxito.

Durante los dos semestres del año 2020, se utilizó esta nueva arquitectura. Sumándole un condimento adicional, dada la situación de distanciamiento social preventivo. Solo se contaba con un dispositivo Raspberry, que se instaló y se mantuvo online durante la cursada. No se pudo acceder a los laboratorios de la universidad, así que el trabajo se realizó con una red de uso particular, conexión hogareña. Se utilizó el mismo Raspberry Pi con un servicio de DNS dinámico.

Además se implementó una Wiki, para que los alumnos tuvieran acceso a todo el material teórico, de forma más accesible [7].

A los alumnos se les asignó un usuario para que pudieran conectarse en forma remota, para realizar sus prácticas. De esta forma podían trabajar, sin necesidad de instalar nada en sus computadoras. Esto fue una ventaja

adicional, siendo una materia de primer año, la mayoría no tenía manejo del sistema operativo GNU/ Linux, pudieron tener una primera aproximación, sin necesidad de instalarlo, ya que esta actividad es incumbencia de otra materia. Brindándole una primera experiencia, que podrán capitalizar en la siguiente materia Sistemas Operativos y Redes.

A su vez los docentes tuvieron acceso a todas las cuentas y cuando surgió una consulta fue posible acceder a los ejercicios resueltos por los alumnos para asesorarlos o despejar dudas. De la misma forma realizaron las entregas de los TPs finales y fue posible realizar coloquios donde los alumnos presentaban su solución y podían explicarla, pudiendo trabajar directamente sobre su código. En esta etapa trabajaron en grupos y cada grupo tuvo una cuenta común, lo que les facilitaba realizar un trabajo colaborativo. *“El uso instructivo de grupos pequeños para que los estudiantes trabajen juntos y aprovechen al máximo el aprendizaje propio y el que se produce en la interrelación. Para lograr esta meta, se requiere planeación, habilidades y conocimiento de los efectos de la dinámica de grupo”* [8]

En cuanto a la enseñanza/aprendizaje, la estrategia fue exitosa, se notó mayor motivación y un aumento de interés por la nueva tecnología.

Algo no proyectado inicialmente fue la posibilidad del trabajo remoto, con el agregado de poder compartir un espacio de trabajo, docentes y alumnos. Generando lo más parecido al trabajo en laboratorio de forma presencial.

El próximo paso es ampliar el uso de esta plataforma en la materia Sistemas Operativos y Redes.

Disponer de un laboratorio de Raspberry Pi, dado su bajo costo y la posibilidad de accederlo remotamente, este es un beneficio

adicional para los alumnos que no disponen de computadora con gran capacidad de procesamiento. Podrán acceder a los dispositivos y trabajar de forma remota.

Una vez afianzada la nueva tecnología y con alumnos ya familiarizados con ella, continuar con la etapa de investigación sobre estos dispositivos, que dan amplias posibilidades, abriéndose muchas ramas de investigación posible.

FORMACIÓN DE RECURSOS HUMANOS

Los miembros del equipo poseen formación en informática y electrónica. Ambos están cursando el Ciclo de Licenciatura en Tecnologías Digitales para la Educación, en la UNLA

BIBLIOGRAFIA

1. Ausubel-Novak-Hanesian (1983). Psicología Educativa: Un punto de vista cognoscitivo. 2º Ed. TRILLAS México
2. Robert G. Plantz. Introduction to Computer Organization: ARM Assembly Language Using the Raspberry Pi. Online, 2018. [URL:http://bob.cs.sonoma.edu/IntroCompOrg-RPi/intro-co-rpi.html](http://bob.cs.sonoma.edu/IntroCompOrg-RPi/intro-co-rpi.html).
3. Andrew S. Tanenbaum and Todd Austin. Structured Computer Organization. PearsonPrentice-Hall, New Jersey 07458, 6 edition, 2013. URL:https://goo.gl/N2YQc3.
4. Tecno. 9 de noviembre 2020. *Infobae*. Qué se puede esperar del evento donde Apple presentaría sus nuevas Mac con chip propio <https://www.infobae.com/america/tecno/2020/11/09/que-se-puede-esperar-del-evento-donde->

- [apple-presentaria-sus-nuevas-mac-con-chip-propio/](#)
5. VILLENA GODOY, ASENJO PLAZA, CORBERA PEÑA. Prácticas de ensamblador basadas en Raspberry pi. (2015) Universidad de Malaga. España. <https://riuma.uma.es/xmlui/handle/10630/10214>
 6. Rodríguez Sebastián, Chichizola Franco, Rucci Enzo. (2018) *Análisis del Uso de un Cluster de Raspberry Pi para Cómputo de Alto Rendimiento*. (Tesis de grado). Universidad Nacional de La Plata. Buenos Aires. Argentina
 7. Semken, Vargas, Velcic, Rojas Paredes, Lopez Holtman. (2020) Wiki Sistemas. <http://wikisistemas.servehttp.com/>
 8. Johnson & Johnson, D. (1998). *Cooperation in the classroom* (7a ed.). Interactionbook Company.

PRÁCTICAS DE FORMACIÓN DEL PENSAMIENTO COMPUTACIONAL EN LA ENSEÑANZA SUPERIOR MEDIADA POR TECNOLOGIA

Mg. Ing. Luciana Gabriela Terreni

Instituto Sedes Sapientiae, Universidad Nacional de la Patagonia Austral

E-mail: luciterreni@gmail.com

RESUMEN

El pensamiento computacional es una forma de pensar que no es exclusiva de los programadores sino que consiste en la resolución de problemas, el diseño de los sistemas, y la comprensión de la conducta humana haciendo uso de los conceptos fundamentales de la informática. En una posmodernidad atravesada por la tecnología, los ciudadanos necesitan construir competencias que les permitan afrontar los desafíos y situaciones que el contexto presenta. Es por ello que el desarrollo y formación del pensamiento computacional debe promoverse desde los espacios educativos de manera transversal, en especial los de enseñanza superior. Por lo expuesto se analiza un caso de estudio en la enseñanza superior desde un enfoque exploratorio para detectar y describir las prácticas tendientes a formar este tipo de pensamiento en la Tecnicatura en Análisis y Desarrollo de Software del Instituto de Profesorado Sedes Sapientiae.

Palabras claves: pensamiento computacional, educación superior, prácticas pedagógicas.

CONTEXTO

Este trabajo se enmarca dentro del proyecto de trabajo intercátedra propuesto por el Instituto Superior Sedes Sapientiae para la Tecnicatura en Análisis y Desarrollo de Software.

Desde dos espacios curriculares de la tecnicatura se trabaja desde el enfoque docencia-investigación-extensión y se ha planteado este proyecto que indaga sobre las prácticas o estrategias que los docentes proponen para formar una competencia

transversal como es el pensamiento computacional en la carrera antes mencionada.

1. INTRODUCCIÓN

El término pensamiento computacional ha ganado notoriedad en los últimos años, siendo utilizado para hacer referencia a técnicas y métodos de resolución de problemas donde intervienen e integran la experiencia y los saberes relacionados con la programación de computadoras. La aplicación del pensamiento computacional no solo se realiza en problemas informáticos sino que se puede aplicar mas ampliamente para razonar y trabajar sobre otros tipos de situaciones y áreas de conocimiento. Es una metodología de resolución de problemas que se puede automatizar (Zapata-Ros, 2015). El pensamiento computacional conjuga el pensamiento ingenieril, el científico y el lógico matemático (Wing, 2008) e implica el desarrollo de un mayor nivel de abstracción para resolver problemas concretos de diseño de sistemas y otros que requieren soluciones automatizadas de pensamiento algorítmico y programación. Entendiendo el pensamiento computacional como un proceso cognitivo que implica un razonamiento lógico aplicado a la resolución de problemas, sus elementos son los siguientes (CAS, 2015): capacidad de pensar de forma algorítmica, capacidad de pensar en términos de descomposición, capacidad de pensar en generalizaciones, identificando y haciendo uso de patrones, capacidad de pensar en términos abstractos y elección de buenas representaciones, la

capacidad de pensar en términos de evaluación.

Estas capacidades se construyen por medio de técnicas como el análisis, diseño, aplicación, reflexión y programación, las cuales se implementan mediante actividades propuestas por el docente.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El Instituto Sedes Sapientiae es una institución de enseñanza superior que dicta 4 profesorado y 2 tecnicaturas, tiene una vasta experiencia en la educación mediada por tecnología y mediante entornos virtuales, implementando desde 2009 una modalidad de enseñanza de aula virtualizada (Giordano, 2015) bajo la plataforma Moodle.

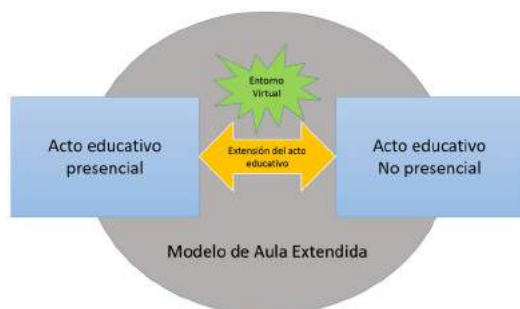


Figura 1. Modelo de aula virtualizada de IPSS. Elaboración propia.

La Tecnicatura en Análisis y Desarrollo de Software del IPSS tiene una duración de tres años y 32 espacios curriculares. Es el caso de estudio seleccionado, más específicamente las cohortes 2019, 2020 y 2021 y en torno al mismo, se han planteado los siguientes objetivos de investigación:

- Descubrir cuales son las prácticas o estrategias que los docentes proponen para formar el pensamiento computacional.
- Clasificar hacia que capacidad y técnicas apuntan las acciones propuestas.

- Describir las particularidades de las estrategias propuestas.
- Formular un modelo de formación de pensamiento computacional en la enseñanza superior.

El abordaje metodológico de la investigación se corresponde con un enfoque exploratorio de diseño mixto aplicado a un caso de estudio tomando como unidad de análisis a los docentes.

El diseño cuantitativo y el cualitativo se han aplicado de manera secuencial. Se definieron claramente dos etapas y la segunda derivó de los resultados de la primera.

En primera instancia, a todos los docentes de la carrera mencionada se les proporcionó un cuestionario por cada asignatura o espacio que tenían a cargo para detectar quienes proponían actividades o e-actividades para la formación del pensamiento computacional.

De los docentes que respondieron el cuestionario se seleccionó para la instancia de entrevista en profundidad a aquellos que desarrollaban actividades tendientes a construir y fortalecer el pensamiento computacional de manera frecuente y/o significativa.

Los hallazgos en las entrevistas fueron registrados en formato de audio digital para luego ser transcritos, tabulados y analizados a la luz del marco teórico y las categorías de análisis establecidas.



Figura 2. Metodología aplicada al caso de IPSS. Elaboración propia.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto se encuentra en etapa de procesamiento y triangulación de los datos sistematizados. De forma pre-eliminar se observa que los docentes mencionan reiteradamente actividades como casos de estudio, resolución de problemas mediante pseudocódigo y formulación de proyectos como formas de construcción del pensamiento computacional.

Se espera poder clasificar las menciones de actividades realizadas por los docentes según a las capacidades y técnicas a las que apuntan.

Acerca de las particularidades del proceso de formación del pensamiento computacional, de las expresiones vertidas por los docentes surge pre-eliminarmente que:

- Las secuencias inician mediante el formato tradicional de exposición en la clase presencial para luego generarse una secuencia de trabajo no presencial, es decir un modelo de aula extendida o virtualizada (Zangara, 2008)
- El alumno adquiere un rol activo en las actividades para el desarrollo del PC.
- El docente asume un rol de facilitador de recursos para la resolución de la actividad propuesta.
- Las actividades propuestas integran tecnologías de algún tipo y además de promover el PC, propician espacios de construcción de competencias digitales (Adell, 2013)
- El entorno virtual institucional es el escenario unificador entre el trabajo presencial y no presencial, donde se proponen las actividades y se dialoga en torno a ellas (Terreni, Vilanova, Varas, 2019).

Esta investigación permitirá, una vez realizado el análisis de la totalidad de los resultados, desarrollar la diagramación y especificación de un modelo de formación de pensamiento computacional que pueda extrapolarse a otros casos, lo cual representaría una contribución al abordaje de esta temática transversal a todos los campos. Un esbozo de este modelo se presenta en la figura 3.



Figura 3. Modelo de Formación de PC.
Elaboración propia.

4. FORMACIÓN DE RECURSOS HUMANOS

El co-diseño del proyecto fue realizado por los docentes de las cátedras intervinientes que además de diagramar el mismo, formaron a los alumnos en la recolección y procesamiento de los datos. Por otro lado, en la medida que avanzan las distintas fases del proyecto se han realizado instancias de difusión para promover la investigación y la extensión dentro de las cátedras del IPSS.

5. BIBLIOGRAFÍA

Adell, J. (Productor). (2012). Competencia digital. Disponible en: https://www.youtube.com/watch?v=yZBe1-J_cNQ

CAS (2015). Pensamiento Computacional. Guía para profesores. Computing At School. Recuperado de <http://www.codemas.org/wp-content/uploads/2016/04/Pensamiento-computacional-Guía-para-profesores.pdf> [Enlínea 31/10/19]

Giordano, O. (2015). La formación de competencias digitales de estudiantes de profesorado universitarios: La estrategia de e-actividades en un modelo de aula extendida. Tesis de maestría. Universidad Nacional de La Plata. Facultad de Informática. En Memoria Académica. Disponible en: <http://www.memoria.fahce.unlp.edu.ar/tesis/te.1214/te.1214.pdf>

Ros, M. (2015). Pensamiento computacional: una nueva alfabetización digital. RED, Revista de Educación a Distancia, 46(4).

Terreni, L., Vilanova, G., & Varas, J. (2019). Desarrollo de competencias digitales en propuestas pedagógicas en ambientes mediados. Informes Científicos Técnicos - UNPA, 11(3), 61-87. <https://doi.org/10.22305/ict-unpa.v11.n3.797>

Wing, J. (2006). Computational thinking. Communications of the ACM, 49(3), pp. 33-35. Recuperado de <https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf> [Enlínea 29/10/19] Wing, J. (primavera, 2011). Computational thinking - What and why? The Link, (6), pp. 20-23. Recuperado de https://www.cs.cmu.edu/sites/default/files/11-399_The_Link_Newsletter-3.pdf Zapata-

Zangara, A. (2008). Conceptos básicos de educación a distancia o ... "las cosas por su nombre". Disponible en: <http://es.scribd.com/doc/15679132/Conceptos-basicos-de-educacion-a-distancia-olas-cosas-por-su-nombre-A-Zangara>

PRÁCTICAS DE PROGRAMACIÓN GRUPALES EN EL AULA ESTRATEGIAS DIDÁCTICAS PARA EL DESARROLLO DEL PENSAMIENTO COMPUTACIONAL EN LOS PRIMEROS CURSOS DE PROGRAMACIÓN DEL CICLO INICIAL UNIVERSITARIO

Lic. Natalia Colussi & Dra. Pamela Viale

colussi@fceia.unr.edu.ar pamela@fceia.unr.edu.ar

Departamento de Ciencias de la Computación

Escuela de Ciencias Exactas y Naturales

Facultad de Ciencias Exactas, Ingeniería y Agrimensura

Universidad Nacional de Rosario

RESÚMEN

Nuestra intención con el siguiente artículo es presentar una línea de investigación en desarrollo la cual busca mejorar el proceso de enseñanza y aprendizaje de la programación en carreras afines a las Ciencias de la Computación a través de la Metodología de Enseñanza Basada en Proyectos, las actividades grupales, y las estrategias de resolución de problemas basadas en el Pensamiento Computacional.

Palabras Claves: enseñanza de la programación por grupo, pensamiento computacional, aprendizaje basado en proyectos, primer año universitario.

CONTEXTO

El presente proyecto está radicado en la Facultad de Ciencias Exactas, Ingeniería y Agrimensura dependiente de la Universidad Nacional de Rosario. La investigación conforma un proyecto bienal 2020-2021, NRO: 80020190100255UR, aprobado por la Universidad Nacional de Rosario, vigente con resolución del CS en trámite.

1. INTRODUCCIÓN

Las integrantes de este proyecto, somos docentes de las materias Programación I y Programación agrupadas en una única cátedra del ciclo básico de la Licenciatura en Ciencias de la Computación (LCC), la Licenciatura en Matemática, y el Profesorado en Matemática, pertenecientes a la Facultad de Ciencias Exactas, Ingeniería y Agrimensura (FCEIA), dependiente

de la Universidad Nacional de Rosario (UNR). Puntualmente estamos a cargo del segundo dictado de dichas materias, conocido popularmente como el redictado. Los segundos dictados de las materias de primer año de la LCC surgieron hace ya más de una década en la UNR, como parte de las medidas de retención de los estudiantes ante la no aprobación de las primeras materias disciplinares, en el marco de las políticas de estímulo y promoción de las carreras informáticas en el país. La complejidad, en general de la inserción al sistema universitario, y en lo particular, a cuestiones curriculares de la carrera, genera una cantidad significativa de alumnos que necesitan volver a cursar las materias del primer y segundo cuatrimestre de primer año.

En la cátedra de Programación I y Programación se imparten conocimientos referidos al Pensamiento Computacional [Phi09,Win06], haciendo fundamental hincapié en los principios de programación [L13, FFFK01, SBL10]. La adquisición y construcción del conocimiento disciplinar para los estudiantes de primer año de la LCC, LM y PM no se produce para todos al mismo tiempo. Los estudiantes provienen de diferentes especialidades y contextos en su nivel medio, conviviendo así, en un mismo espacio áulico, alumnos con experiencias, intereses y expectativas dispares. En un extremo encontramos a aquellos estudiantes que llegan como un lienzo en blanco, sin ningún conocimiento previo sobre algoritmia y programación y, en el extremo opuesto, aquellos que provienen de escuelas técnicas especializadas, y en el medio queda, otro gran

grupo de estudiantes que tienen algún tipo de conocimiento previo en el área, adquirido de forma autodidacta, pero no en un espacio educativo formal. De igual modo surgen heterogeneidades propias asociadas a las carreras de base [DGPC14].

La realidad nos describe entonces una cátedra masiva, con más de doscientos inscriptos, heterogénea, del ciclo inicial universitario, con los problemas propios de esta etapa universitaria. Ocurre entonces que muchas veces una gran cantidad de estudiantes no logran aprobar o regularizar este primer curso de programación y asisten entonces al segundo dictado que se ofrece al siguiente cuatrimestre. Este nuevo dictado representa una segunda oportunidad para incorporar habilidades y/o competencias no alcanzadas en la primera instancia.

Decidimos entonces en el año 2017 proponer una estrategia de abordaje de la problemática del redictado que permita atacar las dificultades de fondo de los estudiantes, no repetir el modelo previo, e innovar para lograr modificar estructuras de aprendizaje y enseñanza que no era suficientes para lograr que los estudiantes aprendan a programar y aprueben en consecuencia el curso.

Indagamos y categorizamos entonces los problemas del estudiantado, los cuales podríamos principalmente vincularlos a: i) la no adaptación al cursado y las metodologías de evaluación, ii) a la dificultad en la incorporación de nuevos conceptos, herramientas y lenguajes de programación, iii) a la no integración a la comunidad académica y espacios educativos, el entorno social entre pares, la relación con la comunidad docente y científica, como así también la adaptación a la vida universitaria. En lo referido a cuestiones técnicas, los estudiantes en su mayoría presentan dificultades para: a) formular un problema y diseñar soluciones algorítmicas, b) descomponer un problema en subproblemas, c) representar los datos, d) modelar, analizar y organizar los datos de forma

abstracta, e) generalizar, generar y automatizar casos de prueba, etc.

Como eje central de la propuesta para el redictado estuvo entonces trabajar fuertemente en que los estudiantes apoyen su proceso de aprendizaje en desarrollar e incorporar habilidades y conceptos necesarios para la resolución de problemas a partir de los pilares que promueve el Pensamiento Computacional, el cual favorece el ejercicio de una serie de destrezas que incluye: i) confianza al trabajar con la complejidad, ii) persistencia al trabajar con problemas difíciles, iii) tolerancia a la ambigüedad, iv) habilidad para lidiar con problemas abiertos y cerrados; y v) habilidad para comunicarse y trabajar con otros para lograr una meta en común y soluciones. Se implementó en la cátedra del redictado la realización de dos proyectos grupales de programación en el marco de un “Plan de Trabajo Didáctico para el Aula” (PTDA) [CVJCYT19] utilizando la Metodología de Aprendizaje y Enseñanza Basada en Proyectos [SJ16,SB12].

El PTDA tuvo así su origen en la necesidad de redefinir el proceso de enseñanza-aprendizaje utilizado hasta el momento en la cátedra del redictado, para brindar un mejor marco de contención y solución a los problemas que presentaba el estudiantado del curso. Esta iniciativa requirió de la reelaboración de la presentación, ejercitación y evaluación de contenidos propios de la materia. La realización de dos proyectos grupales permitieron a los estudiantes ejercitar y poner en práctica temas disciplinares, y ser evaluados por los docentes a lo largo de todo el proceso de desarrollo de los mismos. Además, se complementa la evaluación de los proyectos con uno o dos exámenes parciales sobre conceptos que no se encuentran incluidos en los mismos. Buscamos que no haya superposición de temas para no evaluar conceptos ya evaluados, y evitar que las instancias de evaluación con parciales sean innecesariamente extensas.

Actualmente estamos finalizando el cuarto período de aplicación de esta práctica de

enseñanza-aprendizaje en el aula, y hemos participado con publicaciones en distintos eventos de divulgación científica relacionados con la temática, como por ejemplo, las Jornadas Didácticas de la Programación 2019 (JaDiPro) [CVJDP19], las Jornadas de Ciencias de la Computación (JCC)[CVJCC17, CVJCC18, CVJCC19], y las Jornadas de Ciencia y Técnica [CVJCyT19], presentando dicha propuesta y alguno de sus resultados preliminares.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El trabajo de esta investigación se encuentra centrado entorno a los siguientes aspectos:

- Identificar detalladamente los problemas que manifiestan tener los estudiantes durante el primer curso de programación tanto en la faceta disciplinar como vincular con el medio académico.
- Implementar y adaptar las estrategias de enseñanza y aprendizaje activas, grupales, y colaborativas como la metodología de Enseñanza y Aprendizaje Basada en Proyectos dentro de las cátedras Programación I y Programación.
- Desarrollar e investigar sobre distintas estrategias didácticas conjuntas y combinadas para abordar las problemáticas técnicas vinculadas al aprendizaje de la programación, como así también, a la inclusión de todos los estudiantes en su heterogeneidad en el proceso.
- Desarrollar e integrar actividades que fomenten aspectos motivacionales en los estudiantes, las cuales brinden el acercamiento entre pares dentro de la carrera, reconocimiento dentro de la comunidad académica, y el ímpetu de continuar y progresar dentro de la carrera; brindando así herramientas que perduren en el tiempo y sean fundacionales para mantener el proceso de aprendizaje y así la permanencia y el avance dentro de la universidad.
- Trabajar con experiencias que permitan desarrollar las denominadas *soft-skill* (habilidades o competencias no medibles o difíciles de medir, de transmisión fundada en al experiencia) técnicas o disciplinares, como por

ejemplo, entender procesos de refinamiento y mejora de diseño de funciones, resolución de problemas, como aquellas *soft-skill* vinculadas más a aspectos profesionales como por el ejemplo las comunicacionales e interpersonales, desarrollo de la creatividad, y la capacidad de imaginar soluciones a problemas originales, desarrollo de capacidad crítica y objetiva ante otras soluciones, valorar respuestas disímiles a la propia, etc.

3. RESULTADOS OBTENIDOS/ESPERADOS

Cuando nos propusimos diseñar esta experiencia para el cursado del redictado de la materia Programación I y Programación nos trazamos ciertas metas y objetivos a alcanzar en pro de comenzar a transitar un camino de mejoras en la cátedra a fin de ayudar a los estudiantes tanto en la adquisición del conocimiento como integrarlos a la comunidad universitaria siendo ellos parte activas de ambos procesos.

Los objetivos de esta línea de investigación podemos clasificarlos en tres categorías:

a) Aptitudes de Programación:

(Primer Proyecto del Alumnado)

- Asimilar e incorporar los elementos esenciales de la programación.
- Reflexionar sobre las ventajas de descomponer un problema en subproblemas.
- Reconocer patrones y proponer abstracciones.
- Modularizar el problema usando funciones.

(Segundo Proyecto del Alumnado)

- Trabajar en forma incremental a partir de lo ya realizado, evolucionar desde el primer proyecto hacia uno más complejo.
- Favorecer la aprehensión de los contenidos, competencias y habilidades antes logrados.
- Reutilizar y rediseñar el código para adaptarlo a las nuevas necesidades de la segunda parte del proyecto.

b) Trabajo Grupal

- Permitir una construcción del conocimiento a partir del trabajo de todos, y la puesta en común de experiencias y saberes propios. La diversidad del grupo tiene que ser un punto a favor del mismo para el crecimiento de este a partir de la interrelación de las áreas de intereses individuales.
- Programar de a pares: para encarar un proyecto que presenta una dificultad superior a la ejercitación de la práctica de la materia.
- Favorecer el proceso de transformación del grupo en un verdadero equipo de trabajo mediante propuestas que favorezcan a la búsqueda conjunta de soluciones, posibles mejoras, detección de errores, distinguir particularidades y abstraer generalidades del problema a resolver.
- Fortalecer el vínculo entre compañeras y compañeros.

c) Integración a la comunidad Universitaria

- Favorecer la integración con el resto de la comunidad educativa afín.
- Ayudar a la permanencia en la facultad.
- Generar un espacio propio para la divulgación de los Proyectos I y II en el marco de las Jornadas de Ciencias de la Computación (JCC), actividad del Departamento de Ciencias de la Computación organizada por estudiantes y docentes de la carrera LCC cada año.

Los datos que analizaremos corresponden al relevamiento y las estadísticas de los cuatro años donde pudimos aplicar la estrategia en el aula, siendo este último año 2020 un año de experiencia virtual de la estrategia de trabajo lo cual nos permitió observar la fortaleza de la misma ya que no tuvimos que realizar mayores modificaciones y se pudo mantener el mismo espíritu de trabajo. La experiencia de estos cuatro años fue más que satisfactoria. Los estudiantes lograron cumplir los objetivos antes mencionados, alcanzando una alta participación

en la propuesta de los proyectos. En promedio, aproximadamente el 80% de los alumnos que participan regularmente del cursado realiza los proyectos, logran promover y regularizar la materia y rinde el examen final de la misma aprobando la materia.

Los estudiantes manifiestan en los cuestionarios que realizamos en cada proyecto los aspectos positivos que les dejó la realización de los mismos y cómo éstos los ayudaron a incorporar conocimiento, mantener la motivación alta para mejorar en cada etapa de su desarrollo, y relacionarse con el resto de los compañeros. Los datos del año en pandemia, nos muestran cifras similares a las anteriores, utilizamos para la exposición de los proyectos un sitio web denominado “Vidriera de Exposición de Proyectos”

(<https://sites.google.com/view/vidriera-proyectos-fceia/inicio?authuser=0>) donde expusimos los videos que los estudiantes entregaron a modo de divulgación y comunicación de su trabajo para con sus compañeros y el resto de la comunidad. Se generó una página web para cada grupo, junto con el video que cada grupo presentó y un formulario para que los estudiantes de otros grupos puedan observar de manera crítica y objetiva el trabajo realizado por los compañeros. La actividad fue completada por todos los estudiantes, ellos recibieron la misma con mucho entusiasmo y nos manifestaron a su vez que les permitió en la virtualidad conocer a sus compañeras y compañeros a los cuales desconocían completamente valorando aún más la experiencia. Parte de esta información también puede ser observada en la presentación en la Jornada “Compartiendo Experiencias de la FCEIA” realizada de manera virtual disponible online (<https://www.youtube.com/watch?v=UUCokZUnXbo&feature=youtu.be>, hora 2:06:11) donde junto a nuestra presentación se pueden escuchar las voces de los estudiantes con su visión y su vivencia sobre lo realizado.

Al completar todo el proyecto de investigación esperamos mostrar más resultados

y observar también cuestiones vinculadas a la permanencia en la carrera de aquellos estudiantes que cursaron los primeros cursos del redictado mediante encuestas que serán realizadas este 2021 y analizadas para divulgación en el transcurso del 2022 y estarán vinculadas a las primeras experiencias del año 2017.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está constituido por las autoras de este artículo, con dos cargos de Profesor Adjunto con Dedicación Simple. Conformamos un grupo de investigación en desarrollo buscando consolidarse en el tiempo. Este año, durante el cursado de la materia, se encuentra previsto la incorporación de algún auxiliar más a la cátedra como así también estudiantes adscriptos de las carreras Licenciatura en Ciencias de la Computación, Licenciatura en Matemática, y Profesorado en Matemática. Nuestra intención para con ellos es brindarles la capacitación y formación necesaria para comprender la metodología de trabajo, y para que puedan realizar tareas de acompañamiento a los estudiantes que cursan el redictado de forma tal que ellos puedan desde su propia experiencia realizar un apoyo a los nuevos estudiantes en el desarrollo de los proyectos de la cursada. Del mismo modo, la experiencia de acompañamiento y seguimiento de los proyectos en la cursada entrelazada con los lineamientos del proyecto didáctico del aula, les permitirá a los auxiliares y estudiantes adscriptos conocer en profundidad la estrategia de Aprendizaje Basada en Proyectos. También hemos diagramado un trabajo en conjunto con otras universidades, para compartir y enriquecer la experiencias al poder replicarlas en otros espacios académicos. Se espera que el desarrollo de esta línea de investigación contribuya a la formación de más recursos humanos en el campo de las Ciencias de la Computación y la Educación.

5. BIBLIOGRAFÍA

[L13] P. E. Martinez Lopez. Las Bases de la Programación. Publicado electrónicamente por la

Universidad Virtual de Quilmes, La Plata, 1Ed., 12/ 2013.

[BM05] T. Beaubouef and J. Mason. Why the high attrition rate for computer science students: Some thoughts and observations. SIGCSE Bull., 37(2):103-106, June 2005.

[CVJCYT19] N. Colussi and P. Viale. Actividades de Programación Grupales para Primer año de la Licenciatura en Ciencias de la Computación - Experiencias Didácticas en el Aula. Versión Extendida. In Jornadas de CyT de la UNR, (JorCyT), 2019.

[CVJCC17, CVJCC18, CVJCC19] N. Colussi and P. Viale. “¡Racketeando! Dibujos y Arte en Racket (2017), “¡Racketeando! Racket Logos y Diseños en Racket (2018). “¡Seguimos Racketeando! Foto- Racket- Gramas” (2019). Proyectos de Programación Grupales usando el Lenguaje Racket”. (JCC 2017, 2018, y 2019).

[DGPC14] G. Dapozo, C. Greiner, G. Pedrozo Petrazzini, and J. Chiapello. Vocaciones tic. ¿qué tienen en común los alumnos del nivel medio interesados por carreras de informática? In IX Congreso de Tecnología en Educación & Educación en Tecnología, pages 128-139, 2014.

[SB12] P. Sanchez , C. Blanco. Implantación de una metodología de aprendizaje basada en proyectos para una asignatura de Ingeniería del Software. Actas XVIII JENUI 2012, Ciudad Real, Universidad Nacional de Cantabrias, España. 2012

[SJ16] José Sanchez. ¿Qué dicen los estudios sobre el aprendizaje basado en proyectos?. Revista de Actualidad Pedagógica. 2016.

[Win06] J. M. Wing. Computational thinking. Commun. ACM , 49(3):33-35, March 2006.

[Phi09] P. Phillips. Computational thinking: A problem solving tool for every classroom. 2009.

[FFFK01] M. Felleisen, R. B. Findler, M. Flatt, and S. Krishnamurthi. How to Design Programs: An Introduction to Programming and Computing. MIT Press, USA, 2001.

[SBL10] Stephen Bloch. Picturing Programs, An Introduction to Computer Programming. College Publications, ISBN 978-1-84890-015-8, 2010.

Rediseño de una propuesta didáctica para enseñar a programar

Gustavo ASTUDILLO¹, Silvia BAST¹, Leandro CASTRO¹, Yamila MINETTI²

¹ Departamento de Matemática/FCEyN/UNLPam; ² FCH/UNLPam
astudillo@exactas.unlpam.edu.ar, silviabast@exactas.unlpam.edu.ar,
castro.leandro@exactas.unlpam.edu.ar; minettiyamila@humanas.unlpam.edu.ar

RESUMEN

En la primera etapa del proyecto de investigación “Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática”, se identificaron y categorizaron distintas propuestas de enseñanza de la programación. Esta investigación, permitió identificar algunas estrategias didácticas que podrían ser conjugadas con las que se implementaban en el dictado de la asignatura Introducción a la Computación. Sumado a lo anterior, el grupo de investigadores participó en la capacitación a docentes de nivel medio en el marco de un convenio firmado entre la UNLPam y la Fundación Sadosky.

Ambas actividades (investigación y docencia) dieron lugar a la redefinición de la propuesta de enseñanza en las prácticas de Introducción a la Computación. Conjugando la propuesta de Sadosky, basada en didáctica por indagación, y el enfoque de Thomson para la resolución de problemas (ya utilizado en la cátedra) se logró una propuesta superadora que dio soporte a las clases virtuales en el contexto de la pandemia por la COVID-19.

Palabras clave: propuesta de enseñanza, resolución de problemas, programación

CONTEXTO

El grupo de investigación GrIDIE¹ (Grupo de Investigación y Desarrollo en Innovación Educativa) enfoca, desde 2005, su investigación en tecnologías informáticas aplicadas en educación. Desde 2018, el grupo impulsa dos proyectos de investigación:

“Aprendizaje de las ciencias con tecnologías educativas” e “Incorporación de estrategias innovadoras en los procesos de enseñanza y de aprendizajes de informática”. Este último cuenta con cuatro líneas de investigación: (i) la medición y evaluación de IDEs para robótica educativa, (ii) la definición de criterios de evaluación que permitan identificar las posibilidades de la programación tangible para el aprendizaje de nociones básicas de programación, (iii) la revisión/evaluación de propuestas didácticas para la enseñanza/aprendizaje de la programación en función de su enfoque pedagógico/didáctico, y (iv) la implementación de una propuesta didáctica para el aprendizaje de la programación en el ámbito Universitario.

Los proyectos se desarrollan y son financiados por la Facultad de Ciencias Exactas y Naturales - UNLPam y fueron aprobados por resolución 27/18 CD-FCEyN.

1. INTRODUCCIÓN

En este artículo se presentan avances realizados durante el 2020 en “la implementación de una propuesta didáctica para el aprendizaje de la programación en el ámbito Universitario” del proyecto de investigación “Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática”.

Existe a nivel mundial, y en Argentina en particular, un auge de las propuestas didácticas para la enseñanza y el aprendizaje de la programación (Astudillo et al., 2019; Astudillo & Bast, 2020), así como también políticas que vienen acompañando este proceso. Entre estas últimas; en 2015, el Consejo Federal de Educación (CFE) estableció “que la enseñanza

¹ Facultad de Ciencias Exactas y Naturales-UNLPam

y el aprendizaje de la ‘Programación’ es de importancia estratégica en el Sistema Educativo Nacional” (Resolución CFE N. 263, 2015 Art. 1). En 2017, se lanza el Plan Nacional Integral de Educación Digital (PLANIED) (Resolución N. 1536-E, 2017), con la misión de “integrar la comunidad educativa en la cultura digital, favoreciendo la innovación pedagógica, la calidad educativa y la inclusión socio educativa” (art. 2). Como corolario de este proceso, en 2018, el CFE aprueba el documento Núcleos de Aprendizaje Prioritarios (NAP) para educación digital, programación y robótica, con el fin de “facilitar la integración del acceso y dominio de las tecnologías de la información y la comunicación en los contenidos curriculares indispensables para la inclusión en la sociedad digital” (Resolución CFE N. 343/18, 2018, p. 1). Esto último implica la incorporación efectiva de la temática al currículum.

A lo anterior se deben sumar las iniciativas de la Fundación Sadosky, que promueven el estudio de la programación en las escuelas argentinas y contribuye a la difusión de las carreras relacionadas con las Ciencias de la Computación. Para ello lleva adelante diversos programas, entre estos, Vocaciones en TIC y Program.ar (Dapozo et al., 2016). Particularmente, “Program.AR es una iniciativa que trabaja para que el aprendizaje significativo de Computación esté presente en todas las escuelas argentinas.” (Fundación Sadosky, 2020). El compromiso de la mencionada fundación se hace presente también en la línea de capacitación denominada “La Programación y su Didáctica” que se ofrece a docentes de nivel medio y es dictados por docente de Universidades Nacionales (de la cual participó la UNLPam en 2020²).

Tanto la investigación como la capacitación sentaron las bases para rediseñar nuestra propuesta didáctica. A continuación, se presenta la motivación, parte de la base teórica

que sustenta la propuesta y la secuencia didáctica.

1.1 Motivación

Los estudiantes que ingresan a la carrera Profesorado en Computación (FCEyN-UNLPam) participan de un Taller de Introducción a la Programación (TIP). Se trata de un grupo heterogéneo con mayoría de varones, de entre 18-30 años que, por lo general, se encuentran por primera vez con conceptos de programación (aunque en los últimos dos años esta tendencia se está revirtiendo). El TIP se enfoca, fuertemente, en los conceptos básicos de la programación (secuencia, estructuras de control, expresiones, datos simples y variables), dejando en un segundo plano la resolución de problemas (Astudillo et al., 2016). Por tal razón se plantea la incorporación de estrategias de resolución de problemas desde el inicio de la cursada de Introducción a la Computación.

Si bien enseñar es un término muy utilizado y que, en principio, es muy conocido por todos los docentes, conviene echar luz sobre el mismo antes de continuar. En la revisión sistemática, realizada por los autores sobre las propuestas de enseñanza/aprendizaje de la programación a nivel país (Astudillo & Bast, 2020) se encontró que, en algunos casos, una falta de precisión en términos pedagógicos (propósitos vs objetivos, bases pedagógicas que sustenten la propuesta didáctica, entre otras) y una asociación demasiado directa entre enseñanza y aprendizaje (por ejemplo, “enseñanza-aprendizaje”).

Dicen Basabe y Cols (2007) “De modo general, puede definirse a la enseñanza como un intento de alguien de transmitir cierto contenido a otra persona. Es una definición sencilla que sólo indica el tipo de actividad que puede designarse como «enseñanza» sin especificar nada acerca de las acciones de los

² Lista de Universidades que participaron en la convocatoria: <http://program.ar/wp-content/uploads/2015/04/Resultados-Convocatoria-2019-2020-Línea-B.pdf>

participantes, los recursos utilizables y los resultados esperables” (p. 126).

Entonces ¿Qué se debe tener en cuenta cuando se diseña una propuesta de enseñanza? En primer lugar, no hay una relación de tipo causal que permita asumir que la enseñanza conduce necesariamente al aprendizaje (Basabe & Cols, 2007). Otro aspecto importante es responder cómo, cuándo, dónde, para quiénes se llevará adelante la propuesta de enseñanza y cuáles son los saberes previos con los que cuentan los destinatarios de la propuesta (Anijovich et al., 2009).

1.2 La propuesta didáctica

La propuesta se basa en la resolución de problemas o desafíos. Pero ¿Qué es un problema en este contexto? “una situación sólo puede ser concebida como un problema en la medida en que existe un reconocimiento de ella como tal problema, y en la medida en que no dispongamos de procedimientos de tipo automático que nos permitan solucionarla de forma más o menos inmediata, sino que requieren de algún modo un proceso de reflexión o toma de decisiones sobre la secuencia de pasos a seguir. Esta última característica sería la que diferenciase un verdadero problema de situaciones similares como pueden ser los ejercicios.” (Pérez Echeverría & Pozo, 1994, p. 17).

La resolución de este tipo de situaciones problema implica, entonces, utilizar de modo estratégico técnicas ya conocidas (Pozo y Postigo, 1993 citado en Pozo et al., 1994). En este sentido dice Martínez López (2016) “Toda vez que se busca solucionar algún problema, es necesario contar primero con alguna idea de cómo encarar dicha solución, o sea, qué elementos disponer para la solución y de qué manera. Esto en programación se conoce como estrategia de solución: ¿qué cosas considerar a la hora de realizar la solución? ¿Cuáles son los componentes que interactuarán en la solución para obtener la respuesta deseada, y ¿de qué manera lo harán?” (p. 9).

La “solución de un problema exige una comprensión de la tarea, la concepción de un plan que nos lleve hacia la meta, la ejecución del mencionado plan y, por último, un análisis que nos lleve a determinar si hemos alcanzado o no la meta” (Pérez Echeverría & Pozo, 1994). En concordancia con lo anterior, desde la cátedra, se viene trabajando con el enfoque de Thomson quien propone ayudar a los estudiantes a programar a través de un ciclo que incluye: comprender el problema; diseñar el programa; escribir el programa y finalmente mirar hacia atrás (o "reflexión") (Thompson, 1997).

A este enfoque se lo complementa/integra con el de la Fundación Sadosky³. En la misma, basada en didáctica por indagación, la estrategia de solución es expresada a través de la división en subtareas y “representa a la forma de pensar composicionalmente y es una herramienta invaluable en el pensamiento de alto orden” (Martínez López, 2016, p. 9). Dicha estrategia es especificada a través del uso de procedimiento (herramienta del lenguaje). Cada solución a un problema (programa) inicia con un único procedimiento con un nombre que refleja la estrategia de solución (el problema como un todo), éste a su vez se irá dividiendo en tantas subtareas como sea necesario (nombrando cada una adecuadamente, iniciando con un verbo que describa la acción, para aportar a la legibilidad). “No debe utilizarse ni un solo comando primitivo antes de poder leer la estrategia de solución en el procedimiento principal y convencerse de que es adecuada; recién entonces es momento de definir cada uno de los procedimientos que expresan a las subtareas” (Martínez López, 2016, p. 11). Las herramientas del lenguaje también incluyen: comandos (acciones) que se expresan mediante verbos, y expresiones (datos) que lo hacen con sustantivos (apoyando la legibilidad) (Martínez López, 2016).

Esta combinación dio origen a la propuesta que se implementó durante 2020 en la cátedra de Introducción a la Computación. Cabe

³ Para detalles de la propuesta pude acceder a

<http://program.ar/material-curso-docentes/>.

aclarar que en la asignatura se hace uso del lenguaje Pascal y del IDE Lazarus para crear los programas que resuelven los desafíos planteados.

La propuesta se resume en:

1. Se presenta el desafío/problema al grupo de estudiantes. El mismo está centrado en el aprendizaje de un concepto y se basa en lo aprendido en desafíos previos.
2. Se pregunta al grupo de estudiantes *¿Qué hay que hacer?* y se registran/discuten las respuestas.
3. Se crea un mapa mental (plan) con base en una imagen del formulario que contiene los distintos objetos que resuelven el problema. La imagen del formulario forma parte del enunciado.
4. Se accede al IDE Lazarus, se confecciona el formulario y se crean los procedimientos que dan funcionalidad a cada objeto. Dentro de los procedimientos se refleja el plan a través de comentarios.
5. Se programa la solución, focalizando en cada uno de los comentarios que especifican el plan.
6. Se prueba el programa y se revisa la solución para conseguir mejoras en la misma.

Se busca que, a partir de este acompañamiento inicial, el grupo de estudiantes se apropie de esta estrategia para resolver problemas y la transfieran a los ejercicios que se proponen a continuación para consolidar los conceptos.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Este proyecto se desarrolla bajo la hipótesis de que es posible definir estrategias innovadoras para la enseñanza y el aprendizaje de temas de Ciencias de la Computación, utilizando distintas tecnologías de la información y la comunicación, que impacten positivamente en la motivación de los

estudiantes y en los diseños didácticos de los docentes.

Esto implica, por una parte, indagar y evaluar distintas tecnologías que permitan innovar y que faciliten el diseño de actividades y recursos educativos; y por otro, identificar las bases pedagógico-didácticas que sustenten y den sentido a la innovación.

Es así que se está trabajando en un proceso de medición y evaluación de la calidad de IDEs que permitan la implementación de actividades con robótica educativa, la definición de criterios de evaluación que permitan identificar las posibilidades de la programación tangible para el aprendizaje de nociones básicas de programación, la revisión de propuestas didácticas para enseñanza de la programación con el fin de identificar las bases pedagógicas que las sustentan y, finalmente, como se mostró en este artículo, se está implementando/evaluando una propuesta de enseñanza de la programación basada en la didáctica por indagación y la resolución de problemas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los avances realizados durante el año 2020 pueden mencionarse:

Se implementó la propuesta de enseñanza de la programación, descrita en este artículo, con buenos resultados iniciales.

Se comenzó con el re-diseño de las Guías Prácticas de la asignatura con vistas a una nueva implementación de la propuesta en 2021.

Se avanzó en la evaluación de las propuestas de enseñanza de la programación, esta investigación comenzó a arrojar los primeros indicios de una falta de descripción de las bases pedagógicas que sustentan la propuesta y algunos errores conceptuales desde la didáctica.

En cuanto a los trabajos futuros:

Se espera completar el trabajo de evaluación de las propuestas de enseñanza de la programación y definir una encuesta que permita validar las hipótesis definidas.

Se continuará trabajando en la medición del impacto de la propuesta didáctica diseñada y haciendo los ajustes necesarios en la misma.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan un investigador formado y tres investigadores en formación. Dos de ellos avanzando sobre sus tesis de doctorado y un tercero en su tesis para alcanzar el grado de magíster.

El proyecto cuenta con tres graduados que se inician en la investigación. Uno de ellos desarrollando su trabajo en el marco de una beca de estímulo a las vocaciones científicas y otro definiendo su proyecto de trabajo final de especialización, para ser presentado para su aprobación en el primer trimestre de 2021.

5. BIBLIOGRAFÍA

Anijovich, R., Mora, S., & Luchetti, E. (2009). *Estrategias de enseñanza: Otra mirada al quehacer en el aula* (Vol. 1). Aique.

Astudillo, G. J., & Bast, S. (2020). Enseñanza y aprendizaje de programación. Hacia un estado del arte. *Virtualidad, Educación y Ciencia*, 11(20), 138-155.

Astudillo, G. J., Bast, S. G., Segovia, D., & Castro, L. (2019). *Revisión de propuestas para la enseñanza de la programación*. XII Congreso de Tecnología en Educación y Educación en Tecnología, San Luis, Argentina.

Astudillo, G. J., Bast, S. G., & Willging, P. A. (2016). Enfoque basado en gamificación para el aprendizaje de un lenguaje de programación. *Virtualidad, Educación y Ciencia*, 7(12), 125-142.

Basabe, L., & Cols, E. (2007). La enseñanza. En A. Camilloni, *El saber didáctico* (1ra ed.). Editorial Paidós.

Resolución CFE N. 263, Pub. L. No. 263 (2015).

Resolución CFE N. 343/18, Pub. L. No. 343 (2018).
https://www.argentina.gob.ar/sites/default/files/res_cfe_343_18_0.pdf

Dapozo, G. N., Petris, R. H., Greiner, C. L., Espíndola, M. C., Company, A. M., & López, M. (2016). Capacitación en programación para incorporar el pensamiento computacional en las escuelas. *TE & ET*, no. 18.
<http://hdl.handle.net/10915/58516>

Fundación Sadosky. (2020). ¿De qué se trata? | Program.AR [Blog]. *Program.AR*.
<http://program.ar/de-que-se-trata/>

Martínez López, P. E. (2016). *Sugerencias para el dictado del curso La programación y su didáctica. Método Program. AR*. Fundación Sadosky. http://program.ar/wp-content/uploads/2015/04/GuiaParaCursoProgram.AR_1.pdf

Resolución N. 1536-E, Pub. L. No. 1536-E (2017).
<https://www.boletinoficial.gob.ar/detalleAviso/primera/161556/20170404>

Pérez Echeverría, M. del P., & Pozo, J. I. (1994). Aprender a resolver problemas y resolver problemas para aprender. En *La solución de problemas* (pp. 14-50). Santillana.

Pozo, J. I., Pérez, M. d., Domínguez, J., Gómez, M., & Postigo, Y. (1994). *La solución de problemas*. Santillana Madrid.

Thompson, S. (1997). Where do I begin? A problem solving approach in teaching functional programming. En H. Glaser, P. Hartel, & H. Kuchen (Eds.), *Programming Languages: Implementations, Logics, and Programs* (pp. 323-334). Springer.
<https://doi.org/10.1007/BFb0033853>

REDISEÑO EDUCATIVO PARA LA ENSEÑANZA Y APRENDIZAJE DEL CÁLCULO NUMÉRICO

Angel R. Barberis, Lorena E. del Moral, Jorge A. Silvera, Eusebio Méndez, Nicolás Rojas
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

RESUMEN

La programación de computadoras integra un conjunto de actividades que demanda del estudiante, no solo esfuerzo y dedicación, sino también otras capacidades como la resolución de problemas y el pensamiento algorítmico. En los inicios de la vida universitaria, los alumnos experimentan una amplia gama de dificultades y deficiencias, que sumado a una metodología de enseñanza desactualizada de una asignatura hacen que el arte de aprender a programar sea una de las primera y más desafiante tarea que enfrentan los estudiantes de informática en los primeros años de carrera. Las dificultades se extienden e impactan fuertemente en otras asignaturas que contemplan a la programación como uno de sus objetivos en la práctica académica. Para revertir este panorama es necesario investigar y desarrollar nuevas metodologías de enseñanza, aprendizaje y prácticas de la programación en un ambiente centrado en el alumno. El presente trabajo expone algunas actividades y resultados de investigación alcanzados en el marco de un proyecto que busca formalizar una metodología de enseñanza y aprendizaje de la programación algorítmica, bajo el enfoque activo, participativo y centrada en el alumno.

Palabras claves:

Metodología de enseñanza del Cálculo Numérico, Aula Invertida, Scrum como metodología para el entrenamiento en la Programación de computadoras.

CONTEXTO

El proyecto se centra en el rediseño de las estrategias de enseñanza y aprendizajes tradicionales, en el que se busca un proceso que permita organizar y desarrollar nuevas actividades educativas, cuya reestructuración satisfaga las necesidades formativas de los estudiantes en el nuevo mundo de la Sociedad de la Información. El rediseño del proceso de formación se realiza en el contexto de la asignatura Programación Numérica que se dicta en el segundo año de la carrera Licenciatura en Análisis de Sistemas en la Universidad Nacional de Salta, y contempla la adecuación de las estrategias de enseñanza tanto para las clases prácticas como para las teóricas. Se propone por un lado, el desarrollo de clases teóricas con una estrategia didáctica centradas en el alumno y basada en el aprendizaje por descubrimiento en un contexto de Aula Invertida, que les permita a los discentes una interacción dinámica, activa y participativa en clases aúlicas; por el otro, el desarrollo de clases prácticas centrada en el aprendizaje cooperativo y colaborativo, facilitando así la creación de un ambiente en el que los estudiantes se entrenan en la programación de computadoras bajo el enfoque del desarrollo ágil de software a través de Scrum (en lugar de realizar meras practicas computacionales).

El presente trabajo se lleva a cabo dentro del marco del proyecto de investigación Nro: 2.536/19 “Rediseño educativo para el aprendizaje de cálculo numérico”, aprobado en el año 2.019 por el consejo de investigación de la Universidad Nacional de Salta.

1. INTRODUCCION

Durante los últimos años se han identificado problemas intrínsecos vinculados al cursado de materias como Cálculo Numérico y Programación de Numérica, relacionados con una práctica débil en la programación de aplicaciones en asignaturas previas, que influía severamente en el normal cursado de las mismas. Además, se detectaron otros inconvenientes vinculados al perfil psicológico que provocaba una falta de motivación, de interés por el aprendizaje e integración activa a las clases, tanto teóricas como prácticas. El impacto de estos inconvenientes derivaba en el abandono del cursado de la asignatura, o bien un bajo rendimiento académico.

El arte de la programación de computadoras es una tarea compleja y difícil de abordar académicamente [1]. La complejidad del Proceso Educativo de la Programación radica en que éste demanda la interacción de habilidades tanto del profesor como de los alumnos, y exige la garantía de que el educador propicie un ambiente cooperativo y colaborativo para desarrollar en el discente otras habilidades como las psico-cognitivas [2] y trabajo en equipo [3], entre otras, necesarias para el abordaje de problemas multidisciplinarios [4]. Por lo tanto, es de suma importancia contar con una estrategia metodológica de enseñanza y aprendizaje que propicie un ambiente de trabajo en grupo, en el que se pueda fomentar tanto, habilidades sociales como de comunicación, haciendo del hábito de ayudar, compartir, colaborar y cooperar, una norma inexcusable en el aula.

En la era del desarrollo tecnológico, no sólo importa la disponibilidad de la información, el conocimiento y los medios para comunicarla, sino también, el modo en que ellos puedan ser aplicados en prácticas reales. El desarrollo de habilidades inherentes a la programación de computadoras (creatividad [5], autoeficacia [6], resolución

de problemas [7], razonamiento [8], trabajo en equipo [3], etc.), desarrolla en el alumno capacidades multifacéticas que le permiten enfrentar problemas interdisciplinarios de diferentes grados de dificultad, que sumado a un buen entrenamiento, adquiere la experiencia de un buen programador. Las estimulaciones cognitivas inherentes a la resolución de problemas (actividades de exploración, análisis y búsqueda de soluciones) [7], estimulan un proceso de aprendizaje, que favorece el desarrollo mental, colocan en primer plano las destrezas de investigación, los entrena en la generación de soluciones, y con ello, los estudiantes se encaminan hacia el mayor desafío de dobligar las capacidades de programación de computadoras. La actividad de resolución de problemas en programación necesita además de la habilidad técnica para sintetizar o resumir una solución [9]. Esta habilidad junto al trabajo en equipo se puede desarrollar mediante la práctica constante en el aula, a través de una estrategia pedagógica adecuada. Bajo esta línea de pensamiento, y la oportunidad del trabajo en equipos para la adquisición de habilidades tanto cognitivas y situadas [10] como las de resolución de problemas [11]; y las competencias relacionadas con la programación de computadores [12] es que surge la necesidad reestructurar los modos de enseñanza para contemplar grupos de estudios que interactúan en la resolución de problemas con integración de las tecnologías existentes bajo el enfoque de Aula Invertida.

El Aula Invertida es uno de los llamados enfoques y metodologías de aprendizaje activo en la educación superior del siglo XXI, donde se impulsa y promueve la participación del estudiante en la clase presencial, y de forma autónoma y desestructurado, en actividades previas diseñadas por el Docente. Es un modo de enseñanza contrapuesto a los de aula tradicional donde el estudiante asiste a escuchar clases magistrales.

Algunos investigadores, visualizan al Aula Invertida, como una estrategia pedagógica que se focaliza en la importancia del uso del

tiempo de clase para la construcción del significado, más que para la transmisión de información [13]. Así mismo, promueve el aprendizaje teórico de la currícula de una asignatura fuera del ámbito de la clase (aprendizaje diferido), y procura consolidar los procesos de enseñanza y aprendizaje durante la clase presencial en el aula, a través de actividades experienciales y de resolución de problemas (aprendizaje activo) [14]. El estudiante no accede a los contenidos teóricos desde una posición totalmente pasiva mediante una clase magistral del docente, si no que antes de la clase presencial, puede acceder a ellos mediante estructuras tecnológicas como los entornos virtuales a través de diferentes recursos (videos, audios, presentaciones interactivas, animaciones, textos, páginas web, etc.) [15], con la libertad de hacerlo en el momento y lugar que deseen, y a su propio ritmo. En este contexto, las clases teóricas presenciales pueden ser aprovechadas para una participación más activa, en el que los estudiantes analizan, aplican, evalúan e incluso crean situaciones significativas de aprendizaje; tendiendo, en todos los casos, al desarrollo de habilidades cognitivas de orden superior considerados en la taxonomía de Bloom [16], y al trabajo cooperativo y colaborativo. Esto no implica que no se puedan plantear también actividades y trabajo colaborativo por medios virtuales (hay numerosos recursos digitales para propiciarlo), aunque se dejen para esa instancia, lo que implique habilidades de orden inferior. Así, el docente no solo puede propender al aprendizaje activo, sino también al desarrollo del pensamiento crítico y creativo del estudiante, empleando metodologías basadas en el aprendizaje en equipo, basadas en problemas y en estudios de caso [17]. Esto hace que el rol del docente se reconceptualice como mediador o guía, y no solamente como expositor [18]. Luego de las clases presenciales, se pueden derivar actividades post-clase, donde el estudiante manteniendo su rol activo, transfiere y aplica los conocimientos logrados [19].

Una de las metodologías activas que se favorece en la clase invertida, es el aprendizaje cooperativo. El aprendizaje cooperativo consiste en trabajar juntos para alcanzar objetivos comunes, maximizando el propio aprendizaje y el de los demás [13]. Si este concepto es aplicado en las clases teóricas, la estrategia didáctica de Aula Invertida se ve favorecida al desarrollarse en clase casos de estudios, métodos y técnicas de aplicabilidad en forma conjunta y en cooperación con los grupos de estudios. El aprendizaje cooperativo en el aula, privilegia el desarrollo de actividades y de la comunicación, intensificando y diversificando la participación de los alumnos en clases. Al mismo tiempo, exige la bidireccionalidad necesaria del proceso, entre, el que guía y orienta la actividad y el aprendiz [20]. El trabajo en equipo propiciado por el aprendizaje cooperativo, y la estrategia pedagógica de Aula Invertida constituye una alternativa que promete una enseñanza centrada en el alumno; con mejoras sustanciales en la dinámica de clases, lo que provocaría en el estudiante una acción activa, comprometida, de mayor participación e interés por la asignatura [21].

Para una exitosa consolidación del aprendizaje durante las clases teóricas, se estima necesario desarrollar objetos de aprendizajes adecuados para la correcta asimilación de los conceptos. Los Objetos de Aprendizajes (OA) es un contenido educativo digital cuya finalidad última es el aprendizaje del usuario y que, en sí mismo, constituye o puede llegar a constituir, mediante su integración con otros objetos más simples (texto, imágenes, audio, video, etc.), un material educativo multimedia [22]. Básicamente, un OA es un conjunto de recursos digitales, autocontenible y reutilizable, con un propósito educativo y constituido por al menos tres componentes internos: contenidos, actividades de aprendizaje y elementos de contextualización [23]. La característica de ser autocontenible, implica que sus elementos internos deben tener sentido por sí mismos y ser autosuficientes, para el logro del objetivo de

aprendizaje para el cual fue desarrollado. Por esta razón, puede usarse tanto en la enseñanza presencial como en procesos de educación virtual o a distancia [24]. Esta doble usabilidad, facilita el estudio independiente de los estudiantes, contribuyendo así, a la formación de su autonomía frente al conocimiento, uno de los mayores objetivos educativos, que exige con mayor firmeza la educación de seres humanos capaces de aprender por sí solos.

Por otro lado, para las clases prácticas, se implementa una estrategia metodológica que permite a las cátedras de Programación Numérica y Cálculo Numérico entrenar a sus alumnos en el desarrollo ágil de software como un enfoque superador a la mera práctica de la programación de lenguajes, permitiendo al mismo tiempo, promover acciones cooperativas y colaborativas en equipos de trabajos, que les facilita la adquisición de habilidades propias del programador profesional. Básicamente, la estrategia consiste en la apropiación de los conceptos del Aprendizaje Cooperativo Basado en Problemas, diseñadas para separar y resaltar aspectos importantes de la programación y la resolución de problemas, bajo el marco referencial de Scrum. En pocas palabras, se trata de llevar la realidad laboral al aula, en un ambiente simplificado y controlado por los docentes.

El *Aprendizaje Basado en Problemas* (ABP), resulta del proceso de trabajar sobre la comprensión y la resolución de problema, donde el problema es un elemento importante en el proceso de aprendizaje. El marco teórico del ABP establece características [25], que combinadas con las actividades de la metodología del desarrollo ágil de software de Scrum [26], propicia una ambiente apto para desarrollar un sistema de aprendizaje en el que el único protagonista es el alumno.

Las actividades investigadas y que se desarrollan en clases prácticas persiguen tres objetivos básicos, por el cual, se las agrupan en tres categorías [27]. El primero de los objetivos, busca fomentar la inclusión para

que el proceso de formación sea equitativo y esté al alcance de todos. Las dificultades de aprendizajes en asignaturas previas, provocan en un porcentaje de alumnos falta de motivación que los conduce a una auto marginación del grupo activo de la clase, con el consiguiente abandono del cursado. Las actividades que se desarrollan bajo este objetivo son:

- *Realizar Evaluación Diagnóstica Inicial.*
- *Realizar Adecuación Pedagógica de las estrategias de enseñanza.*
- *Identificar programadores entusiastas.*
- *Organizar Grupos de Trabajos Estudiantiles con un programador entusiasta (en lo posible).*
- *Fomentar tutorío de pares.*

Con el segundo de los objetivos se busca equilibrar formalismo teórico-práctico con pragmatismo para el mejoramiento académico en la realización de las prácticas. Uno de los mayores inconvenientes que se ha detectado en la enseñanza curricular de las asignaturas de Programación de Numérica y Cálculo Numérico, es que las clases (teórica y prácticas) no eran lo suficientemente pragmática, provocando que ante la falta de motivación, los alumnos se limitasen a la asimilación memorista, y no tengan un panorama más amplio de la aplicabilidad práctica de los nuevos conceptos. Las actividades que apoya el segundo objetivo son:

- *Diseñar Casos de Estudios que favorezcan la rápida asimilación de conceptos.*
- *Proporcionar el marco de actuación cooperativa de Scrum.*
- *Entrenar para la socialización.*
- *Entrenar para un desempeño ágil en la programación de aplicaciones.*
- *Entrenar para enfrentar situaciones de contingencia.*

El tercer objetivo es, disponer de las acciones evaluativas, como verdaderas herramientas de adecuación de la metodología a las características pedagógicas de los alumnos. Las actividades que dan soporte al objetivo son:

- *Analizar resultados de la evaluación diagnóstica inicial.*
- *Realizar evaluación de proceso y formativo.*
- *Realizar evaluación final y formativa.*
- *Realizar autoevaluación docente y continua.*
- *Evaluar la efectividad de la aplicación estratégica de la Metodología en cada período lectivo.*

Los detalles de cada actividad se describe en [27].

3. LINEAS DE INVESTIGACION y DESARROLLO

Los principales ejes temáticos que se investigan son los siguientes:

- » Actividades pedagógicas que propicien un ambiente de participación activa y comprometida que facilite el aprendizaje significativo, en un contexto de Aula Invertida.
- » Actividades pedagógicas centradas en el alumno, que propicie la adquisición de habilidades para la resolución de problemas, el desarrollo del pensamiento algorítmico, en un contexto de desarrollo ágil de software.

4. FORMACION DE RECURSOS HUMANOS

La estructura del equipo de investigación es de 5 (cinco) miembros incluidos el Director.

Uno de sus miembros desarrolla la tesis de la Especialidad en Tecnología Multimedia para Desarrollos Educativos; posgrado que se dicta en la Universidad Nacional de Córdoba.

Otro de sus miembros, desarrolla la tesis de la Especialidad en Psicopedagogía Institucional; posgrado que se dicta en la Universidad Nacional de Salta.

Por otro lado, uno de los miembros desarrolla la tesina para acceder al título de Técnico Universitario en Programación; carrera que se dicta en la Facultad de Cs. Exactas de la Universidad Nacional de salta.

4. RESULTADOS OBTENIDOS / ESPERADOS

La estrategia metodológica basada en el entrenamiento de la programación ágil en lugar de la mera práctica de contenido, significó para las clases prácticas, un cambio radical de pensamientos y de las formas de abordar las guías de trabajos prácticos. Con el tutorero de pares se logró motivar e incluir a los alumnos auto marginados al grupo activo de las clases.

En cuanto a las clases teóricas, se siguen diseñando OA en un contexto de Aula Invertida. Se ha realizado una prueba piloto con algunos temas de la currícula obteniéndose resultados muy alentadores, tanto del punto de vista humano como académico. Se espera que al finalizar el proyecto se tenga una metodología de enseñanza depurada que le permita a las cátedras de Programación Numérica y Cálculo Numérico mejorar el rendimiento académico y disminuir la tasa de abandono.

5. BIBLIOGRAFIA

- [1] Sarpong K. A.-m., Arthur J. K. and Owusu Amoako P. Y. (2013). Causes of Failure of Students in Computer Programming Courses: The Teacher - Learner Perspective. *International Journal of Computer Applications (IJCA)*. Vol. 77 (12):27-32. doi: 10.5120/13448-1311
- [2] Milne I. and Rowe G. (2002). Difficulties in Learning and Teaching Programming -- Views of Students and Tutors. *Education and Information Technologies*. Vol. 7 (1):55-66. doi: 10.1023/a:1015362608943
- [3] Sancho-Thomas P., Fuentes-Fernández R. and Fernández-Manjón B. (2009). Learning teamwork skills in university programming courses. *Computers & Education*. Vol. 53 (2):517-531. doi: 10.1016/j.compedu.2009.03.010
- [4] Roberts F. S. (2011). The Challenges of Multidisciplinary Education in Computer Science. *Journal of Computer Science and Technology*. Vol. 26 (4):636-642. doi: 10.1007/s11390-011-1164-1

- [5] Hershkovitz A., Sitman R., Israel-Fishelson R., Eguiluz A., Garaizar P. and Guenaga M. (2019). Creativity Inside and Outside Programming Learning. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. pp. 293-298. Association for Computing Machinery. Tempe, Arizona, USA.
- [6] Tsai M.-J., Wang C.-Y. and Hsu P.-F. (2018). Developing the Computer Programming Self-Efficacy Scale for Computer Literacy Education. *Journal of Educational Computing Research*. Vol. 56 (8):1345-1360. doi: 10.1177/0735633117746747
- [7] Kotovsky K. (2003). Problem Solving – Large/Small, Hard/Easy, Conscious/Nonconscious, Problem-Space/Problem-Solver: The Issue of Dichotomization. *The Psychology of Problem Solving*. pp. 373-384. Cambridge University Press. Cambridge. doi: 10.1017/CBO9780511615771.013
- [8] Fox R. and Farmer M. (2011). The Effect of Computer Programming Education on the Reasoning Skills of High School Students. *Paper of The 2011 International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS'11)*. July 18-21, 2011. WorldComp 2011 Proceedings.
- [9] López-Cruz O., Mora A. L., Sandoval-Parra M. and Espejo-Gavilán D. L. (2017). Teaching Computer Programming as Knowledge Transfer: Some Impacts on Software Engineering Productivity. *Proceedings of Trends and Applications in Software Engineering*. pp. 145-154. Springer International Publishing. Cham.
- [10] Díaz Barriga F. (2003). Cognición situada y estrategias para el aprendizaje significativo. *Revista Electrónica de Investigación Educativa (Redie)*. Vol. 5 (2):1-13. <https://redie.uabc.mx/redie/article/view/85>
- [11] Corral-Lage J., Ipiñazar-Petralanda I. and Ipiñazar-Petralanda I. (2015). Aplicación del aprendizaje basado en problemas en la asignatura contabilidad financiera superior: ventajas y desventajas. *Tendencias Pedagógicas*. Vol. 23:45-60.
- [12] De Buenaga Rodríguez M., Ortega Ortiz de Apodaca M., Monsalve Piqueras B. and Mata Ortega M. (2004). Desarrollo de competencias generales en los estudios de Informática: la experiencia en la Universidad Europea de Madrid. *Jornadas de Innovación Universitaria*.
- [13] Abío G., Alcañiz M., Gómez-Puig M., Rubert G., Serrano M., Stoyanova A. and Vilalta-Bufi M. (2017). El aula invertida y el aprendizaje en equipo: dos metodologías para estimular al estudiante repetidor. *Revista d'Innovació Docent Universitària (RIDU)*. Vol. 9:1-15. doi: 10.1344/RIDU2017.9.1
- [14] Mok H. N. (2014). Teaching Tip: The Flipped Classroom. *Journal of Information Systems Education*. Vol. 25 (1):7-12.
- [15] Espinoza Pastén L. M. and Araya Cortés A. A. (2019). Clase invertida y aprendizaje cooperativo en postgrado: una experiencia en Chile. *Educere*. Vol. 23 (75):477-486.
- [16] Gilboy M. B., Heinerichs S. and Pazzaglia G. (2015). Enhancing student engagement using the flipped classroom. *Journal of nutrition education and behavior*. Vol. 47 (1):109-114. doi: 10.1016/j.jneb.2014.08.008
- [17] Betihavas V., Bridgman H., Kornhaber R. and Cross M. (2016). The evidence for 'flipping out': A systematic review of the flipped classroom in nursing education. *Nurse Education Today*. Vol. 38:15-21. doi: 10.1016/j.nedt.2015.12.010
- [18] Jensen J. L., Kummer T. A. and Godoy P. D. d. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE life sciences education*. Vol. 14 (ar5):1-12. doi: 10.1187/cbe.14-08-0129
- [19] Abeysekera L. and Dawson P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research & Development*. Vol. 34 (1):1-14. doi: 10.1080/07294360.2014.934336
- [20] Ferreiro Gravié R. (2006). *Estrategias didácticas del aprendizaje cooperativo*. 2 Ed. Editorial Trillas, Mexico.
- [21] Fortanet van Assendelft de Coningh C., González-Díaz C., Mira Pastor E. and López Ramón J. (2013). Aprendizaje cooperativo y flipped classroom. Ensayos y resultados de la metodología docente pp. 1151-1162. Universidad de Alicante. Vicerrectorado de Estudios, Formación y Calidad. <http://hdl.handle.net/10045/43329>
- [22] Marzal M. Á., Prado J. C. and Burgoa E. R. (2015). Objetos de aprendizaje como recursos educativos en programas de alfabetización en información para una educación superior de posgrado competencial. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*. Vol. 29 (66):139-168. doi: 10.1016/j.ibbai.2016.02.029
- [23] Chiappe Laverde A., Segovia Cifuentes Y. and Rincón Rodríguez H. Y. (2007). Toward an instructional design model based on learning objects. *Educational*

Technology Research and Development. Vol. 55 (6):671-681. doi: 10.1007/s11423-007-9059-0

[24] Chiappe Laverde A. (2009). Acerca de lo Pedagógico en los Objetos de Aprendizaje-Reflexiones Conceptuales hacia la Construcción de su Estructura Teórica. *Estudios Pedagógicos*. Vol. XXXV (1):261-272.

[25] Barrows H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*. Vol. 1996 (68):3-12. doi: 10.1002/tl.37219966804

[26] Alaimo M. (2013). *Proyectos Ágiles con Scrum*. Flexibilidad, aprendizaje, innovación y colaboración en contextos complejos. 1 Ed. (np. 126) Kleer Agile Coaching & Training, Buenos Aires, Argentina.

[27] Barberis A. R., Del Moral L. E., Silvera A. J. and Méndez E. (2020). Metodología para un Entrenamiento Pedagógico de la Programación de Computadoras. *Proceedings of 8º Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2020)*. UTN Facultad Regional San Francisco. Córdoba, Argentina.

Ingeniería de Software

APLICACIÓN DE METODOLOGÍAS PARA CONTRIBUIR CON LA CALIDAD DEL SOFTWARE EN SISTEMAS DE GESTIÓN Y SISTEMAS CRÍTICOS

Irrazabal, Emanuel Dapozo; Gladys N.; Greiner, Cristina;
Company, Ana M.; Mascheroni, Agustín, Acevedo, Joaquín

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
{eirrazabal, gndapozo, cgreiner, eirrazabal, mascheroni }@exa.unne.edu.ar,
anamacom@gmail.com

RESUMEN

Este es el cuarto y último año del proyecto F018-2017; una continuación de los proyectos F07-2009 y F10-2013, ambos enfocados en modelos, métodos y herramientas para la calidad del software. Este tercer proyecto aborda los temas emergentes en el área de la calidad de software, en particular, aspectos vinculados con la validación de requerimientos software, la gestión de los procesos administrativos mediante herramientas software y la entrega continua de productos software. El equipo de trabajo busca promover y generar métodos y herramientas que contribuyan a mejorar la calidad del proceso y del producto software.

En particular, se está trabajando en el desarrollo de modelos de prueba continua dentro de la disciplina de entrega continua, en la eficiencia de los sistemas en organismos públicos mediante el enfoque de procesos de negocios, la construcción de un modelo de procedimientos para la gestión de requerimientos en entidades agrícolas con una cultura organizacional jerárquica y la validación de requerimientos software. A la par, se está desarrollando el proyecto PI-F17-2017 donde en el último año se han realizado investigaciones en calidad de la arquitectura software para sistemas críticos de altas prestaciones.

Palabras clave: calidad de software,

gestión de procesos de negocio, entrega continua, ingeniería de requerimientos

CONTEXTO

Las líneas de Investigación y Desarrollo presentada en este trabajo corresponden a los proyectos PI-F17-2017 “Análisis e implementación de tecnologías emergentes en sistemas computacionales de aplicación regional” y PI-17F018 “Metodologías y herramientas emergentes para contribuir con la calidad del software”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021.

En este marco se acompaña con el asesoramiento al proyecto PID 7057 sobre "Ingeniería de Software", de la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos (UNER).

1. INTRODUCCIÓN

Las características de un producto software necesitan ser cada vez más complejas para satisfacer a los usuarios [1]. Esto genera avances constantes que se traducen en una gran cantidad de nuevas técnicas y tecnologías relacionadas con el desarrollo de software [2]. Pero el desafío sigue siendo llevar adelante un proceso de desarrollo que termine en un plazo acordado con unos atributos de calidad esperados, como, por ejemplo la mantenibilidad o la usabilidad.

En este sentido, la calidad del software está formada por diferentes factores y apoyados por herramientas que automatizan su medición. Por lo tanto, las nuevas tecnologías y técnicas emergentes muchas veces no tienen el conjunto de herramientas necesarias para la medición y mejora de la calidad del producto software resultante [3]. Por ello la necesidad de investigar en el uso y adecuación de las tecnologías emergentes incorporando aspectos de calidad del software.

Siguiendo con este mismo enfoque y con el objetivo de lograr una mayor satisfacción del cliente, es crítico para las organizaciones entregar productos de calidad de manera aún más rápida. Esto dio lugar al surgimiento de un nuevo enfoque denominado “Entrega Continua de Software”, más conocido en inglés como Continuous Delivery (CD). En este enfoque los equipos mantienen la producción de software en ciclos cortos de tiempo, asegurando que el producto pueda ser lanzado de manera fiable en cualquier momento [4] [5].

Nuevamente, uno de los principales problemas de estos enfoques estaría en la calidad del producto software. Ésta puede disminuir, dado que, al realizarse los despliegues del sistema con mayor frecuencia, aparecen más defectos en el producto [5]. Por tanto, es esencial desarrollar un enfoque de priorización de los diferentes aspectos en la calidad del producto software, teniendo en cuenta la forma de trabajo actual de las empresas de desarrollo software.

Para adoptar este enfoque, se utiliza el concepto de Tubería de Despliegue (DP - Deployment Pipeline), un estándar para automatizar el proceso de CD [6].

Desde el enfoque de los procesos de construcción también existen desafíos, no solamente desde el sector privado sino también en el sector público. La gestión de procesos de negocio se basa en la idea de que cada producto es el resultado de un conjunto de actividades que se realizan a fin de obtener dicho producto. Por este

motivo, la correcta y eficiente gestión de los procesos de negocio es un aspecto importante para la productividad de toda organización, ya que permite identificar las tareas, el orden de ejecución de las mismas y las personas responsables de realizarlas [7]. En las entidades privadas con una estructura piramidal también se ven los mismos problemas, que pueden ser atacados mediante el ajuste de los procedimientos y el desarrollo de personalizaciones, especialmente en la obtención y validación de los requerimientos.

En la línea de Ingeniería de Software para sistemas críticos, cada vez es más común trabajar con sistemas de propósitos dedicados, especialmente en aplicaciones como la de los procesos industriales, la automotriz, o la aviónica. En particular, ciertas aplicaciones son usadas en entornos críticos de tal manera que los fallos podrían provocar pérdidas financieras o incluso pérdida de vidas humanas [8]. Como respuesta a esto existen marcos regulatorios que estipulan la necesidad de demostrar la seguridad del sistema construido. Respecto del software embebido los estándares principales en sistemas críticos provienen de la normativa IEC 61508 – parte 3. En particular, en el apartado 7.4.3 de la norma IEC 61508-3 se especifican las buenas prácticas al construir la arquitectura del software en términos de actividades, documentación, especificación integral de cada módulo de la arquitectura y uso de buenas prácticas de programación. Actualmente existen diferentes estudios que presentan arquitecturas software para sistemas críticos, pero no emergen arquitecturas estándar por nivel de seguridad integral y orientadas a cubrir la mayor cantidad de dominios de problemas.

Por lo tanto, las líneas de trabajo del proyecto tienen que ver con características de calidad de calidad del producto software y del proceso de desarrollo, tanto en sistemas tradicionales como en sistemas críticos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se describen las siguientes líneas de investigación y desarrollo:

Entrega continua

Se está realizando el desarrollo de un modelo de mejora y evaluación de pruebas continuas de software teniendo en cuenta los principales problemas presentados por la literatura científica y la industria. Como primera medida se han realizado revisiones sistemáticas de la literatura y como segundo paso el desarrollo de encuestas a grandes empresas mundiales.

El paso siguiente fue la creación de un primer modelo formado por distintas etapas de pruebas basándose en la Tubería de Despliegue. Actualmente el modelo se está validando en un proyecto argentino a nivel nacional y en 10 proyectos internacionales con equipos ágiles.

Gestión de requerimientos y de procesos de negocio

La validación de requerimientos es un proceso continuo en el proyecto de desarrollo de software con el fin de asegurar que los requerimientos elicidados sean representaciones exactas de las necesidades y expectativas de los usuarios. Esta actividad contribuye a mejorar la calidad de los requerimientos, reducir costos, tiempos y riesgos en el desarrollo de software. En este proyecto se aborda el estudio de técnicas y herramientas libres para la validación de requerimientos de software con el objetivo de contribuir con información que apoye la actividad de los profesionales del sector de desarrollo de software, en cuanto a lograr la integridad de la especificación de los requerimientos de software [9].

Actualmente se está trabajando en el desarrollo de procedimientos para la gestión de requerimientos tanto en el sector público como en el sector privado.

Respecto del sector público el objetivo es lograr un prototipado rápido de aplicaciones con alto contenido de procedimental, pasos secuenciales y controles cruzados de información apoyado por esquemas visuales de procesos de negocio. una de las herramientas más utilizadas es Bonita [10], En esta línea de investigación también se busca realizar un modelo de trabajo para mejorar la gestión del proceso administrativo de las instituciones públicas. Por un lado, el modelado mediante la notación BPM y por otro lado el desarrollo de prototipos rápidos mediante la herramienta antes descrita.

Finalmente, para el sector privado se están realizando pruebas piloto de desarrollos personalizados para la gestión ágil de requerimientos software. En particular, los modelos propuestos buscan mejorar las validaciones en entornos de empresas tradicionales.

Arquitectura software para sistemas críticos

Esta línea de trabajo ha construido un modelo de procesos para la gestión de requerimientos de software en sistemas críticos, la comparación de patrones de arquitectura de sistemas críticos respecto de la mantenibilidad, el análisis de estrategias de arquitectura, la gestión de la calidad del producto software y, finalmente, un modelo de mejora continua basada en ISO 9001. Este año se trabajará en el diseño de componentes software de acuerdo con las buenas prácticas de la norma IEC 61508-3.

3. RESULTADOS

OBTENIDOS/ESPERADOS

En el marco de este proyecto y respecto de la línea de gestión de procesos se lograron los siguientes objetivos:

En la temática sobre investigación sobre validación de requerimientos, se realizó un análisis comparativo de las herramientas para determinar la performance de su

aplicación en base a determinados criterios de evaluación establecidos previamente [9] y se evalúan técnicas para la validación de requerimientos en entornos de trabajo para el desarrollo de software [11]

En [12] se realizó una revisión sistemática sobre las técnicas, métodos o herramientas disponibles para la validación de los requerimientos software desde el punto de vista de las metodologías ágiles. Los artículos analizados fueron adquiridos desde las fuentes SCOPUS, ACM y IEEE. Se han analizado 40 trabajos donde se puede ver como existe una continua búsqueda para lograr el objetivo de maximizar la calidad, ya sea de requerimientos o de procesos, utilizando diferentes técnicas y modelos, buscando el orden y calidad así como la entrega continua y reducción de tiempos y trabajo en equipo fusionando o creando técnicas, modelos, procedimientos, los cuales utilizan herramientas para la gestión de dirección de proyectos, gestores de incidencia y modelado evidenciando que los trabajos que han fusionado técnicas han podido ser probados y validados en entornos empresariales. Finalmente, para la mejora de procesos de gestión de requerimientos en la Administración Pública se está trabajando en el diseño del Estudio de Caso en un ente autárquico de la provincial de Corrientes y en su puesta en marcha.

En [13] se desarrolló un marco metodológico aplicable al desarrollo de software en el Sistema de Información Universitaria (SIU), que permita la construcción de un proceso de pruebas continuas, así como la propuesta de un marco tecnológico para implementar y dar soporte a dicho proceso de pruebas continuas. La validación del modelo de procesos para pruebas continuas en el SIU fue realizado sobre el proyecto SIU-Araí. Para lograr la implementación del modelo, se seleccionó una serie de procesos, etapas y actividades. Su implementación permitió obtener retroalimentación con el desarrollo real del proyecto, reflejando el impacto que

produce la automatización en la ejecución y validación de pruebas de software. El proyecto SIU-Araí mejoró la forma de realizar el despliegue (antes una tarea manual de 30 minutos y ahora una tarea de 2 minutos de fácil automatización), introdujo nuevas pruebas de software específicas sobre los módulos y expuso errores de desarrollo de manera temprana en el ciclo desarrollo del proyecto.

En la línea de Ingeniería de Software para la Entrega Continua se desarrolló el modelo de Continuous Testing Improvement Model (CTIM) y la herramienta de evaluación en línea EvalCTIM

La validación fue realizada mediante la técnica de Investigación-Acción con cuatro casos de estudio y un ciclo inicial teórica.

En la línea de Ingeniería de Software para Sistemas Críticos se realizó un estudio secundario amplio y sistematizado sobre las arquitecturas software aplicadas en el dominio de los sistemas críticos, el nivel de seguridad alcanzado y las herramientas utilizadas para lograrlo. Para ello se llevó a cabo una revisión sistemática de la literatura para identificar estudios publicados desde enero de 1999 a diciembre de 2019 sobre arquitecturas software para sistemas críticos. Como resultado se identificaron los tipos de arquitectura más utilizados de acuerdo al nivel de seguridad pretendido. Asimismo, se encontró evidencia de estudio en diferentes dominios de aplicación, con especial hincapié en las normativas automotrices e industriales [14].

FORMACIÓN DE RECURSOS

HUMANOS

En esta línea del Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 3 docentes investigadores, 1 tesista de doctorado, 1 becario de investigación de pregrado, 3 tesistas de maestría que finalizaron el cursado en el año 2018 y 1 tesista de maestría de la Universidad Nacional de Entre Ríos.

4. REFERENCIAS

- [1] M. Lehman, "On Understanding Laws, Evolution and Conservation in the Large Program Life Cycle", *J. of Sys. and Software*, vol. 13, pp. 213-221, 1980.
- [2] Randell, Brian. "Fifty Years of Software Engineering-or-The View from Garmisch." arXiv preprint arXiv:1805.02742 (2018).
- [3] Kong, Pingfan, et al. "Automated testing of android apps: A systematic literature review." *IEEE Transactions on Reliability* 99 (2018): 1-22.
- [4] L. Chen, "Continuous Delivery: Huge Benefits, but Challenges Too" in *IEEE Software* 03/2015. V. 32(2).
- [5] J. Humble and D. Farley. "Continuous delivery: reliable software releases through build, test, and deployment automation", 1st ed. Boston, US: Pearson Education, 2010.
- [6] O. Prusak. "Continuous Testing: The Missing Link in the Continuous Delivery Process". *Blaze Meter*. 2015. <https://blazemeter.com/blog/continuous-testing-missing-link-continuous-delivery-process>
- [7] Patricia Bazan. "Implementación de Procesos de Negocio a través de Servicios aplicando Metamodelos, Software Distribuido y Aspectos Sociales". Tesis presentada para obtener el grado de doctor en ciencias informáticas. 2015.
- [8] International Electro-technical Commission IEC, "IEC 61508: Functional safety of electrical/electronic/ programmable electronic safety-related systems." 199
- [9] Santana S.R., Perero L.R., Delduca A.G., Dapozo G.N. (2020) Evaluation of Open Source Tools for Requirements Management. In: Pesado P., Arroyo M. (eds) *Computer Science – CACIC 2019*. CACIC 2019. Communications in Computer and Information Science, vol 1184.
- [10] <https://es.bonitasoft.com/plataforma-bonita>
- [11] Sonia R. Santana, Lucrecia R. Perero, Amalia G. Delduca, Gladys N. Dapozo. Evaluación de técnicas para la validación de requerimientos en entornos de trabajo para el desarrollo de software. XXII Workshop de Investigadores en Ciencias de la Computación.
- [12] Pereyra Coimbra, R., Carruthers, J., Pinto, J., & Irrazábal, E. (2020). Personalización de técnicas ágiles en el desarrollo de software para la obtención de requisitos de calidad en Pymes: una revisión sistemática de la literatura. *Electronic Journal of SADIO (EJS)*, 19(2), 76-93. Recuperado a partir de <https://publicaciones.sadio.org.ar/index.php/EJS/article/view/169>
- [13] Trabajo Final de Maestría en Tecnologías de la Información Diseño y construcción de una metodología de pruebas continuas aplicable al modelo de desarrollo de software dentro del Sistema de Información Universitaria Autor: Lic. Sergio Fabián Vier
- [14] Joaquín Acevedo, Andrea Lezcano, Juan Pinto Oppido, Emanuel Irrazábal. Revisión sistemática de la literatura sobre implementación de arquitecturas software para sistemas críticos, CACIC 2020.

Aplicaciones Móviles 3D y Realidad Virtual

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 análisis de performance y consumo de energía de las aplicaciones 3D, como así también en el desarrollo de aplicaciones 3D inmersivas, a través del uso de Realidad Virtual (RV). También se realizan desarrollos relacionados con Internet de las Cosas, mediante aplicaciones móviles que utilizan sensores inteligentes.

Palabras claves: Dispositivos Móviles - Aplicaciones 3D - Aplicaciones Multiplataforma - M-Learning – Performance – Consumo de energía – Realidad Virtual – Internet de las Cosas

▪ Contexto

Esta línea de Investigación forma parte del proyecto (2018-2021) “*Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos. Mejora de proceso.*”, 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.

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado.

▪ 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].

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 [2] [3] [4].

Sin embargo, muchos de los potenciales usuarios de estas aplicaciones educativas pueden no disponer de dispositivos de última generación. Por esta razón, resulta de vital importancia realizar un análisis en profundidad de los parámetros que inciden en la performance final de una aplicación 3D [5] y

en el consumo de energía que dichas aplicaciones generan, el cual es generalmente alto [6] [7].

Los sistemas de domótica comprenden un conjunto de módulos cuyo objetivo en común es la automatización de las viviendas y de sus funciones. Los objetos cotidianos, tales como electrodomésticos, cuentan con conexión a Internet mediante la integración de sensores y/o dispositivos (Internet de las Cosas, o IoT por sus siglas en inglés) [8] [9].

Una aplicación móvil 3D de domótica facilita la interacción con los objetos conectados a una vivienda, mediante controles visuales, cómodos e intuitivos para el usuario.

▪ 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
- Performance de aplicaciones móviles 3D
- Consumo de energía en aplicaciones móviles 3D
- Realidad Virtual en aplicaciones móviles 3D
- Domótica en aplicaciones móviles 3D

▪ 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 de juego que permiten desarrollar aplicaciones 3D multiplataforma, particularmente para dispositivos móviles [10] [11] [12][13].

- Avanzar en el desarrollo de aplicaciones educativas, teniendo como finalidad enriquecer las experiencias interactivas y motivar el aprendizaje mediante su uso.
- Avanzar en el análisis de performance de ejecución y de consumo de energía de aplicaciones generadas con diferentes frameworks o motores de juego 3D.
- Avanzar en el desarrollo de aplicaciones móviles 3D relacionadas con sensores inteligentes e IoT.
- Se ha ampliado el prototipo móvil R-Info3D [2], una herramienta de aprendizaje sencilla de los conceptos básicos para la construcción de algoritmos en la Facultad de Informática. El intérprete de código fue actualizado, permitiendo el uso de variables, estructuras de control, operadores matemáticos, múltiples robots y paralelismo, entre otros. Además, la aplicación puede ser utilizada con lentes de RV, permitiendo una inmersión completa en el escenario virtual. Figura 1.
- Se ha desarrollado mediante los motores de juego multiplataforma Unity y Unreal Engine prototipos de análisis de performance de ejecución [14] y de consumo de energía de aplicaciones móviles 3D [15]. Figura 2.

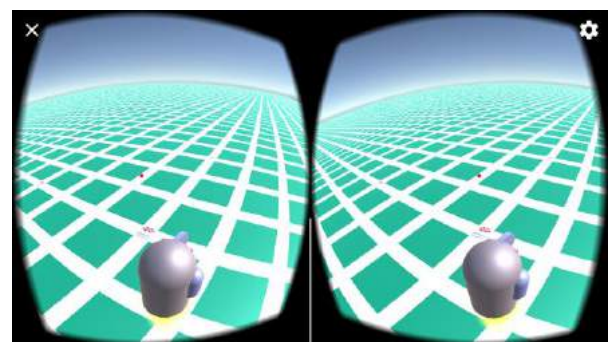


Figura 1. R-Info3D, controlado mediante las lentes de RV.

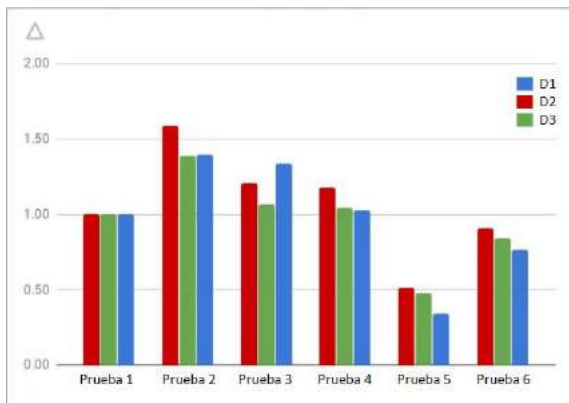


Figura 2. Prototipo de análisis de consumo de energía. Resultados obtenidos.

- Se ha desarrollado un prototipo móvil 3D de domótica. La aplicación facilita la interacción con los objetos conectados a una vivienda, mediante controles visuales, cómodos e intuitivos. El servidor puede estar en una red local o en la nube, permitiendo el control remoto de los dispositivos. Figura 3 y 4.



Figura 3. Prototipo móvil 3D de domótica. Pantalla inicial.



Figura 4. Prototipo móvil 3D de domótica. Control visual de los dispositivos.

- Se está avanzando en el estudio comparativo de los motores de juego multiplataforma Unity y Unreal Engine con el fin de contrastar las ventajas y desventajas de cada motor [16].
- Se está avanzando en el desarrollo de una aplicación móvil interactiva con realidad virtual y virtualidad aumentada para el Club Estudiantes de La Plata.

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.

Referencias

- Linowes J. "Unity Virtual Reality Projects". 2015. ISBN-13: 978-1783988556.
- 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.
- 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.
- 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.
- Akekarat Pattrasitidecha. "Comparison and evaluation of 3D mobile game engines". Chalmers University of Technology. University of Gothenburg. 2014.
- Riaz, M. "Energy consumption in hand-held mobile communication devices: A comparative study". International Conference

- on Computing, Mathematics and Engineering Technologies (iCoMET). 2018. ISBN: 978-1-5386-1370-2/18.
7. Aaron Carroll, Gernot Heiser. "An Analysis of Power Consumption in a Smartphone". USENIXATC'10 Proceedings of the 2010 USENIX conference on USENIX annual technical conference. 2010.
 8. J. C. Montesdeoca Contreras, R. S. Avila Campoverde, J. C. Cabrera Hidalgo and P. E. Vintimilla Tapia, "Mobile applications using TCP/IP-GSM protocols applied to domotic," 2015 XVI Workshop on Information Processing and Control (RPIC), Cordoba, 2015, pp. 1-4.
 9. Protocolo de red abierto MQTT.
<http://mqtt.org/>
 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>
 14. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. Capítulo de Libro: "Performance evaluation of a 3D engine for mobile devices". Libro: "Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790". Springer. ISBN: 978-3-319-75213-6, 978-3-319-75214-3, páginas 155-163. Año 2018.
 15. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. "Análisis de consumo de energía en aplicaciones 3D sobre dispositivos móviles". CACIC 2018. Tandil, Argentina. ISBN: 978-950-658-472-6, páginas 622-630.
 16. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado, Jefferson Perez Altamirano, Martin De la Canal Erbeta. "Aplicaciones Móviles 3D: un estudio comparativo de performance y consumo de energía". CACIC 2020. San Justo, Argentina. (ISBN en trámite).

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.

Luciano Marrero , Pablo Thomas , Ariel Pasini , Rodolfo Bertone , Eduardo Ibáñez ,
Verónica Aguirre , Marisa Panizzi¹, 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 CIC
¹ Universidad de Morón

*{lmarrero, pthomas, apasini, pbertone, eibanez, vaguirre, volsowy, ftesone,
ppesado}@lidi.info.unlp.edu.ar,*
¹marisapanizzi@outlook.com

RESUMEN

Se presenta una línea de investigación que tiene por objeto estudiar las problemáticas actuales que afronta los procesos de Ingeniería de Software y Bases de Datos ante una gran variedad aplicaciones multiplataformas (Web y móvil), que son de uso constante por millones de usuario simultáneamente y de cualquier punto geográfico. En este contexto, las metodologías ágiles, son sin duda, las más utilizadas en el proceso de desarrollo del Software actual. El modelo {relacional de Bases de Datos (Codd 1970) [5], es el modelo predominante de almacenamiento de información. Sin embargo, la idea de considerar que un único modelo de datos pueda adaptarse de forma eficiente a todos los requerimientos, ha sido discutida. Surgen así, otros motores de Bases de Datos que poseen implementaciones propias no relacionales y se denominan Bases de Datos NoSQL (No solo SQL). Estas Bases de Datos son un

complemento y/o alternativa a las Bases de Datos Relacionales en el proceso de diseño de la información. En la actualidad, los proveedores de Bases de Datos No Relacionales ofrecen implementaciones alternativas en la nube. Con esto se cuenta con un conjunto de herramientas importantes para el desarrollo de aplicaciones móviles. Esto implica, tener en cuenta aspectos que hace algunos años no eran considerados, tales como, movilidad, geolocalización, generación de grandes volúmenes de información y la diversidad de los dispositivos electrónicos involucrados [1, 2, 9, 10].

Palabras claves: Metodologías ágiles, Bases de Datos Relacionales, Bases de Datos NoSQL, Bases de Datos en la nube, Aplicaciones Móviles.

CONTEXTO

Esta línea de Investigación forma parte del proyecto (2018-2021) “Metodologías, técnicas y herramientas de ingeniería de software en

escenarios híbridos. Mejora de proceso.”, en particular del subproyecto “Ingeniería de Software para escenarios híbridos”, del Instituto de Investigación en Informática LIDI (III-LIDI) de la Facultad de Informática UNLP, acreditado por el Ministerio de Educación de la Nación.

Hay 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 utilizan los recursos de Hardware y Software disponibles en el III-LIDI para diseñar, desarrollar y probar diferentes soluciones a problemáticas relacionadas con escenarios a investigar. Como resultado de esto, se espera obtener métricas reales que sirvan como referencia para los investigadores en la comparación de resultados.

Las publicaciones científicas generadas y la transferencia continua de resultados concretos, validan esta línea de investigación.

1. INTRODUCCIÓN

“Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso”, 2018-2021 del Programa de Incentivos, es un proyecto 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 - Gobernanza Digital. Mejora de Procesos.” Y “SP3 - Metodologías y

herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos [19].

Este artículo 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, Bases de Datos NoSQL y Bases de Datos en la nube, con énfasis en los escenarios híbridos y las nuevas aplicaciones Web móvil para cualquier dispositivo informático.

La disponibilidad de la información en todo momento (internet), las nuevas tecnologías móviles y el desarrollo de aplicaciones híbridas, han cambiado las necesidades en el mundo informático. Se presentan así, un conjunto de nuevas alternativas y desafíos para la Ingeniería de Software, lo que conlleva al desafío de establecer nuevos procesos de diseño y nuevas arquitecturas de Hardware, cambiando así, la forma de pensar y almacenar los datos. Las Bases de Datos NoSQL representan una alternativa en la evolución del almacenamiento de datos, complementándose con una generación de tecnología móvil Web que debe responder eficientemente a las exigencias del usuario [9, 10, 11 y 12].

Cada vez, son más los recursos o servicios informáticos que poseen variantes en la nube (Cloud Computing). Esta tecnología, permite el ahorro en la administración de servicios y equipos. Es una abstracción de los recursos tecnológicos en donde se puede acceder a la información independientemente del punto geográfico en el que se encuentre el dispositivo. [4, 5, 8, 15 y 16].

En la actualidad, no solo se debe responder de manera óptima al número de usuarios finales, sino que también, cada usuario ha aumentado

sus exigencias al momento de utilizar y/o catalogar a una aplicación Web o móvil Web. Esto supone que la escalabilidad de los recursos y el rendimiento se han convertido en auténticos retos para la Ingeniería de Software y los proveedores de Bases de Datos.

En este contexto, las tecnologías de Bases de Datos No Relacionales (NoSQL), son un complemento o alternativa para satisfacer las demandas actuales. Las bases de datos NoSQL, surgen como respuesta para la administración de grandes volúmenes de información. Son altamente escalables y no respetan estrictamente las propiedades ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad) [4 y 5]. NoSQL propone un sistema llamado BASE (Básicamente Disponible, Estado Suave y Consistencia Eventual) [10, 13 y 14]. Existe una gran variedad de motores de Bases de Datos NoSQL que en general, se pueden catalogar en cuatro categorías de almacenamiento no estructurado de información: Clave-Valor, Documental, Orientado a Columnas y Orientado a Grafos. Herramientas como MongoDB (documental), Apache Cassandra (orientada a columnas), Redis (clave-valor) y Neo4j (grafos) son ejemplos de esta nueva generación de motores de Bases de Datos [9, 12 y 18].

Además, existen proveedores que brindan alternativas de Bases de Datos NoSQL en la nube.

Cloud Firestore, es un motor de Base de Datos en la nube con almacenamiento documental que pertenece a un conjunto de servicios que brinda Google para el desarrollo de aplicaciones Web y móvil.

MongoDB Atlas, es una Base de Datos en la nube con almacenamiento documental. Es proporcionada por MongoDB, es decir, que se

encuentra operada y mantenida por MongoDB. Este servicio incluye de forma automática, la configuración de los servidores y de todo el entorno necesario para la Base de Datos. Además, es compatible con diferentes proveedores de servicios en la nube, como son: Amazon Web Services (AWS), Google Cloud Platform (GPC) y Microsoft Azure.

DataStax (Astra), es otro ejemplo de un servicio de Bases de Datos en la nube, en este caso impulsado por Apache Cassandra [7, 11, 13 y 17].

Como parte de las investigaciones realizadas, el análisis del proceso de despliegue del Software es otro aspecto de estudio. Las PyMES en Argentina representan casi el 80% de la industria del software y dada la necesidad de ser competitivas deben mejorar sus métodos y procesos de trabajo. 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. Ante estas dificultades, se propone realizar un modelo de proceso de despliegue de sistemas de software que permita a las PyMES mejorar la ejecución del proceso de despliegue.

La comunicación y la sincronización del trabajo continúa siendo un pilar fundamental para el éxito de un proyecto. La utilización de repositorios de información, por ejemplo, GIT, 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 [3, 6 y 16].

Todas las particularidades previamente mencionadas conducen a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software y la gestión de Bases de Datos como disciplina.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para escenarios híbridos.
- Investigar los distintos tipos de almacenamiento no estructurado de información (documental, orientado a columnas, clave-valor y grafos).
- Investigar motores de Bases de Datos No Relacionales (NoSQL). MongoDB, Apache Cassandra, Redis, Neo4j, entre otros.
- Investigar Bases de Datos en la nube. Cloud Firestore (Google), MongoDB Atlas (MongoDB), DataStax Astra (Apache Cassandra), entre otras.
- Investigar nuevos conceptos de tipos de Bases de Datos, como, por ejemplo: NewSQL y Bases de Datos de Serie Temporales.
- Desarrollo de casos de estudio, pruebas de comparación y rendimiento de Bases de Datos Relacionales, No Relacionales y Bases de Datos en la nube.
- Modelo de procesos para el despliegue / puesta en producción de sistemas de software
- Repositorios GIT.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados esperados/obtenidos se pueden resumir en:

- Capacitación continua de los miembros de las líneas de investigación.
- Estudio y análisis de Bases de Datos no relacionales.
- Estudio y análisis de Bases de Datos en la nube.
- Analizar, comparar y determinar escenarios para los distintos tipos de almacenamiento no estructurado de información.
- Comparar y analizar resultados para diversos casos de estudio entre Bases de Datos relacionales, Bases de Datos NoSQL y Bases de Datos en la nube.
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT.
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles.

Algunas de las transferencias realizadas por el III-LIDI relacionadas con este proyecto, entre otras, son:

- Aplicación móvil para la comunidad de la Facultad de Informática de la UNLP con información sobre horarios de finales, planes de estudio, calendario académico, las clases en tiempo real y las últimas novedades.
- Diseño de Bases de Datos para Instituciones Provinciales.
- Diseño y Gestión de Sistemas de congresos, utilizado por la RedUNCI y por otras entidades.

- Sistemas de Gestión Administrativa de Instituciones Universitarias (Sistema de inscripción y seguimiento de alumnos de la Facultad de Informática, Sistema de gestión administrativa de asignaturas, entre otros).

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. Ingeniería de Software. Un Enfoque Práctico. Séptima Edición. Roger S. Pressman. McGraw-Hill Interamericana Editores, S.A. 2010. ISBN: 978-607-15-0314-5
2. Ingeniería de Software Teoría y Práctica. Pflieger Shari Lawrence. Pearson / Prentice Hall. 2002. ISBN: 9789879460719
3. Ingeniería de Software. Novena Edición. Ian Sommerville. Addison Wesley / Pearson. 2011. ISBN: 978-607-32-0603-7
4. Ingeniería de Software Clásica y Orientada a Objetos. Sexta Edición. Stephen R. Schach. Mc Graw Hill Interamericana Editores S.A. 2006. ISBN: 970-10-5636-1.
5. 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).
6. Administración de Proyectos. Guía para el Aprendizaje. Francisco Rivera Martínez, Gisel Hernández Chávez. Prentice Hall / Pearson. 2010. ISBN: 978-607-442-620-5.
7. NoSQL: modelos de datos y sistemas de gestión de bases de datos. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). <http://sedici.unlp.edu.ar/handle/10915/67258>.
8. Utilización de NoSQL para resolución de problemas al trabajar con cantidades masivas de datos. Róttoli, Giovanni, López Nocera, Marcelo, Pollo Cattaneo María Florencia. 2015. XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015). <http://sedici.unlp.edu.ar/handle/10915/45514>.
9. 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 Olsow, Fernando Tesone, Patricia Pesado. XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020). Universidad Nacional de la Patagonia Austral (UNPA). El Calafate, Santa Cruz (Mayo 2020). <http://sedici.unlp.edu.ar/handle/10915/104026>.
10. Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales. Luciano Marrero, Verena Olsow, Fernando Tesone, Pablo Thomas, Lisandro Delia y

- Patricia Pesado. XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020). Universidad Nacional de La Matanza, del 5 al 9 de octubre del 2020. <https://cacic2020.unlam.edu.ar/es-ar/>
11. Aspectos de ingeniería de software y bases de datos para el desarrollo de sistemas de software en escenarios híbridos. Patricia Mabel 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). Universidad Nacional de San Juan (UNSJ). (Abril 2019). <http://sedici.unlp.edu.ar/handle/10915/77088>.
 12. Un estudio comparativo de bases de datos relacionales y bases de datos NoSQL. Pesado Patricia Mabel, Thomas Pablo, Delia 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. <http://sedici.unlp.edu.ar/handle/10915/91403>.
 13. NoSQL A Brief Guide To The Emerging World of Polyglot Persistence. Pramod J. Sadalage y Martin Fowler. Pearson Education. 2013. ISBN: 978-0-321-82662-6.
 14. Data Modeling with NoSQL Database. Ajit Singh, Sultan Ahmad. ISBN 978-1072978374 (2019).
 15. Aplicaciones para Dispositivos Móviles. Estrategias y enfoques de desarrollo. Thomas Pablo Javier, et. al.. XXI Workshop de Investigadores en Ciencias de la Computación (San Juan 2019). ISBN: 978-987-3984-85-3
 16. Which Change Sets in Git Repositories Are Related?. Ramadani, J., Wagner, S. EEE 2016. International Conference on Software Quality, Reliability and Security (Viena, Austria.)
 17. Synchronization and replication in the context of mobile applications. STAGE, A. (2005 Joint Advanced Student School Course 6: Next-Generation User-Centered).
 18. A Performance Optimization Scheme for Migrating Hive Data to Neo4j Database. Publicado en 2018 International Symposium on Computer, Consumer and Control (IS3C). ISBN: 978-1-5386-7036-1. <https://ieeexplore.ieee.org/document/8644938>.
 19. III-LIDI: <http://weblidi.info.unlp.edu.ar/wp/proyectos/investigacion/>

Atender aspectos lingüístico-cognitivos en la captura de términos del contexto

Graciela D.S. Hadad¹, Jorge H. Doorn¹, María C. Elizalde¹, Ignacio A. García Ravlic¹, Damián Casfuz¹

¹ Escuela de Informática, Universidad Nacional del Oeste
ghadad@uno.edu.ar, jdoorn@uno.edu.ar, melizalde@uno.edu.ar, gnachoxp@gmail.com,
dcasafuz@gmail.com

RESUMEN

La construcción de requisitos involucra procesos cognitivos intensivos del ingeniero de requisitos, sin embargo, los abordajes corrientemente en uso no incluyen ninguna consideración al respecto. Para una elicitación eficaz, es indispensable entablar una buena comunicación entre los involucrados y utilizar técnicas apropiadas de recolección de hechos y de registro de éstos, y además en la elaboración de la información capturada se debería alcanzar una comprensión precisa y objetiva de lo elicitado. Una comunicación efectiva requiere un lenguaje común entre los involucrados. Es frecuente, a tal efecto, la creación de un glosario con términos del contexto de aplicación. A pesar de disponerse de heurísticas suficientemente afianzadas en la creación de glosarios, se ha observado un nivel preocupante de subjetividad en dicho proceso, lo que indudablemente exige estudiar qué aspectos lingüístico-cognitivos pueden introducir distorsiones en los términos incorporados al glosario y en sus definiciones. Entonces, en la etapa del presente proyecto se propone diseñar heurísticas que atiendan con mayor eficacia la incorporación de términos en el glosario, favoreciendo su completitud y consistencia, tratando de mitigar aspectos subjetivos del ingeniero de requisitos.

Palabras clave: Ingeniería de Requisitos, Elicitación, Glosarios, Lingüística Cognitiva, Pragmática, Procesamiento de Lenguaje Natural.

CONTEXTO

La línea de investigación que se presenta es parte del proyecto de investigación “Ajustes lingüístico-cognitivos a la construcción de

glosarios” que se desarrolla en la Universidad Nacional del Oeste.

1. INTRODUCCIÓN

En todo proceso de desarrollo de software los requisitos existen, aun cuando estos no se presenten en una forma explícita. En estos casos, pueden estar empotrados en modelos o artefactos menos formales, como una simple lista de servicios a desarrollar en una iteración, o incluso pueden permanecer solo en la mente de las personas, cuando el usuario trasmite sus necesidades al ingeniero, quien luego las vuelca a líneas de código. En cualquiera de estos casos, la comunicación es un punto de suma relevancia para acordar apropiadamente los requisitos del software [Nuseibeh00]. Es por ello que, principalmente en proyectos con equipos de trabajo numerosos o con una diversidad de usuarios, se debe formalizar un lenguaje común que facilite el entendimiento entre las partes sobre las necesidades actuales y futuras, y qué deberá proveer el sistema de software. Esto involucra crear un glosario con términos específicos empleados en el contexto de aplicación [Leite93]. Un objetivo principal de este glosario es que sus términos (denominados en lingüística símbolos) sean utilizados en los modelos, documentos, informes, e incluso en el código, construidos durante el proceso de desarrollo de software, como así también que sean internalizados por el equipo de desarrollo, de manera de usarlos eficazmente en las comunicaciones orales con los clientes y los usuarios, durante entrevistas, reuniones, negociaciones, presentaciones y validaciones, entre otras actividades.

La creación de glosarios en la Ingeniería de Requisitos es una práctica relativamente usual

[Kovitz98] [Aranda08] [Nonyelum12] [Mighetti16], aunque hay pocos autores que brindan guías para su construcción [Kovitz98] [Leite04] [Hadad09]. Un modelo léxico que se ha difundido por su capacidad expresiva es el Léxico Extendido del Lenguaje (LEL) [Leite93]. A pesar de sucesivas mejoras en las heurísticas de construcción del LEL, basadas en el manejo de homónimos, reconocimiento de nominalizaciones de verbos [Litvak17], uso de procesamiento de lenguaje natural [Escudero20] y tratamiento de jerarquías de términos [Kaplan17], entre otras, se ha observado que persisten problemas de completitud, de ambigüedad, y de inconsistencias en el nombre de los términos y en sus definiciones. Evidencias de ello se han reflejado en trabajos de comparación de métodos de inspección sobre el LEL [Sebastián16] [Hadad19], trabajos para estudiar su nivel de completitud [Doorn03] [Doorn08] y un trabajo posterior [Doorn19] donde se observó además la aparición de términos en el glosario que no estaban presentes en el contexto de aplicación. Este hecho muestra que en realidad sólo se está observando la punta de iceberg, ya que quien construye el glosario interpreta y elabora lo que lee, escucha u observa, incorporando su propia impronta. Esto condujo a una revisión de las heurísticas existentes, habiéndose observado que las mismas están más orientadas a mejorar la construcción de los modelos, pero que omiten indicar cómo se elabora la información adquirida. En otras palabras, las heurísticas indican qué se debe poner en un modelo, por ejemplo, qué tiempo verbal utilizar, cómo debe ser el estilo de las frases, pero poco colaboran acerca de cómo establecer la pertenencia de términos a determinadas categorías taxonómicas (jerarquías) [Vivas13], cómo distinguir relaciones temáticas [Vivas13] entre términos para obtener definiciones más completas, cómo interpretar las nominalizaciones [Alexiadou01], o cómo descubrir frases nominales empotradas en frases verbales, entre muchas otras posibilidades que involucran tanto aspectos cognitivos como lingüísticos.

Es debido a ello que se ha comenzado a estudiar cómo se podrían incorporar términos a léxicos con una mirada lingüístico-cognitiva, de manera tal de definir heurísticas que atemperen los problemas detectados. La definición de estas heurísticas no es una tarea sencilla dado que las mismas deberán guiar al ingeniero de requisitos en el desarrollo de las actividades cognitivas involucradas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el presente proyecto de investigación se intenta poner foco en el abordaje de la elicitación de información y su posterior elaboración y modelado, considerando los aportes de las ciencias cognitivas, particularmente en relación con la lingüística, de una manera pragmática.

Para ello es importante considerar los aspectos cognitivos de los involucrados en el proceso de Ingeniería de Requisitos, ya que es posible que, debido a los propios procesos de percepción, atención y memoria [Sternberg09] del ingeniero de requisitos, el mismo altere la información realmente recibida en función de los recuerdos que tiene o las asociaciones que realice. La atención permite seleccionar eficazmente la información relevante recibida [Lupón12]. En [Schmidt90], el autor expone que no se puede aprender aquello a lo que no se le ha prestado atención, y entonces un procesamiento consciente en la captación de información es un facilitador del aprendizaje. Trabajos posteriores en Lingüística Aplicada [Ellis94] [Gilakjani11] han estudiado y, en gran medida, apoyado esta teoría sustentada en el aprendizaje de un segundo lenguaje o en la adquisición de vocabulario. Es posible que algunos mecanismos propuestos desde la Lingüística Aplicada puedan aplicarse a la incorporación de términos a un glosario.

Por otro lado, desde la Semántica, se diferencia el significado conceptual o denotativo del significado asociativo o connotativo, el cual depende de los modelos mentales y experiencias previas de las personas [Muñoz17]. Mientras que la Lingüística Cognitiva, en particular la

Semántica Cognitiva, estudia la diferencia entre el significado literal de un concepto y su significado figurado (por ejemplo, metáforas y metonimias) como procesos mentales [Cuenca99]. Por otro lado, desde la Semántica, algunas relaciones semánticas deben ser consideradas en la definición de términos, tales como sinonimia, polisemia, hiponimia, hiperonimia, meronimia y holonimia, entre otras [Salvador85]. Por tal motivo, al construir glosarios se debe tener en cuenta tanto aspectos semánticos, como pragmáticos y cognitivos.

En lo que hace a este trabajo de investigación, se está asumiendo que los factores lingüístico-cognitivos tienen una influencia de relevancia en cualquier actividad donde haya una fuerte participación e interacción entre seres humanos, por lo que el proceso de requisitos se enmarca en un contexto de estas características y debe consecuentemente tener en consideración dichos factores de una manera proactiva.

Asimismo, se considera que algunas herramientas computacionales que proveen hoy en día el Procesamiento de Lenguaje Natural [Vilares05] [Padró12] y la Inteligencia Artificial en ese ámbito [Huk15] [Kunets18], pueden dar un valioso apoyo en mejorar la efectividad de actividades de elaboración de la información capturada.

El foco del presente trabajo es diseñar heurísticas que ayuden al ingeniero de requisitos a concentrar su atención en hechos propios del contexto de aplicación, despojándose de sus preconcepciones, es decir, con una visión objetiva frente a lo que está conociendo de ese contexto. Por otro lado, se planea diseñar y construir una herramienta que examine documentos de texto con información proveniente principalmente de entrevistas, y que, partiendo de unas pocas palabras o frases relevantes, utilizadas como semillas, proponga nuevas palabras o frases candidatas, en base a reglas de co-ocurrencia y combinación de roles gramaticales, manteniendo rastros de cada palabra o frase candidata a su uso en la documentación origen. Estos rastros permitirán que la herramienta proponga definiciones para cada

una de estas palabras o frases seleccionadas, o al menos estimule al ingeniero de requisitos a prestarles atención. Esto colaborará por un lado en reducir la invención de símbolos y la incorporación de definiciones basadas en conocimientos previos del ingeniero de requisitos y por el otro en mejorar la completitud.

3. RESULTADOS

OBTENIDOS/ESPERADOS

En base a factores cognitivos y lingüísticos identificados previamente, se han definido cambios en el proceso de construcción del LEL. La principal modificación elimina la elaboración de una lista inicial de términos y se aboca a que el ingeniero de requisitos se concentre en seleccionar unos pocos términos y los defina, y en sucesivas iteraciones seleccione otro grupo de términos. Esta modificación del mecanismo de selección de términos intenta atemperar los inconvenientes relacionados con la inhabilidad del ser humano para atender muchas cuestiones simultáneamente. La noción esencial no es eliminar esa lista de términos, sino construirla incrementalmente, procurando concentrarse siempre en unos pocos términos.

El otro cambio se refiere al mecanismo de definición de los términos. Se intenta minimizar la introducción de términos no usados por los usuarios, o la alteración de los mismos. En ese sentido se propone reducir, tanto como sea posible la “edición” de los términos utilizando en forma expresa técnicas de “cortar y pegar” a partir de documentos que contengan transcripciones de entrevistas a usuarios u otra documentación textual capturada.

En la Figura 1 se muestra el proceso propuesto de construcción del LEL.

Se realizó un primer experimento para comparar el proceso existente frente al nuevo proceso con las modificaciones antes mencionadas. Para ello dos sujetos entrenados en el modelo LEL y disponiendo de una descripción del proceso a seguir, recibieron un documento con la descripción de un sistema de transporte público de bicicletas para una ciudad, y construyeron una

versión del LEL, de manera independiente, aplicando las heurísticas del proceso asignado (el existente o el nuevo).

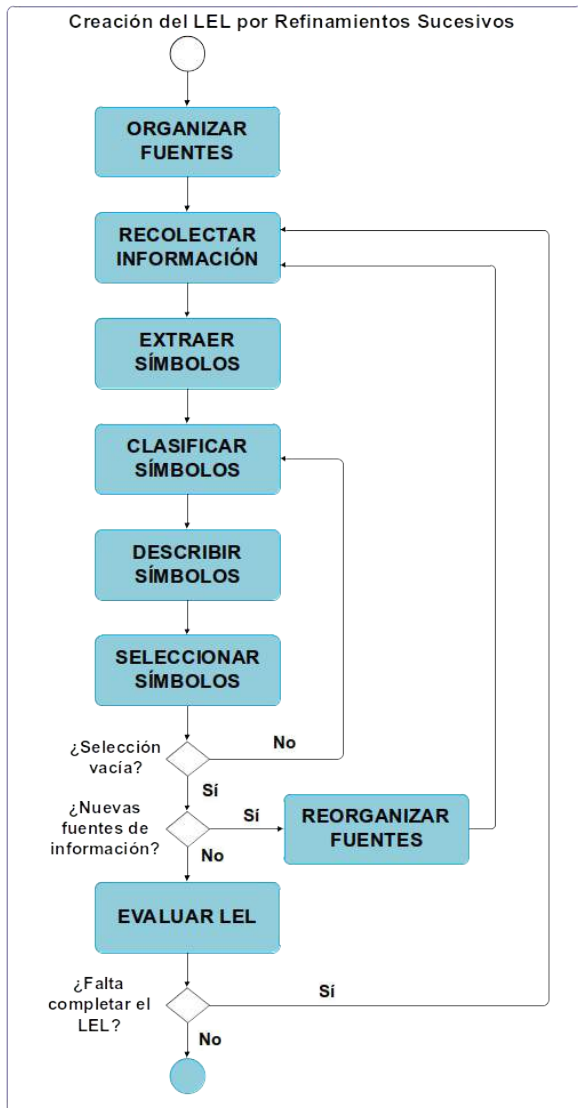


Figura 1. Nuevo Proceso de Construcción del LEL

Los resultados iniciales muestran una mejora en la calidad de los términos identificados y de sus definiciones, usando el proceso propuesto. Se observó que con el proceso existente se identificaron 17 símbolos correctos de 22 definidos, mientras que con el nuevo proceso se identificaron 25 símbolos correctos de 28 definidos. Es decir, hubo una mejora en cuanto a mayor cantidad de términos correctos identificados 83% del total estimado frente al 57% detectado con el proceso existente, con una reducción

significativa en la cantidad de omisiones de símbolos con el proceso propuesto. Aquí se denomina como correctos a aquellos símbolos que fueron ratificados en las verificaciones. En cuanto a las definiciones de los términos en el LEL, con el proceso existente se identificaron 30 porciones de texto no presentes en la documentación origen, mientras que se redujo a 11 las porciones de texto no presentes en la fuente aplicando el nuevo proceso.

Falta aún un estudio más detallado de los resultados, aunque ya en base a estos resultados preliminares, se realizarán algunos ajustes a las heurísticas del nuevo proceso y en próximos pasos se espera repetir el experimento con otros sujetos y las heurísticas ajustadas. Además, se incluirán al proceso nuevas mejoras relacionadas con otros factores lingüístico-cognitivos a atender, entre ellos una heurística para organizar el acceso a las fuentes de información, en función del tipo de información que brindan, en cuanto a la vigencia de la información (actual - formal) y su nivel de granularidad (abstracta - detallada). En caso de fuentes de información humanas, se analizará el tipo de narrativa que utilizan (orientada a la proximidad u orientada a taxonomías), y así definir un tratamiento diferenciado de dicha información.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto participan tres investigadores, y dos alumnos de grado.

En el marco del presente proyecto y del anterior, se finalizó un trabajo final de Especialización en Ingeniería de Software [Escudero20] sobre el uso de procesamiento de lenguaje natural en la construcción del modelo léxico.

5. BIBLIOGRAFÍA

- [Alexiadou01] Alexiadou, A. (2001). Functional Structure in Nominals. Nominalization and ergativity. Amsterdam, The Netherlands: John Benjamins Publishers.
- [Aranda08] Aranda, G.N., Vizcaíno, A., Cechich, A., Piattini, M. (2008). A Methodology for Reducing Geographical Dispersion Problems during Global

- Requirements Elicitation, 11th Workshop on Requirements Engineering, pp.117-127.
- [Cuenca99] Cuenca, M.J., Hilferty, J. (1999). Introducción a la lingüística cognitiva. Grupo Planeta (GBS).
- [Doorn03] Doorn, J., Ridao, M. (2003). Completitud de Glosarios: Un estudio experimental. VI Workshop on Requirements Engineering, pp. 317-328.
- [Doorn08] Doorn, J., Ridao, M. (2008). Completeness Concerns in Requirement Engineering. En: Encyclopedia of Information Science and Technology, Second Edition, IGI Global.
- [Doorn19] Doorn, J., Hadad, G., Elizalde, M.C., García, A.R.G., Carnero, L.O. (2019). Críticas Cognitivas a Heurísticas Orientadas a Modelos. 22nd Workshop on Requirements Engineering, Recife, Brasil.
- [Ellis94] Ellis, N.C. (1994). Consciousness in Second Language Learning: Psychological Perspectives on the Role of Conscious Processes in Vocabulary Acquisition. AILA Review, Vol. 11, pp. 37-56.
- [Escudero20] Escudero, J.F. (2000). Aplicación de la Inteligencia Artificial en la Ingeniería de Requisitos, Trabajo Final de Especialización en Ingeniería de Software, Universidad Católica Argentina.
- [Gilakjani11] Gilakjani, A.P., Ahmadi, M.S. (2011). Role of Consciousness in Second Language Acquisition. Theory and Practice in Language Studies, Vol. 1, No. 5, pp. 435-442.
- [Hadad09] Hadad, G.D.S., Doorn, J.H., Kaplan, G.N. (2009). Creating Software System Context Glossaries. En: Encyclopedia of Information Science and Technology, Second Edition, IGI Global, Information Science Reference, EEUU, pp.789-794.
- [Hadad19] Hadad, G.D.S., Sebastián, A. (2019). Inspecciones para Mejorar la Calidad de Modelos en Lenguaje Natural. XXV Congreso Argentino de Ciencias de la Computación, pp. 737-746.
- [Huk15] Huk, M., Kwiatkowski, J., Konieczny, D., Kędziora, M., Mizera-Pietraszko, J. (2015). Context-sensitive text mining with fitness leveling Genetic Algorithm. IEEE 2nd International Conference on Cybernetics (CYBCONF), pp. 342-347.
- [Kaplan17] Kaplan, G.N., Doorn, J.H. (2017). Jerarquías Naturales en el Contexto del Proceso de Requisitos. 20th Workshop on Requirements Engineering, Buenos Aires, pp. 441-454.
- [Kovitz98] Kovitz, B.L. (1998). Practical software requirements: a manual of content and style. Greenwich, CT: Manning Publications Co.
- [Kunanets18] Kunanets, N., Levchenko, O., Hadzalo, A. (2018). The Application of AntConc Concordancer in Linguistic Researches. IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, Vol. 2, pp. 144-147.
- [Leite04] 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, EEUU, capítulo 8, pp.169-199.
- [Leite93] Leite, J.C., Franco, A.P.M. (1993). A strategy for conceptual model acquisition, IEEE International Symposium on Requirements Engineering. San Diego, IEEE Computer Society Press, pp. 243-246.
- [Litvak17] Litvak, C.S., Hadad, G.D.S., Doorn J.H. (2017). Nominalizations in Requirements Engineering Natural Language Models. En: Encyclopedia of Information Science and Technology, Fourth Edition, IGI Global.
- [Lupón12] Lupón, M., Torrents, A., Quevedo, L. (2012). Procesos cognitivos básicos, Apuntes de Psicología en Atención Visual.
- [Mighetti16] Mighetti, J.P., Hadad, G.D.S. (2016). A Requirements Engineering Process Adapted to Global Software Development, CLEI Electronic Journal, Vol. 19, No. 3, paper 7.
- [Muñoz17] Muñoz-Basols, J. et al. (2017). Portal de Lingüística Hispánica, Ministerio de Educación, Cultura y Deporte-España y Universidad de Oxford. <http://hispaniclinguistics.com/>
- [Nonyelum12] Nonyelum Ogwueleka, F. (2012). Requirement elicitation problems in software development – A case study of a GSM service provider, Indian Journal of Innovations & Developments, Nigeria, Vol. 1, No. 8, pp.599-605.
- [Nuseibeh00] Nuseibeh, B., Easterbrook, S. (2000). Requirements Engineering: A Roadmap. Future of SE Track 2000, Limerick, Irlanda, pp.35-46.
- [Padró12] Padró, L., Stanilovsky, E. (2012). FreeLing 3.0: Towards Wider Multilinguality, Language Resources and Evaluation Conference, Estambul, pp. 2473-2479.
- [Salvador85] Salvador, G. (1985). Semántica y Lexicología del Español, Madrid, Paraninfo.
- [Schmidt90] Schmidt, R.W. (1990). The Role of Consciousness in Second Language Learning. Applied Linguistics, Vol. 11, No. 2, pp. 129-158.
- [Sebastián16] Sebastián, A., Hadad, G.D.S. (2016). Experimento Controlado en la Inspección de un Léxico mediante Mapas Conceptuales. 3er Congreso Argentino de Ingeniería, Resistencia, pp. 2738-2752.
- [Sternberg09] Sternberg, R.J., Sternberg, K. (2009). Cognitive Psychology, 6° edición, CENGAGE Learning, Boston, pp. 185-227.
- [Vilares05] Vilares Ferro, J. (2005). Aplicaciones del procesamiento del lenguaje natural en la recuperación de información en español, Tesis doctoral, Departamento de Computación, Universidade da Coruña.
- [Vivas13] Vivas Radakoff, J., García Coni, A.G. (2013). La Memoria Semántica. Organización y Desarrollo, Curso Programa de Altos Estudios Universitarios – IAEU, 56 pág.

Clasificando información en sitios de CQA

Valeria Zoratto, Gabriela Aranda, Nadina Martinez Carod, Alejandra Cechich,
Carina Noda, Mauro Sagripanti

Grupo de Investigación en Ingeniería de Software del Comahue (GIISCo)

<http://giisco.uncoma.edu.ar>

Facultad de Informática. Universidad Nacional del Comahue

Buenos Aires 1400, (8300) Neuquén

Contacto: {vzoratto, gabriela.aranda, nadina.martinez}@fi.uncoma.edu.ar

RESUMEN

La cantidad de información disponible en Web crece día a día. En particular, los servicios de Community Question Answering (CQA) se han convertido en una forma popular de búsqueda de información en línea, donde los usuarios pueden interactuar e intercambiar conocimientos en forma de preguntas y respuestas. La información que contienen los foros de discusión y las CQA son muy valiosas para usuarios con los mismos intereses y más aún, cuando tienen un problema similar, ya que estos sitios son colaborativos, donde se presentan discusiones e intercambio de ideas sobre un tema específico. Esta información merece ser consultada y estructurada. y la recuperación de información (IR) es una tarea esencial para lograr estos objetivos.

Nuestro proyecto se enfoca en extraer y analizar la información que contienen los foros de discusión ya que estos sitios tienen base de conocimiento lo suficientemente completa para ser utilizada. Pero, para que dicha información sea de utilidad, debemos definir estrategias para clasificar las soluciones disponibles, y obtener de ellas las más confiables.

Para ello, nuestro objetivo principal es crear una herramienta que clasifique automáticamente la información que

contienen los foros de discusión, utilizando diferentes estrategias de recuperación a nivel de hilo/foro, de clasificación a nivel de post y además, teniendo en cuenta la red de usuarios.

Palabras clave

Recuperación de Información, calidad de datos, Foros de discusión, CQA.

CONTEXTO

Nuestro proyecto se enmarca en el Programa “Desarrollo de Software Basado en reuso – Parte II”, de la Universidad Nacional del Comahue, a realizarse en el periodo 2017-2021, el cual extiende al Programa “Desarrollo de Software Basado en reuso” realizado en el período 2013-2016. Dicho Programa está compuesto por tres subproyectos los cuales coinciden en el tratamiento del desarrollo de software basado en reuso, pero desde aspectos diferentes: orientado a dominios, orientado a servicios y orientado a foros de discusión. El proyecto actual, denominado “reuso de Conocimientos en Foros de Discusión – Parte II”, continúa la línea de investigación enfocada en la recuperación de información y de conocimiento disponible en foros de discusión técnicos.

1. INTRODUCCIÓN

Ante la necesidad de procesar y reutilizar la información en grandes volúmenes de datos surge, en la década de 1950 [1], la Búsqueda y Recuperación de Información (del inglés Information Search and Retrieval (ISR)). Desde entonces, este campo de investigación fue creciendo logrando grandes contribuciones.

Se destacan dos grandes grupos de investigaciones, por un lado están las investigaciones que se enfocan en la recuperación de documentos específicos mientras que otros han desarrollado técnicas para generación automática de tesauros (lista de sinónimos, en conjunto con lista de antónimos, etc.) para su uso en distintos tipos de consultas.

En general el proceso comienza con una búsqueda del usuario en el sistema, y las respuestas que retorna poseen diferentes grados de relevancia. En particular, los foros de discusión disponibles en la Web sobre temáticas relacionadas al desarrollo y mantenimiento de software, contienen un amplio conocimiento sobre diferentes problemáticas recurrentes.

Los foros de discusión además son herramientas colaborativas accesibles a todos los usuarios pero no todos permiten que cualquier usuario realice consultas, para ello deben estar registrados en dicho foro (la mayoría de los foros cumplen con estas características). Esto permite la generación constante de información, por lo que hacer un análisis de dicha información es algo deseable y valioso [2].

Una de las características que distingue a los foros de discusión es que su interacción es asincrónica, es decir, que no se necesita estar conectados al mismo tiempo para obtener la información que solucione un problema. Un usuario de la comunidad realiza una pregunta y espera que otro usuario conteste a su consulta, de allí surgen diferentes respuestas que pueden ser de ayuda para el usuario que

pregunta o no. Incluso en la gran mayoría de las veces los participantes no se conocen personalmente, pero sí a través de sus nombres, alias o avatares (representaciones gráficas que se asocian a usuarios para identificarse). Un ejemplo de estos foros pueden ser Yahoo Answers! (YA)¹. YA es un sitio web de preguntas y respuestas impulsado por la comunidad o un mercado de conocimiento de Yahoo!, que permite a sus usuarios tanto formular preguntas como responderlas. Para hacerlo, el usuario tiene que tener una cuenta Yahoo!². Otro foro a destacar es Stackoverflow³ el cual está referido específicamente a problemas en entornos informáticos. Ofrece además la posibilidad de agregar código como respuesta a una pregunta, e incluye información de los usuarios, por ejemplo, la reputación, que, cuantos más puntos tenga el usuario, más cosas puede hacer en la comunidad.

Esto implica que al trabajar con foros, como cada uno tiene una estructura diferente, se complejiza la tarea de recopilación de información y del análisis a realizar sobre ella.

De acuerdo al permiso que se le otorga a los participantes dentro de un foro, se pueden distinguir 3 (tres) tipos bien definidos: *los públicos*, donde todos los participantes pueden comunicarse o leer mensajes escritos por el resto sin necesidad de registrarse; los foros *protegidos* donde es necesario registrarse para luego poder enviar mensajes. Por último en los foros *privados* se exigen ciertas restricciones para participar y utilizar la información. Como el proyecto se centra en la información contenida en los foros de discusión, es necesario utilizar foros de discusión *públicos* o *protegidos*, donde se pueda acceder a la información de los mismos sin necesidad de registración.

Un foro de discusión está compuesto por

¹ <https://answers.yahoo.com>

² <https://yahoo.com>

³ <https://stackoverflow.com/>

múltiples hilos. Los hilos son creados por usuarios que tienen algún problema o duda respecto a un tema. El usuario abre un hilo con una consulta inicial y a partir de ese momento, los usuarios de la comunidad podrán debatir en función del tema y del problema enunciado. En base a este comportamiento es que luego, para obtener conocimiento proveniente de los hilos de discusión se utilizan diferentes técnicas y estrategias para establecer cuáles de las posibles soluciones obtenidas de los foros pueden ser relevantes para consultas sobre problemas similares.

El proyecto realiza, por un lado, el tratamiento del texto contenido en los hilos dentro del foro y por otro lado analiza la red de usuarios que existe en él.

Para ello se ha continuado con el enfoque de Elsas & Carbonelli [6], que fueron unos de los primeros en revisar las estrategias de recuperación de hilo. Utilizan la estructura de los hilos separando la pregunta inicial del resto del hilo, tratándolo como un par <pregunta, hilo>. Cong et al. [5] extraen pares de <pregunta, respuesta> utilizando un enfoque basado en grafos no supervisados. Al igual que Cao et al. [7] se enfocan en extraer contextos y respuestas para preguntas, asumiendo preguntas ya identificadas formando tuplas con formato <pregunta, respuesta, contexto>.

Otra orientación que se ha analizado es la de clasificar o estructurar temas mediante jerarquías, como el enfoque de Nicoletti [17] o el de Helic et al. [8]. O bien el trabajo de Gottipati et al. [15] que aplican técnicas de minería de texto para extraer conocimientos de un foro de discusión mediante la generación de resúmenes basados en temas.

También se ha estudiado la satisfacción del usuario, como lo hace Liu et al. [12], intenta predecir si el autor de la pregunta estará

satisfecho con las respuestas enviadas por los participantes de la comunidad, al igual que Agichtein et al. [13].

Existen además propuestas de generación de algoritmos de ranking basados en la calidad de los atributos, como en la investigación de Kuna et al. [4], o el enfoque que presenta Bathia y Mitra [9] que, a partir del análisis de expertitud de los usuarios, detectan niveles de conocimiento de los comentarios, para clasificar con mayor valoración los hilos en los cuales intervienen personas expertos o con altos conocimientos.

Por otro lado, la investigación de Hecking et al. [10] que combina varias de las técnicas mencionadas anteriormente, ya que analiza la estructura social y semántica de los foros de discusión en cursos MOOC en términos de intercambio de información y roles de usuario.

En base a estos antecedentes, nuestro proyecto tiene como objetivo principal hacer reuso de la información existente en foros de discusión de la Web, haciendo uso no solo de la información textual de los hilos sino que teniendo en cuenta además a los usuarios, utilizando la red que se forma con las distintas interacciones que tienen con el resto de la comunidad y analizando los roles que cumplen, para poder detectar usuarios expertos y darle un mayor peso a las respuestas candidatas de una pregunta. Además, se ha experimentado tanto con la aplicación de algoritmos de análisis de lenguaje natural como de aprendizaje automático. Ya que el análisis del lenguaje natural permite analizar el tipo de fragmento dentro de un hilo de discusión [11]. Teniendo esto en cuenta, nuestro proyecto está enfocado en determinar un ranking de soluciones posibles, y cada línea de investigación dentro del proyecto lo hace desde ópticas diferentes con resultados favorables en su mayoría, permitiendo la

extensión de algunas líneas de avance y la elaboración de nuevas líneas a favor del objetivo del proyecto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación se denomina “Reuso de Conocimientos en Foros de Discusión – Parte II” y está enmarcado dentro del Programa de Investigación “Desarrollo de Software Basado en Reuso – Parte II”, con período de vigencia 2017-2021.

El programa mencionado extiende la investigación realizada durante el programa denominado “Desarrollo de Software Basado en Reuso”, realizado en el período 2013-2016. Respecto a este proyecto en particular, el objetivo es extender los estudios realizados sobre reuso de conocimiento en foros de discusión técnicos, incorporando la definición de métodos y algoritmos de recomendación para la asistencia inteligente a usuarios en la búsqueda de soluciones a preguntas frecuentes. Por otra parte, el programa está conformado por otros dos subproyectos que profundizan en las temáticas de Reuso Orientado al Dominio y Reuso Orientado a Servicios.

Dicho programa está desarrollado por el Grupo de Ingeniería de Software de la Universidad Nacional del Comahue, (GIISCo), formado por docentes y estudiantes de la Facultad de Informática de la Universidad Nacional del Comahue, junto con asesoría y colaboración de otras universidades. En particular, este proyecto es desarrollado en colaboración con la Facultad de Ciencias Exactas de la Universidad Nacional del Centro de la Provincia de Buenos Aires. Aunque el objetivo del Grupo GIISCo es brindar soporte en investigación y transferencia de tópicos relacionados con la Ingeniería de Software, el proyecto también involucra a docentes pertenecientes a otras áreas de la Facultad, como Programación y Teoría de la Computación, lo que permite

abordar la investigación desde ópticas diferentes, enriqueciendo el desarrollo con un trabajo conjunto y colaborativo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Considerando que el objetivo de nuestro proyecto es la realización de un recomendador de hilos de discusión teniendo en cuenta el ranking de las soluciones favorables, podemos mencionar los resultados obtenidos hasta el momento y los esperados a partir de ellos.

A partir del 2013 se investiga un modelo de calidad para foros de discusión en base a modelos de calidad de datos e información en la Web y estándares para la calidad de datos software. En dicho modelo se determinaron métricas para medir la calidad de información contenida en un hilo y estándares para la calidad de datos software [14], que fueron validadas mediante encuestas [16]. Con el propósito de mejorar los resultados obtenidos en [23] se propone una variación del peso de cada métrica por medio de un sistema de parametrización ad-hoc.

Otra línea de investigación trabaja con el procesamiento del texto de los hilos de discusión, para ello se implementó una herramienta para la recuperación de información de foros de discusión técnicos y su análisis mediante un conjunto preliminar de métricas de calidad, del cual se propone un ranking de soluciones posibles para una pregunta [18].

Para poder manipular la información de foros, se trabajó en el análisis de textos, utilizando la herramienta Lucene⁴ [19] con mecanismos personalizados para las *stopwords* (palabras que no aportan significancia) propias del dominio, haciendo tratamientos de recuperación de información para lenguaje específico de Java [20]. Se continuó con la

⁴ <https://lucene.apache.org>

utilización de sinónimos, mediante el uso de la base de datos léxica en inglés WordNet⁵ y un analizador morfológico como Stanford POS Tagger⁶ [21, 22].

Otra de las líneas de investigación se enfoca en el rol de los usuarios activos de un foro (los que participan compartiendo opiniones y experiencias). Para ello se trabajó con una estrategia para determinar la jerarquía de roles determinados por el nivel de conocimientos de los participantes en los hilos de acuerdo a los posts realizados [24].

Teniendo en cuenta los resultados obtenidos hasta el momento, se continúa trabajando, por un lado analizando la satisfacción del usuario que pregunta a partir de la propuesta de Liu et al. [25] y de Agichtein [13], por otro lado se está trabajando en las respuestas de calidad siguiendo la investigación de Burel et al. [26].

Otra línea en marcha se enfoca en el rol de los usuarios activos de un foro (los que participan compartiendo opiniones y experiencias). Bajo esta premisa, se han estudiado las propuestas [9, 10] y se está trabajando en una tesina, a partir de una estrategia empírica basada en la observación de hilos de discusión obtenidos de la web.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto se encuentra conformado por docentes de diferentes áreas debido a su naturaleza multidisciplinaria. En particular en el área de Ingeniería en Sistemas, Programación, y Teoría de la Computación. Las personas que forman parte del proyecto, tanto como colaboradores, asesores o integrantes son:

- Dos docentes investigadores del Departamento de Programación, con dedicación exclusiva, ambos con Doctorado en Informática.

- Un docente investigador del departamento de Programación con beca del CONICET para realización de doctorado.
- Tres docentes con dedicación simple, uno de ellos del Departamento de Ingeniería de Sistemas y dos del Departamento de Programación.
- Una profesora adjunta, asesora local, con dedicación exclusiva del Departamento de Teoría de la Computación
- Una docente investigadora externa, perteneciente al Instituto Superior de Ingeniería del Software (ISISTAN) de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), con experiencia en Sistemas de Recomendación y Recuperación de Información. Doctora en Ciencias de la Computación.
- Seis estudiantes de la carrera de Licenciatura en Ciencias de la Computación realizando sus tesis dentro del proyecto

De esta manera, se van incorporando actividades para extender líneas de investigación al proyecto inicial con nuevos enfoques.

5. BIBLIOGRAFÍA

- [1] Singhal,. Modern information retrieval: A brief overview. IEEE Data Eng. Bull., 2001, vol. 24, no 4, p. 35-43
- [2] S. Gottipati, D. Lo, and J. Jiang, Finding relevant answers in software forums, in 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011), Lawrence, KS, USA, November 6-10, 2011, pp. 323332, 2011.
- [3] Chen, W., & Persen, R. (2009, June). A Recommender System for Collaborative Knowledge. In AIED (pp. 309-316).
- [4] Kuna, H. D., Rey, M., Martini, E., Solonezen, L., & Sueldo, R. (2013). Generación de un algoritmo de Ranking para Documentos Científicos del Área de las Ciencias de la Computación. En el XVIII Congreso Argentino de Ciencias de la Computación.

⁵ <https://wordnet.princeton.edu/>

⁶ <https://nlp.stanford.edu/software/tagger.shtml>

- [5] G.Cong, L. Wang, C. Lin, Y. Song, and Yueheng (2008). Finding question-answer pairs from online forums. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '08). Association for Computing Machinery, 467–474.
- [6] J.L. Elsas and J. G Carbonell, "It pays to be picky: an evaluation of thread retrieval in online forums", in Proceedings of the 32nd international ACM SIGIR (2009), pp. 714—715.
- [7] Cao, Y., Yang, W. Y., Lin, C. Y., & Yu, Y. (2011). A structural support vector method for extracting contexts and answers of questions from online forums. *Information processing & management*, 47(6), 886-898.
- [8] Helic, D., & Scerbakov, N. (2003). Reusing discussion forums as learning resources in wbt systems. In *IASTED International Conference Computers and Advanced Technology in Education*, (Rhodes, Greece) (p. 223).
- [9] S. Bhatia and P. Mitra. Classifying user messages for managing web forum data. In Z. G. Ives and Y. Velegrakis, editors, *WebDB*, pages 13-18, 2012
- [10] T. Hecking, I. Chounta, and H. U. Hoppe. Investigating social and semantic user roles in MOOC discussion forums. In *LAK*, pages 198-207. ACM, 2016.
- [11] A. Tigelaar, R. Op Den Akker and D. Hiemstra, Automatic summarisation of discussion fora, *Natural Language Engineering*, ISSN 1469-8110, Vol 16, Issue 02, pp. 161-192, 2010.
- [12] Liu, Bing. *Web data mining: exploring hyperlinks, contents, and usage data*. Springer Science & Business Media, 2007.
- [13] Agichtein, E., Liu, Y., & Bian, J. (2009). Modeling information-seeker satisfaction in community question answering. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 3(2), 1-27.
- [14] G. Aranda, N. Martínez Carod, P. Faraci, A. Cechich. Hacia un framework de evaluación de calidad de información en foros de discusión técnicos. ASSE 2013.
- [15] Gottipati, S., Shankararaman, V., & Ramesh, R. (2019, October). TopicSummary: A Tool for Analyzing Class Discussion Forums using Topic Based Summarizations. In *2019 IEEE Frontiers in Education Conference (FIE)* (pp. 1-9). IEEE.
- [16] N.Martínez Carod, G. Aranda. Análisis de la información presente en foros de discusión técnicos. In *CACIC 2013*, pp. 847- 856, 2013.
- [17] M. Nicoletti, S. Schiafino, and D. Godoy. Mining interests for user profiling in electronic conversations. *Expert Syst. Appl.*, 40(2):638-645, Feb. 2013..
- [18] G. Aranda, N. Martínez-Carod, S. Roger, P. Faraci, and A. Cechich. Una herramienta para el análisis de hilos de discusión técnicos. In *CACIC 2014*, pages 803 - 812, 2014.
- [19] V. Zoratto, G. Aranda, S. Roger, A. Cechich, Análisis de estrategias para clasificar contenidos en foros de discusión: Un caso de estudio ASSE 2015, pp. 176-190.
- [20] V. Zoratto, G. Aranda, S. Roger, A. Cechich, Analyzing Discussion Forums Threads About Java Programming Language Usage, *Electronic Journal of SADIO*, 2016.
- [21] Zoratto, V., Martínez Carod, N., Otermin, F., & Aranda, G. N. (2017). Análisis de estrategias para clasificar contenidos en foros de discusión. In *XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017)*.
- [22] G. Aranda, V. Zoratto, N. Martínez Carod, Sandra Roger, F. Otermín, A. Cechich. Clasificación de contenido de hilos de discusión mediante análisis sintáctico y morfológico. In *CICCSI 2018*, pp 35-44, 2018.
- [23] N. Martínez Carod, V. Zoratto, G. Aranda, P Faraci. Aplicación de Métricas de calidad a hilos de discusión. In *CICCSI 2018*, pp 151-160, 2018.
- [24] N.Martínez Carod, G. Aranda. Valeria Zoratto, Christian Murray. Una propuesta para clasificación de roles de usuarios en foros de discusión técnicos. In *CACIC 2019*, pp. 836- 845, 2019.
- [25] Liu, Y., Bian, J., & Agichtein, E. (2008, July). Predicting information seeker satisfaction in community question answering. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (pp. 483-490).
- [26] Burel, G., Mulholland, P., & Alani, H. (2016, April). Structural normalisation methods for improving best answer identification in question answering communities. In Proceedings of the 25th International Conference Companion on World Wide Web (pp. 673-678).

CONTEXTOS DIGITALES para ASISTENCIA de LOS CIUDADANOS: ENFOQUES de EXPERIENCIA de USUARIO

Adriana MARTIN, Gabriela GAETAN, Viviana SALDAÑO, Claudia CARDOZO, Micaela VEGA, Hernán SOSA, Silvia VILLAGRA, Alejandra CARRIZO

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; gabrielagaetan@yahoo.com.ar; vsaldanio@gmail.com; claudia_yoryi@yahoo.com.ar; vegamicaela97@outlook.es; hassio_09@hotmail.com; svillagra@uaco.unpa.edu.ar; acarrizo@uaco.unpa.edu.ar.

RESUMEN

Un punto de interés que debería permanecer activo en el gobierno y sus instituciones, es encontrar y sostener en el tiempo, formas efectivas para canalizar experiencias satisfactorias en la interacción con sus ciudadanos. La eXperiencia de Usuario (UX), es un enfoque que se refiere básicamente al conjunto de factores y elementos relativos a la interacción del usuario con un entorno, cuyo resultado es la generación de una percepción (experiencia) positiva o negativa de ese entorno. La UX es satisfactoria cuando la interacción con ese entorno, soluciona problemas reales a la gente real. Trasladar este enfoque UX a la interacción del ciudadano, puede contribuir a desarrollar plataformas y productos que den respuesta a las necesidades de las personas, ofreciéndoles soluciones en vez de crearles nuevos problemas.

Desde el presente Proyecto de Investigación (PI), los integrantes de Grupo de Investigación y Formación en Ingeniería de Software (GIFIS), están enfocados en mejorar los contextos de interacción para contribuir con experiencias satisfactorias del ciudadano.

Palabras clave: *Experiencia de Usuario (UX) | Enfoque UX | Experiencia del Ciudadano | Usabilidad | Accesibilidad | Realidad Aumentada | Sistemas Aumentativos y Alternativos de Comunicación (SAAC)*

CONTEXTO

La mayoría de los expertos en diseño de UX, trabajan en el sector privado, donde el usuario

cuya experiencia se pretende mejorar, es un cliente. Sin embargo, considerar el diseño de UX en el sector de gobierno y sus instituciones, puede tener un rol clave para ofrecer servicios a la medida de sus ciudadanos.

Don Norman y Jakob Nielsen [4], indiscutidos referentes de la Usabilidad, establecen que la UX comprende todos los aspectos de la interacción del usuario final con la empresa y/o organización, sus servicios y sus productos. Referirnos a la UX, implica considerar todos los aspectos vinculados a la interacción del usuario y su alcance es muy amplio, debido a que lograr la satisfacción general del usuario es un objetivo que aplica a todas y cada una de las acciones que ejecuta y/o roles que desempeña una persona en su vida diaria. Por lo tanto, aplicar el enfoque UX al desarrollo de plataformas y productos destinados a la interacción entre los ciudadanos y el gobierno y sus instituciones, implica el desafío de: (i) canalizar de manera adecuada las necesidades de las personas para evitar el divorcio entre las partes, y con base en (i), (ii) proponer contextos adecuados que permitan superar conflictos y derribar barreras, propiciando una experiencia satisfactoria.

GIFIS, es un grupo de investigación perteneciente al Instituto de Tecnología Aplicada (ITA), de la Unidad Académica Caleta Olivia (UACO)¹, Universidad Nacional de la Patagonia Austral (UNPA)². En este momento, estamos iniciando la ejecución del segundo año del Proyecto de Investigación (PI) N° 29/B256, Período: 2020-2022, denominado: “Contextos

¹ <<http://www.uaco.unpa.edu.ar/uaco/>>

² <<http://www.unpa.edu.ar/>>

Digitales para Asistencia de los Ciudadanos: Enfoques de Experiencia de Usuario”, dirigido por la Dra. Adriana Martín y la Mg. Gabriela Gaetán. Desde 2012, y en el marco de varios proyectos de investigación consecutivos y relacionados, los integrantes de GIFIS han estado trabajando desde el diseño, evaluación y desarrollo de productos Web y móviles, para incluir propiedades de usabilidad y accesibilidad Web que mejoren la experiencia de los usuarios. Este trabajo ha incluido diferentes dominios y perfiles de usuarios de interés.

En este contexto, el PI N° 29/B256, se planteó y desarrolla como una continuación del PI 29/B222 (2018-2020): “*Diseño y Evaluación de Experiencia de Usuario para Multi-Dispositivos*”, enfocándose en aplicar estos antecedentes en mejorar y enriquecer la UX en contextos desarrollados para satisfacer la interacción ciudadana.

1. INTRODUCCIÓN

Diseñar UX implica un proceso complejo y nada trivial para crear productos significativos y relevantes a los usuarios destinatarios de dichos productos. Este proceso requiere tener en consideración diversos aspectos, tales como: (i) identificación y marca del producto --para crear lealtad y compromiso en el usuario, (ii) diseño --para definir lo que debe transmitir y hacer sentir al usuario, (iii) usabilidad y accesibilidad web --para lograr fácil acceso e interacción en el uso del producto, y (iv) funcionalidad --para proveer lo que espera recibir el usuario del producto, entre otras consideraciones. El enfoque UX es una herramienta poderosa al desarrollo de productos, ya que su objetivo es aumentar la satisfacción del usuario. Las técnicas y métodos de diseño de UX se vienen aplicando con éxito en la creación de estrategias para competir y atraer clientes en el sector privado. También, se han destinados esfuerzos de investigación para definir y aplicar de manera completa y efectiva este valioso enfoque [5][6][7][8][9]. Y además, se observa en el mundo una tendencia a aplicar el enfoque UX para propiciar la interacción de los ciudadanos con plataformas y productos ofrecidos por el gobierno y sus instituciones [1][2][3].

Como ya señalamos, la mayoría de los expertos UX trabajan en el sector privado, para mejorar cuestiones de negocio. En este sector, el usuario cuya experiencia se pretende mejorar, es un cliente. Sin embargo, como ya señalamos, el verdadero desafío es aplicar el enfoque UX en el desarrollo de plataformas y productos para el gobierno y sus instituciones. En este ámbito, el usuario cuya experiencia se pretende mejorar, es un ciudadano.

El enfoque UX ha cobrado relevancia internacional; este hecho se pone de manifiesto en la proliferación de jornadas y congresos, consultoras y grupos asesores, como así también numerosos estudios que plasman las acciones y prácticas recomendadas para mejorar la experiencia del ciudadano. Por ejemplo, en [14], la unión política y económica de los estados Árabes que bordean el golfo (GCC), presenta “*Government’s Citizen Experience Ecosystem*”, un ciclo de 3 componentes básicos a aplicar si el objetivo es atender a la experiencia del ciudadano. En [15], el Centro de Estudios Políticos y Constitucionales de Madrid, publica un compendio, denominado “*Participación Ciudadana: Experiencias Inspiradoras en España*”, destinado a transmitir experiencias de ciudadanos españoles, que son promovidas desde lo autonómico, provincial y local, para garantizar una participación ciudadana relevante y sostenible. Como otro ejemplo, en [16], la organización sin fines de lucro “*Centre for Public Impact*”, ofrece valiosos repositorios y propicia proyectos y grupos de trabajo con los gobiernos, servidores públicos y otros agentes de cambio, convirtiendo la idea de experiencia del ciudadano en acciones concretas para que el sector público y gobierno se reinventen y funcionen para todos. Estos ejemplos, son sólo una pequeña muestra de cómo ha germinado en el mundo, la problemática de aplicar el enfoque UX en la experiencia del ciudadano.

Desde esta perspectiva del usuario ciudadano, GIFIS, estableció las bases del PI N° 29/B256 (2020-2022), denominado: “*Contextos Digitales para Asistencia de los Ciudadanos: Enfoques de Experiencia de Usuario*”. El grupo de investigación ha estado trabajando desde hace varios años y desde los PI-UNPA 29/B167 (2014-2016), 29/B194 (2016-2018), y 29/B222

(2018-2020), sobre la problemática de identificar las expectativas y necesidades de diferentes grupos de usuarios para aplicar el enfoque UX en la evaluación y desarrollo de productos. A partir de este intenso trabajo de investigación, han surgido contribuciones, de las cuales podemos citar [11][12][13][17][18][19][20][21], que se corresponden al primer año de ejecución del presente PI. El objetivo es contribuir a la creación de experiencias enriquecidas, considerando propiedades de calidad en el diseño de las interfaces e interacciones, pero también, aplicando técnicas: (i) innovadoras --como la Realidad Aumentada (RA) y, (ii) específicas --como los Sistemas de Símbolos Pictográficos para personas con discapacidad comunicativa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El Objetivo General de investigación y desarrollo del PI N° 29/B256 es: “*Aplicar Enfoques de Diseño y Evaluación de Experiencia en Contextos Digitales para Asistir al Ciudadano*”. Para alcanzar este Objetivo General, se establecieron los siguientes Objetivos Específicos:

OE1. Proponer estrategias de solución para problemas de UX en aplicaciones móviles (Apps) de Gobierno.

OE2. Diseñar un Ecosistema de Contenido para el dominio Universitario.

OE3. Construir soluciones de software orientadas a Patrimonio Cultural y Turismo.

OE4. Desarrollar un Sistema Alternativo y Aumentativo de Comunicación (SAAC) para el ámbito de la Salud.

Estos Objetivos Específicos, definen en GIFIS líneas de investigación y desarrollo, dentro de las cuales cada uno de los integrantes están trabajando para alcanzar sus objetivos de investigación y/o formación particulares.

3. RESULTADOS OBTENIDOS/ ESPERADOS

A continuación, se presenta una breve descripción del trabajo realizado durante 2020, alineada a los Objetivos Específicos planteados en la Sección 2.:

- Enmarcado en OE1, en la contribución [17] se ha realizado una evaluación de Usabilidad aplicando principios heurísticos, para determinar la experiencia del ciudadano en la interacción con las Apps de Gobierno. El trabajo se basa en los comentarios de tiendas de aplicaciones, a los efectos de identificar los motivos que llevan a malas experiencias, para revertirlas canalizando de manera adecuada las necesidades de las personas.
- Dentro de OE2, se está aplicando los conocimientos adquiridos en Ecosistemas de Contenido, plasmados en [9], a un caso real del dominio de la Educación Superior: el sitio Web móvil de la UNPA-UACO. A tal efecto: (i) se evaluaron los contenidos que representa la columna vertebral de un servicio móvil para la comunidad Universitaria y, (ii) se incluyeron durante el desarrollo las propiedades de Usabilidad y Accesibilidad Web. En este momento se está desarrollando la implementación.
- Inmerso en OE3, se está trabajando en incorporar los beneficios de la Realidad Aumentada (RA) a contextos digitales del dominio Patrimonio Cultural y Turismo. En [18][21], se ha estado revisando el estado-del-arte de las plataformas y herramientas de RA --aprendiendo el uso y realizando pruebas, para determinar capacidades y ventajas y desventajas de las mismas. El objetivo es desarrollar productos de soporte al patrimonio cultural de la provincia de Santa Cruz (zona norte), Patagonia Argentina. Aplicando técnicas de RA en contextos interactivos, se busca propiciar la divulgación del valioso material histórico producido en UNPA-UACO.
- En OE4, se está trabajando sobre el concepto de Sistemas Aumentativo y Alternativo de Comunicación (SAAC), para desarrollar una aplicación móvil en dispositivo *Tablet*, la cual proveerá soporte específico al proceso de comunicación e interacción, entre los agentes de salud y los pacientes con discapacidad comunicativa. El desarrollo incorpora un conjunto de pictogramas del Centro Aragonés para la Comunicación

Aumentativa y Alternativa (ARASAAC)³, que son propiedad del Gobierno de Aragón, distribuidos bajo Licencia Creative Commons BY-NC-SA⁴. A tal fin, se ha obtenido previamente el permiso de uso respectivo. La aplicación está dirigida a las áreas de guardia e internación del Hospital Zonal Caleta Olivia, para asistir en el proceso de atención de los pacientes con Necesidades de Comunicación Complejas (NCC). En [19], se presentan los avances del proyecto, el cual se encuentra en la etapa final de desarrollo e implementación.

4. FORMACIÓN DE RECURSOS HUMANOS

En la actualidad, GIFIS tiene 8 integrantes: 4 docentes/investigadores UNPA, 1 docente/investigador invitado externo UTN-FRTDF, 1 alumno de postgrado UNPA, 1 alumno de grado UNPA, 1 alumno de pre-grato UNPA. A continuación, se enumeran los proyectos de formación durante 2020:

- 2 maestrandas, se encuentran desarrollando la Maestría en Informática y Sistemas (MIS) UNPA. Ambos han reunidos los créditos y requisitos de investigación solicitados por el plan de estudios, y con diferente grado de avance, se encuentran trabajando para presentar sus proyectos de tesis en el marco de OE1 y OE2, respectivamente.
- 1 alumno de grado, se encuentra finalizando su Proyecto de Final de Carrera para la Ingeniería en Sistemas UNPA. Su trabajo de investigación y desarrollo se está llevando a cabo en el marco de OE4. Se planea defender este proyecto entre Marzo-Abril de 2021.
- 1 alumno de grado, finalizó su Beca de Iniciación a la Investigación para Estudiantes de Grado y Pre-Grado UNPA y está elaborando el Informe Científico-Técnico (ICT-UNPA), respectivo. El trabajo de investigación y desarrollo, se esta llevando a cabo en el marco de OE3 y continuará en 2021 con una nueva Beca de Iniciación a la Investigación para Estudiantes de Grado y Pre-Grado UNPA.

Los integrantes docentes/investigadores de GIFIS, están en permanente actualización, para dar soporte a los proyectos de formación desde la dirección/ co-dirección. En 2020, se ha asistido a charlas y se han tomado cursos de capacitación en RA, y en 2021, se comenzó a realizar pruebas implementando estos conocimientos adquiridos. Además, para los integrantes más jóvenes de GIFIS, se está organizando la asistencia a cursos de Escritura de Documentos Científico-Técnicos en Inglés, que será ofrecido en 2021 en el Instituto de Tecnología Aplicada UNPA-UACO.

5. AGRADECIMIENTOS

A la UNPA por el soporte al Proyecto de Investigación (PI) N°: 29/B256, Período: 2020-2022, Denominado: “*Contextos Digitales para Asistencia de los Ciudadanos: Enfoques de Experiencia de Usuario*”.

A la Profesora Mg. Patricia Sampaoli, por aportar su valioso material sobre el patrimonio cultural de la provincia de Santa Cruz (Zona Norte).

A las Profesoras Iris Rasgido e Irene Morales y todo su equipo de profesionales especialistas, por contribuir con sus bastos conocimientos en pedagogía, fonoaudiología y discapacidad.

6. REFERENCIAS

- [1] Zamzami, K. F. The Key Service Feature of M-Government Based on Interactive User Experiences. IEEE Access. Vol. 7, DOI: 10.1109/ACCESS.2019.2924136, 2019.
- [2] Skarlatidou, A., Ponti, M., Sprinks, J., Nold, C., Haklay, M., Kanjo, E. User Experience of Digital Technologies in Citizen Science. Journal of Science Communication. Vol. 18(1), pp.1-8, DOI: 10.22323/2.18010501, 2019.
- [3] Adarsha, A., Reader, K., Erban, S. User Experience, IoMT, and Healthcare. Transactions on Human-Computer Interaction. Vol. 11(4), pp. 264-27, DOI: 10.17705/1thci.00125, 2019.
- [4] Norman, D., Nielsen, J. “The Definition of User Experience (UX)” <<https://www.nngroup.com/articles/definition-user-experience/>> Accedido 2020.

³ <<https://arasaac.org/>>

⁴ <<https://creativecommons.org/licenses/by-nc-sa/3.0/es/>>

- [5] Hartson, R. & Pyla, P. S. "The UX Book: Process and Guidelines for Ensuring a Quality User Experience". Elsevier; 2012.
- [6] Gothelf, J. Lean UX: Applying Lean Principles to Improve User Experience. Ed: O'Reilly; 2013. ISBN: 978-1-449-31165-0.
- [7] Levin, M. Designing Multi-Device Experiences: An Ecosystem Approach to User Experiences Across Devices. Ed: O'Reilly; 2014. ISBN: 978-1-449-34038-4.
- [8] Budiu, R.; Nielsen, J. "User Experience for Mobile Applications and Websites." NN/g Nielsen Norman Group; 3rd Edition; 2015.
- [9] Maoli, L. "Fixing Bad UX Designs." Packt Publishing Ltd. 1st Edition; 2018.
- [10] Gaetán, G., Villagra, S. Creación de Contenido Web Móvil con Base en Ecosistemas de Contenido. Revista Pistas Educativas del Tecnológico Nacional de México. vol. 41(133), pp. 231-244; ISSN: 2448-847X. <<http://www.itcelaya.edu.mx/ojs/index.php/pistas/article/view/2124/0>>, 2019.
- [11] Cardozo, C., Martín, A., Saldaño, V., Gaetán, Gabriela. Una Propuesta para Mejorar la Experiencia de los Adultos Mayores con las Redes Sociales. Revista Tecnología, Ciencia y Educación. Centro de Estudios Financieros (CEF) /Universidad a Distancia de Madrid (UDIMA). N° 16, pp. 113-142, ISSN 2444-250X e ISSN-e 2444-2887. <<https://tecnologia-ciencia-educacion.com/index.php/TCE/article/view/445>>, 2020.
- [12] Vidal, P., Martín, A. Experiencia de Usuario + Web Responsivo: Un Estudio desde la Perspectiva de un Enfoque Integrado. ICT-UNPA, Vol. 12(1), pp. 49-75, ISSN: 1852-4516, DOI: <<https://doi.org/10.22305/ict-unpa.v12.n1.703>>, 2020.
- [13] Martín, A., Gaetán, G., Saldaño, V., Cardozo, C., Carrizo, A., Sosa, H., Piñero, R., Villagra, S. La Experiencia de Usuario en el Contexto de los Multi-Dispositivos. XXII Workshop de Investigadores en Ciencias de la Computación (WICC); pp. 532-542. ISBN: 978-987-3714-82-5. <<http://sedici.unlp.edu.ar/handle/10915/104046>>, 2020.
- [14] EY Services. Driving Citizen Centricity: An Ecosystem to Help GCC Governments Sustainably Deliver Better Citizen Experiences. <[https://www.ey.com/Publication/vwLUAssets/EY-what-is-driving-gcc-governments-to-deliver-better-citizen-experiences/\\$FILE/EY-what-is-driving-gcc-governments-to-deliver-better-citizen-experiences.pdf](https://www.ey.com/Publication/vwLUAssets/EY-what-is-driving-gcc-governments-to-deliver-better-citizen-experiences/$FILE/EY-what-is-driving-gcc-governments-to-deliver-better-citizen-experiences.pdf)>, 2017.
- [15] Güemes, C., Resina, J., y Cruz-Rubio. C. Participación Ciudadana: Experiencias Inspiradoras en España. Centro de Estudios Políticos y Constitucionales. Gobierno de España. eISBN: 978-84-259-1843-1. <<http://www.cepc.gob.es/publicaciones/libros/colecciones?IDP=2738>>, 2018.
- [16] Center for Public Impact. A BCG Foundation. Research & Conversation → Articles: Citizen Experience. <<https://www.centreforpublicimpact.org/>>, 2015-2018.
- [17] Carrizo, A., Gaetán, G. Gobierno Móvil en Argentina: Análisis de Experiencia de Usuario basado en Comentarios de Tiendas de Aplicaciones. Revista ICT-UNPA, vol. 12(1), pp.18-48, ISSN: 1852-4516. DOI: <<https://doi.org/10.22305/ict-unpa.v12.n1.702>>, 2020.
- [18] Cardozo, C.; Martín, A.; Saldaño, V., Gaetán, G. Análisis de Herramientas de Realidad Aumentada para Crear Experiencias de Turismo. VI Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, pp. 710-714, ISBN: 978-987-3714-88-7 (1ª Ed.), 2021.
- [19] Sosa, H., Martín, A., Saldaño, V., Gaetán, G. Desarrollando una Aplicación SAAC Móvil para el Hospital Zonal de Caleta Olivia. VI Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, pp. 726-730, ISBN 978-987-3714-88-7 (1ª Ed.), 2021.
- [20] Martín, A., Gaetán, G., Saldaño, V., Cardozo, C., Vega, M., Carrizo, A., Villagra, S., Sosa, H. Enriqueciendo la Experiencia de Usuario en Dominios para Asistencia de los Ciudadanos. VI Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, pp. 750-754, ISBN 978-987-3714-88-7 (1ª Ed.), 2021.
- [21] Vega, M., Gaetán, G., Martín, A., Saldaño, V. Realidad Aumentada para Mejorar Experiencias de Turismo Cultural. VI Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, pp. 787-791, ISBN 978-987-3714-88-7 (1ª Ed.), 2021.

Desarrollo de incumbencias en el modelado de la vista estática

Fernando Pincirolí, Laura Zeligueta, Marcelo Palma,
Iván Cappello y Elías Motta

Instituto de Investigaciones
Facultad de Informática y Diseño
Universidad Champagnat
Mendoza, Argentina, +54 (261) 424-8443
pincirolifernando@uch.edu.ar; zeliguetalaura@uch.edu.ar; mpalma@civitas.com.ar;
ivancappello.e@gmail.com; eliasmotta1@gmail.com

Resumen

En este proyecto se está trabajando en el establecimiento de los lineamientos para la fase de requisitos de software del ciclo de vida del desarrollo de software (SDLC), poniendo especial énfasis en la vista estática y con el empleo del paradigma de orientación a aspectos, de manera de obtener los beneficios que ofrece [1].

Este proyecto es el quinto de la misma línea de investigación de desarrollo de software orientado a aspectos (AOSD) que se vienen desarrollando en forma consecutiva desde el año 2013 en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat y que continúan de un proyecto que fuera desarrollado en la Facultad Regional Mendoza de la Universidad Tecnológica Nacional.

Palabras clave: orientación a aspectos, aspectos tempranos, requisitos de usuario, requisitos funcionales, separación de incumbencias, encapsulamiento de incumbencias, composición de incumbencias, resolución de conflictos, AOP4ST.

Contexto

Dentro de las líneas de investigación del Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat, se encuentra la línea de “Ingeniería de Software”. En ella, y a lo largo de poco más de siete años, se llevaron a cabo los siguientes proyectos dedicados al desarrollo de software orientado a aspectos (AOSD):

- “*Procesos de desarrollo de software de calidad basados en aspectos*”, realizado en conjunto con la UTN Facultad Regional Mendoza durante 2011 y 2012 [2].
- “*Definición de criterios para la detección temprana de aspectos en el modelado de negocios y el desarrollo de los requisitos*” llevado a cabo por completo en la Universidad Champagnat, desde 2013 a 2016 y presentado en WICC 2015 [3].
- “*Modelado de procesos de negocio orientados a aspectos con BPMN*”, de la Universidad Champagnat y realizado en los años 2016 a 2018, y que fue presentado en WICC 2016 [4] y 2017 [5].
- “*Ingeniería de requisitos orientada a aspectos en AOP4ST*”: este proyecto abordó la fase de requisitos de usuario y se desarrolló entre 2018 y 2020 [6].

El proyecto actual, que se denomina “Ingeniería de requisitos de software orientada a aspectos en AOP4ST”, se inició en julio de 2020 y ya se produjeron algunas publicaciones: [1] [7] [8]. Cuenta con el financiamiento de la Universidad Champagnat y de la empresa Aconcagua Software Factory S.A., una de las principales fábricas de software de la Argentina.

Introducción

AOP4ST, sigla derivada de Aspect-Oriented Process for a Smooth Transition, es un proceso marco para el SDLC. En su versión actual, está enfocado en las fases tempranas, early aspects [9], compuesta por diferentes modelos: de negocio, de requisitos de usuario y de requisitos de software. Este último considera tres vistas: funcional, estática y de estados.

El acrónimo AOP4ST corresponde Aspect-Oriented Process for a Smooth Transition, que es un proceso marco para el desarrollo de software que pretende [8]:

- a) Ofrecer un proceso marco, no específico y liviano, de modo que permita su empleo con diferentes modelos del ciclo de vida del desarrollo del software (SDLC).
- b) Cubrir las etapas tempranas del SDLC, desde el modelado de negocio hasta la obtención de una especificación de requisitos completa y coherente.
- c) Emplear herramientas y técnicas estándares, de amplia difusión en la industria, para facilitar su adopción inmediata.
- d) Emplear notaciones estándares, para lograr modelos y especificaciones comprensibles y no ambiguas, que puedan contar con soporte de

herramientas de software disponibles en el mercado.

- e) Desarrollar los productos intermedios de las etapas tempranas del SDLC empleando el paradigma de la orientación a aspectos.
- f) Obtener las incumbencias en forma progresiva a lo largo de todos los modelos.
- g) Mantener la separación de incumbencias a lo largo de todos los modelos.
- h) Mantener la trazabilidad bidireccional de las incumbencias de punta a punta.
- i) Obtener las incumbencias en forma natural a lo largo de todos los modelos, de manera de no afectar la obtención de los objetivos de cada modelo.

El presente proyecto de investigación pretende establecer lineamientos para la vista estática del modelo de requisitos de software a partir de la vista funcional realizada en trabajos anteriores, dejando la vista de estados como próximo objeto de estudio. Procura presentar una visión interna del dominio del problema al descomponerlo en las clases específicas que lo constituyen, a diferencia del modelo de casos de uso, que presenta una visión externa del sistema [7].

Considerando las actividades propias del enfoque orientado a aspectos, mantiene la separación de incumbencias considerando los mismos paquetes que existen en la vista funcional del modelo de requisitos y elaborando dentro de ellos los diagramas de clases que materializan la vista estática.

En cuanto a la composición aplica un procedimiento sintáctico que combina las

clases que inicialmente poseen el mismo nombre en las diferentes incumbencias. El total de los atributos, operaciones y relaciones es el resultado de la unión de todos estos elementos desde todas las incumbencias individuales y que terminan apareciendo por completo, y solo una vez, en la composición resultante. El proceso presenta distintas estrategias según se trate de incumbencias pares e impares [8].

En cuanto a la resolución de conflictos establece algunas consideraciones previas y preventivas tales como criterios de diagramación y nomenclatura.

Finalmente, propone buenas prácticas de manera de asegurar los objetivos de la fase del ciclo de vida en cuestión, la calidad del producto software en desarrollo y la obtención de los beneficios que son la razón del empleo del paradigma de orientación a aspectos.

Estos beneficios apuntan al desarrollo de un producto de software final, como así también de los productos intermedios que permiten producirlo, más modular, mantenible, reusable, extensible, comprensible, etc. [8], al administrar en forma separada las incumbencias que están desparramadas y enredadas en cada uno de los niveles de abstracción a lo largo del ciclo de vida completo.

Desde ya que los principales aportes a los requisitos tempranos (early aspects) con el empleo de modelado estático, obtenidas mediante nuestro estudio de mapeo sistemático [1] y otros estudios de mapeo complementarios [10] [11] [12] [13], son

fuentes obligadas de referencia y de las que nos nutrimos en este proyecto.

Líneas de Investigación, Desarrollo e Innovación

Los proyectos de investigación que se vienen desarrollando en nuestra Facultad presentan cuatro ejes de investigación, donde los tres primeros se ven integrados con el cuarto:

1. Modelado de procesos de negocio orientados a aspectos.
2. Separación y composición de incumbencias con resolución de conflictos.
3. Especificación de requisitos y gestión de incumbencias con casos de uso e historias de usuario.
4. Procesos de desarrollo de software orientados a aspectos.

Resultados y Objetivos

Se consideraron resultados obtenidos en proyectos anteriores como insumos para el trabajo actual, como así también la línea de trabajo que sigue AOP4ST. Tras haber detectado las primeras incumbencias pares e impares en el modelo de negocio [4] [5], con el empleo del enfoque asimétrico tradicional de la orientación a aspectos [14] [15] y el enfoque simétrico que se propuso más tarde [16], se sigue el desarrollo y la administración de los requisitos de usuario, que permiten enriquecer el modelo con nuevas incumbencias [17]. Estos requisitos de usuario son especificados mediante elementos vistas: funcional, estática y de estados [3].

La vista estática constituye el objeto de estudio de nuestro proyecto de

investigación actual. Tiene como objetivo el profundizar los aportes que ella debe realizar a AOP4ST desde su rol en el proceso marco completo.

Hasta el momento, estos son los siguientes resultados alcanzados y los objetivos pendientes:

Detección y separación de incumbencias: desde el modelo de negocio y hasta el modelo de requisitos de software en su vista funcional fue posible mantener separadas las incumbencias. Pero al entrar en la vista estática, se observa que la mayoría de las clases involucradas en las funcionalidades (incumbencias) que se mantuvieron separadas hasta aquí, están presentes en más de una funcionalidad. Dicho de otra manera, los elementos de las vistas funcional y estática se relación de manera “muchos a muchos”, con lo que aparece nuevamente el enredo. Así, para poder seguir manteniendo la separación de las incumbencias en la vista estática, se consideró el concepto de “módulo de caso de uso” [18], que contiene un diagrama de clases para cada incumbencia en forma separada y que, por supuesto, luego se deberán componer.

Composición de incumbencias: en este momento nos encontramos trabajando en este punto, teniendo en cuenta el proceso que define AOP4ST [8]. Dado que hay diferentes formas de realizar la composición del modelo de clases, se siguieron las alternativas de mezcla y sobrescritura propuestas por Clarke y Baniassad [15], a las que en incorporamos algunas variaciones y a las que sumamos una tercera solución, de selección, que comprobamos que ofrece resultados más alentadores. Durante este proceso es posible la aparición de nuevas clases que podrían suscitar nuevas incumbencias. Esto obliga a que las vistas estática y

funcional se elaboren en forma iterativa e incremental. La composición desde lo semántico, que es la parte más difícil de resolver y a la que se debe prestar la mayor atención, requiere un estudio más profundo y que está pendiente en nuestra agenda de trabajo.

Resolución de conflictos: se está trabajando fuertemente en este punto, con cuatro analistas elaborando los mismos modelos por separado para analizar luego los conflictos que se generen, diseñar reglas de trabajo que apunten a reducirlas, volver a elaborar los mismos modelos y, finalmente, evaluar el grado de reducción de los conflictos y la consecuente necesidad de ajustar el conjunto de reglas.

Formación de Recursos Humanos

El proyecto promueve la formación del equipo de profesores, alumnos y egresados de la Universidad Champagnat.

Además, ya se han defendido tesis y se están elaborando otras sobre el cuerpo temático de esta investigación:

1. Tesis de Fernando Pincioli, Doctorado en Ciencias Informáticas, Universidad Nacional de San Juan, con la que se obtuvo el grado de Doctor en Ciencias de la Informática en noviembre pasado.
2. Tesis de Marcelo Palma, Maestría en Negocios y Tecnología, Universidad Champagnat.
3. Tesis de Jerónimo Vargas, Maestría en Negocios y Tecnología, Universidad Champagnat.
4. Tesis de Gustavo Albino, Maestría en Ingeniería de Software, Universidad Nacional de San Luis.
5. Tesina de Elías Motta, Licenciatura en Sistemas de Información, Universidad Champagnat.

6. Tesina de Javier Amutio, Licenciatura en Sistemas de Información, Universidad Champagnat.

Referencias

- [1] Pinciroli, F., Barros Justo, J.L. and Forradellas, R. (2020). Systematic Mapping Study: on the coverage of aspect-oriented methodologies for the early phases of the software development life cycle. *Journal of King Saud University, Computer and Information Sciences*, Elsevier.
- [2] Pinciroli, F. (2011). Consideraciones para un proceso de desarrollo de software de calidad orientado a aspectos. Sexto Encuentro de Investigadores y Docentes de Ingeniería, EnIDI 2011, San Rafael.
- [3] Pinciroli, F. (2015). AOP4ST - Aspect-Oriented Process for a Smooth Transition. XVII Workshop de Investigadores en Ciencias de la Computación, WICC 2015, Salta.
- [4] Pinciroli, F. and Zeligueta, L. (2016). El modelo de negocio en AOP4ST. XVIII Workshop de Investigadores en Ciencias de la Computación, WICC 2016, Concordia, pp. 423-426.
- [5] Pinciroli F. and Zeligueta, L. (2017). Modelado de negocios orientado a aspectos con AOP4ST. XIX Workshop de Investigadores en Ciencias de la Computación, WICC 2017, Buenos Aires.
- [6] Pinciroli, F. and Palma M. (2020). Desarrollo de requisitos en “Aspect-Oriented Process for a Smooth Transition”. XXII Workshop de Investigadores en Ciencias de la Computación, WICC 2020, El Calafate.
- [7] Pinciroli, F. (2019). Modeling the Static View in Aspect-Oriented Software Development. III Congreso Internacional de Ciencias de la Computación y Sistemas de Información, CICC SI 2019, Mendoza.
- [8] Pinciroli, F. (2019). Proceso marco orientado a aspectos en las etapas tempranas del ciclo de vida del desarrollo de software para una transición en la industria. Tesis del Doctorado en Ciencias de la Informática, Universidad Nacional de San Juan, 2019.
- [9] Rashid, A., Sawyer, P., Moreira, A. and Araújo, J. (2002). Early aspects: a model for aspect-oriented requirements engineering. *IEEE Joint Int. Conference on Requirements Engineering*.
- [10] Magableh, A., Shukur, Z. and Ali, N.M. (2013). Systematic review on aspect-oriented UML modeling: A complete aspectual UML modeling framework. *Journal of Applied Sciences*, vol. 13, no. 1, pp. 1–13.
- [11] Magableh, A.A. (2017). Systematic review of aspect-oriented formal method. *International Journal of Computers and Applied Technology*, vol. 56, no. 2, pp. 132–140.
- [12] M. Guessi, L., Oliveira, B.R. and Nakagawa, E.Y. (2010). Modeling Aspect-oriented Software Systems Using UML: A Systematic Mapping. *CLEI Electronic Journal* vol.14, no.1, p.3.
- [13] M. Guessi, L. Oliveira, B.R. and Nakagawa, E.Y. (2011). Extensions of UML to Model Aspect-oriented Software Systems. *CLEI Electronic Journal* vol.14, no.1, pp. 1-18.
- [14] Katara, M. and Katz, S. (2007). A Concern Architecture View for Aspect-Oriented Software Design. *Software and Systems Mod.*, vol. 6, no. 3, pp. 247–265.
- [15] Clarke, S. and Baniassad, E. (2005). *Aspect-Oriented Analysis and Design. The Theme Approach*. Addison-Wesley, Boston, 2005.
- [16] Harrison, W., Ossher, H. and Tarr, P. (2002). Asymmetrically vs. Symmetrically Organized Paradigms for Software Composition. *IBM Research Report RC22685*.
- [17] Pinciroli, F. (2018). Requisitos de usuario y gestión de la demanda en AOP4ST. XX Workshop de Investigadores en Ciencias de la Computación, WICC 2018, Buenos Aires, pp. 511-515.
- [18] Jacobson, I. and Ng, P. (2005). *Aspect-oriented software development with use cases*. Addison-Wesley, 2005.

Despliegue de sistemas de software: estado de la práctica actual en PyMES de Argentina

Marisa Daniela Panizzi, Agustín Gabriel Hodes, Felipe Ortiz, Pablo Vazquez, Rodolfo Alfredo Bertone

¹ Departamento de Ingeniería en Sistemas de Información. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires. Medrano 951 (C1179AAQ), C.A.B.A, Argentina.

² Instituto de Investigación en Informática (III-LIDI). Facultad de Informática. Calles 50 y 120 - La Plata - Bs. As. - Argentina

³ Instituto de Tecnología e Ingeniería. Universidad Nacional de Hurlingham. Av. Vergara 2222 (B1688GEZ) - Villa Tesei - Bs. As. Argentina.

marisapanizzi@outlook.com, hodesagustin@gmail.com, ortizfd@gmail.com, vazpablo@gmail.com, pbertone@lidi.info.unlp.edu.ar

Resumen

La línea de investigación de este trabajo se enmarca en fortalecer el proceso de despliegue de sistemas de software en Pequeñas y Medianas Empresas (PyMES) de Argentina. En Argentina, las PyMES ocupan aproximadamente el 80 % del sector de la industria del software, requieren incrementar la capacidad y calidad de sus procesos para lograr mayor competitividad. Esta línea promete realizar aportaciones al proceso de despliegue por tratarse de un proceso crucial en la construcción de sistemas de software, de éste dependen el cierre del proyecto, así como también la aceptación del producto software por parte del cliente.

Con el propósito de evidenciar el estado actual de la práctica sobre el proceso de despliegue de sistemas de software en PyMES de Argentina, se propone la realización de un estudio exploratorio basado en el método de encuesta.

Palabras clave: proceso de despliegue, sistemas de software, PyMES, encuesta.

Contexto

La línea de investigación que se reporta en este artículo es financiada parcialmente por 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 Secretaria 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

Las PyMES ocupan aproximadamente el 80 % del sector de la industria del software en Argentina [1]. El notable desarrollo queda demostrado en el último Informe Argentina Productiva – Economía del Conocimiento [2]. Esto constituye un eslabón fundamental, en el sector, para el país y refuerza la necesidad de llevar adelante iniciativas que contribuyan con el desarrollo y mejora de competitividad de dichas empresas. A nivel internacional se presenta la misma situación en la industria del software, las PyMES ocupan una gran porción de la industria en varios países [3]. Se observa que en los últimos años las PyMES han surgido muy rápidamente y, en la mayoría de las economías en desarrollo, el sector está dominado por pequeñas y recientes jóvenes [4].

Las PyMES se han dado cuenta de que es fundamental para su negocio mejorar sus procesos y métodos de trabajo, pero carecen del conocimiento y los recursos para hacerlo. La única forma de contribuir al éxito de los proyectos, por tanto, es definir, implementar y estabilizar los procesos de desarrollo [5].

El proceso de despliegue de sistemas de software se considera crucial dentro del ciclo de vida de desarrollo de software, de éste dependen el cierre del proyecto y la aceptación del producto software por parte del cliente. Los usuarios informan con frecuencia

de errores mientras se implementa el sistema de software [6][7]. Algunos softwares empresariales pueden tardar meses, en desplegarse por completo, hay ausencia de actividades automatizadas y el tiempo de ejecución del proceso suele ser largo [6]. Los problemas anteriores se presentan una vez finalizado el despliegue, al mismo tiempo que se está utilizando el sistema de software, por lo que los problemas que surgen deben resolverse en la fase de mantenimiento; esto conduce a un esfuerzo innecesario en términos de recursos económicos y humanos [6][7].

Todas las consideraciones anteriores nos llevaron a definir diferentes objetivos investigación a largo plazo orientados a robustecer el proceso de despliegue de sistemas de software para ayudar a las PyMES a ejecutar el despliegue de sistemas empresariales.

El grupo de investigación presenta como desafío dar respuestas a las siguientes preguntas de investigación (PI):

PI1: ¿Se logra cubrir la vacancia de un modelo de proceso de despliegue de sistemas de software, que integre las actividades y sus tareas, las técnicas, las herramientas, los artefactos, los roles y sus competencias?

PI2: ¿Se fortalece el proceso de despliegue mediante la identificación de un conjunto de riesgos, así como los procedimientos para su mitigación?

PI3: ¿Se mejora la calidad del proceso de despliegue mediante la determinación de métricas específicas para el mismo?

Líneas de Investigación, Desarrollo e Innovación

La línea de investigación aborda los problemas que se presentan en el proceso de despliegue de sistemas de software en PyMES de Argentina y dar respuesta a las PI presentadas en la sección anterior.

En respuesta a la PI1, se considera la sistematización del proceso de despliegue de sistemas de software mediante la definición de un modelo de proceso que permita:

- La ejecución de **actividades y tareas**.
- El uso de **plantillas orientadoras**.
- La asignación de **roles específicos** que cuenten con las competencias necesarias para ejecutar el despliegue.
- El uso de **herramientas** para automatizar algunas de las actividades del proceso.
- Su **acoplamiento** a las metodologías de desarrollo de software que utilicen las PyMES.

Antes de comenzar con la definición del modelo, se realizó un mapeo sistemático de la literatura (en inglés, Systematic Mapping Study o SMS) [8]. Actualmente se están analizando los resultados de la encuesta realizada a las PyMES de Argentina sobre el proceso de despliegue de sistemas de software con el propósito de conocer de que manera se encuentra la práctica actual. Como respuesta a la PI2 se pretende definir un conjunto de riesgos para el proceso de despliegue de sistemas de software, así como también los procedimientos para su mitigación. Antes de comenzar con la definición del conjunto de riesgos, se realizó un SMS [9] que permitió cotejar y sistematizar la evidencia empírica para identificar (determinar) cuáles de las metodologías, métodos y estándares que abordan la gestión de riesgos en proyectos de desarrollo de software, son las más estudiadas por la comunidad científica y las más utilizadas en la industria del software. Una vez logrado esto, mediante el método DESMET [10] se logró evidenciar las metodologías, métodos y estándares que soportan riesgos para el proceso de despliegue. Se definieron los riesgos y sus procedimientos de mitigación, finalmente la solución se validó mediante dos estudios de caso [11] en PyMES de desarrollo de software de Argentina [12].

Como respuesta a la PI3, se pretende definir un conjunto de métricas para el proceso de despliegue de sistemas de software que permiten mejorar la calidad del proceso. La revisión de metodologías de desarrollo realizada hasta el momento ha indicado la ausencia de métricas específicas para el despliegue, como aún no se ha finalizado esta

revisión no se puede confirmar dicha afirmación [13].

Resultados y Objetivos

En esta línea de investigación en progreso se han logrado una serie de **resultados** que se detallan a continuación:

- a) Académicos, se han logrado dos trabajos de especialidad en Ingeniería en Sistemas de Información y una tesis de Maestría en Ingeniería en Sistemas de Información, recientemente depositada para la defensa.
- b) Producción Científica: se ha presentado la línea de investigación en eventos científicos de alcance nacional (WICC¹ 2017, WICC 2019, WICC 2020, CACIC² 2016, CACIC 2018, CACIC 2019, CACIC 2020 y en el ámbito internacional, CIACA³ 2017, SEKE⁴ 2017, CibSE⁵ 2019, CibSE 2020 y ICAETT⁶ 2019. Además, se lograron tres publicaciones en Springer y en el Brazilian Journal of Development.
- c) Proyectos de Investigación: ha finalizado el PID UTNBA4347 titulado: "Impacto del factor peopeware en el proceso de implantación de sistemas informáticos" (período 2017-2019). Se inicia el PID UTNBA6576 titulado: "Estudio del proceso de implantación de sistemas informáticos en el contexto industrial de la República Argentina" (período 2020-2022).
- d) Formación en investigación: el grupo de investigación se encuentra en un proceso de aprendizaje constante de métodos de investigación de ingeniería de software experimental, revisiones sistemáticas [14], estudios de casos [11] y encuestas [15][16].

El grupo de investigación se plantea como **próximos objetivos**:

- Analizar los resultados de la encuesta realizada a 195 participantes con el propósito de conocer el estado actual de la práctica para luego continuar con el diseño y refinamiento del modelo de proceso de despliegue.
- Avanzar con el SMS sobre las métricas para el proceso de despliegue de sistemas de software.
- Presentar los avances realizados en InGENIO⁷ 2021, CibSE 2021 y CACIC 2021.

Formación de Recursos Humanos

El grupo se encuentra conformado por un Director, dos tesistas de Maestría, un becario de investigación y un asesor científico-tecnológico.

Se estima la formación de un Doctor en Ciencias Informáticas y un Magister en Ingeniería en Sistemas de Información. Así como también la continuación de la formación de un becario de investigación.

Referencias

[1] OPSSI. Reporte anual 2018 sobre el Sector de Software y Servicios Informáticos de la República Argentina (2018). Disponible en <https://www.cessi.org.ar/opssi>.

[2] Informe Argentina Productiva – Economía del Conocimiento. Ministerio de Producción y Trabajo. Presidencia de la Nación. Secretaría de la Transformación Productiva (2019). Diciembre de 2019. Disponible en:

<https://biblioteca.produccion.gob.ar/buscar/?fid=16>.

[3] Hisham M. Abushama. PAM-SMEs: process assessment method for small to medium enterprises. Software: Evolution and Process, 28, pp. 689 –711 (2016).

[4] Sharma P., Sangal A.L. Building a hierarchical structure model of enablers that affect software process improvement in

⁷ Congreso Latinoamericano de Ingeniería

¹ Workshop de Investigadores en Ciencias de la Computación

² Congreso Argentino de Ciencias de la Computación.

³ Conferencia Iberoamericana de Computación Aplicada.

⁴ International Conference on Software Engineering and Knowledge Engineering.

⁵ Congreso Iberoamericano en Ingeniería de Software.

⁶ Congreso Internacional sobre Avances en Nuevas Tendencias y Tecnologías.

software SMEs-A mixed method approach. *Computer Standards & Interfaces*, 66, pp. 1–23 (2019).

[5] Ianzen A., Mauda E.C., Paludo M.A., Reinehr S., Malucelli A. Software process improvement in a financial organization: an action research approach. *Computer Standard & Interfaces*, 36, pp 54–65 (2013).

[6] Subramanian N. The software deployment process and automation. *CrossTalk*, 30(2), pp. 28-34 (2017).

[7] Carrizo D., Sanchez L. Benchmarking to Adopt an Asap-Based Methodological. *Guideline for Software Systems Deployment*. In: *Proceedings of the 30th International Business Information Management Association Conference, (IBIMA '17): Sustainable Economic development, Innovation Management, and Global Growth*, pp. 183-193 (2017).

[8] Panizzi M., Genero M., Bertone R. Software system deployment process: A systematic mapping study. In *Proceedings of 23rd Iberoamerican Conference on Software Engineering (CIbSE 2020)*. Curitiba, Brazil, November 9 - 13, 2020.

[9] Felipe Ortiz, Mauricio Dávila, Marisa Panizzi, and Rodolfo Bertone. State of the Art Determination of Risk Management in the Implantation Process of Computing Systems. *Advances in Emerging Trends and Technologies*. Vol 1066, Springer, pp. 23-32. Springer Nature Switzerland AG (2019).

[10] Kitchenham B., Linkman S., Law D.T. DESMET: A method for evaluating software engineering methods and tools. Keele University (1996).

[11] Runeson P., Höst M., Rainer A., Regnell B. (2012). *Case study research in software engineering: guidelines and examples*. Wiley Publishing, Hoboken.

[12] Ortiz F., Panizzi M., y Bertone R. Risk refinement in the deployment process of software systems: a case study. En las *Actas del XXVI Congreso Argentino de Ciencias de la Computación - CACIC 2020*. Universidad Nacional de La Matanza, 5 al 9 de octubre. ISBN en trámite.

[13] Marisa Panizzi, Rodolfo Bertone, Alejandro Hossian. *Proceso de implantación de sistemas informáticos -identificación de vacancias en metodologías usuales*. Portugal. Vilamoura, Algarve, Portugal. 2017. Libro de Actas de la V Conferencia Iberoamericana de Computación Aplicada (CIACA 2017). Portugal: IADIS PRESS. 2017. 978-989-8533-70-8.

[14] Kitchenham, B. y Chartes, S. (2007). *Guidelines for performing systematic literature reviews in Software engineering*, Keele University, EBSE-2007-01.

[15] Jefferson Seide Molléri, Kai Petersen, Emilia Mendes. An empirically evaluated checklist for surveys in software engineering. *Information and Software Technology* 119 (2020) 106240.

[16] Genero, M., Piattini, M., & Cruz Lemus, J. A. (2014). *Métodos de investigación en Ingeniería del Software*. Madrid: Ra-Ma S.A. Editorial y Publicaciones (2014).

ENFOQUE METODOLÓGICO DIRIGIDO POR MODELOS PARA DESARROLLOS DE INTERNET/WEB DE LAS COSAS

Edgardo Belloni^{1a} Diego A. Godoy^{1b} Juan de Dios Benítez^{1c} Nelson Acosta^{2d}

¹Centro de Investigación en Tecnologías de la Información y Comunicaciones (CITIC),
Universidad Gastón Dachary (UGD)

²Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA)

^{a,b,c}{ebelloni, diegodoy, juan.benitez}@citic.ugd.edu.ar ^dnacosta@exa.unicen.edu.ar

RESUMEN

Se presenta un proyecto de investigación y desarrollo que tiene por objetivo general el de elaborar un enfoque metodológico dirigido por modelos para soportar la generación (semi)automática de aplicaciones en el dominio de Internet/Web de las Cosas.

Palabras clave: Ingeniería Web Dirigida por Modelos; Internet/Web de las Cosas.

CONTEXTOS

El proyecto se enmarca dentro de un paradigma de investigación aplicada y de desarrollo tecnológico, teniendo como contexto disciplinar el área de Ingeniería Web Dirigida por Modelos (MDWE, *Model-Driven Web Engineering* [1]), y como dominio de aplicación el de los desarrollos basados en Internet / Web de las Cosas (I/WoT, *Internet / Web of Things* [2] [3]).

Asimismo, el proyecto:

- Se encuentra registrado -con código A11004, y misma denominación que el título de este documento- por la Secretaría de Investigación y Desarrollo de UGD; teniendo como instrumento de adjudicación la R.R. UGD 39/A/20, correspondiente al 11° concurso de proyectos de I+D con evaluación externa de la institución.
- Está radicado, y se ha iniciado recientemente su ejecución, en CITIC-UGD; articulando líneas de investigación que se desprenden del proyecto denominado “*Tecnologías para desarrollos sostenibles de ciudades inteligentes*”, registrado y acreditado en la misma institución (R.R. UGD 44/A/19; código A10003).

1. INTRODUCCIÓN

Esencialmente, el proyecto aborda la selección y extensión de un enfoque metodológico en el área de MDWE, con el propósito de aportar una solución parcial a un problema ingenieril abierto -u oportunidad recurrente de mejora- en dicha área [1]: brindar soporte de generación de aplicaciones en tecnologías de Ingeniería Web emergentes. En particular; y, a este último respecto, se aborda la generación de aplicaciones en plataformas avanzadas en el dominio de I/WoT [2] [3].

Ciertamente, hoy en día se reconoce que los entornos de aplicación de I/WoT son prácticamente infinitos, y que su impacto en la sociedad está siendo ya, seguramente, enorme [2]. No obstante; y, justamente, las problemáticas de heterogeneidad y portabilidad entre las numerosas plataformas de I/WoT que han surgido y evolucionado en los últimos años [13][14][16][17][18][19], representa una gran oportunidad de adoptar enfoques de MDWE para dicho dominio tecnológico. En tal sentido, el colectivo de desarrolladores que podría beneficiarse del enfoque metodológico y tecnológico integrado que se propone, resulta potencialmente amplio.

Marco disciplinar

El paradigma ingenieril de desarrollo dirigido por modelos [4], prescribe el uso de modelos como artefactos principales del proceso de desarrollo. Los modelos se construyen a distintos niveles de detalle y perspectiva, e inicialmente con independencia de la plataforma de destino. Mientras que, las implementaciones específicas en determinadas tecnologías, se generan de manera (semi) automática por medio de transformaciones sistemáticas de los modelos construidos.

Esencialmente, en dicho paradigma, los modelos deben especificarse conforme a tipos y un metamodelo; modelo del modelo, que especifica la sintaxis y semántica de elementos o abstracciones que lo componen. De esta manera, se soporta el proceso de transformación entre modelos. Proceso, basado en reglas, que define operaciones para el paso de un modelo origen a otro de destino, y que se sustenta en las correspondencias que se prescriben entre tipos del modelo origen y los del destino; especificados en sus correspondientes metamodelos.

Este paradigma promueve la reducción del esfuerzo inherente al desarrollo de aplicaciones complejas, así como, la mejora de la calidad y eficiencia de tales desarrollos. En particular, se posibilita la obtención de ventajas respecto de: mejoras en la comunicación entre desarrolladores y con los usuarios; incremento de productividad de desarrolladores; mejora en la adaptación de desarrollos provocados por cambios tecnológicos y en requisitos [5] [6]. Asimismo, tales ventajas se sostienen, en: postergar toma de decisiones de carácter tecnológico; generación consistente de artefactos de desarrollo; y, reutilización de modelos y transformaciones entre modelos soportados.

Teniendo en cuenta que dichas ventajas podrían ser aprovechadas en su aplicación al ámbito del desarrollo Web, surge entonces el área disciplinar de MDWE [1]. Área en la que se han realizado numerosos esfuerzos de investigación y desarrollo, durante las tres décadas que abarca sus inicios hasta el presente. Contexto y espacio temporal en el que han surgido diferentes propuestas de nuevas metodologías, lenguajes de modelado, y herramientas para facilitar su aplicación, así como también, criterios de evaluación de su calidad [7] [1]. Específicamente, entre los varios enfoques de MDWE que han surgido y evolucionado durante las últimas décadas, se incluyen entre los más difundidos los de: UWE [8]; NDT [9]; e, IFML [10].

A pesar de los muchos beneficios que presenta la adopción de enfoques de MDWE, así como, los logros que se han alcanzado en dicha área; la comunidad científica de referencia ha

señalado problemas abiertos y necesidades aún no cubiertas en el área [1]. Entre tales problemáticas y necesidades se incluyen:

- Aspectos de organización arquitectónica;
- Necesidad de Evaluaciones empíricas de enfoques en MDWE (esencialmente, en cuanto a productividad y calidad de aplicación por parte de desarrolladores);
- Necesidades de soporte para: Requerimientos No Funcionales; Agilidad; Usuarios finales; y, Tecnologías de Ingeniería Web emergentes (*Internet/Web of Things*, *Semantic Web*, y *Social Web applications*, entre otras).

En el sentido expuesto; y, como ya se adelantó, en este trabajo se aborda el problema ingenieril abierto de extender un enfoque metodológico - ya existente en el área de MDWE- para cubrir la necesidad recurrente de dar soporte a tecnologías emergentes. En este caso, se adopta como dominio tecnológico de aplicación el que se reseña a continuación.

Dominio tecnológico de aplicación

El dominio tecnológico de I/WoT se sustenta en la interconexión digital de “cosas”, es decir, objetos (físicos y/o virtuales) con Internet [2], o más específicamente la Web [3], permitiendo así que estos objetos se comuniquen e intercambien datos entre sí. Entre tales objetos se puede incluir a: dispositivos móviles, equipamiento del hogar o industrial, vehículos, y hasta incluso productos consumibles y de vestimenta. Cada objeto que forma parte de una infraestructura de I/WoT es unívocamente identificable, y puede contener mini computadores, sensores y actuadores, dependiendo de su propósito específico; ofreciendo una amplia gama de posibilidades para facilitar la vida de las personas apoyando o automatizando muchas tareas cotidianas.

En la última década, han surgido numerosas plataformas de I/WoT para soportar desarrollos en tal dominio [11] [12]. Las soluciones más evolucionadas proporcionan una plataforma de desarrollo de aplicaciones integrada, comprendiendo: la administración de dispositivos, almacenamiento y administración de datos, *frameworks* de desarrollo de aplicaciones basados en *APIs*, así

como también, de distribución e implantación de aplicaciones [11] [12]. Entre las plataformas más destacadas, actualmente, se pueden citar a las de: *Amazon Web Services IoT* [13]; *FIWARE* [14]; *OpenMTC* [15]; *SmartThings* de *Samsung* [16]; *IBM Watson IoT Platform* [17]; *Azure IoT Hub* de *Microsoft* [18]; y, *Google IoT* [19].

Mayormente, el desarrollo de plataformas de I/WoT ha tenido lugar en nichos industriales, comerciales y tecnológicos particulares, difiriendo -a consecuencia de la falta de una estandarización- en su conceptualización, tecnologías de soporte y terminología. En tal sentido, el gran número de plataformas de I/WoT existentes resultan muy heterogéneas; aun cuando respecto de su funcionalidad todas soportarían, esencialmente, lo mismo: permitir conectar diferentes dispositivos, objetos (físicos y/o virtuales) o “cosas”; almacenar, acceder y procesar los datos generados por tales dispositivos; y, utilizar la información o el conocimiento adquirido de procesar tales datos para monitorear, controlar y/o actuar -de manera automatizada- sobre distintos bienes de uso, procesos, etc.

La heterogeneidad de estas plataformas causa un problema para todo desarrollador que tiene que seleccionar una de ellas. Decidir cual plataforma de I/WoT utilizar para un desarrollo, demanda, tanto, tiempo, como, conocimientos técnicos para ser capaz de entender y comparar diferentes conceptualizaciones; dado que cada una utiliza tecnologías y terminologías diferentes. Adicionalmente, como sucede con toda tecnología emergente, la rápida evolución y obsolescencia -no existiendo aun una estandarización de conceptualización y terminología- de plataformas para desarrollos basados en I/WoT, muchas veces se promoverá tener que portar aplicaciones, ya funcionando en una plataforma, a otra plataforma emergente, para, por ej. acceder a servicios más eficientes a menores costos.

En el sentido expuesto, las problemáticas de heterogeneidad y portabilidad entre plataformas de I/WoT, representa una oportunidad de adoptar enfoques de MDWE. Tal oportunidad refiere a la posibilidad de

reutilizar modelos (de conceptos en el dominio, de análisis y diseño, o bien, específicos de plataformas tecnológicas), portando (semi)automáticamente sistemas complejos a otras plataformas; cuando aquella elegida originalmente para su implementación se considere obsoleta, discontinuada, menos eficiente o más demandante en costos, que otras emergentes.

2. LÍNEA DE INVESTIGACIÓN Y RESULTADOS ESPERADOS

La línea principal que se desarrolla, tiene por **objetivo general** -la **acción central** alrededor de la cual se articula el proyecto:

- Elaborar un enfoque metodológico de Ingeniería Web Dirigida por Modelos para soportar la generación (semi)automática de aplicaciones en el dominio de Internet/Web de las Cosas.

Mientras que los **objetivos específicos** -concebidos como **logros o resultados parciales**, que contribuyen a la concreción del objetivo general- son los siguientes:

- Compilar tecnologías y plataformas, abstracciones relevantes, aspectos ortogonales y patrones, buenas prácticas, problemas abiertos y desafíos, inherentes al desarrollo de aplicaciones de I/WoT.
- Definir un metamodelo que integre abstracciones y aspectos relevantes en el dominio de desarrollos de I/WoT.
- Analizar críticamente enfoques metodológicos existentes en el área de MDWE, para determinar su grado de soporte al desarrollo de aplicaciones en el dominio de I/WoT.
- Diseñar un enfoque metodológico y desarrollar un marco tecnológico para su soporte, que emplee el metamodelo definido, y que extienda de forma adecuada otro enfoque ya existente en MDWE.
- Evaluar la validez y aplicabilidad del enfoque propuesto.

Para alcanzar dichos objetivos, la línea comprende, tanto, investigación documental, como, investigación experimental.

Investigación Documental

Esta investigación se funda en la búsqueda,

revisión, compilación, y reseñas sistemáticas de conceptos, soluciones y tecnologías existentes en el contexto disciplinar y dominio de aplicación referidos; basándose en el proceso sistemático de revisión delineado por Kitchenham et al. [20], como así también, en revisiones sistemáticas multi vocales [21]. Asimismo, es inherente a esta investigación, un análisis crítico de las soluciones y tecnologías compiladas -identificando puntos fuertes y débiles, determinando problemas y oportunidades viables de ser tratados, y cuáles de estos se afectarán al desarrollo específico del trabajo a desarrollar.

Investigación experimental

A partir de la selección de un enfoque de MDWE, el cual resulte viable de complementar para abordar la resolución de alguno de sus problemas abiertos o aspectos no tratados respecto del desarrollo de aplicaciones en el dominio de I/WoT, se propondrá una solución inicial sujeta a validación, revisión y ajuste sistemáticos.

Sintéticamente, entonces, se prescribe un ciclo iterativo disciplinar compuesto por las fases de: Análisis y Especificación de Requisitos, Proposición y Especificación de Soluciones, Prototipación, Experimentación con Casos de Estudio cubriendo escenarios típicos en el dominio de aplicación, Revisión de resultados, y Ajuste de acuerdo a las lecciones aprendidas a partir de los resultados. Asimismo, se prevé:

- Validar la propuesta a través del juicio de expertos en la disciplina y dominio. Esto involucrará el delineamiento e instrumentación de encuestas o entrevistas estructuradas, así como también, grupos focales.
- Analizar y mejorar de forma continua el enfoque propuesto a partir de retroalimentaciones que se puedan compilar en la divulgación de actividades y resultados parciales.
- Identificar requisitos básicos para -y, evaluar la calidad del- marco de soporte tecnológico para el enfoque metodológico propuesto; basándose en criterios y características deseables, comúnmente aceptados por la comunidad académica y de práctica profesional de referencia.

Delimitación

A priori, se delimitará el alcance del proyecto a escenarios típicos de ciudades inteligentes soportados por I/WoT.

Otros resultados esperados

Por otra parte, se espera promover procesos de transferencia, favoreciendo alianzas estratégicas, la innovación y el aprendizaje organizacional, y las acciones de apropiación de conocimiento en las organizaciones -en la región de radicación del proyecto- potenciales beneficiarias de resultados y desarrollos de esta investigación.

El abordaje de temáticas de I+D+i como las de este proyecto, se alinea con el aprovechamiento potencial de la gran oportunidad de desarrollo económico y social en nuestra provincia (Misiones), que representa la constitución de empresas de base tecnológica para el desarrollo de aplicaciones en el dominio de I/WoT. Considerando, en particular, el emergente proyecto del gobierno provincial delineado y en vías de implantación concreta por parte de Marandú Comunicaciones SE [22]. Dicha empresa de servicios tecnológicos, ha iniciado el despliegue e implantación de una red de IoT [23]; contando con su propia Cloud abierta asociada. Red, que se prevé de cobertura integral para nuestra provincia, que permitirá disponer de distintos tipos de sensores que podrán instalarse tanto en ciudades como en zonas rurales. Estos sensores servirán, de acuerdo a lo que prevé Marandú, para medir tráfico vehicular, ocupación de puestos de estacionamiento, creación de contenedores de basura inteligentes, medir polución, contaminación, temperatura, humedad, variables inherentes a la calidad del agua, entre otros aspectos diversos, en contextos de materialización y aplicación de ciudades, agro o industria, inteligentes.

3. FORMACIÓN DE RECURSOS

El equipo de docentes-investigadores, constituido para este proyecto, radicado en CITIC-UGD, incluye: un Doctor en Tecnologías de la Información y Comunicaciones y Magister en Ingeniería de Software; un Ingeniero de Sistemas y

Maestrando en Tecnologías de la Información; y, un Ingeniero en Telecomunicaciones, Maestrando en Redes de Datos. Adicionalmente, el proyecto cuenta con la colaboración de un tutor externo; Doctor en Ingeniería Informática y docente-investigador perteneciente a UNCPBA.

Por otra parte, en el contexto de este proyecto, se prevé el desarrollo de: un Trabajo Final de Maestría en Tecnologías de la Información; y, dos Trabajos Finales de Carrera en Ingeniería en Informática.

4. BIBLIOGRAFÍA

- [1] G. Rossi, M. Urbieta, D. Distant, J. M. Rivero and S. Firmenich, "25 Years of Model-Driven Web Engineering. What we achieved, What is missing," *CLEI Electronic Journal*, vol. 19, no. 3, pp. 5-57, 2016.
- [2] P. Asghari, A. M. Rahmani and H. H. S. Javadi, "Internet of Things applications: A systematic review," *Computer Networks*, vol. 148, pp. 241-261, 2019.
- [3] W3C, "Web of Things (WoT) Description," [Online]. Available: <https://www.w3.org/TR/wot-thing-description/>.
- [4] S. Kent, "Model Driven Engineering," in *Integrated Formal Methods. Lecture Notes in Computer Science*, vol. 2335, Springer, 2002, pp. 286-298.
- [5] C. Pons, R. S. Giandini y G. Pérez, Desarrollo de software dirigido por modelos. Conceptos teóricos y su aplicación práctica, EDULP | McGraw-Hill Educación, 2010.
- [6] M. Brambilla, J. Cabot and M. Wimmer, Model-Driven Software Engineering in Practice, 1st ed., Morgan & Claypool Publishers, 2012.
- [7] K. Wakil and D. N. A. Jawawi, "Model Driven Web Engineering: A Systematic Mapping Study," *e-Informatica Software Engineering Journal*, vol. 9, no. 1, p. 107-142, 2015.
- [8] LMU - Institute for Informatics, "UWE – UML-based Web Engineering," [Online]. Available: <http://uwe.pst.ifi.lmu.de/>.
- [9] IWT2, "NDT Metodology," 2004-2014. [Online]. Available: <http://iwt2.org/en/group-activities/research/results/ndt-metodology/>.
- [10] Object Management Group, Inc., "IFML: The Interaction Flow Modeling Language," 1997-2018. [Online]. Available: <https://www.ifml.org/>.
- [11] M. Weinberger, M. Köhler, D. Wörner and F. Wortmann, "Platforms for the Internet of Things –An Analysis of Existing Solutions," in *5th Bosch Conf. on Systems and Software Engineering*, 2014.
- [12] J. Guth, U. Breitenbücher, M. Falkenthal, P. Fremantle and O. Kopp, "Detailed Analysis of IoT Platform Architectures: Concepts, Similarities, and Differences," in *Internet of Everything*, Springer, 2018, pp. 81-101.
- [13] Amazon Web Services, Inc., "AWS IoT," [Online]. Available: <https://aws.amazon.com/es/iot/>.
- [14] FIWARE Foundation, "FIWARE: The Open Source Platform for Our Smart Digital Future," [Online]. Available: <https://www.fiware.org/>.
- [15] Open MTC, "Open MTC in a nutshell," [Online]. Available: <http://www.openmtc.org/>.
- [16] SmartThings Inc., "Samsung SmartThings," [Online]. Available: <https://www.smartthings.com/>.
- [17] IBM, "Watson IoT Platform," [Online]. Available: <https://www.ibm.com/ar-es/business-operations/iot-platform>.
- [18] Microsoft, "Azure IoT Hub," [Online]. Available: <https://azure.microsoft.com/es-es/services/iot-hub/>.
- [19] Google, "Cloud IoT Core," [Online]. Available: <https://cloud.google.com/iot-core?hl=es>.
- [20] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele University and Durham University, 2007.
- [21] V. Garousi, M. Felderer and M. V. Mäntylä, "The need for multivocal literature reviews in software engineering," in *20th Int. Conf. on Evaluation and Assessment in Software Engineering*, 2016.
- [22] Marandú Comunicaciones SE, "Marandú Comunicaciones SE," 2021. [Online]. Available: <https://marandu.com.ar/>.
- [23] noticiadel6.com, "[Artículo periodístico] Misiones ya cuenta con la primera antena Sigfox para la Internet de las Cosas," 2020. [Online]. Available: <https://www.noticiadel6.com/>.

Enfoques y Tendencias en el Desarrollo de Aplicaciones para Dispositivos Móviles.

Pablo Thomas , Lisandro Delia , Leonardo Corbalán ,

Juan Fernández Sosa , Fernando Tesone , Verónica Aguirre , Verena Olsowy ,

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

{pthomas, ldelia, corbalan, jfernandez, ftesone, vaguirre, volsowy, ppesado}@lidi.info.unlp.edu.ar

RESUMEN

Se presenta una línea de investigación y desarrollo centrada en el estudio de aspectos de Ingeniería de Software aplicados a las diferentes tendencias en el desarrollo de Aplicaciones para Dispositivos Móviles.

Palabras claves: Plataformas para Dispositivos Móviles - Aplicaciones Nativas- Aplicaciones Híbridas – Aplicaciones Interpretadas – Aplicaciones por Compilación Cruzada – Aplicaciones Web Progresivas – Offline First. – Instant App

CONTEXTO

Esta línea de Investigación forma parte del Proyecto (2018-2021) “*Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso*”, 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.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa y Latinoamérica.

1. INTRODUCCIÓN

Los avances tecnológicos en la comunicación móvil, junto con los avances en el campo de la computación han dado lugar a la computación móvil, que puede definirse como un entorno de cómputo con movilidad física que brinda al usuario la capacidad de acceder a datos, información u otros objetos lógicos desde cualquier dispositivo en cualquier red, al mismo tiempo que va moviéndose, cambiando su ubicación geográfica [1].

Las características específicas de la computación móvil plantean nuevos desafíos en el desarrollo de software para este tipo de dispositivos. La necesidad de tratar con diversos estándares, protocolos y tecnologías de red; las capacidades limitadas, no sólo en cuanto a procesamiento, sino también, en cuanto a consumo, tamaño físico y capacidad de almacenamiento de los dispositivos; las restricciones de tiempo impuestas por un mercado altamente dinámico y la existencia de distintas plataformas de hardware y software son sólo algunas de las dificultades a las que se enfrentan los desarrolladores en esta área.

Las aplicaciones se generan en un entorno dinámico e incierto. En su mayoría se trata de aplicaciones pequeñas, no

críticas, destinadas a un gran número de usuarios finales que son liberadas en versiones rápidas para poder satisfacer las demandas del mercado. En otros casos las aplicaciones son de mayor tamaño, a veces para brindar movilidad a una parte de la funcionalidad de un sistema más grande, mientras que otras veces son el único punto de interacción del sistema.

Algunas iniciativas como *offline first* proponen nuevas maneras de abordar el desarrollo de aplicaciones móviles tomando como requerimiento no funcional la posibilidad de seguir brindando un servicio incluso cuando el dispositivo pierde conectividad con la red móvil.

Todas las particularidades previamente mencionadas hacen que el desarrollo de software para dispositivos móviles difiera considerablemente del tradicional. Ello conduce a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software como disciplina, acompañando este proceso de desarrollo tecnológico.

La existencia de una enorme competencia en el mercado de las aplicaciones móviles, promovida en parte por la proliferación de dispositivos móviles y el número de negocios que están migrando sus servicios a Internet, revela la importancia que implica para las empresas el posicionamiento en este mercado. Para conseguirlo, es necesario reducir al máximo el tiempo de desarrollo de las aplicaciones y al mismo tiempo hacer que éstas se ejecuten en el mayor número de dispositivos posible.

Este propósito se ve obstaculizado por la excesiva fragmentación de hardware y software existente, originada por el alto número de dispositivos distintos, con sus propios sistemas operativos y plataformas de desarrollo.

La alternativa más costosa para hacer frente a esta problemática consiste en el desarrollo específico de aplicaciones de manera nativa en cada una de las

plataformas existentes, utilizando entornos de desarrollo, lenguajes de programación y tecnologías propias de cada plataforma.

El desarrollo nativo de aplicaciones móviles posee ciertas ventajas como la posibilidad de acceder sin limitaciones a todas las características del dispositivo (cámara, GPS, acelerómetro y agenda, entre otras), el alto rendimiento, interfaces gráficas consistentes con el resto de la plataforma, la posibilidad de trabajar sin acceso a Internet y de correr procesos en segundo plano. Estas aplicaciones pueden distribuirse a través de las tiendas en línea correspondientes. Sin embargo, el precio de todas estas ventajas es alto: no es posible reutilizar el código fuente entre plataformas diferentes, el esfuerzo se multiplica y se elevan los costos de desarrollo, actualización y distribución de nuevas versiones.

Como alternativa al desarrollo nativo en múltiples plataformas, se presentan diversos enfoques de desarrollo que procuran optimizar la relación costo/beneficio compartiendo la misma base de código fuente entre las versiones para las distintas plataformas [2]. La construcción de aplicaciones Web Móviles constituye un ejemplo representativo de este enfoque. Estas aplicaciones se diseñan para correr dentro de un navegador, se desarrollan con tecnologías web bien conocidas (HTML, CSS y JavaScript), no necesitan adecuarse a ningún entorno operativo móvil específico; su puesta en marcha es rápida y sencilla.

Las desventajas de las aplicaciones Web Móviles recaen sobre su rendimiento. Los tiempos de respuesta se dilatan afectados por la interacción cliente-servidor y las restricciones de seguridad impuestas a la ejecución de código por medio del navegador limitan el acceso a todas las capacidades del dispositivo. Además, la experiencia de usuario dista de ser similar a la de las aplicaciones nativas, resultando menos atractiva para el usuario final.

El enfoque híbrido es una alternativa que posee también la ventaja de estar basado en tecnologías web estándar (HTML, Javascript y CSS) pero, a diferencia del anterior no funciona dentro de un navegador, sino en un contenedor web especial con mayor acceso a las características del dispositivo a través de una API específica.

Las aplicaciones híbridas permiten la reutilización de código en las distintas plataformas, el acceso al hardware del dispositivo, y la distribución a través de las tiendas de aplicaciones. Sin embargo, conservan algunas de las desventajas de las aplicaciones Web Móviles: la utilización de componentes no nativos en la interfaz perjudica la experiencia de usuario, y la ejecución se ve ralentizada por la carga asociada al contenedor web.

El enfoque interpretado se presenta como una alternativa en donde las aplicaciones son traducidas en su mayor parte a código nativo, mientras que un resto se interpreta en ejecución. Se implementan de forma independiente de las plataformas utilizando diversas tecnologías y lenguajes, tales como Javascript, Typescript y XML, entre otros.

La obtención de interfaces nativas constituye una de las principales ventajas de este tipo de aplicaciones, mientras que la definición de nuevas componentes suele tener un alto grado de complejidad debido a la necesidad de definir abstracciones compatibles con diferentes plataformas.

Finalmente, las aplicaciones generadas por compilación cruzada también constituyen un tipo de desarrollo multiplataforma. Estas aplicaciones se compilan de manera nativa creando una versión específica de alto rendimiento para cada plataforma destino.

Un nuevo concepto ha surgido en los últimos años denominado Aplicaciones Web Progresivas (PWA por sus siglas en inglés). Una PWA es una aplicación web que utiliza las últimas tecnologías

disponibles en los navegadores para ofrecer en dispositivos móviles una experiencia lo más parecida posible a la de una aplicación nativa.

Los objetivos que persiguen las PWA son: lograr el mayor rendimiento posible en dispositivos móviles, que la aplicación cargue de manera casi instantánea, que la interfaz de usuario se parezca lo máximo posible a una nativa, que se pueda trabajar sin conexión (*offline first*) y que se puedan enviar notificaciones a los usuarios, como en una aplicación nativa.

Una nueva tecnología, desarrollada recientemente por Google, denominada TWA (Trusted Web Activities) o Actividades Web de Confianza, permite integrar una PWA con una aplicación Android. Las TWA son ejecutadas desde un APK y distribuidas desde Google Play Store. Muestran en pantalla completa un navegador web dentro de una aplicación Android sin mostrar la interfaz del navegador.

Las TWA representan un punto de inflexión para los desarrolladores que ahora pueden distribuir sus PWA en Google Play Store que ha dejado de ser una tienda exclusiva de aplicaciones nativas.

Desde 2017 los desarrolladores de Android tienen una nueva opción para hacer llegar sus apps a los usuarios finales. El concepto de *Instant App* permite ejecutar una funcionalidad específica de una aplicación sin necesidad de instalar la App completa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Enfoques de desarrollo de Aplicaciones para Dispositivos Móviles.
- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones Nativas en Android [3].

- Aplicaciones Nativas en iOS [3].
- Aplicaciones Web Móviles.
- Aplicaciones Móviles Híbridas (PhoneGap [4], Ionic [5]).
- Aplicaciones Móviles Interpretadas (Appcelerator Titanium [6], NativeScript [7]).
- Aplicaciones Móviles generadas por compilación cruzada (Xamarin [8], Corona [9]).
- PWA, Instant App y Offline First.
- TWA (Trusted Web Activities),
- 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.
- Experiencia de usuario en Aplicaciones Móviles generadas con distintos enfoques de desarrollo.
- Internet de las Cosas (IoT), Internet de Todo (IoE) y aplicaciones de sensado móvil o sensado urbano.
- Aplicaciones Móviles y su utilización en el proyecto de *Smart City*.
- El rol de los dispositivos móviles en aplicaciones de *crowdsourcing* o colaboración abierta distribuida.
- El rol de las aplicaciones móviles como herramientas para hacer frente a la pandemia de COVID-19

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos/esperados se pueden resumir en:

- Se ha estudiado el impacto que tiene el enfoque de desarrollo elegido sobre el rendimiento de las aplicaciones construidas. Se consideraron las plataformas iOS y Android junto a varios *frameworks* de desarrollo multiplataforma. Los resultados y

conclusiones fueron publicados en [10] y [11].

- Se desarrolló "*Informática UNLP*", una aplicación móvil para la comunidad de la Facultad de Informática de la Universidad Nacional de La Plata [12]. Esta aplicación contribuye a la mejor comunicación alumno/docente. Entre otras funciones incluye una cartelera virtual ajustable a las preferencias de los alumnos y utiliza tecnología de realidad aumentada para proveer información sobre la ocupación de las aulas en tiempo real. Las primeras versiones fueron liberadas para Android pero actualmente también está disponible para iPhone y iPad.
- En [13] se estudiaron las implicaciones de la elección de un enfoque de desarrollo determinado sobre la eficiencia energética de las aplicaciones generadas. Se contemplaron aplicaciones con alta carga de procesamiento, reproducción de video y acceso a imágenes generadas con diversos enfoques de desarrollo.
- En [14] se han publicado los resultados de un estudio sobre el impacto que tienen 9 tecnologías distintas de desarrollo de aplicaciones móviles sobre 23 características de interés para la Ingeniería de Software.
- En [15] se estudió el modo en que distintos *frameworks* de desarrollo multiplataforma afectan el tamaño de la aplicación construida. Los resultados son relevantes dado que el espacio disponible se ha convertido en un recurso crítico para muchos usuarios.
- Se ha cuantificado el impacto que tiene la elección del *framework* de desarrollo sobre tres de los requerimientos no funcionales más demandados por los usuarios de dispositivos móviles: rendimiento, consumo de energía y uso de espacio de almacenamiento. Los resultados y conclusiones se publicaron

en [16] y constituyen una guía para los desarrolladores al momento de elegir el *framework* más adecuado a sus expectativas.

- Las Aplicaciones Web Progresivas o PWA, por sus siglas en inglés, fueron analizadas en [17]. Dada su estrecha relación con las aplicaciones web móviles fueron comparadas con estas últimas explicitando diferencias y similitudes. Se concluyeron una serie de ventajas y desventajas de un enfoque respecto del otro.
- Se desarrolló “InnovApp”, una PWA que permite realizar una visita virtual guiada a la muestra de Ciencia y Tecnología 2019 llevada a cabo en el Centro de Innovación y Transferencia Tecnológica (CIyTT) de la Facultad de Informática de la UNLP
- Se han estudiado distintas estrategias para enfrentar la pandemia de COVID-19 que incluyen tecnología móvil. Para ello se analizaron las aplicaciones promovidas por los gobiernos de 22 países distintos como herramientas tecnológicas para combatir la Pandemia. Los resultados y conclusiones se publicaron en [18].
- Se continúa con el desarrollo de nuevas características de “*Informática UNLP*”, una aplicación multiplataforma en continua evolución y crecimiento.
- Se continúa con el estudio de PWA y el análisis de los alcances de la tecnología TWA (Trusted Web Activities). TWA es la estrategia implementada por Google para integrar una PWA con una aplicación Android que puede ser entonces distribuida desde Google Play Store. Actualmente se está realizando un análisis detallado de la combinación PWA+TWA, teniendo en cuenta características no funcionales, de desarrollo, y de la gestión de proyectos de software.
- Se esperan resultados concretos al estudiar las posibilidades que surgen de

la utilización de los sensores de los dispositivos móviles en aplicaciones de Internet de las Cosas (IoT), Internet de Todo (IoE) y particularmente aplicaciones de sensado móvil, también conocido como sensado urbano.

- Se examinará la utilidad de las Aplicaciones Móviles en el contexto de las ciudades inteligentes, haciendo foco especialmente en aquellas que implementan la colaboración abierta distribuida (*crowdsourcing*).
- Se analizarán ventajas y desventajas de las “*Instant App*” de Android. Resulta de interés su estudio y posterior comparación con las aplicaciones nativas tradicionales.
- Se Profundizará en el estudio sobre la aplicación de la tecnología móvil como herramienta para mejorar algún aspecto de la lucha contra la pandemia COVID-19.
- Se examinarán las distintas formas de persistencia de datos en aplicaciones móviles en función del enfoque/*framework* de desarrollo utilizado.
- Se Promoverá el avance sostenido y continuo de la formación de los miembros involucrados en esta línea de investigación.

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. Hongxing Li, Guochu Shou, Yihong Hu, Zhigang Guo. *Mobile Edge Computing: Progress and Challenges*. 2016

- 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud). Oxford UK.
2. Spyros Xanthopoulos, Stelios Xinogalos, ***A Comparative Analysis of Cross-platform Development Approaches for Mobile Applications***, BCI' 2013, Greece
3. Tracy, K.W., ***Mobile Application Development Experiences on Apple's iOS and Android OS***, Potentials, IEEE, 2012.
4. <http://phonegap.com/>
5. <https://ionicframework.com/>
6. <http://www.appcelerator.com/>
7. <https://www.nativescript.org/>
8. <http://xamarin.com/>
9. <https://coronalabs.com/>
10. Delia L., Galdamez N., Corbalan L., Thomas P., Pesado P., ***Un Análisis comparativo de rendimiento en Aplicaciones Móviles Multiplataforma***, CACIC 2015, UNNOBA Junín, Octubre 2015.
11. Delia L., Galdamez N., Corbalan L., Thomas P., Pesado P. ***Approaches to Mobile Application Development: Comparative Performance Analysis*** SAI Computing Conference (SAI), 2017. Londres, del 18 al 20 de Julio de 2017.
12. Fernández Sosa J., Cuitiño A., Thomas P., Delia L., Caseres G., Corbalán L., Pesado P. ***"Informática UNLP" la App de la Facultad de Informática***. CACIC 2017. La Plata, del 9 al 13 de Octubre de 2017.
13. L. Corbalan, J. Fernandez Sosa, A. Cuitiño, L. Delía, G. Cáseres, P. Thomas, P. Pesado. ***"Development Frameworks for Mobile Devices: A Comparative Study about Energy Consumption"***. 5th IEEE/ACM International Conference on Mobile Software Engineering and Systems. Gothenburg, Sweden. 27 y 28 de Mayo del 2018.
14. L. Delia, P. Thomas, L. Corbalan, J. Fernandez Sosa, A. Cuitiño, G. Cáseres, P. Pesado. ***"Development Approaches for Mobile Applications: Comparative Analysis of Features"*** SAI - Computing Conference 2018. London, United Kingdom. 10 al 12 de Julio de 2018.
15. J. Fernández Sosa, P. Thomas, L. Delía, G. Cáseres, L. Corbalán, F. Tesone, A. Cuitiño, P. Pesado. ***"Mobile Application Development Approaches: A Comparative Analysis on the Use of Storage Space"***. XV Workshop de Ingeniería de Software - XXIV Congreso Argentino de Ciencias de la Computación CACIC 2018. Tandil, Argentina. 8 al 12 de octubre de 2018.
16. Corbalán, L., Thomas, P., Delía, L., Cáseres, G., Sosa, J. F., Tesone, F., & Pesado, P. (2019, June). A Study of Non-Functional Requirements in Apps for Mobile Devices. In *Conference on Cloud Computing and Big Data* (pp. 125-136). Springer, Cham
17. Aguirre, V., Ortu, A., Delía, L. N., Thomas, P. J., Corbalán, L. C., Cáseres, G., & Pesado, P. M. (2019). PWA para unificar el desarrollo Desktop, Web y Mobile. In *XXV Congreso Argentino de Ciencias de la Computación (CACIC) (Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019)*.
18. Fernandez Sosa, J., Aguirre, V., Delía, L., Thomas, P., Corbalán, L., & Pesado, P. M. (2020). COVID-19: un análisis comparativo de Apps. In *XXVI Congreso Argentino de Ciencias de la Computación (CACIC) (Universidad Nacional de La Matanza, 5 al 9 de octubre de 2020)*.

Estrategia de evaluación de la reusabilidad de los conjuntos de datos de los portales de Infraestructuras de Datos Espaciales.

María A. Barrera¹, Carlos Salgado²

¹Departamento de Informática
Facultad de Tecnología y Ciencias Aplicadas
Universidad Nacional de Catamarca
Maximio Victoria N° 55 – (4700) Catamarca - Argentina
Tel.: +54-383- 4435112 – int 168
marita.latingeo@gmail.com

²Departamento de Informática
Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – Argentina
Tel.: +54-266-4520300– Int. 2101
csalgado@unsl.edu.ar

Resumen

La información geoespacial se ha transformado en un bien de consumo primario en nuestra sociedad. A nivel mundial, las organizaciones gubernamentales han contribuido a esta realidad a través de las Infraestructuras de Datos Espaciales (IDE), promoviendo el desarrollo económico, estimulando una mejor gobernanza y fomentando la sostenibilidad ambiental. (Vilches-Blázquez, Ballari, 2020)

Las IDE, son aplicaciones web que se utilizan para buscar y acceder a información geográfica (IG) y servicios geográficos asociados a través de Internet (Inierto M., Núñez et al, 2014). Están definidas por un conjunto de políticas, tecnologías y acuerdos institucionales que aprovechan la provisión y uso de datos y servicios espaciales estandarizados.

Sin embargo, se ha prestado poca atención al dominio de la información geoespacial, que está llamando cada vez más la atención de los tomadores de decisiones que buscan mejorar los

modelos de decisión considerando factores espacio-temporales. En este contexto, esta investigación realiza un análisis de la información publicada en los geoportales IDE, con el objeto de medir el grado de reutilización de sus conjuntos de datos, para esto se propone una adaptación de la métrica MELODA, teniendo en cuenta la naturaleza específica de los geodatos.

Palabras Claves: Datos Abiertos, Datos Geoespaciales, Infraestructura de Datos Espaciales, Métricas e Indicadores, Meloda.

Contexto

La presente investigación está contextualizada en el trabajo colaborativo entre dos grupos de investigación: Los integrantes del 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 N° P-03-2020; y, por otra parte, el grupo de investigación interviniente, del proyecto de investigación: “Las Tic al Servicio del Dato Abierto: Situación Actual, Conceptualización e Iniciativas de Apertura de Información Pública”, del departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca.

El trabajo colaborativo entre las universidades aportará dimensiones analíticas prácticas, condición indispensable para integrar y comparar la información obtenida a partir de los proyectos que genere cada grupo. En este contexto, el grupo de la Universidad Nacional de Catamarca aporta su conocimiento y experticia en lo referente al análisis de los Datos Abiertos Gubernamentales y desde la Universidad de San Luis, el grupo de investigación proveerá el conocimiento y la experiencia necesaria para definir y adaptar conceptos y modelos de calidad a este nuevo escenario de las ciudades modernas. Su experiencia en definir modelos, métodos, métricas e indicadores.

Introducción

La disponibilidad, accesibilidad y el intercambio de información georeferenciada son requerimientos que cobran paulatinamente mayor importancia para una sociedad interesada en incorporar este tipo de información como apoyo fundamental en los procesos de planificación y toma de decisiones, tanto a nivel local como regional e inclusive global (Rix et al., 2011). Esta utilización de conjuntos de datos espaciales conlleva múltiples abordajes para el diseño y construcción de herramientas tecnológicas dirigidas hacia este fin (Rajabifard & Williamson, 2001).

Algunos autores resumen la idea anterior diciendo que “la información geoespacial se ha transformado en un bien de consumo primario en nuestra sociedad. A nivel mundial, las organizaciones gubernamentales han contribuido a esta realidad a través de Infraestructuras de Datos Espaciales (IDE), promoviendo el desarrollo económico, estimulando una mejor gobernanza y fomentando la sostenibilidad ambiental”. (Vilches-Blázquez, Ballari, 2020).

Las IDE son aplicaciones web que actúan como un punto de entrada a la información Geográfica (IG) distribuida y disponible para ser compartida. Esta IG proveniente de instituciones oficiales, es publicada, con frecuencia, bajo estándares y normas, asegurando su interoperabilidad y uso, ofreciendo como mínimo algunas funcionalidades básicas, tales como: visualización de mapas, localización y descarga de datos y servicios, para el procesamiento de datos. La idea de publicar estos datos para los ciudadanos, es que éstos puedan consultarlos, descargarlos o incluirlos en aplicaciones externas que modelan, analizan, muestran y/o generan datos, información y productos geoespaciales de valor agregado en múltiples formatos de archivo para algunos propósitos de información (por ejemplo, visualización) o toma de decisiones (es decir, a través del análisis), lo que abre nuevas posibilidades para generar valor, tanto en el ámbito social como en el económico.

El costo de producir esa información es bastante elevado, por lo que ahora el desafío es demostrar sus capacidades y la reutilización que se hace de esos datos. En este sentido, y en lo que respecta al diseño que presentan los portales IDE, se puede observar que no responde a las necesidades de los distintos perfiles de usuarios que necesitan acceder en busca de información, lo que determina un uso y

reuso bastante limitado de los mismos y, en algunos casos, que solo sean utilizados por profesionales muy especializados de IG. Además, aunque estos datos se encuentran potencialmente abiertos, muchos de ellos resultan difíciles de encontrar, desorganizados y en un formato que no es el adecuado para análisis espacial, es decir, estos datos no son “técnicamente” abiertos, por lo que resulta necesario hacer un análisis de la efectiva utilidad de los datos ofrecidos por una IDE.

Por otra parte, no se cuenta con estándares específicos de análisis de calidad que describan las características que deben tener los recursos que componen los conjuntos de datos geográficos disponibles en una IDE para ser efectivamente reutilizados, ya que hay que tener en cuenta que no es igual el acceso a los mismos y su difusión, que su reutilización.

La presente investigación realiza el análisis de la información publicada en los geoportales, con el objeto de medir el grado de reutilización de sus conjuntos de datos, para esto se propone una adaptación de la métrica MELODA (<http://www.meloda.org>), teniendo en cuenta la naturaleza específica de los geodatos, su topología y su localización espacial y tomando como caso de prueba la IDE de la República Argentina (<https://www.idera.gob.ar>).

MELODA (Metric for releasing open data) es una métrica de evaluación que permite calificar la información y evaluar el grado de reutilización de los datos públicos y el valor de la información publicada en los portales de DA. Ayuda a los organismos productores a hacer que todos los datos que publiquen sean buenos para la reutilización. También ayuda a los consumidores de open data a encontrar los mejores datos a reutilizar con el objetivo de crear nuevos productos

y servicios de calidad. (Albella A. et al, 2014).

Si bien es cierto que ya comienzan a observarse iniciativas y guías de buenas prácticas que han tomado este nuevo reto y han planteado algunas estrategias para mejorar el nivel de reuso de la información geográfica abierta, aún hay muchos interrogantes que siguen sin resolverse.

En el presente trabajo se pretende realizar un análisis general de la información publicada en los portales de IDE, con el objeto de medir el grado de reutilización efectiva de sus conjuntos de datos, teniendo en cuenta estándares técnicos, acceso legal, formatos de datos, visualización, su utilización en la recuperación de información, y las exigencias que las mismas tienen por parte de organismos internacionales de normalización para asegurar su calidad e interoperabilidad, para que las IDE sean herramientas ciertas que mejoren el acceso y reutilización de los datos abiertos geoespaciales a toda la ciudadanía.

Abella, Ortiz-de-Urbina-Criado y De Pablos-Heredero (2014; 2017a) han elaborado la métrica MELODA que permite evaluar el grado de reutilización de los datos abiertos. MELODA analiza seis dimensiones: Estándares Técnicos, Acceso a la Información, Marco Legal, Modelo de Datos, Geolocalización de la Información y Actualización en Tiempo Real. Cada una de estas dimensiones consta de cinco niveles mediante los cuales se obtiene un porcentaje de calidad de las mismas, dicho porcentaje se utiliza posteriormente en una fórmula que proporcionará el nivel de reusabilidad de los conjuntos de datos disponibles en los portales abiertos.

Esta última métrica se ha utilizado por otros autores como Vicente-Paños y Jordán-Alfonso (2017) para estudiar el

acceso a la información del sector público y su reutilización en las comunidades autónomas de España. Estos autores observaron que cuando se calificaban a través de Meloda los portales de datos abiertos de esas comunidades, las puntuaciones más bajas se obtenían en las dimensiones de información geolocalizada e información en tiempo real, por lo que estimamos que esta métrica es la que más se adecúa como base de análisis para la evaluación de los datos disponibles en una IDE.

Líneas de investigación y desarrollo

La presente es una investigación que se desprende del Proyecto: "Las TIC al Servicio del Dato Abierto: Situación actual, conceptualización e iniciativas de apertura de información Pública".

La naturaleza de este estudio es la investigación descriptiva con dos propósitos principales. En primer lugar, se ocupa de la explicación de las motivaciones, procesos y dificultades asociados a la adopción de métricas de evaluación de reusabilidad a nivel general para cualquier portal de DA. El segundo propósito apunta a describir los procedimientos que una IDE realiza para la apertura de los conjuntos de datos que pone a disposición en su geoportal. Por lo tanto, el método cualitativo resulta ser el enfoque más apropiado para esta investigación. Este enfoque ayuda a comprender mejor los fenómenos y obtener un conocimiento exhaustivo sobre el estudio de los procesos y los participantes. Con respecto a la obtención de información se prevé recolección de muestras mediante examen de características de las administraciones públicas que ofrecen a la ciudadanía sus DA geoespaciales a

través de las Infraestructuras de Datos Espaciales.

En cuanto a la métrica utilizada para el análisis de los datos, se ha seleccionado a MELODA. Esta métrica se utiliza realizando un análisis de los conjuntos de datos, en seis dimensiones: Estándares Técnicos, Acceso a la Información, Marco Legal, Modelo de Datos, Geolocalización de la Información y Actualización en Tiempo Real. Cada una de estas dimensiones consta de cinco niveles mediante los cuales se obtiene un porcentaje de calidad de las mismas, dicho porcentaje se utiliza posteriormente en una fórmula que proporcionará el nivel de reusabilidad de los conjuntos de datos disponibles en los portales abiertos.

Resultados y Objetivos

El objetivo principal planteado para la ejecución del trabajo fue: Evaluar la reusabilidad de los conjuntos de datos disponibles en las Infraestructura de Datos Espaciales. En este sentido, nos planteamos redefinir las dimensiones e indicadores de la métrica MELODA, y de su fórmula aplicada para el cálculo, presentando una propuesta cuantitativa para evaluar la reusabilidad de los geodatos disponibles en una IDE, teniendo en cuenta la naturaleza específica de la IG.

Entre las nuevas dimensiones propuestas, podemos destacar aquellas que tienen que ver con el análisis de los conjuntos de datos, teniendo en cuenta aspectos tales como: estándares técnicos, acceso legal, formatos de datos y metadatos, sistemas de referencia espacial, visualización en servicios de mapas, su utilización en la recuperación de información, además de las exigencias que las mismas tienen por parte de organismos internacionales de normalización, y que tienen relación

directa con los servicios web que ofrecen las IDE a través de sus geoportales, tales como el visualizador o servicio de mapas, servicio de fenómenos, servicios de coberturas, servicios de catálogos, etc., ya que estos son los componentes más utilizados para la realización de los análisis de datos espaciales en estas infraestructuras y los que son requeridos por profesionales y desarrolladores de nuevas aplicaciones.

El análisis de los conjuntos de datos, teniendo en cuenta las dimensiones propuestas nos dará un porcentaje de calidad de las mismas, dicho porcentaje se utiliza posteriormente en una nueva fórmula de cálculo que proporcionará el nivel de reusabilidad de los conjuntos de datos disponibles en los portales abiertos de una IDE.

Formación de Recursos Humanos

El primer autor se desempeña como docente de la carrera de Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca. Mientras que el segundo autor es docente de la Facultad de Ciencias Físico-Matemáticas y Naturales de la Universidad Nacional de San Luis. La temática abordada en este proyecto de investigación en desarrollo, dio lugar al desarrollo de una tesis para optar al título de Magister en Ingeniería en Software. El mismo se desarrolla en la Facultad de Ciencias Físico-Matemáticas y Naturales de la Universidad Nacional de San Luis, en donde el Mg. Ing. Carlos Salgado, se desempeña como director de dicha Tesis. Además, se trabaja en el desarrollo de tesinas de grado y trabajos finales de carrera en el Marco de la Ingeniería informática y la Licenciatura en Ciencias de la Computación de la FCFMyN de la UNSL y la carrera de

Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca

Bibliografía

- Vilches-Blázquez Luis M. y Ballari Daniela (2020): Unveiling the diversity of spatial data infrastructures in Latin America: evidence from an exploratory inquiry, *Cartography and Geographic Information Science*, DOI: 10.1080/15230406.2020.1772113; <https://doi.org/10.1080/15230406.2020.1772113>
- Zarazaga-Soria, F.J. et al. (2012) “Hacia las Infraestructuras de Datos Abiertos Espaciales” - Actas de las III Jornadas Ibéricas de Infraestructuras de Datos Espaciales (JIIDE'2012).
- Iniesto M., Núñez et all. (2014). “Introducción a las Infraestructuras de Datos Espaciales”. <http://publicacionesoficiales.boe.es>. NIPO: 162-14-022-0 –
- Metric for releasing open data - MELODA “<http://www.meloda.org>
- Infraestructura de Datos Espaciales de la República Argentina - IDERA (2020) <https://www.idera.gob.ar/>
- Abella, Aberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen (2019). “Meloda 5: A metric to assess open data reusability”. *El profesional de la información*, v. 28, n. 6, e280620. <https://doi.org/10.3145/epi.2019>.
- Vicente-Paños, A.; Jordán-Alfonso, A. (2017). “Acceso a la información pública y su reutilización en las comunidades autónomas: evaluación de la reutilización de datos abiertos”. *El profesional de la información*, v. 26, n. 3, pp. 381-391.
- Infraestructura de Datos Espaciales de la República Argentina - IDERA (2020) <https://www.idera.gob.ar/>

Estudios en torno a la Accesibilidad Web. Aporte regional en el NEA

Sonia I. Mariño, Pedro L. Alfonzo, Romina Y. Alderete, Verónica K. Pagnoni, Maria L. Gronda,
José Bordón

*Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste,
Corrientes, Argentina*
simarinio@yahoo.com, plalfonzo@hotmail.com

*Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste,
Corrientes, Argentina*

RESUMEN

Se sintetizan las propuestas que involucran la definición de métodos, procedimientos y herramientas en el marco de actividades de I+D+I. El propósito es innovar desde lo metodológico y procedimental para lograr mejoras sustantivas en torno a la Accesibilidad Web. En los casos de estudio, se seleccionan situaciones problemáticas del contexto identificados por los participantes, quienes disponen de conocimiento sobre las mismas.

Palabras clave: recursos humanos, métodos, procedimientos, AW, herramientas TIC.

CONTEXTO

En el marco de proyectos de I+D+I acreditados por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste, denominados: “TI en los sistemas de información: modelos, métodos y herramientas” y “Sistemas informáticos. Modelos, métodos y herramientas”, en que se inscriben becas, formación de docentes-investigadores y Trabajos Finales de Maestría, se promueve la identificación de problemáticas en el contexto de desarrollo informático y profesional de los implicados con miras a proponer soluciones que aporten a la región.

1. INTRODUCCIÓN

La Universidad, una institución intensiva en conocimiento, se constituye en un agente promotor de innovaciones en un contexto complejo en el que continuamente emergen nuevos requerimientos y desafíos [1, 2, 3].

La RedUNCI incorpora en el documento de Recomendaciones Curriculares entre las cuestiones sociales y profesionales la Accesibilidad web [4]. Otras evidencias que permiten tratar un abordaje social tecnológico se menciona en Barchini [5] quien sostiene que la disciplina Informática es una disciplina bio-psico-socio-tecno-cultural. Estos referentes, entre otros, permiten proponer un abordaje en torno a cuestiones de responsabilidad social que se puede relacionar con la Responsabilidad Social Universitaria como se menciona en [6, 7].

Siguiendo a [8] “los estudios empíricos permiten crear conocimiento que pueden mejorar la práctica”. Por lo expuesto, en una sociedad del conocimiento en que las TIC emergieron para suplantar numerosas actividades, particularmente ocasionada por la Pandemia causada por el COVID-19 es menester asegurar la calidad de los contenidos a la web.

La Accesibilidad Web (AW), es un aspecto de la calidad del software. Se refiere al acceso universal a la Web, independientemente del tipo de hardware, software, infraestructura de red, idioma, cultura, localización geográfica y capacidades de los usuarios [9, 10].

En [11] se expuso una reseña de métodos, procedimientos y herramientas con miras a contribuir al desarrollo regional y en [12] y [13] el abordaje desde la AW, enfatizando la formación de recursos humanos que contribuyen desde espacios universitarios al Sector de Servicios y Sistemas Informáticos.

Particularmente, se sintetizan los avances en torno a la AW, temática que implica una mirada desde la TIC para la sociedad, como estrategia para el acceso a la información por

diferentes colectivos. Se enfatiza que estas experiencias permiten ampliar la categoría “métodos y aplicaciones prácticas” propuesta en [14].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Siguiendo lo expuesto en [11] las líneas de I+D+I precisadas para el logro de las propuestas consisten en:

- Revisión de la literatura [12, 13]. Se identificaron y seleccionaron antecedentes disciplinares [10, 15] y casos de aplicaciones similares de los tratados en cada uno de los proyectos que se sintetizan en este artículo. Avanzar en esta estrategia, que produce información actualizada, facilita la definición y alcance de los proyectos y la innovación de la producción en función al contexto al que está dirigido.
- Identificación de aspectos metodológicos de la Ingeniería del Software, la calidad del software y su tratamiento en propuestas de AW. Los estándares y propuestas [10, 15].
- Definición y adaptación de métodos y procedimientos para la realización de los estudios basados en [16, 17, 18, 19]. Derivada de la revisión disciplinar se opta por aquellos enfoques metodológicos, normativas y estándares que integrados aportan a mejoras en el contexto en que emergen y en que se validan.
- Identificación de herramientas TIC, privilegiando las comprendidas en la filosofía Open Source, tanto para la construcción de soluciones informáticas ante los problemas delimitados y como en la evaluación de la AW.

3. RESULTADOS OBTENIDOS/ESPERADOS

A continuación se sintetizan los proyectos presentados con miras a profundizar actividades de I+D, iniciarse en becas de formación de grado y producciones de trabajos finales de maestría o tesis. En estas propuestas se integran las líneas de trabajo mencionadas.

En [20] se propone la construcción de un procedimiento para implementar la creación de Objetos de Aprendizaje (OA) en un entorno e-learning, en particular seleccionado como contexto de validación una asignatura de finalización de estudios universitarios de grado. Se avanzó en el estudio y selección de estándares para la construcción de OA, y de herramientas de autor que facilitan su elaboración. En particular se indagará en cómo incorporar principios de AW a fin de asegurar el acceso a los contenidos y así apoyar procesos educativos significativos. Esta producción también se corresponde con el Trabajo Final de Maestría en Tecnologías de la Información.

En [21] se avanza en la elaboración de un marco de trabajo ajustado a los estándares de AW universales factible de ser aplicado a entornos virtuales de enseñanza-aprendizaje destinados a la formación continua. Este marco se centra en el papel primordial del usuario así como en el uso de métricas, considerando puntualmente la accesibilidad visual [16, 21, 22, 23]. La evaluación con los usuarios es fundamental pues ayuda a identificar problemas en cómo se están aplicando las soluciones técnicas [23]. En tanto el uso de métricas permite una valoración más objetiva haciendo foco en las variables de mayor importancia en la AW [16]. En particular la propuesta se validará en un contexto de educación superior no universitaria.

La inclusión digital en un tema central en contextos universitarios. En [15] se describe la automatización de la auto-percepción de capacidades visuales de estudiantes universitarios a través de SIU-KOLLA. La encuesta original se adaptó a cuestiones identificadas en el contexto de implementación. Actualmente, se disponen de información generada en los ciclos lectivos 2019 [15] y 2020. Esta propuesta puede ser adaptada y transferida a otros contextos dada la importancia que reviste la accesibilidad en los entornos virtuales actualmente. La solución sintetizada se enmarca en una Tesis de Maestría en Ingeniería del Software.

En estudios de calidad de productos software, por ejemplo los centrados en la AW, es relevante la aplicación de métricas que orientan el cumplimiento de los principios y los criterios en los productos software objeto de estudio. Diversas herramientas como TAW [24], Examinator [25], entre otras [26], proporcionan medidas del cumplimiento. Además, es viable la aplicación de otras a fin de obtener una visión general respecto a que principios requieren de mayor atención o cuáles criterios deben contemplarse en mayor medida a fin de disminuir las limitaciones – total para ver, grave para ver, de los miembros superiores, para comprender, derivadas de la edad– de acceso a los contenidos. Como se observa estas limitaciones están en consonancia con la definición del alcance de la AW.

En [16, 27, 28, 29, 30] se incluyen algunas evaluaciones a partir de evidencias empíricas que aportan al conocimiento, experiencia con miras a determinar los niveles de accesibilidad.

En [27] se describe una propuesta que integra la evaluación de la AW en la Ingeniería del Software Basada en la Evidencia (ISBE), específicamente en la etapa 2.

En [30] se presenta un aporte al ámbito académico, a través del desarrollo de una herramienta software parametrizable orientada a la gestión de recursos físicos, como ser las aulas y laboratorios de informática, considerando en etapas temprana de su construcción criterios relacionados con la AW, en particular con accesibilidad visual y de esta manera contribuir a la inclusión social de los e-ciudadanos. En su desarrollo se aplicó un proceso iterativo que al finalizar cada incremento verifica el cumplimiento de los criterios de calidad considerados.

Una experiencia de mantenimiento correctivo, aspecto tendiente a asegurar la calidad del software, en particular la AW y de esta manera generar productos tecnológicos de calidad y permitir la inclusión digital de los individuos independientemente de sus capacidades, se describe en [31].

Las propuestas sintetizadas precedentemente ilustran distintos casos de estudio, desarrollo e investigación aplicada en que se profundizan temas de la disciplina Informática, en particular se aborda la calidad de los productos software desde la Accesibilidad Web. Se aborda a partir del estudio, análisis, adaptación y análisis de diversos modelos, métodos y herramientas con fines de aportar mejoras sustanciales a los contextos en los cuales se identificaron las necesidades de mejora, en particular respecto a la accesibilidad a los contenidos.

Los procesos de diseño, desarrollo y evaluaciones tratados en estas propuestas son soportados por procesos de gestión de conocimiento que involucran la captura, almacenamiento, procesamiento y análisis con fines de producir información oportuna para apoyar la toma de decisiones mediadas por TIC.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de este proyecto se desarrollan y desarrollaron becas de grado, tesis de maestría, formación de graduados y de docentes-investigadores. El conocimiento generado, adquirido y consolidado a través de estas propuestas, comprende a los recursos humanos de los proyectos de I+D+I acreditados por la Universidad y que realizan investigaciones aplicadas con la finalidad de profundizar sus conocimientos y plasmarlos en productos tecnológicos que contribuyen directamente a la mejora en sus contextos laborales.

REFERENCIAS

- [1] I. Nonaka, H. Takeuchi. “The Knowledge-creating company. How japanese companies create the dynamics of innovations”. Oxford University Press, New York.1995
- [2] S. I. Mariño y M. V. Godoy, Gestión del conocimiento y sistemas informáticos. Una propuesta para las organizaciones del siglo XXI. XX Workshop de Investigadores en Ciencias de la

- Computación, 2018, pp. 703-705.
- [3] A. H. Mazorodze y S. Buckley, Knowledge management in knowledge e intensive organisations: Understanding its benefits, processes, infrastructure and barriers, *South African Journal of Information Management*, 2019, 21(1)
- [4] RedUNCI. Documento de Recomendaciones Curriculares de la RedUNCI2014-2015. Disponible en: <http://redunci.info.unlp.edu.ar>, 2014.
- [5] G. E. Barchini, Informática. Una disciplina bio-psico-socio-tecno-cultural, *Revista Ingeniería Informática*, Edición 12, abril 2006, Disponible en <http://inf.udec.cl/~revista/ediciones/edicion12/articulo%2012-3.pdf>
- [6] S. I. Mariño, P. L. Alfonzo y M. V. Godoy. La Responsabilidad Social centrada en la accesibilidad web. *European Scientific Journal*, 16(30), 35-50, 2020.
- [7] S. I. Mariño, P. L. Alfonzo y M. V. Godoy. Capítulo: Accesibilidad web. Un aporte de responsabilidad social universitaria, Libro: A Ciencia da Computacao e o desenvolvimento de conteneúdo tecnológico relevante para a sociedade Editorial: Atena, 2020. DOI 10.22533/at.ed.683202003
- [8] M. Género, “Ingeniería del software basada en la evidencia,” Escuela Superior de Informática, Universidad de Castilla-La Mancha, 2016. Material en diapositiva
- [9] S. Luján Mora. Accesibilidad Web, Definición de accesibilidad web, 2020. Disponible en: <http://accesibilidadweb.dlsi.ua.es/?menu=definicion>
- [10] Consorcio World Wide Web (W3C). Disponible en: <http://www.w3c.es/>
- [11] S. I. Mariño, M. V. Godoy, P. Alfonzo, M. E. Sánchez, S. N. Ivaniszyn, V. Pagnoni, R. Alderete, M. Fernández, G. de Los Reyes, J. M. Bordón, Diseño de métodos, procedimientos y herramientas. Aportes al desarrollo regional, Workshop de Investigadores en Ciencias de la Computación 2020, WICC 2020
- [12] S. I. Mariño, P. L. Alfonzo y V. K. Pagnoni. “Accesibilidad web visual. Formación de RRHH para fomentar su desarrollo en la región NEA”. XXI Workshop de Investigadores en Ciencias de la Computación - WICC, 2019, Argentina.
- [13] S. I. Mariño, P. L. Alfonzo, V. K. Pagnoni, L. Gronda, C. Galain Garcia, J. Maidana, J. P. Gallardo, V. Blanco y M. V. Godoy, “Formación Universitaria en torno a la Accesibilidad Web. Algunas experiencias en la FaCENA-UNNE”. 48 Jornadas Argentinas de Informática – JAIIO, 2019, Argentina.
- [14] G. Barchini y M. Sosa, “La informática como disciplina científica. Ensayo de mapeo disciplinar,” *Revista de Informática Educativa y Medios Audiovisuales*. Año 1, vol. 1, no. 2, pp.1-11, 2004.
- [15] Web Content Accessibility Guidelines (WCAG) 2.0. Disponible en: <http://www.w3.org/TR/WCAG20/>
- [16] SEDIC (s.f.) Accesibilidad Web: Técnicas y herramientas para evaluar la accesibilidad web. Disponible en: <http://www.sedic.es/autoformacion/accesibilidad/9-tecnicas-herramientas.html>.
- [17] E. Serrano Mascaraque, A. Moratilla Ocaña y I. Olmeda Martos, Métrica para la evaluación de la accesibilidad en Internet: propuesta y testeo. *Revista Española de Documentación Científica*, 33(3), 378–396, 2010.
- [18] B. A. Kitchenham y S. L. Pfleeger. Personal opinion surveys. En: F. Shull, J. Singer, Sjoberg, D.I.K. (eds.) *Guide to advanced empirical software engineering*. Springer, London, 2008, pp. 63–92.
- [19] M. Género, J. A. Cruz-Lemus y M. Piattini. *Métodos de Investigación en Ingeniería del Software*. RaMa. 2014. pp. 49-62.
- [20] R. Y. Alderete, Proceso para la creación de Objetos de Aprendizaje para asignaturas de finalización de carrera, Trabajo Final de Maestría en Tecnologías de la Información, Universidad Nacional del Nordeste, Dir. S. I. Mariño 2019.

- [21] V. K. Pagnoni, (2021) Marco de trabajo de evaluación de la accesibilidad web para entornos virtuales de enseñanza-aprendizaje Tesis de Maestría en Ingeniería de Software: (en desarrollo).
- [22] V. K. Pagnoni y S. I. Mariño, Calidad de contenidos en dominios de educación. Evaluación de la accesibilidad Web mediada por validadores automáticos, EDMETIC, 10(6), 2019, pp. 107-127.
- [23] J. M. Bordón, Sistema informático para la gestión de espacios físicos en ámbitos de la Educación Superior, Beca pregrado, 2019, Directores. P. L. Alfonzo, S. I. Mariño.
- [24] TAW. Test de Accesibilidad Web. Disponible en: <https://www.tawdis.net/>
- [25] Benavidez, C. Libro Blanco de eXaminator, Disponible en: http://examinator.ws/info/libro_blanco_examinator.pdf
- [26] Test de Accesibilidad. Disponible en: <https://play.google.com/store/apps/details?hl=es&id=com.google.android.apps.accessibility.auditor>
- [27] S. I. Mariño y P. L. Alfonzo, “Evidencias de Accesibilidad Web en la generación de sitios. Propuesta de un método” *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, no. 23, pp. 52-60, 2019.
- [28] S. I. Mariño y P. L. Alfonzo. “Calidad en uso en plataformas educativas: estudio centrado en la Accesibilidad Web”. XXV Congreso Argentino de Ciencias de la Computación - CACIC, 2019, Argentina.
- [29] S. I. Mariño, P. L. Alfonzo y M. V. Godoy, Medidas de accesibilidad web. Aplicación en una plataforma educativa, *European Scientific Journal*, 16(1), 11-22, 2020.
- [30] J. M. Bordón, P. L. Alfonzo y S. I. Mariño. “Criterios de calidad del software. Un enfoque de Accesibilidad Web vinculado a la gestión de recursos físicos”. III Congreso Internacional de Ciencias de la Computación y Sistemas de Información – CICCSI, 2019. Argentina.
- [31] S. I. Mariño, J. P. Gallardo y P. L. Alfonzo. “Mantenimiento del software, aspectos aplicados en sitios web accesibles”. III Congreso Internacional de Ciencias de la Computación y Sistemas de Información – CICCSI 2019.

EVALUACIÓN DEL ATRIBUTO APRENDIZAJE EN APLICACIONES WEB

José L. Andrada¹, Ana Funes², Aristides Dasso²

¹Universidad Nacional de La Rioja, La Rioja, Argentina
joseluisandrada2002@gmail.com

²SEG / Departamento de Informática / Facultad de Ciencias Físico-Matemáticas y Naturales /
Universidad Nacional de San Luis
Ejército de los Andes 950, D5700HHW San Luis, Argentina
+54 (0) 266 4520300, ext. 2126
{afunes, arisdas}@unsl.edu.ar

RESUMEN

En este trabajo, presentamos los objetivos, lineamientos generales y resultados esperados de una línea de investigación sobre la creación de modelos de evaluación de la característica “Aprendizaje” en aplicaciones web. Dicha línea de investigación forma parte integral del desarrollo de modelos de evaluación de sistemas complejos. Considerando que la evaluación de la estructura y metodología de implementación del Aprendizaje de un sistema de software, que debería incluir los puntos necesarios para que el sistema pueda ser fácilmente aprendido, implica una evaluación de un sistema complejo, es que esta investigación tiene como objetivo la creación, puesta a punto y aplicación de diversos modelos que permitan obtener indicadores del nivel alcanzado en la implementación de medidas que contribuyan a facilitar el aprendizaje de aplicaciones web.

Comprobar el atributo Aprendizaje es fundamentalmente valioso para aplicaciones y sistemas complejos a los que los usuarios acceden con frecuencia, no obstante saber cuán rápido los usuarios pueden prosperar en el uso de su interfaz es valioso incluso para sistemas objetivamente simples.

En esta primera etapa, se aborda el estudio y la aplicación de la metodología para el desarrollo de dichos modelos de evaluación siguiendo los lineamientos del método Logic

Score of Preference (LSP) [3].

Asimismo, se toma como referencia para la creación del modelo, estándares reconocidos como los de la International Organization for Standardization (ISO). Así, de acuerdo con la ISO/IEC 25010 [4], el atributo Aprendizaje es uno de los seis componentes de calidad de la Usabilidad (o Capacidad de Uso); otros son la Inteligibilidad (Appropriateness Recognizability), Operatividad (Operability), Protección Frente a Errores de Usuarios (User Error Protection), Estética de la Interfaz de Usuario (User Interface Aesthetics) y la Accesibilidad (Accessibility), como puede observarse en la figura 1, en donde se muestran todas las dimensiones de la calidad del software de acuerdo a dicha norma. También se siguen los lineamientos dados en la norma ISO 9241-110: Ergonomics of Human-system Intereaction [8] y la parte 210: Human-centred design for interactive systems de la ISO 9241-210 [9]. Ambas sirven de guía a las organizaciones para, entre otras cosas, formular e implementar estrategias de desarrollo a lo que dan en llamar ‘Learnability’ del software. Estos atributos pueden ser aumentados con desarrollos propios, enriqueciendo los requisitos considerados en la evaluación del Aprendizaje. Asimismo continuamos con el estudio de nuevas características a incluir en esta investigación con el fin de ajustar el modelo adecuadamente.

Palabras clave: Facilidad de Aprendizaje. Learnability de sistemas web. Usabilidad.

Atributos de Calidad del Software. Métodos de Evaluación. Logic Score of Preference (LSP).



Figura 1

CONTEXTO

El trabajo de investigación aquí presentado se encuentra enmarcado dentro del ámbito del SEG (Software Engineering Group), de la Universidad Nacional de San Luis, ejecutándose dentro de una de las líneas de investigación del Proyecto de Incentivos código 22/F222 “Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software de Calidad”, dirigido por el Dr. Daniel Riesco y co-dirigido por el Dr. Roberto Uzal. El mismo se encuentra acreditado con evaluación externa y financiamiento de la Universidad Nacional de San Luis.

INTRODUCCIÓN

En una organización, una vez elegido un proceso de desarrollo de software resulta imperioso contar con un modelo que permita, dentro de dicho proceso, la evaluación de cuán fácil resulta el aprendizaje del uso del software por parte de los usuarios. Este modelo debe permitir conocer y controlar los atributos que contribuyen a que un producto sea amigable, fácil de usar y de aprender.

Si bien la construcción de modelos de evaluación de sistemas complejos, entre los que se encuentran los sistemas de software, constituye una necesidad importante, no es una tarea sencilla. Múltiples aspectos deben ser considerados en esta tarea, teniendo en cuenta los aspectos funcionales del sistema, por ejemplo, aspectos que hacen a la facilidad

de entendimiento tales como la facilidad de lectura, familiaridad, legibilidad visual, predictibilidad, entre muchos otros. Todos estos aspectos van a definir cuán rápida es la velocidad de aprendizaje para que los usuarios puedan emplear todas las características del sistema en el menor tiempo posible y con la mayor profundidad.

Por lo tanto, para una organización preocupada en la calidad de sus productos de software, conocer y controlar el atributo Aprendizaje de un sistema de software, en desarrollo o ya desarrollado, resulta necesario contar con estándares así como con herramientas apropiadas para evaluar el grado de adecuación con dichos estándares.

En este sentido, existen en la literatura y en las web múltiples propuestas. En [11] se elabora un Learnability Attributes Model (Modelo de Atributos de Aprendizaje)¹, basado en lo indicado en la ISO/IEC 25010: 2011 [10]. Se descompone el atributo de calidad Learnability en sub ítems. En el artículo se presenta, además, un caso de estudio.

Manuela Unsöld, en su tesis de posgrado [13], discute varias definiciones de ‘Learnability’ así como métodos de evaluación de este atributo de calidad del software; finalmente, construye un framework

¹ Learnability no es una palabra del idioma inglés; la hemos traducido simplemente como ‘Aprendizaje’. Algunos autores lo traducen como Facilidad de Aprendizaje, otros como Capacidad de Aprendizaje.

para evaluar y clasificar los distintos métodos presentados.

Robillard [12] realiza una encuesta con el objetivo de encontrar las causas que hacen que una API sea difícil de aprender a usarla. Identifica obstáculos que agrupa en cinco categorías: Recursos; Estructura; Antecedentes; Medioambiente Técnico; Proceso.

Por nuestra parte, en este trabajo de investigación, nos proponemos como objetivo construir un modelo que, siguiendo las pautas del método LSP y aquellas indicadas en las normas ISO antes mencionadas, permita construir un modelo que sirva para evaluar el atributo de calidad “Aprendizaje” de un sistema de software.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Este trabajo se desarrolla enmarcado dentro de una línea de investigación del SEG (Software Engineering Group) de la Universidad Nacional de San Luis; creemos que el mismo reviste un gran interés y potencial de desarrollo. En este sentido, cabe aclarar que se trata de una extensión de una línea de investigación más amplia y consolidada dentro del grupo, que se ocupa de la aplicación y desarrollo de técnicas de evaluación de atributos de sistemas de software.

Dentro de los diversos modelos de evaluación antes desarrollados en este contexto, parte del esfuerzo ha estado dirigido hacia la evaluación de atributos de calidad del software, como por ejemplo la evaluación de la Accesibilidad de aplicaciones web (ver, por ejemplo, [6], [7]) o aspectos relacionados con su Seguridad (ver [5], [1], [2]). También, el método ha sido propuesto para el desarrollo de una herramienta para la evaluación de carreras de grado [4]. En el trabajo presente, en particular, nos enfocamos en otro de los atributos de la calidad del software, el Aprendizaje, para producir un modelo que sirva para conocer cuán fácil resulta aprender a usar un sistema de software. Esto se realiza

teniendo en cuenta normas establecidas, como la ISO 25010 y sus complementarias así como aspectos considerados en las normas ISO 9241-110 e ISO 9241-210.

Para construir el modelo de evaluación final, el Árbol de Requerimientos constituye el primer paso en el desarrollo de dicho modelo, de acuerdo a las pautas indicadas por el método LSP. Esta estructura jerárquica permite que cada sub ítem pueda también ser evaluado de forma parcial y a su vez agregado bajo una de las funciones de agregación provistas en el método, brindando de esta manera no solo un indicador global del atributo Aprendizaje sino también un conjunto de indicadores parciales para cada uno de las sub características o sub ítems en el árbol.

A modo de ilustración, en la Tabla 1 se muestra una primera versión de la estructura del Árbol de Requerimientos. Como puede observarse, se ha descompuesto el atributo “Aprendizaje” en tres sub ítems, cada uno de los cuales a su vez también ha sido descompuesto hasta llegar a las hojas del árbol que son aquellos atributos que deben ser medidos ya sea por medio de una métrica directa o indirecta.

Tabla 1. Requerimientos de Adecuación al Aprendizaje

1. Facilidad de entendimiento.
1.1. Legibilidad visual.
1.1.1. Adecuación de fuente.
1.1.2. Adecuación de la visualización textual.
1.1.3. Disposición.
1.2. Facilidad de lectura.
1.2.1. Agrupación Cohesiva de la Información.
1.2.2. Densidad de información.
1.3. Familiaridad
1.3.1. Consistencia de formato
1.3.2. Internacionalización
1.3.3. Metáfora
1.4. Ahorro de esfuerzo
1.4.1. Acciones mínimas
1.4.2. Auto-descripción
1.4.3. Complejidad de la

Tabla 1. Requerimientos de Adecuación al Aprendizaje
información
1.4.4. Organización de la Información
1.5. Orientación al usuario
1.5.1. Calidad de los mensajes de actualización
1.5.2. Calidad de los mensajes de aviso
1.5.3. Retroalimentación inmediata de los controles
1.6. 1.6 Navegabilidad
1.6.1. Soporte a búsqueda interna
1.6.2. Clickabilidad
1.6.3. Interconectividad
1.6.4. Alcanzabilidad
2. Facilidad de aprendizaje
2.1. Predictibilidad
2.1.1. Nombres de enlaces significativos
2.1.2. Etiquetas significativas
2.1.3. Controles significativos
2.1.4. Contenido multimedia significativo
2.2. Potencialidad
2.2.1. Determinación de acciones posibles
2.2.2. Determinación de acciones prometedoras
2.3. Retroalimentación informativa
2.3.1. Progreso explícito de las transacciones
2.3.2. Contexto explícito del usuario
2.4. Completitud de las tareas
2.5. Comandos utilizados
3. Facilidades de ayuda
3.1. Completitud de la ayuda online
3.2. Documentación Multi-usuario
3.3. Completitud del mapa de sitio
3.4. Calidad de los mensajes de asesoramiento
3.5. Uso de la documentación.

RESULTADOS Y OBJETIVOS

Una medida elevada del atributo Aprendizaje contribuye a la Usabilidad y a la

calidad general de un producto de software. Esto trae aparejado una rápida incorporación del sistema, lo que se traduce en bajos costos de capacitación. Además, puede resultar en una alta satisfacción por parte del usuario porque éste se sentirá seguro de sus habilidades. Por lo cual, medir el nivel del atributo Aprendizaje es fundamentalmente valioso en los productos software. En consecuencia, el objetivo principal, que nos hemos planteado en este trabajo es proponer un modelo de evaluación para el atributo de calidad “Aprendizaje” de un sistema de software; dicho modelo es desarrollado por medio de la aplicación del método genérico de evaluación de sistemas complejos LSP. En el modelo que se propone, el mismo se encuentra definido conforme a las normas ISO 25010 la cual descompone la característica Usabilidad en sub-características y atributos medibles a los que se les asocian métricas definidas genéricamente, donde uno de estos atributos en el Aprendizaje; también se basa en características aportadas por las normas ISO 9241-110 e ISO 9241-210.

FORMACIÓN DE RECURSOS HUMANOS

Dentro del SEG (Software Engineering Group), en el ámbito de la Universidad Nacional de San Luis, en el que se ejecuta el Proyecto de Incentivos código 22/F222 “Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software de Calidad”, se vienen llevando a cabo numerosas tesis de grado y de posgrado.

En este sentido, creemos que la línea de investigación aquí descripta, la cual es una extensión de una línea más amplia sobre aplicación y desarrollo de modelos de evaluación de sistemas de software, seguirá dando sus frutos, tanto en publicaciones nacionales e internacionales como en la formación de recursos humanos (2 tesis de maestría presentadas, más una tesis de maestría finalizada lista para su defensa). Asimismo, de momento, se ha encarado la posibilidad de la ejecución de una nueva tesis

de maestría basada en los objetivos que aquí nos hemos propuesto.

REFERENCIAS Y BIBLIOGRAFÍA

- [1] Dasso, Aristides y Funes, Ana. “Threat and Risk Assessment Using Continuous Logic”, *Encyclopedia of Organizational Knowledge, Administration, and Technologies*, 1st. edition. IGI Global. Aceptado para su publicación en 2020.
- [2] Dasso, Aristides, Funes, Ana, Montejano, Germán, Riesco, D. Uzal, Roberto, Debnath, N. “Model Based Evaluation of Cybersecurity Implementations”. ITNG 2016. Las Vegas, Nevada, USA, 11-13 abril 2016. In S. Latifi (ed.), *Information Technology New Generations, Advances in Intelligent Systems and Computing* 448. DOI: 10.1007/978-3-319-32467-8_28. Springer International Publishing, Switzerland 2016.
- [3] Dujmović, Jozo. “Soft Computing Evaluation Logic. The LSP Decision Method and Its Applications”. Wiley, IEEE Press. Hoboken, NJ : John Wiley & Sons, 2018
- [4] Funes, Ana, Berón, Mario, Dasso, Aristides. “Soporte en la Toma de Decisiones para la Evaluación de Carreras de Grado”. 5° Congreso Nacional de Ingeniería Informática / Sistemas de Información (5° CoNaIISI 2017), 2 - 3 de Noviembre, Universidad Tecnológica Nacional Facultad Regional Santa Fe. Santa Fe, Argentina, 2017. *Actas de CoNaIISI 2017*, pp.861-870.
- [5] Funes, Ana, Dasso, Aristides, Montejano, Germán, Riesco, Daniel. “A SAMM-based model for assessing Cybersecurity Implementations”, *actas de CoNaIISI 2018*, 29 y 30 de Noviembre de 2018, Mar del Plata, Buenos Aires, Argentina.
- [6] Gallardo, Cecilia, Funes, Ana y Ahumada, Hernán. Soporte para la Medición y Evaluación de la Accesibilidad al Contenido en Aplicaciones Web. *Anales de ASSE 2019. Simposio Argentino de Ingeniería de Software (48 JAIIO) - ISSN 2451-7593* pp. 56-70. 2019
- [7] Gallardo, Cecilia, Funes, Ana. Un Modelo para la Evaluación de la Calidad de la Accesibilidad al Contenido Web, *Actas de CONAIIISI 2015*, ISBN: 978-987-1896-47-9. <http://conaiisi2015.utn.edu.ar/memorias.html>
- [8] ISO 9241-110, International Standard. 2006-04-01. Ergonomics of Human-system Interaction. Part 110: Dialogue Principles.
- [9] ISO 9241-210 International Standard. 2010-03-15 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems
- [10] ISO/IEC 25010:2011 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuARE) — System and software quality models
- [11] Rafique, Irfan, Jingnong, Weng, Yunhong, Wang, Qanber Abbasi, Maissom, Lew, Philip, Wang, Xinran. Evaluating Software Learnability. A Learnability Attributes Model. 2012 International Conference on Systems and Informatics (ICSAI 2012)
- [12] Robillard, Martin P. What Makes APIs Hard to Learn? Answers from Developers. November/December 2009 *IEEE Software*.
- [13] Unsöld, Manuela. Measuring Learnability in Human-Computer Interaction. Master’s thesis at Universität Ulm. Faculty of Engineering, Computer Science and Psychology, Databases and Information Systems Department. Version from September 3, 2018

EVALUACIÓN Y MEJORA DE NARRATIVAS DIGITALES BASADAS EN DATOS

Emanuel Irrazabal; Andrea Lezcano Airaldi; Joaquín Acevedo Duprato

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
ISIS
{eirrazabal, alezcano}@exa.unne.edu.ar

RESUMEN

Esta línea se enmarca en el Proyecto F018-2017; una continuación de los proyectos F07-2009 y F10-2013, ambos enfocados en modelos, métodos y herramientas para la calidad del software de Universidad Nacional del Nordeste. Este proyecto aborda los temas emergentes en el área de la calidad de software, en particular, aspectos referidos a la visualización de la información aplicando técnicas narrativas o de *storytelling* y a las buenas prácticas asociadas a ella. Esta línea es apoyada por CONICET a través de la Beca Interna Doctoral, con una duración de 60 meses. Se pretende generar métodos y herramientas que permitan evaluar y mejorar la calidad de las visualizaciones en los productos software.

En particular, se está trabajando en la definición de un enfoque para la evaluación de la calidad y mejora asistida de visualizaciones aplicando técnicas narrativas. Esto ha incluido el desarrollo de un estudio de caso y de revisiones sistemáticas de la literatura.

Palabras clave: calidad de software, data storytelling, narrativa basada en datos, visualización de la información.

CONTEXTO

Las líneas de Investigación y Desarrollo presentada en este trabajo corresponden al proyecto PI-17F018 “Metodologías y herramientas emergentes para contribuir con la calidad del software”, acreditado por

la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021 y a la Beca Interna Doctoral de CONICET 2021 – 2025.

1. INTRODUCCIÓN

La narrativa digital o storytelling basada en datos es una estrategia popular y efectiva para transmitir información, siendo su propósito estimular la atención y la energía de la audiencia (o engagement) a través de emociones [1]. Una de las definiciones más utilizadas de storytelling es: “el arte de comunicar ideas a través de historias”[2]. En el contexto actual, donde las posibilidades de toma de datos de distintos sistemas software ha tenido un crecimiento explosivo (Big Data), la visualización de datos proporciona un medio valioso por permitir un análisis asistido (por ejemplo, mediante controles de sumarización, división de los datos, búsquedas contextuales, o gráficos interactivos) [3]. Más aún, dado que las narrativas digitales muchas veces se utilizan para la toma de decisiones, es importante considerar las perspectivas de los usuarios finales en la construcción de dichas narrativas. Esta interrelación entre: i) las fuentes de datos que dan soporte a una narrativa digital, y ii) los intereses, preferencias y necesidades de los destinatarios de la misma, determina en gran medida la calidad de las narrativas a construir. La no consideración de estos aspectos suele ocasionar problemas relacionados con: baja calidad del

contenido visual, pobre utilización de las fuentes de datos, ineficiencia en el desarrollo del contenido y, en general, poca entrega de valor.

En 2001, Gershon y Page propusieron por primera vez el uso del *storytelling* para la visualización[4], mientras que Segel y Heer [5] (en 2010) introducen el concepto de visualizaciones narrativas basadas en datos analizando una gran cantidad de ejemplos. En el contexto actual de desarrollos ágiles, uno de los problemas recurrentes es la poca calidad del contenido visual [6]. Por ello, la implementación de las buenas prácticas de *storytelling* en la construcción de un sistema y su posterior evaluación, pueden tener un impacto significativo en la calidad del producto resultante y en la entrega de valor por parte del sistema a los usuarios finales [7][8].

Se han realizado varios estudios que incluyen las prácticas y pautas generales a seguir para crear visualizaciones efectivas [1][3], [10] - [13] o [14]. Si bien se han propuesto algunos criterios para evaluaciones parciales, no se encuentran trabajos exhaustivos, herramientas que asistan a ello o propuestas sistemáticas de modelos para evaluar las narrativas basadas en datos. Entre los criterios descritos en la literatura pueden mencionarse: la aceptación de la historia por parte de los usuarios[15], el impacto de la historia [16], y qué tanto pueden ser recordados los puntos claves [17], el mensaje y el contenido [18]. Otro criterio importante es qué tanto se mantiene la atención de la audiencia durante la entrega de una historia [3] y su grado de compromiso, aunque hay múltiples definiciones, desde ver hasta interactuar, analizar o tomar decisiones [19].

Desde el punto de vista de cómo medir estos criterios, existen estudios que buscan cuantificarlos a partir de los datos estadísticos en los soportes digitales de las visualizaciones, principalmente los sitios Web. Así, por ejemplo, la cantidad de clics determinaría el nivel de audiencia o el

compromiso [1]. Otros trabajos estudian y adaptan las evaluaciones para medir el tiempo de permanencia en una página o la cantidad de visitas del mismo usuario [20]. En [21], el nivel de impacto también se cuantifica a partir de la cantidad de unidades de historia recordadas.

De acuerdo con lo expuesto anteriormente, la hipótesis de trabajo plantea que: "la definición de un modelo de evaluación para narrativas digitales basadas en datos que asimile los esfuerzos parciales existentes y los materialice en un entorno semiautomático de asistencia, puede mejorar el valor entregado de las narrativas respecto sus usuarios finales". En este caso, el valor entregado se relaciona directamente con el soporte para la toma de decisiones por parte de los usuarios. Se prevé una implementación del modelo vía métricas y refactorizaciones en función de dichas métricas, tomando analogías de otros dominios de desarrollo de software, tales como: requerimientos [22] diseño detallado [23] y documentación [6].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo principal de la investigación es "la definición de un enfoque para la evaluación de la calidad y mejora asistida de narrativas digital basadas en datos". Este objetivo se descompone en tres partes, atendiendo tanto aspectos metodológicos como tecnológicos:

1. Identificar un conjunto de acciones (o patrones) comúnmente utilizados en el diseño de narrativas digitales efectivas.
2. Elaborar un modelo de evaluación de la calidad de una narrativa digital, y particularmente sus formas de visualización en el marco de sitios Web.
3. Desarrollar herramientas para asistir en la aplicación del modelo de evaluación, y adicionalmente asistir en la provisión (semiautomática) de recomendaciones de mejora para el

equipo encargado del diseño de las visualizaciones.

La contribución esperada del trabajo consiste en la provisión de guías accionables (es decir, mediante un entorno semiautomatizado) para evaluar la calidad de las narrativas digitales basadas en datos y orientadas a la toma de decisiones, y por ende mejorar el proceso de construcción de estas.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el marco de este proyecto y respecto de la línea de gestión de procesos se llevó a cabo un estudio de caso comparando con el objetivo de fue determinar los beneficios de incluir las buenas prácticas de visualización de datos en el desarrollo de un sistema realizado en el contexto actual de crisis sanitaria con tiempos límite restringidos. Para ello, se presentó el estudio de caso realizado sobre 16 gráficos de un sistema de información que gestiona el seguimiento del aislamiento y los permisos de circulación en la cuarentena por la pandemia de COVID-9.

Los resultados del estudio evidencian que el uso de técnicas de narrativa basada en datos contribuye a facilitar el proceso de toma de decisiones mediante la implementación de buenas prácticas de visualización que aumentaron la comprensión y la memorabilidad de los gráficos. Asimismo, se comprueba que la causa por la cual no se aplicaron las buenas prácticas de visualización fue la falta de conocimiento, no la falta de tiempo. En ausencia de estas habilidades, se utilizan los ajustes predeterminados de las herramientas de desarrollo disminuyendo su comprensión.

Para este estudio se llevó a cabo el test de alfabetización visual [24] en el cual los participantes obtuvieron puntajes significativamente mayores cuando se trataba de visualizaciones que implementaban buenas prácticas,

demostrando una mayor comprensión con este tipo de gráficos.

Por otro lado, se está realizando una revisión sistemática de la literatura con el objetivo de caracterizar los modelos de evaluación de visualizaciones reportados al momento. En este contexto, se analizaron una serie de estudios secundarios sobre problemáticas similares, resumidos en la Tabla 1.

Tabla I: Estudios secundarios analizados por categoría

Categoría	# de RSL
Criterios de evaluación	3
Estrategias de evaluación	6
Problemas asociados al uso de visualizaciones	6
Dominios de aplicación de las visualizaciones	4

Para esta línea de trabajo se espera desarrollar un modelo de medición de calidad de visualizaciones a partir de la evaluación de las buenas prácticas implementadas.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo del Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 3 docentes investigadores, 1 becario interno CONICET y 1 becario de investigación de pregrado.

5. REFERENCIAS

- [1] Nussbaumer Knaflic, C. *Storytelling with Data: A data visualization guide for business professionals*. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2015.
- [2] Schreyögg, C. *Knowledge Management and Narratives: Organizational Effectiveness Through Storytelling* (9783503090297). 2006.
- [3] Kosara, R., Mackinlay, J. *Storytelling: The Next Step for Visualization*.

- Computer, Vol 46, Issue 5, 44-50 (2013)
- [4] Gershon, N., Page, W. What storytelling can do for information visualization. *Communications of the ACM*, 44(8):31-37, 2001.
- [5] Segel E., Heer. J. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139 – 1148, Dec. 2010.
- [6] Nicoletti, M., Schiaffino, S., Díaz Pace, J. A.: An optimization-based tool to support the cost-effective production of software architecture documentation. *J. Softw. Evol. Process.* 27(9): 674-699 (2015).
- [7] Baker, R. *Agile UX Storytelling*. 2017.
- [8] Zimmerman, B. Applying Tufte's principles of information design to creating effective Web sites. *ACM SIGDOC Annu. Int. Conf. Comput. Doc. Proc.*, pp. 309–317, 1997, doi: 10.1145/263367.263406.
- [9] Krausz, R. M., et al. E-Health in crisis management-emergency response to COVID 19 in Canada: Platform Development and Implementation. *JMIR Public Health and Surveillance* (2020).
- [10] Kosara, R. An Argument Structure for Data Stories. *Proc. Eurographics/IEEE VGTC Symposium on Visualization (EuroVis)*, 2017.
- [11] Tong, C. et al., "Storytelling and visualization: A survey," *VISIGRAPP 2018 - Proc. 13th Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl.*, vol. 3, pp. 212 224, 2018, doi: 10.5220/0006601102120224.
- [12] Boy, J., Detienne, F., Fekete, J. Storytelling in information visualizations: Does it engage users to explore data?. *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2015-April, pp. 1449–1458, 2015.
- [13] Tufte, E. R. *The Visual Display of Quantitative Information*. 1983.
- [14] Ram D., et al. Case Report: Rapid Development of Visualization Dashboards to Enhance Situation Awareness of COVID-19 Telehealth Initiatives at a Multi-Hospital Healthcare System. *Journal of the American Medical Informatics Association* (2020).
- [15] Lee, B., Kazi, R., Smith, G. SketchStory: Telling more engaging stories with data through freeform sketching. *IEEE Trans. on Visualization and Computer Graphics (Proc. of InfoVis)*, 19(12):2416–2425, 2013.
- [16] Borkin, M., Vo, A., Bylinskii, A., Isola, P., Sunkavalli, S., Oliva, A., Pfister, H. What makes a visualization memorable? *IEEE Trans. on Visualization and Computer Graphics (Proc. of InfoVis)*, 19(12):2306–2315, 2013.
- [17] Kim, K., Boy J., Lee S., Yi, J., Elmqvist, N. Towards an open visualization literacy testing platform, 2014.
- [18] Bateman, S., Mandryk, R., Gutwin, C., Genest, A., McDine, D., Brooks, C. Useful junk?: The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 2573–2582, 2010.
- [19] Mahyar, S., Kim, S., Kwon, B. Towards a taxonomy for evaluating user engagement in information visualization. In *Proc. of the IEEE VIS 2015 Workshop Personal Visualization*, 2015.
- [20] Saraiya, P., North, C., Lam, V., Duca, K. An insight-based longitudinal study of visual analytics. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 12(6):1511–1522, 2006.

- [21] Borkin, M., Bylinskii, Z., Kim, N., Bainbridge, C., Yeh, C., Borkin, D., Pfster, H., Oliva, A. Beyond memorability: Visualization recognition and recall. *IEEE Trans. on Visualization and Computer Graphics*.
- [22] Rago, A., Marcos, C., Diaz-Pace, J. A.: Identifying duplicate functionality in textual use cases by aligning semantic actions. *Software and Systems Modeling* 15(2): 579-603 (2016).
- [23] Vidal, S., Berra, I, Zulliani, S., Marcos, C., Díaz Pace, J. A.: Assessing the Refactoring of Brain Methods. *ACM Trans. Softw. Eng. Methodol.* 27(1): 2:1-2:43 (2018)
- [24] 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, 2017, doi: 10.1109/TVCG.2016.2598920.

Exploración de Técnicas de Machine Learning para Migración de Sistemas Legados hacia Microservicios

Guillermo Rodríguez¹ Leonardo Da Rocha Araujo¹ Fabio Rocha²
Rodrigo Pereira dos Santos³

¹ ISISTAN (UNICEN-CONICET), Tandil, Argentina.

² UNIT, Aracaju, Sergipe, Brasil.

³ UNIRIO, Rio de Janeiro, Brasil.

{guillermo.rodriguez, leonardo.araujo}@isistan.unicen.edu.ar, fabio.gomes@souunit.com.br,
rps@uniriotec.br

RESUMEN

Actualmente, cada vez más empresas están adoptando microservicios para modernizar sus productos y tomar ventaja de sus prometedoros beneficios como: agilidad, escalabilidad e integración continua, entre otros. Por un lado, los sistemas basados en microservicios presentan una arquitectura flexible y con alta capacidad de evolución. Sin embargo, por otro lado, hay desafíos técnicos (por ej. automatización de la infraestructura y *debugging* distribuido) y organizacionales (por ej. creación de equipos de trabajos *cross-functional*) que necesitan ser abordados.

Lamentablemente, migrar una arquitectura orientada a microservicios no es una tarea simple. En este proceso, los servicios pueden escalar más eficientemente y los ciclos de entregas se acortan debido al continuo despliegue. Normalmente, estas decisiones de diseño quedan sujetas a la intuición de desarrolladores y/o arquitectos, pero carecen de un análisis sistemático que les facilite la evaluación de alternativas y toma de decisiones. En este contexto, las técnicas de *machine learning* podrían contribuir a facilitar la exploración de diferentes alternativas de descomposición de arquitecturas de software en microservicios.

Palabras clave: *Arquitecturas de Microservicios, Sistemas Legados, Migración, Aprendizaje Automático, SOA.*

CONTEXTO

La propuesta emerge del proyecto PICT Joven (PICT-2018-01456, Migración de arquitecturas monolíticas hacia microservicios y contenedores) otorgado por la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), a través del Fondo para la Investigación Científica y Tecnológica (FONCyT). El mismo se desarrolla dentro del Instituto Superior de Ingeniería de Software Tandil (ISISTAN-CONICET) de la Facultad de Ciencias Exactas (EXA), Universidad Nacional del Centro de la provincia de Buenos Aires (UNCPBA), y unidad ejecutora del CONICET.

Colaboran con el proyecto miembros de Universidades extranjeras:

- Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Rio de Janeiro, Brasil.
- Universidade Tirandentes (UNIT), Sergipe, Brasil.

1. INTRODUCCIÓN

En el entorno dinámico del mundo de hoy, las aplicaciones de software necesitan responder rápidamente a las demandas del entorno y ser lo más ágiles posible. Las aplicaciones necesitan estar alineadas con necesidades de negocio cambiantes, a las cuales se le suma una diversidad de clientes, como navegadores de escritorio y móviles, como así también aplicaciones móviles nativas y APIs de terceras partes. Dado que es difícil cumplir estos requisitos mediante el uso de aplicaciones monolíticas, el estilo arquitectónico cliente-servidor monolítico está siendo reemplazado por el estilo de microservicios (Rahman, 2018). Cada vez más empresas están adoptando este estilo

arquitectónico para modernizar sus productos y tomar ventaja de beneficios tales como: agilidad, escalabilidad, e integración continua (Di Francesco, 2018). Estas ventajas se perciben como superadoras respecto a arquitecturas orientadas a servicios convencionales.

La definición más aceptada del concepto de microservicios apunta a un enfoque para desarrollar una aplicación vista como un conjunto de pequeños servicios, cada uno corriendo en su propio proceso y comunicándose con recursos ligeros como HTTP (Fowler, 2018). El estilo de Microservicios no sólo trae aparejados beneficios, sino también desafíos. Por un lado, los sistemas basados en microservicios presentan una arquitectura flexible y con alta capacidad de evolución. Sin embargo, hay varios desafíos técnicos (por ej. automatización de la infraestructura y *debugging* distribuido) y organizacionales (por ej. creación de equipos de trabajos *cross-functional*) que necesitan ser abordados. Por ejemplo, las cuestiones técnicas pueden ser que el sistema está altamente acoplado o es difícil para mantener; mientras que cuestiones relacionadas con el negocio pueden ser el largo tiempo de espera para entregar una nueva funcionalidad o baja productividad de los desarrolladores (Di Francesco, 2018). En consecuencia, adoptar un estilo de Microservicios no es un proceso trivial ya sea para desarrollar un sistema *greenfield*, o bien para migrar un sistema legado que presenta cuestiones difíciles de resolver con el paso del tiempo (Newman, 2015). En algunos casos, migrar hacia microservicios representa una buena opción para resolver los problemas existentes del sistema y, al mismo tiempo, mejorar su mantenimiento y la frecuencia de entregas de funcionalidad.

Los microservicios están contruidos sobre la idea de descomponer un gran servicio en un conjunto de servicios dueños de una sola responsabilidad, los cuales se comunican a través de la red. En este proceso, los servicios pueden escalar eficientemente y los ciclos de entregas se acortan debido al continuo despliegue (Ahmadvand, 2016).

Normalmente, estas decisiones de diseño quedan sujetas a la intuición de desarrolladores y/o arquitectos, pero carecen de un análisis sistemático que les facilite la evaluación de alternativas y toma de decisiones. Se han reportado experiencias interesantes de asistencia inteligente a usuarios en la derivación de diseños orientados a objetos a partir de especificaciones de arquitecturas de software orientadas a servicios (Rodríguez, 2018). En consecuencia, las técnicas de asistencia basadas en *machine learning* podrían servir para sistematizar y facilitar la exploración de diferentes alternativas descomposición de arquitecturas de software en microservicios. Para ello, es necesario contar con una base de conocimiento de aspectos estructurales, funcionales y no funcionales de la arquitectura, que evolucione a largo del tiempo y permita mejorar la precisión de las técnicas definidas y aumente la inteligencia del enfoque propuesto.

En este contexto, el objetivo del proyecto I+D es explorar técnicas de machine learning para asistir a los desarrolladores en descomponer (semi-) automáticamente arquitecturas de software en microservicios, considerando tanto aspectos funcionales como no funcionales.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Este proyecto propone un enfoque para migración de arquitecturas monolíticas hacia microservicios y contenedores. El estilo de arquitectura de microservicios está aumentando rápidamente ya que tiene muchas ventajas sobre otros estilos arquitectónicos como la escalabilidad, aislamiento mejorado de fallas (y, por lo tanto, resiliencia) y rendimiento mejorado.

La hipótesis de esta investigación es que existen técnicas de machine learning para detectar microservicios en arquitectura de software monolíticas. De esta manera, es posible identificar relaciones entre las operaciones de un sistema y los estados que toman las diferentes variables variables que

esas operaciones leen o escriben. Entonces es posible visualizar las relaciones entre las operaciones del sistema y las variables reconociendo clusters de relaciones. Asimismo, los aspectos no funcionales contenidos en las especificaciones de microservicios podrían proveer indicios sobre cómo esos microservicios son desplegados en contenedores.

Para corroborar nuestra hipótesis, se proponen los siguientes 4 ejes temáticos:

- Exploración de diferentes alternativas descomposición de arquitecturas de software;
- Visualización de software como asistencia a la descomposición de arquitecturas de software;
- Análisis de especificaciones de APIs REST de aplicaciones basadas en microservicios;
- Integración de arquitecturas de microservicios y contenedores a partir del conocimiento extraído de las especificaciones de los microservicios.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo de este proyecto es desarrollar un enfoque inteligente que asista a los desarrolladores en descomponer (semi-) automáticamente arquitecturas de software en microservicios, considerando tanto aspectos funcionales como no funcionales que contengan indicios sobre cómo se realizará el despliegue de dichos microservicios en contenedores.

Para lograr los objetivos generales descritos anteriormente, se buscará cumplir con un conjunto de objetivos específicos que describimos a continuación. El primer objetivo específico es definir técnicas de machine learning que permitan asistir a los desarrolladores en la exploración de diferentes alternativas descomposición de arquitecturas de software. Para llevar adelante este objetivo es necesario diseñar una metodología que permita vincular las fases de la migración de una arquitectura hacia microservicios. Esta metodología facilitará a

los arquitectos (semi-) automáticamente extraer, transformar, y re-diseñar la arquitectura conservando propiedades claves. Finalmente, para concluir el objetivo se apuntará a construir una base de conocimiento que permita capturar, sistematizar, recuperar conocimiento arquitectónico para evaluar alternativas de descomposición en microservicios. En relación a este objetivo, hemos hallado evidencias en el rol que juegan las técnicas de machine learning en el desarrollo de software orientado a servicios:

- G. Rodríguez, Á. Soria and M. Campo, “AI-based Web Service Composition: A Review” IETE Technical Review, Taylor & Francis, 2015, ISSN: 0256-4602. In Press.
- Rodríguez, G., Soria, Á., Teyseyre, A., Berdun, L., & Campo, M. (2016, September). Unsupervised learning for detecting refactoring opportunities in service-oriented applications. In International Conference on Database and Expert Systems Applications (pp. 335-342). Springer, Cham.
- Rodríguez, G., Soria, Á., & Campo, M. (2016). Artificial intelligence in service-oriented software design. *Engineering Applications of Artificial Intelligence*, 53, 86-104.
- Rodríguez, G., Díaz-Pace, J. A., & Soria, Á. (2018). A Case-based Reasoning Approach to Reuse Quality-driven Designs in Service-Oriented Architectures. *Information Systems*.

El segundo objetivo específico es definir técnicas de machine learning y procesamiento de lenguaje natural para extraer información de aspectos funcionales y no-funcionales a partir de especificaciones de microservicios. Para ello, se creará un meta-modelo que sea capaz de estandarizar las especificaciones de microservicios y tecnología de contenedores con el fin de facilitar la extracción y sistematización de conocimiento. Finalmente, a dicho meta-modelo se le incorporarán aspectos no funcionales, como atributos de calidad y restricciones de despliegue, que puedan proveer soporte para un despliegue

(semi)-automático de los microservicios. En relación a este objetivo, hemos hallado evidencias en el rol que juegan las técnicas de machine learning para análisis de datos, texto y captura sistemática de conocimiento:

- Scott, E., Rodríguez, G., Soria, Á., & Campo, M. (2014). Are learning styles useful indicators to discover how students use Scrum for the first time?. *Computers in Human Behavior*, 36, 56-64.
- Scott, E., Rodríguez, G., Soria, Á., & Campo, M. (2016). Towards better Scrum learning using learning styles. *Journal of Systems and Software*, 111, 242-253.
- Rodríguez, G., Armentano, M., Soria, Á., & Corengia, E. (2020). Evaluation of Markov Models for Architecture Conformance Checking. *IEEE Latin America Transactions*, 18(01), 43-50.
- Vallejos, S., Araujo, L. D. R., Rodríguez, G., Berdun, L., & Toscani, R. (2021). Planificación de IA basada en preferencias para la composición de servicios web. *IEEE Latin America Transactions*, 100(1e).
- Araujo, L. D. R., Rodríguez, G., Vidal, S., Marcos, C. & dos Santos Pereira, R. (2021). An Empirical Analysis on OpenAPI Topic Exploration and Discovery to Support the Developer Community. *Science of Computer Programming*. Elsevier (En evaluación).

Los antecedentes del grupo están vinculados con las investigaciones personales que se han desarrollado vinculadas a la temática propuesta.

El Dr. Guillermo Rodríguez es Investigador Adjunto de CONICET y ha participado en numerosos artículos relacionados a las técnicas de Machine Learning, Educación en Ingeniería de Software, Métodos Ágiles y Arquitecturas Orientadas a Servicios. Es docente de la UNICEN y además se desempeña como Prof. Asociado en UADE y Prof. Titular en CAECE.

El Dr. Fabio Rocha es Profesor Adjunto de la Universidad Tirantes, Scrum Master certificado e investigador con vasta

experiencia en Ingeniería de Software y Microservicios. Es Master en Ciencias de la Computación por la Universidade Federal de Sergipe, y Doctor en Educación por la Universidade Tiradentes. Es coordinador del Laboratorio de Computación Avanzada e Inteligencia Artificial – ITP.

El Dr. Dos Santos es Profesor Asistente en el Departamento de Informática Aplicada de UNIRIO y Jefe del Laboratorio de Ingeniería de Sistemas Complejos (Lab ESC), liderando un equipo de 20 estudiantes. Sus intereses son Ingeniería de Sistemas Complejos (especialmente ecosistemas de software y sistemas de sistemas) y Educación en Ingeniería de Software.

La Dra. Marcos es Doctora en Ciencias de la Computación por la UNCPBA y es Profesora Titular de la UNCPBA con amplia experiencia en Ingeniería de Software y Métodos Ágiles para el Desarrollo de Software.

El Dr. Vidal es docente e investigador de la UNCPBA especializado en evolución y mantenimiento del software. Es Master en Ingeniería de Sistemas y Doctor en Ciencias de la Computación por la UNCPBA.

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo se muestra en la siguiente tabla:

Apellido y nombre	Título	Cargo	Funciones
Rodriguez, Guillermo	Dr.	Prof. UNCPBA	Director
Vidal, Santiago	Dr.	Prof. UNCPBA	Integrante
Teyseyre, Alfredo	Dr.	Prof. UNCPBA	Integrante
Da Rocha Araujo, Leonardo	Lic.	Becario CONICET	Integrante
Catarina, Osmar Felix	Lic.	Becario CONICET	Colaborador
Marcos, Claudia	Dra.	Prof. UNCPBA	Colaborador
Rocha, Fabio	Dr.	Prof. UNIT	Colaborador
Pereira dos	Dr.	Prof.	Colaborador

Santos, Rodrigo		UNIRIO	
Mateos, Crisitan	Dr.	Prof. UNCPBA	Colaborador

El proyecto contempla la participación docente de Facultades de Ciencias Exactas de la UNCPBA, la Universidad Tiradentes, y la UNIRIO; y alumnos de las tres. Por otro lado, se encuentran en desarrollo las siguientes tesis de posgrado:

- Tesis de Doctorado en Ciencias de la Computación. El Lic. Da Rocha Araujo es becario de CONICET y doctorando del doctorado de Ciencias de la Computación de la Facultad de Ciencias Exactas-UNCPBA. Dirección: Guillermo Rodríguez y Rodrigo Pereira dos Santos.
- Tesis de Doctorado en Ciencias de la Computación. El Lic. Catarina es becario de CONICET y futuro doctorando del doctorado de Ciencias de la Computación de la Facultad de Ciencias Exactas-UNCPBA. Dirección: Guillermo Rodríguez y Santiago Vidal.

En relación al presente proyecto se han finalizado las siguientes tesis de grado en Ciencias de la Computación, y Sistemas de Información:

- Um Survey sobre experiências industriais na utilização de Microserviços com Contêineres: Bárbara, Dawitt y Azevedo, Igor. Director: Dr. Fabio Rocha. Finalizado 2020.
- Microservice Architecture: Migrating A Point-of-Sale Application: Dos Santos Moura, Alex. Director: Dr. Fabio Rocha. Finalizado 2020.
- A Novel Unsupervised Learning Approach for Assessing Web Services Refactoring: Hamme, Brian y Listorti, Luciano. Director: Dr. Guillermo Rodríguez. Co-director: Dr. Cristian Mateos. Finalizado 2019.

5. BIBLIOGRAFÍA

- Ahmadvand, M., & Ibrahim, A. (2016). Requirements reconciliation for scalable and

secure microservice (de) composition. IEEE International Requirements Engineering Conference Workshops.

- Di Francesco, P., Malavolta, I., & Lago, P. (2018). Migrating towards Microservice Architectures: an Industrial Survey. IEEE International Conference on Software Architecture, ICSA 2018, Seattle, USA.

- Fowler, M. & Lewis, J. (2018). Microservices a definition of this new architectural term. URL: <http://martinfowler.com/articles/microservices.html>

- Newman, S. (2015). Building microservices: designing fine-grained systems. O'Reilly Media, Inc.

-Rahman, M. and Gao, L. (2015). A reusable automated acceptance testing architecture for microservices in behavior-driven development. IEEE Symposium on Service-Oriented System Engineering (SOSE).

- Rodríguez, G., Díaz-Pace, J. A., & Soria, Á. (2018). A Case-based Reasoning Approach to Reuse Quality-driven Designs in Service-Oriented Architectures. Information Systems.

HACIA LA DEFINICIÓN TEMPRANA DE REQUISITOS DE CALIDAD DE DATOS PARA PRODUCTOS DE INFORMACIÓN EN CONTEXTO ÁGIL

Carrizo Claudio*, Javier Saldarini*, Silvana Armando*, Julio Trasmontana*, Angélica Caro#, Carlos Salgado+, Albero Sanchez+, Mario Peralta+

*Facultad Regional San Francisco – Universidad Tecnológica Nacional
Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
{cjcarrizo77, saldarinijavier, silvana.armando, juliotrasmontana}@gmail.com

#Departamento de Ciencias de la Computación y Tecnologías de la Información - Facultad de Ciencias Empresariales, Universidad del Bío-Bío
Casilla 447, 3780000, Chillán, Chile
{mcaro}@ubiobio.cl

+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

Los Sistemas de Información (SI) permiten procesar datos en bruto para obtener como resultado Productos de Información (PIs), por lo cual es fundamental que dichos datos tengan un alto nivel de calidad, para finalmente poder obtener como resultado productos de información de calidad, que permitan apoyar los procesos de negocio y la toma de decisiones. En relación a los procesos de desarrollo de SI, resulta importante considerar la incorporación de aspectos asociados con la calidad de los datos en fases tempranas, específicamente en fase de requisitos. En procesos de desarrollo donde se utiliza la metodología ágil SCRUM, comúnmente no se definen requisitos que estén relacionados con la calidad de los datos, y enfocados hacia los productos de información. El propósito de esta línea de investigación es proponer un método para la definición y especificación temprana de Requisitos de Calidad de Datos, basado en el concepto de PIs y en la Familia de estándares ISO/IEC 25000. El uso de dicho método está orientado en el contexto de la

Metodología Ágil SCRUM y pretende ser un aporte en pos de garantizar la calidad de datos de los productos de información, cuyo valor es directamente transferible a los usuarios de los SI.

Palabras clave: Requisitos de Calidad de Datos – Productos de Información - Procesos de Desarrollo Ágil – SCRUM

CONTEXTO

El presente trabajo se enmarca dentro de los siguientes proyectos 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 N° 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.

- Gestión de Proyectos de Software: Los Modelos de Calidad como Soporte a los Procesos y Productos Software. Este proyecto fue financiado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, y homologado bajo el código SIPPBSF0008185.

1. INTRODUCCIÓN

En la actualidad las organizaciones gestionan sus procesos de negocio a través de una variedad de SI, los cuales no sólo permiten llevar adelante el desarrollo, seguimiento y control de sus procesos y actividades, sino que también permiten brindar información de valor para tomar decisiones oportunas. El propósito principal de todo SI es brindar información a los usuarios dentro de una organización, partiendo del procesamiento de datos de entrada en bruto, para obtener finalmente como salida “Productos de Información”. Según (R. Wang, 1998), un Producto de Información puede definirse como el resultado del procesamiento de un conjunto de datos, el cual tiene un valor que es directamente transferible a los usuarios. Por ejemplo, una factura es un PI y corresponde al resultado que se obtiene del procesamiento de las compras mensuales realizadas por un cliente (A. Caro et. al, 2013).

Dada la importancia de poder obtener productos de información de los SI que tengan el nivel de calidad deseado (Y. Álvarez, 2015), que no sólo apoyen los procesos de negocio de la organización, sino también la toma de decisiones (Z. Houhamdi, B. Athamena, 2019), es necesario garantizar la calidad de los datos, los cuales al ser procesados se transformarán en información de calidad. Según ISO/IEC 25012:2008, la calidad de datos puede definirse como el “Grado en

que las características de los datos satisfacen necesidades implícitas y establecidas cuando son usados en condiciones específicas”.

Si bien existen autores que abordan la calidad de los datos desde diferentes perspectivas (Wang & Strong, 1996; Pipino et. al, 2002), existe la Familia de Normas ISO/IEC 25000, también conocida como SQUARE (Systems and software engineering-Systems and software Quality Requirements and Evaluation), la cual propone un marco para la definición de requerimientos de la calidad de software/datos y evaluación de la calidad del software/datos, apoyados por un proceso de medición de la calidad de software/datos. El uso de esta serie de normas está enfocado hacia adquirentes, desarrolladores, entre otros. En lo que respecta a la definición de requisitos de calidad, SQUARE provee la Norma ISO/IEC 25030:2019, la cual ofrece un marco que permite guiar la definición de Requisitos de Calidad de Datos, por medio de un proceso de definición y análisis, en el que se utilizan respectivamente la Norma ISO/IEC 25012:2008, que permite definir un modelo general con características de calidad de datos, y la Norma ISO/IEC 25024:2015, que ofrece las medidas o métricas para cada una de las características del modelo de calidad de datos.

Según lo antes mencionado, para poder obtener productos de información de calidad, resulta importante considerar la incorporación de aspectos relacionados con la calidad de los datos en fases tempranas del proceso de desarrollo de los SI, más específicamente en fase de desarrollo de requisitos. Si bien existen trabajos que consideran la inclusión de aspectos de calidad de datos en etapas tempranas del proceso de desarrollo de software (A. Caro et. al 2013; A.

Rodríguez y A. Caro, 2013; A. Rodríguez, 2012; C. Guerra-García, 2012; Y. Alonso, 2020), se han encontrado pocas propuestas que aborden la definición de requisitos de calidad de datos en fase de requisitos en un proceso de desarrollo ágil.

Lo descrito anteriormente, nos motivó a iniciar una línea de investigación que permita realizar un aporte desde la calidad de los datos, hacia los productos de información resultantes de los SI, los cuales se desarrollen en el contexto de metodologías ágiles. Aunque actualmente existen varias de ellas en el mercado (Scrum, Extreme Programming, Kanban), se optó por la metodología ágil Scrum, ya que es la más reconocida y utilizada a nivel mundial, y la misma consiste en un conjunto de prácticas y roles que permiten el trabajo de entregas incrementales de un producto. Se basa en una estructura de desarrollo incremental, esto es, cualquier ciclo de desarrollo del producto y/o servicio se desgrana en pequeños proyectos divididos en distintas etapas: análisis, desarrollo y testing. En la etapa de desarrollo se llevan adelante interacciones del proceso, es decir, entregas regulares y parciales del producto final.

De acuerdo a lo expuesto anteriormente, uno de los objetivos principales de la línea de investigación consistió en desarrollar un método basado en SQUARE y aplicado en el contexto del desarrollo de software utilizando la metodología ágil SCRUM, el cual le pueda servir de soporte al rol del Dueño del Producto (DP), para guiar la definición y especificación temprana de Requisitos de Calidad de Datos para Productos de Información.

Para la definición de Requisitos de Calidad de Datos se toma como punto de partida el concepto de Productos de Información, que están asociados a

Requisitos Funcionales especificados a través de Historias de Usuario. Para obtener una especificación de requisitos de calidad de datos para los productos de información, se propone un formato de documento portátil (pdf), el cual fue elegido por su alta compatibilidad.

El método desarrollado está compuesto por 4 actividades principales:

Actividad 1. Definición de Perfiles de Usuarios: El DP debe definir los distintos perfiles de usuarios que participarán en las historias de usuario definidas.

Actividad 2. Definición de Historias de Usuarios: El DP debe definir las historias de usuarios que se consideran importantes o relevantes y que van a estar relacionadas a los productos de información. Cada historia de usuario está asociada con un perfil de usuario (definido en la actividad anterior) y con un rol con respecto a la calidad de datos (consumidor de datos / productor de datos).

Actividad 3. Definición de Productos de Información: el DP debe definir cada uno de los productos de información, los cuales deben tener relación con las historias de usuario definidas. Por cada producto de información se define el nombre, una descripción, las historias de usuarios relacionadas y los atributos que son propios del producto de información.

Actividad 4. Definición de Requisitos de Calidad de Datos para Productos de Información: el DP deberá identificar y definir los requisitos de calidad de datos que son pertinentes para cada producto de información o atributos del producto de información. Los requisitos de calidad de datos se obtienen a partir del uso de las Normas ISO/IEC 25030, ISO/IEC 25012 e ISO/IEC 25024.

Otro de los objetivos propuestos y que está contemplado dentro de la línea de investigación, es el desarrollo de una herramienta, que será implementada en un

entorno web, y que permitirá facilitar el uso y aplicación del método, ofreciendo al DP un asistente que permita llevar adelante cada una de las actividades que propone el método, hasta lograr la definición y especificación de requisitos de calidad de datos para los productos de información. Cabe acotar, que por cada producto de información definido, la herramienta generará una especificación de requisitos de calidad de datos en un documento en formato pdf, el cual luego podrá ser adjuntado a cada una de las historias de usuario definidas.

La principal contribución de esta línea de investigación es ofrecer desde el ámbito de la calidad de datos, un método sistematizado y automatizado, que está basado en SQUARE, y que permite definir y especificar requisitos de calidad de datos, en pos de garantizar la calidad de los productos de información en contexto de procesos de desarrollos ágiles.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de trabajo de esta línea de investigación se detallan a continuación:

- Revisión Sistemática de la Literatura asociada a la temática calidad de datos en general y en contexto de procesos de desarrollo de software en particular.
- Estudio en profundidad de las Normas ISO/IEC 25030, ISO/IEC 25012 e ISO/IEC 25024.
- Estudio en profundidad de las metodologías ágiles existentes en el mercado.

- Elaboración de un método que permita definir en forma temprana requisitos de calidad de datos para productos de información.
- Desarrollo de una herramienta que permita automatizar el método para facilitar su uso y aplicación.

3. RESULTADOS OBTENIDOS/ESPERADOS

Resultados Obtenidos

- Informe con resultado de la revisión sistemática de la literatura sobre la temática calidad de datos y asociada a procesos de desarrollo ágiles.
- Informe con resultado del proceso de revisión y estudio de las normas ISO/IEC 25030, ISO/IEC 25012 e ISO/IEC 25024.
- Informe con resultado de la revisión y estudio de las metodologías ágiles utilizadas actualmente en el mercado.
- Método para definir requisitos de calidad de datos para productos de información en el contexto de desarrollo ágil.

Resultados Esperados

- Desarrollo de una herramienta que permita automatizar el método propuesto.

4. FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación se trabaja desde el Grupo de Investigación "Calidad de Software y Datos", que pertenece a la UTN Facultad Regional San Francisco, Córdoba.

El equipo de trabajo está constituido por

8 docentes investigadores (un Director de Proyecto, un Co-Director y seis docentes investigadores), todos de la especialidad Sistemas de Información. A su vez, se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis, de los Ing. Claudio Carrizo y Javier Saldarini.

5. BIBLIOGRAFÍA

- A. Caro, A. Fuentes, A. Soto. “Desarrollando sistemas de información centrados en la calidad de datos”, Revista Chilena de Ingeniería, vol.21, pp. 54-69, 2013.
- A. Rodríguez y A. Caro, “Especificación de Requisitos de Calidad de Datos en Modelos de Procesos de Negocio: BPiCD una propuesta”, Jornadas Chilenas de Computación, II Workshop on Business Process Management (WBPM). Temuco, Chile. 2013.
- A. Rodríguez, A. Caro, C. Cappiello, I. Caballero, “A BPMN Extension for Including Data Quality Requirements in Business Process Modeling”. International Workshop on Business Process Modeling Notation. Lecture Notes in Business Information Processing, vol.125, pp. 116-125. 2012.
- C. Guerra-García, I. Caballero, R. Testillano, “Elicitation of Specific Requirements of Data Quality during the Web Portal Development”, In Proceedings of the 10th International Workshop on Modelling, Simulation, Verification and Validation of Enterprise Information Systems and 1st International Workshop on Web Intelligence (WEBI-2012), pages 81-93, 2012.
- Extreme Programming. Disponible en <https://www.agilealliance.org/glossary/xp>
- ISO/IEC 25000. Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models.
- ISO/IEC 25030:2019. Systems and software engineering — Systems and software quality requirements and evaluation (SQuaRE). Quality requirements framework.
- ISO/IEC 25012:2008. Software engineering -- Software product Quality Requirements and Evaluation (SQuaRE). Data quality model.
- ISO/IEC 25024:2015. Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE). Measurement of data quality.
- ISO/IEC 25012:2008. Software engineering -- Software product Quality Requirements and Evaluation (SQuaRE). Data quality model.
- Kanbanize. Qué es Kanban: Definición, Características y Ventajas. Disponible en <https://kanbanize.com/es/recursos-de-kanban/primeros-pasos/que-es-kanban>.
- La Metodología ágil más usada: SCRUM. Disponible en <https://www.nextu.com/blog/que-es-scrum/>.
- Leo L. Pipino, Y.W.L. and Richard Y. Wang, “Data Quality Assessment, in COMMUNICATIONS OF THE ACM”, 2002, p. 211 – 218.
- R. Wang, “A Product Perspective on Total data Quality Management”. Communications of the ACM. Vol. 41, Issue 2, pp. 58-65. 1998.
- ScrumAlliance. Disponible en <https://www.scrumalliance.org/>
- Y. Fernanda Torres Alonso, “Especificación de requerimientos de software con un enfoque de calidad de datos”, Universidad Autónoma de San Luis, Potosí, 2020.
- Y. Álvarez, “Análisis de la calidad de datos en fuentes de la suite ABCD”, Trabajo de Diplomatura. Universidad Central “Marta Abreu” de las Villas, Facultad de Matemática, Física y Computación, Carrera de Licenciatura en Ciencias de la Computación, Santa Clara, Cuba, 2015.

Wang & Strong, D.M. (1996), "Beyond accuracy: What data quality means to data consumers". *Journal of management information systems*, 5-33.

Z. Houhamdi, B. Athamena, "Impacts of information quality on decision-making". *Global Business and Economics Review*, Vol. 21, No. 1, 2019.

Hacia un modelo integral de Accesibilidad en Sistemas e-Learning

Iván Balmaceda Castro¹, Carlos Salgado², Mario Peralta², Alberto Sanchez²,
Mariela Fernandez¹, Carlos E. Vera¹

¹SedeRegional Chemical – Departamento de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de La Rioja, La Rioja, Argentina.
{ibalmaceda, mfernandez}@unlar.edu.ar – carloseduardonvn@gmail.com

²Departamento de Informática. Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis, San Luis, Argentina.
{csalgado, mperalta, alfanego}@unsl.edu.ar

RESUMEN

Con el transcurso del tiempo, las actividades humanas, se han ido modificando, desde el ocio hasta el entablar una conversación, se vieron afectadas por la implementación de tecnologías. La educación, no está exceptuada, ya que cada día toman más impulso los sistemas de aprendizaje en línea mediados por tecnología, apuntando a una educación virtual, que se ha consolidado como una gran alternativa en los procesos de enseñanza-aprendizaje. En este sentido, esta línea de investigación plantea la importancia de incorporar condiciones de accesibilidad en los entornos virtuales de aprendizaje, a través de la definición de métricas a partir de la identificación de algunos criterios que permiten cuantificar el grado de satisfacción de los usuarios.

Palabras clave: Diseño Universal, Accesibilidad, e-Learning, Modelos de Calidad.

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 N° 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

El contexto sanitario del año 2020, nos llevó a repensarnos en nuevas formas de educación. El sistema educativo argentino, tuvo que adaptarse a una modalidad de enseñanza y aprendizaje, donde se explotaron medios y dispositivos electrónicos para facilitar el acceso, la evolución y la mejora de la calidad de la educación y la formación, conocido como *e-Learning* [1]. En ello, se utilizan entornos virtuales de aprendizaje que, mediante la utilización de herramientas e instrumentos, los sujetos implicados, logran generar una interacción en el proceso de enseñanza-aprendizaje.

En un mundo con constantes cambios, la discapacidad, ha ido ganando espacios a través de diferentes disciplinas [2]. Mas allá de estos avances, en el campo de las ciencias informáticas, sigue existiendo una carencia en crear sistemas universales y accesibles [3]. Generalmente, se le da un alto grado de prioridad a la parte estética y se deja de lado los conceptos concernientes a la accesibilidad y usabilidad; esta última, se ha visto

reemplazada por herramientas que muchas veces pueden ser innecesarias [4].

Los problemas de accesibilidad y utilidad de los sistemas son cada vez más críticos, como el reconocimiento del número de personas con diversas necesidades. La tecnología se ha diversificado, donde, cada vez, es más necesario utilizarla para participar plenamente en la cotidianidad.

Todo proyecto de ingeniería de software está ligado a la obtención de un producto, proceso o servicio que es necesario generar a través de diversas actividades, que contribuyen a obtener una parte del producto necesario para continuar hacia el producto final, y facilitar la gestión del proyecto.

Hoy en día, pensar en el e-Learning de calidad, hace referencia a su contenido y a su tecnología, y que cumpla con requerimientos mínimos para su correcto funcionamiento, pero ¿Qué pasa si esos requerimientos no son pensados para una persona con discapacidad, visual por ejemplo? ¿Podrá utilizar la misma adecuadamente? Por tal razón, se considera necesario pensarnos en sistemas e-learning accesibles para, de esta manera, asegurar un contenido de calidad.

Comprender los problemas de accesibilidad, influye en la forma en la que las personas utilizan un producto o servicio. Generalmente, una persona con discapacidad posee necesidades específicas con un apoyo diferente, esto quiere decir que no todas, necesitan de lo mismo. Lo que le sirve a una persona puede que no les sirva a todas las personas con esa misma discapacidad o a personas con discapacidades distintas.

Las necesidades de accesibilidad de los usuarios varían a lo largo del tiempo y en los distintos contextos de uso. Estas, están referidas en relación con las características o atributos necesarios para que un sistema sea accesible. A su vez, estos se transforman en requisitos de usuario teniendo en cuenta el

contexto de uso, prioridades de los usuarios, compensaciones con otros requisitos y limitaciones del sistema [5].

La investigación en HCI (del inglés Human Computer Interaction) se centra en las interacciones entre las personas y los sistemas computacionales, que tienen lugar dentro del ciclo de vida del usuario, aplicando las teorías cognitivas y modelos de la psicología, la sociología, ergonomía y la antropología en los que se basa el diseño, evaluación e implementación de interfaces interactivas [6].

Entonces, es de gran importancia que, para lograr una mayor experiencia de usuario, se debe tener en cuenta la accesibilidad de los contenidos educativos. Lo que conlleva a una adaptación de los componentes y los contenidos de manera que puedan ser utilizados por personas con limitaciones. En ello, [7] plantea analizar la especificación ISO 9001:2015 [8] y aplicarla a un modelo para generar un espacio de mejora continua, con el fin de poder realizar, en la posterioridad, una certificación. Por su parte, en [9] se propone un test heurístico en el que se realiza una evaluación heurística de accesibilidad, las cuales están centradas en la inspección del cumplimiento de un conjunto de heurísticas de usabilidad por parte de un grupo de expertos sobre un producto software [10], obteniendo un conjunto de recomendaciones para mejorar el desarrollo, el desempeño y la accesibilidad del sitio web.

Por ello, consideramos necesario identificar las características del contexto de enseñanza, las emociones y los sentimientos de los estudiantes, quienes se desenvuelven como los usuarios y beneficiarios de los recursos tecnológicos, y que, a través de los sentimientos, que pueden resultar positivos o negativos son determinantes en la generación de entornos más accesibles y usables, que pueden ser descritos como los elementos del contexto de uso: usuarios, tareas, equipos, y los entornos físicos y sociales en los que se usa

un producto o está destinado a ser utilizado. [11][12]

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El Diseño Universal se constituye como un paradigma y una concepción filosófica “orientada a alcanzar la accesibilidad en los entornos, espacios, servicios, bienes, objetos y dispositivos, de modo que sean utilizables por la mayor cantidad de personas sin necesidad que se adapten o especialicen para sectores poblacionales determinados” [13]. Este, es considerado como la base de todo diseño y no se antepone a un diseño particularizado para un usuario concreto.

En base a los trabajos de investigación presentadas en WICC 2019 [14] y WICC 2020 [3], se continúa trabajando en esta línea de investigación en la definición de modelos y métodos de evaluación de calidad, concernientes a la Usabilidad, la Experiencia de Usuario y Accesibilidad en e-Learning.

En ese sentido, se estima fundamental trabajar en la mejora de las condiciones de accesibilidad de los contenidos educativos digitales que se muestran a través de las plataformas de aprendizaje. Teniendo en cuenta las mismas pautas que establece el consorcio del W3C, ya que los materiales educativos en línea operan de la misma forma que las páginas web [15].

3. RESULTADOS ESPERADOS/OBTENIDOS

El objetivo general de la presente línea de investigación consiste en un modelo que incorpore para cada etapa aspectos de accesibilidad que se enfoquen o permitan mejorar la calidad en los entornos virtuales de aprendizaje. En ello, se plantean los siguientes

objetivos particulares, que se han ido logrando y otros en los que se está trabajando:

- Definir el estado del arte en la temática.
- Identificar distintos modelos para evaluar la accesibilidad en los Entornos Virtuales de Enseñanza y Aprendizaje.
- Reconocer fortalezas y debilidades de los modelos analizados.
- Proponer un modelo que incorpore accesibilidad, y un conjunto de heurísticas y/o checklist que permitan validar la accesibilidad.
- Validación del modelo propuesto aplicado a casos de estudio de sistemas e-Learning como, por ejemplo, el Campus Virtual de la UNLAR y de la UNSL.

Para la definición del modelo, se está utilizando el Modelo de Proceso de referencia de ISO/IEC 40180 [16] Como también diversos estándares y buenas prácticas de accesibilidad para personas con discapacidad.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación se da en un trabajo en conjunto entre la Universidad Nacional de La Rioja (UNLaR) y la Universidad Nacional de San Luis (UNSL).

En dicha línea, se está trabajando en el Proyecto de Tesis de la Maestría en Ingeniería de Software (Plan Ord. 05/2010-CD) del Lic. Iván Balmaceda Castro, Docente del Departamento Académico de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de La Rioja.

Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Computación. Y trabajos finales de carrera de la Ingeniería en Informática e Ingeniería en Computación de la Facultad de Ciencias Físico-Matemáticas y Naturales. Del mismo modo, Trabajos Finales de la Tecnicatura en Informática y Licenciatura en Sistemas de Información de la Universidad Nacional de

San Luis y del Departamento Académico de Ciencias Exactas, Físicas y Naturales Universidad Nacional de La Rioja.

5. BIBLIOGRAFÍA

1. Adam, M. R., Vallés, R. S., & Rodríguez, G. I. M. (2013). E-learning: características y evaluación. *Ensayos de economía*, 23(43), 143-159.
2. Balmaceda, C. I., Salgado, C., Peralta, M., Sánchez, A., Saldarini, J., Carrizo, C., ... & Ferreyra, G. UNA PROPUESTA DE EVALUACIÓN DE LA EXPERIENCIA DE USUARIO CON DISCAPACIDAD EN SISTEMAS E-LEARNING. *Jornadas de Ciencia y Tecnología 2020 "50 aniversario"*, 151.
3. Balmaceda Castro, I., Salgado, C. H., Peralta, M., & Sanchez, A. (2020). Heurística de Evaluación de la Experiencia de Usuario en Sistemas e-Learning. In *XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz)*.
4. Lasso Guerrero, J. (2013) Ergonomía en el diseño web: Usabilidad de sitios web dedicados al comercio electrónico en Buenos Aires. pp.5,47,9 Recuperado de https://www.palermo.edu/dyc/maestria_diseno/pdf/tesis.completas/101-Lazzo-juan.pdf
5. ISO/IEC Guide 71/2014. Guide for addressing accessibility in standards, 2. Ginebra, Suiza. 2014.
6. Litwin, E. (2003). Los desafíos y los sinsentidos de las nuevas tecnologías en la educación. entrevista portal EDUC. AR.
7. Russo, C., Sarobe, M., Saenz, M., Alonso, N., Pérez, D., Tessore, J. P., ... & Ramón, H. D. (2014). Calidad, usabilidad y accesibilidad en los entornos virtuales de enseñanza-aprendizaje. In *Workshop de Investigadores en Ciencias de la Computación (mayo 2014, Ushuaia)*
8. ISO 9001:2015 Quality management systems – Requirements.
9. Oliveros, D. P., Chanchi, G. E., & Vidal, M. I. (2019). Propuesta de un test heurístico de accesibilidad para sitios web basados en la norma NTC 5854. *Revista Ibérica de Sistemas e tecnologías de Información*, (E17), 170-182.
10. Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414).
11. Rivero, I., Gómez-Zermeño, M.G. y Abrego Tijerina, R. (2013). Tecnologías educativas y estrategias didácticas: criterios de selección. *Revista de Educación y Tecnología*, 3, 190-206.
12. Fuentes Morán, Y. M. (2013). Usabilidad de los Sistemas de Información en Salud dentro de escenarios de atención crítica: un estudio de los sistemas de historia clínica en IPS de alta complejidad Colombianas (Doctoral dissertation, Universidad Nacional de Colombia).
13. MACE, R.; MUELLER, J.; y STORY, M. (1998). *The Universal Design File. Designing for People for All Ages and Abilities*. NC State University, The Center for Universal Design
14. Balmaceda Castro, I., Salgado, C. H., Peralta, M., Sánchez, A., Fernández, M., Magaquian, J., & Fuentes, N. "Experiencia de usuario en plataforma virtual de aprendizaje." *XXI Workshop de Investigadores en Ciencias de la Computación*, San Juan 2019.
15. Sempertegui, M. M., & Torrez, V. (2013). Accesibilidad en los entornos virtuales de las instituciones de educación superior universitarias. *Virtualidad, Educación y Ciencia*, 4(7), 8-26.
16. ISO 40180:2017 Information Technology – Quality for learning, education and training.

i-QuAGI: Aproximación a un enfoque inteligente para la evaluación de calidad de procesos ágiles de software

Noelia Pinto; César Acuña; Nicolás Tortosa; Gabriela Tomaselli
Centro de Investigación Aplicada a TIC (CInApTIC)
Universidad Tecnológica Nacional, Facultad Regional Resistencia
French 414, Resistencia, Chaco
{ns.pinto;csr.acn;nicotortosa;gabriela.tomaselli}@gmail.com

RESUMEN

Con el objetivo de facilitar la adopción de prácticas ágiles que aseguren la calidad de los procesos de desarrollo de software se ha presentado Agile Quality Framework (AQF) [1][2], un framework que integra un modelo de calidad (QuAM) junto a una herramienta de software (QuAGI) que permite la automatización de dicho modelo y que se adapta a las características de las PYMES. AQF surge, como una plataforma que contribuye con los equipos de desarrollo de software a partir de la evaluación de calidad en proyectos ágiles, considerando como objeto de la medición al proceso de desarrollo independientemente del enfoque ágil seleccionado. Luego de diversas experiencias de validación llevadas a cabo en ambientes reales de producción, los equipos de desarrollo de software participantes han manifestado adaptarse fácilmente a la herramienta de software QuAGI para llevar adelante el seguimiento y evaluación de sus procesos ágiles. Sin embargo, con estas mismas experiencias se ha observado que resulta necesario enriquecer el framework de forma tal de ofrecer una nueva herramienta que permita liberar de trabajo de monitorización manual al grupo de administradores de proyecto y, al mismo tiempo, provea soporte a la toma de decisiones de directivos de las empresas proporcionando recomendaciones automáticas.

Por todo lo expuesto, en este trabajo se presenta la línea de investigación que propone el diseño de un enfoque inteligente que permita recomendar acciones al equipo de forma tal de mejorar los niveles de calidad del proceso ajustando los factores que sean necesarios. Se pretende incorporar al framework AQF una herramienta que dé soporte al equipo de desarrollo, a partir de recomendaciones automáticas que surjan del seguimiento del proyecto ágil y sus actividades, las cuales muchas veces son afectadas por acciones en segundo plano que pasan desapercibidas e impactan negativamente en los niveles de calidad del proceso de desarrollo asociado.

Palabras clave: Calidad del Software, Procesos ágiles de Software, Human Computer Interface; Agente Inteligente

CONTEXTO

El trabajo que aquí se presenta está enmarcado en el proyecto “I-QuAGI: Un enfoque inteligente para la evaluación de calidad de procesos de software ágiles”, que es financiado por la UTN y ejecutado en el Centro de Investigación Aplicada a TIC (CInApTIC) de la Facultad Regional Resistencia, con el código SIPPBRE0008092.

Asimismo, algunas actividades son compartidas con el proyecto de investigación y desarrollo “Evaluación del

impacto de las emociones en la calidad de software desde el punto de vista del usuario” (PID 5517), también financiado por UTN y ejecutado en el CInApTIC.

1. INTRODUCCIÓN

La gestión del conocimiento, en procesos de desarrollo de software, es un campo de la ingeniería de software que se ha ido estudiando en la actualidad, con el fin de aplicarlo en la consecución de la mejora de los procesos de software y por ende, en la calidad del producto final [3]. De hecho, se observa un crecimiento importante en la necesidad de adopción de herramientas de software que, por un lado, permitan la gestión de conocimiento para la innovación y mejora de productos y procesos, y, por el otro, favorezcan la toma de decisiones ejecutivas para la renovación y adaptación de las organizaciones [4].

Además, asociado a la gestión de proyectos de software, de acuerdo al reporte VersionOne [5], los enfoques ágiles han ganado, en los últimos 10 años, gran popularidad para la gestión de proyectos de desarrollo de software, pues ofrece a los equipos el control de requerimientos variables, la gestión efectiva y eficaz de los grupos de trabajo y el involucramiento y empoderamiento del cliente dentro del proyecto.

Sin embargo, no siempre los equipos de desarrollo de software cuentan con los recursos y métodos para afrontar la implementación y adopción rápida y eficiente de un enfoque ágil, debido en primer lugar a hábitos obtenidos de sus métodos tradicionales, y en segundo lugar, por la falta de un verdadero entendimiento y conocimiento de los valores, principios, prácticas y procesos en los cuales se basan los enfoques ágiles [6][7].

En respuesta a ello, se ha desarrollado el framework AQF, con el objetivo de

ofrecer a los equipos una alternativa que permita, no solo el seguimiento de sus proyectos, sino también sea posible evaluar la calidad cuando se opta por trabajar con procesos ágiles de desarrollo de software. Dicho framework se denomina AQF (Agile Quality Framework) y su versión actual está compuesta por un modelo, QuAM (Quality Agile Model) y por una herramienta de software que brinda soporte a dicho modelo, QuAGI (Quality AGile).

A la fecha se han llevado adelante diversas experiencias de validación, que permitieron obtener resultados respecto a la implementación de QuAGI como herramienta de seguimiento de proyectos ágiles y seguir mejorando el framework AQF para lograr, de forma incremental, una mejor herramienta. Sin embargo, con estas mismas experiencias se ha observado que resulta necesario enriquecer el framework de forma tal de ofrecer una nueva herramienta que permita liberar de trabajo de monitorización manual al grupo de administradores de proyecto y, al mismo tiempo, de soporte a la toma de decisiones de directivos de las empresas proporcionando recomendaciones automáticas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Con el proyecto denominado “i-QuAGI: Un enfoque inteligente para la evaluación de calidad de procesos de software ágiles”, se propone el diseño de un enfoque inteligente que permita recomendar acciones a los equipos, de forma tal de mejorar los niveles de calidad del proceso ágil que lleven a cabo en el marco del desarrollo de software.

Se pretende, entonces, incorporar al framework AQF una herramienta que dé

soporte al equipo de desarrollo, a partir de recomendaciones automáticas que surjan del seguimiento del proyecto ágil y sus actividades, las cuales muchas veces son afectadas por acciones en segundo plano que pasan desapercibidas e impactan negativamente en los niveles de calidad del proceso de desarrollo asociado.

A nivel de diseño conceptual, el modelo que se propone está basado en la tecnología multi-agente permitiendo la reacción ante eventos generados a partir del uso de QuAGI por el equipo del proyecto de software y que puedan afectar a la calidad final del proceso ágil subyacente. El modelo incluirá la descripción del comportamiento del sistema multi-agente a través del diseño de una ontología propia, lo que permite la formalización tanto de los pasos del proceso de monitoreo del comportamiento de usuarios en QuAGI así como la información que se requiere administrar por cada proyecto.

Las actividades que se llevarán a cabo en el marco de este proyecto son:

- Ejecución de una Revisión sistemática de la literatura respecto a Agentes inteligentes utilizados en seguimiento de procesos ágiles de desarrollo de software.
- Diseño del enfoque inteligente: características y comportamientos.
- Propuesta de interacción entre el agente inteligente y QuAGI.
- Análisis y validación de resultados,

3. RESULTADOS OBTENIDOS/ ESPERADOS

Si bien existen disponibles en la literatura algunas herramientas que colaboran, a través de sistemas multi-agentes, con equipos de desarrollo de software en la toma de decisiones

relacionados a la gestión de sus proyectos, la mayoría se abocan al seguimiento de un único proyecto. Por ende, al no recopilar datos de proyectos anteriores, se desaprovecha todo el conocimiento previamente adquirido a través de diversas experiencias que atraviesan los equipos de desarrollo de software.

Normalmente, esta experiencia es sólo adquirida por el equipo de desarrollo y cuando éste ya no existe, la información que no esté asociada a la herramienta de seguimiento de proyectos y que se relacione al equipo en sí mismo, se pierde. Esto termina perjudicando a la organización [8] y derivando en lo que se conoce como Amnesia Organizacional (AO).

Con este proyecto, al incorporar una nueva herramienta al framework AQF, se busca ofrecer una alternativa que permita la gestión del conocimiento y la toma de decisiones en el equipo de modo de evitar realizar acciones que repercutan negativamente en el proyecto de desarrollo basado en enfoques ágiles.

El proyecto dará inicio con un mapeo sistemático de la literatura respecto a los desarrollos existentes, de forma tal que esta actividad contribuya al diseño conceptual del modelo del enfoque inteligente, favoreciendo los aspectos que ayuden a incrementar la calidad del proceso ágil de desarrollo asociado a cada proyecto de software.

Cabe destacar que este proyecto representa la continuidad de la línea de investigación "Framework para la evaluación de calidad de procesos ágiles" dirigido por el Dr. César Acuña, por lo que se nutrirá también de resultados correspondientes a dicho trabajo.

El desarrollo de i-QuAgi, constituirá un aporte que beneficiará directamente a las

instituciones que adopten AQF como Framework en sus procesos ágiles de desarrollo de software, ofreciendo la automatización en la identificación de eventos o acciones que perjudiquen la calidad del proceso afectando, directamente, la calidad del producto.

En particular, desde lo disciplinar, el proyecto permitirá generar nuevo conocimiento sobre métodos que faciliten el diseño de enfoques inteligentes y su correspondiente modelado en ambientes de equipos de desarrollo de software; el uso de ontologías para el modelado de las diferentes características de los diversos equipos y proyectos ágiles; así como sobre diseño de interfaces para entornos colaborativos.

4. FORMACIÓN DE RECURSOS HUMANOS

En este sentido, distinguimos dos vertientes:

- 1) *Formación de alumnos y transferencia de resultados de investigación al aula.* Tiene que ver directamente con la vinculación de las actividades y resultados de investigación y su interacción con los alumnos en el aula contribuyendo a su formación e incentivando las vocaciones científicas. En este sentido las actividades y resultados del proyecto se vinculan directamente con las asignaturas electivas del 4to Nivel de Ingeniería en Sistemas de Información (ISI): (i) Calidad del Producto y Proceso de Software y (ii) Técnicas de Desarrollo de Software Ágiles. Asimismo, se compartirán los resultados obtenidos durante el proyecto con la asignatura Ingeniería de Software, obligatoria del 4to nivel, donde los alumnos podrán aplicar de

forma empírica los conocimientos aprendidos durante el curso. Por último, el proyecto se relaciona con la asignatura Inteligencia Artificial perteneciente al 5to nivel de la carrera, puesto que tanto el diseño del enfoque inteligente como el desarrollo del sistema multi-agente requieren el conocimiento de Sistemas Inteligentes y pretenden aportar nuevos conocimientos relacionados con la materia. Cabe aclarar que todas estas asignaturas están a cargo de docentes-investigadores pertenecientes al CInApTIC, al cual pertenece el proyecto aquí descrito. La participación de alumnos de la carrera de ISI posibilitará la formación de Recursos Humanos en esta área de vacancia regional, tanto a nivel científico como profesional.

- 2) *Formación de recursos humanos en investigación.* El proyecto cuenta actualmente con una plaza para Becario de Investigación y Servicios financiada por la Secretaría de Asuntos Universitarios (UTN), una plaza para Becario Alumno de Rectorado (BAR) y una plaza de Beca de Iniciación a la Investigación y Desarrollo (BINID) para graduados financiadas por la Secretaría de Ciencia y Tecnología (UTN). Todas ellas actualmente en proceso de selección de los candidatos, con el objetivo de incorporar y formar jóvenes investigadores (estudiantes, egresados y/o egresadas de Ingeniería en Sistemas de Información) en la temática que el proyecto descripto aborda.

En cuanto a la formación de Postgrado de los docentes investigadores que participan en el proyecto, el equipo de trabajo de esta línea de investigación del CInApTIC está integrado por su Directora y Co Director, en ambos casos Doctores,

con categorías en el Programa de Incentivos y como Docentes Investigadores de UTN. Además, forman parte del equipo dos Docentes Investigadores (Profesionales de Ingeniería en Sistemas de Información), quienes actualmente son estudiantes primer Doctorado en Informática en la Región del Nordeste Argentino, a dictarse en conjunto entre la Universidad Tecnológica Nacional, Facultad Regional Resistencia; Universidad Nacional de Misiones, Facultad de Ciencias Exactas, Químicas y Naturales; y Universidad Nacional del Nordeste, Facultad de Ciencias Exactas y Naturales y Agrimensura.

REFERENCIAS

- [1] Pinto, N., Tortosa, N., Geat, B. C., Ibáñez, L., & Acuña, C. J. (2019). Validation of the reengineering applied on the first version of Agile Quality Framework. *Electronic Journal of SADIO (EJS)*, 18(1), 93-109.
- [2] Pinto, N. S., Acuña, C. J., & Rossi, G. (2020). Framework para la evaluación de calidad de procesos ágiles. *AJEA*, (5).
- [3] Montes, D. P. S., Gaviria, L. A. C., & Franco, O. H. (2018). Gestión del conocimiento en procesos de desarrollo de software: un marco de trabajo para apoyar a las MiPyMEs. *Scientia et Technica*, 23(1), 76-83.
- [4] Rodríguez García, J., Aguilar Romero, M., & Raudales García, N. (2017). Una mirada breve al software para gestión del conocimiento. *I+D Tecnológico*, 13(1), 31-39. Recuperado a partir de <https://revistas.utp.ac.pa/index.php/idtecnologico/article/view/1435>
- [5] 8th Annual State of Agile Report. VersionOne, Mayo 2020. Disponible en <https://stateofagile.com/>.
- [6] Tavares, B. G., da Silva, C. E. S., & de Souza, A. D. (2019). Risk management analysis in Scrum software projects. *International Transactions in Operational Research*, 26(5), 1884-1905.
- [7] Ulrich, D. and Yeung, A. (2019), "Agility: the new response to dynamic change", *Strategic HR Review*, Vol. 18 No. 4, pp. 161-167. <https://doi.org/10.1108/SHR-04-2019-0032>
- [8] Kransdorff, A. (1998) *Corporate Amnesia: Keeping Know-How in the Company*. Butterworth-Heinemann.

IMPLEMENTACION DE SISTEMA DE GESTION DE CALIDAD UTILIZANDO HERRAMIENTAS CASE

Jorge A. Silvera, Daniel A. Arias Figueroa, Valeria A. González, Leopoldo E. Lugones.

Centro de Investigación y Desarrollo en Informática Aplicada (C.I.D.I.A.)

Facultad de Ciencias Exactas – Universidad Nacional de Salta – Salta - Argentina

jsilvera@unsa.edu.ar, daaf@cidia.unsa.edu.ar,
vgonzalez@cidia.unsa.edu.ar, plugones@exa.unsa.edu.ar

RESUMEN

El presente trabajo propone un modelo y prototipo de software resultado de la aplicación de la metodología de desarrollo incremental, utilizando la herramienta *GeneXus*. El mismo se aplicó a la gestión de la calidad a la asignatura “*diseño basado en conocimiento*”, de la Carrera Licenciatura de en Análisis de Sistemas de la Universidad Nacional de Salta (U.N.Sa).

La experiencia demostró que la aplicación de la metodología y la herramienta elegida contribuyeron a la mejora de la calidad en la educación superior en la asignatura.

Palabras claves:

Normas de Calidad ISO 9.001, Sistema de Gestión de Calidad (SGC), GeneXus, Metodología incremental, herramientas CASE.

CONTEXTO

La presente investigación fue desarrollada en el marco del Proyecto de Investigación CIUNSa N° 2278/0, denominado “*Estudio de la familia de normas ISO 9000 y su aplicación a centros educativo*”, del Consejo de Investigación de la Universidad Nacional de Salta, en conjunto con el Centro de Investigación y Desarrollo en Informática Aplicada C.I.D.I.A., en su línea de investigación en calidad.

Como resultado de la ejecución de dicho proyecto de investigación, se realizaron diferentes tareas de aplicación, en particular

mostraremos la metodología incremental y la herramienta *GeneXus*, la cual se viene utilizando en la carrera, desde el año 2009, gracias al convenio por el cual la U.N.Sa se convirtió en socio académico por resolución rectoral RN N° 0811/09. El mismo tiene como objetivo apoyar la formación tecnológica de los estudiantes y proporcionar los mejores caminos para la inclusión de los mismos en el mercado laboral, promoviendo muchos beneficios y ventajas para toda la comunidad universitaria y para la comunidad de desarrolladores *GeneXus*.

El Modelo de Sistema de Gestión de Calidad (SGC) bajo las Normas ISO 9001, se implementó en una primera instancia a la asignatura “*Diseño Basada en Conocimiento*” optativa de cuarto año. Si esta aplicación resulta positiva, la aplicación del estándar ISO se extenderá al resto de las asignaturas que conforman la carrera de la Licenciatura en Análisis de Sistemas, y así a todos los niveles superiores dentro de la Facultad de la universidad.

1. INTRODUCCION

La formación y creación de profesionales altamente capacitados y competentes es el mayor objetivo que busca alcanzar una institución universitaria, ya que así contribuye al mejoramiento de la calidad de uno de los elementos de primera necesidad de la sociedad como es la educación. Bajo la óptica de una organización que presta servicios, la sociedad le exige a las universidades una contraprestación de excelencia, sustentada sobre la base de una educación de calidad

facilitada por docentes y adquiridas por los estudiantes.

Ante un mundo globalizado y en constante desarrollo de la información y las comunicaciones, las nuevas sociedades se exponen a nuevos y exigentes requerimientos haciendo que las necesidades educativas evolucionen y cambien constantemente. Ante esto, las universidades deben instrumentar una permanente actualización y adaptación de sus ofertas académicas para satisfacer las necesidades sociales. Es así como la educación universitaria necesita valerse de ciertos instrumentos que le permita validar que sus prestaciones de servicios responden a los máximos estándares de calidad reconocidos por la sociedad para su aceptación. Estos instrumentos son las Normas ISO 9001, que en su versión 2015 [4], establecen los requisitos que deben reunir y cumplir las organizaciones en general para la obtener el reconocimiento de la calidad en su gestión, con la asistencia de un Modelo de Sistema de Gestión de la Calidad (SGC).

La presente investigación propone entre otras actividades el análisis, diseño e implementación de un software de apoyo al SGC para una asignatura, utilizando como metodología de desarrollo de software incremental junto a la herramienta CASE *GeneXus* [6].

Se espera que el presente trabajo sirva de referencia para cualquier implementación de gestión de la calidad en el ámbito universitario.

La adopción de un SGC es una decisión estratégica para una organización que le ayuda a mejorar su desempeño global y proporcionar una base sólida para las iniciativas de desarrollo sostenible. Los beneficios potenciales para una organización que implementa un SGC basado en la Norma ISO 9001 son:

- La capacidad para proporcionar regularmente productos y servicios que satisfagan los requisitos del cliente, alumnos en nuestro caso.
- Facilitar oportunidades de aumentar la satisfacción del cliente.
- Abordar los riesgos y oportunidades asociadas con su contexto y objetivos.
- La capacidad de demostrar la conformidad con requisitos del SGC.

Los requisitos del SGC especificados en esta Norma son complementarios a los exigidos para los productos y servicios que brinda la organización. El cumplimiento permanente de los requisitos y la consideración constante de las necesidades y expectativas futuras, representa un desafío para las organizaciones en un entorno cada vez más dinámico y complejo. Para lograr estos objetivos, la organización debe considerar necesario adoptar diversas formas de mejora además de la corrección y la mejora continua, tales como el cambio abrupto, la innovación y la reorganización.

Promover la calidad en el proceso de enseñanza-aprendizaje en una asignatura implica mejorar de forma continua sus prácticas, de tal manera que permitan [1]:

- Garantizar los resultados en cuanto a rendimiento académico y deserción de alumnos.
- Asegurar la trazabilidad de los procesos que se desarrollan.
- Posibilitar la mejora continua de la satisfacción de los diferentes actores intervinientes en la Educación.

La gestión de la calidad en la educación debe ser sobre todo flexible y adaptada a las necesidades específicas de las asignaturas, sustentándose en tres elementos:

- Definición de los objetivos que se pretenden alcanzar en función de los

criterios de satisfacción de las diferentes partes interesadas.

- Proceso, que incluye la planificación y desarrollo de actividades tendientes a cumplir los objetivos.
- Los indicadores que medirán el grado en el que se consiguen los objetivos.

La norma ISO 9001 establece que como medida del desempeño del SGC, la organización debe realizar el seguimiento de la información relativa a la percepción del cliente con respecto al cumplimiento sus requisitos por parte de la organización.

Para dar cumplimiento a este requisito la asignatura ha elaborado el procedimiento para el seguimiento y medición de la satisfacción del estudiante, en el cual se define la forma como se da seguimiento a las quejas y sugerencias de los estudiantes y la forma de cómo llevar a cabo la medición de la satisfacción de los mismos.

Se puede definir una metodología para implementar el SGC en el ámbito de una asignatura, utilizando el esquema anterior de manera sistemática, dando lugar a un ciclo de mejora continua, tal y como se representa en la figura 1.

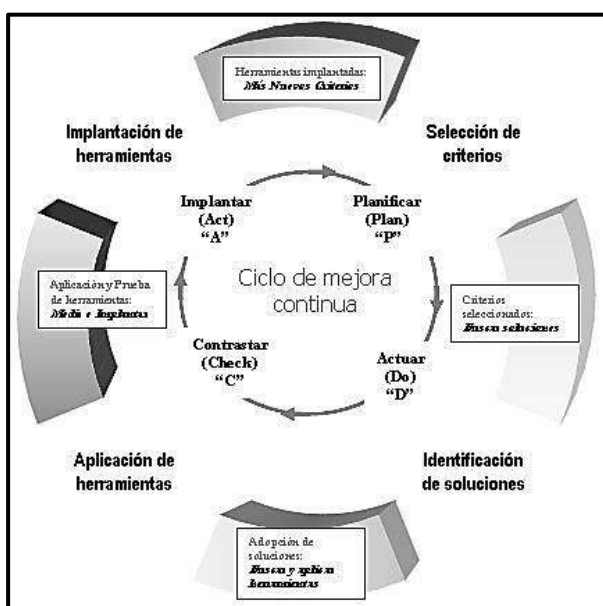


Figura 1- Ciclo PDCA de mejora continua

Edwards Deming popularizó el ciclo PDCA (Planificar, Desarrollar, Comprobar, Actuar), inicialmente desarrollado por Walter Shewhart, que es utilizado extensamente en los ámbitos de la gestión de la calidad. Esta herramienta permitió establecer en la asignatura una metodología de trabajo encaminada a la mejora continua

Un SGC se basa en la administración de muchos documentos, de manera estricta y minuciosa [2] [3]. Por esta razón es fundamental el apoyo de la tecnología informática, es decir que no alcanza con un procesador de textos y un espacio de almacenamiento compartido en un servidor de archivos.

Todo documento relacionado con el sistema de gestión de la calidad debe ser desarrollado a través de un proceso perfectamente documentado. Además, una vez aprobado, debe estar disponible para todos quienes participan en el SGC. A su vez las sucesivas revisiones deben quedar claramente identificadas así como los cambios realizados. También, si los documentos obsoletos se mantienen en el SGC para poder ser consultados, deben quedar claramente identificados como obsoletos para impedir que sean utilizados como vigentes, por error. Por ejemplo, debe impedirse su modificación incluso a personal autorizado para generar documentos.

Esto es claramente el manejo de una base documental asociada a un proceso de decisión y de elaboración conjunta de los que típicamente se realizan con herramientas de trabajo en grupos.

En función a lo expresado, tomando como referencia las especificaciones y directrices investigadas para la aplicación de la familia de norma ISO 9.001:2.008 [5], en el ámbito de la educación, se realizó el análisis y diseño de un SGC, utilizando la metodología

incremental de *GeneXus* [6], logrando generar un prototipo 100% funcional para entorno de software libre en java, con DBMS PostgreSQL, para acompañar una implementación de calidad bajo los requisitos de la norma ISO 9.001:2.015 para la asignatura “Diseño Basado en Conocimiento”. Dicho Sistema de Gestión se encuentra en permanente actualización para aproximarse en la medida de lo posible a los estándares de calidad en el cumplimiento de los objetivos de la asignatura.

Utilizando la herramienta *GeneXus*, se describe la realidad, tratando de entender el problema del usuario, trabajando en un alto nivel conceptual, en vez de realizar tareas de bajo nivel como: diseñar archivos, normalizar, diseñar programas, programar, buscar y eliminar los errores de los programas.

El enfoque de la metodología de desarrollo incremental consiste en construir una aplicación mediante aproximaciones sucesivas, en donde la construcción automática de la base de datos y programas es realizada por la herramienta CASE *GeneXus* tal como se muestra en la figura 2.

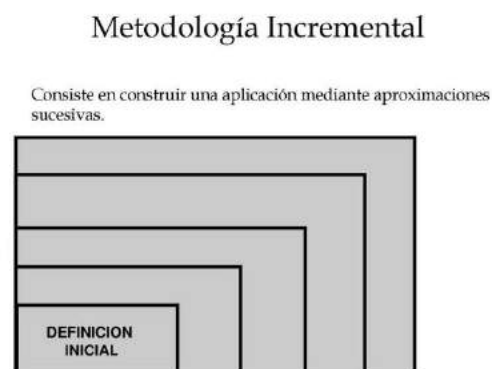


Figura 2- Metodología incremental de *GeneXus*.

Se desarrolló la aplicación con la herramienta CASE *GeneXus*, en donde el primer paso consiste en crear un nuevo proyecto o base de conocimiento, llamada KB (Knowledge Base) o base de conocimiento, en la cual se

describen las visiones de los usuarios, identificando los objetos de la realidad, para luego definirlos mediante objetos *GeneXus*. Con la definición de estos objetos, *GeneXus* puede extraer el conocimiento y diseñar la base de datos y los programas de la aplicación en forma automática, por lo cual la productividad en el desarrollo de aplicaciones aumenta y se disminuye el esfuerzo del mantenimiento de las aplicaciones (datos y programas), gracias al desarrollo incremental y la herramienta CASE *GeneXus*.

En síntesis, los resultados obtenidos a la fecha son los siguientes:

- Una Base de datos de conocimiento (KB) resultado de las interacciones obtenidas de la metodología incremental de *GeneXus*, junto a la especificación de requerimientos, aplicados a los procesos específicos relacionados a la enseñanza y aprendizaje en la asignatura Diseño Basado en Conocimiento.
- Un prototipo 100% funcional de un Sistema de Gestión de Calidad, en ambiente Web con java con DBMS PostgreSQL.
- Indicadores de calidad para la satisfacción del estudiante, y permitir realizar un seguimiento.

2. LINEAS DE INVESTIGACION y DESARROLLO

Los principales ejes temáticos que se están investigando son los siguientes:

- ✓ Tecnología Informática aplicada en Educación.
- ✓ Gestión de Calidad aplicada a la educación superior.
- ✓ Herramientas informáticas para la implementación de un SGC ISO 9.001.
- ✓ Metodología, incremental,
- ✓ Herramientas CASE, *GeneXus*.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el presente trabajo se ha abordado la gestión de los sistemas de calidad desde la perspectiva de las actividades de gestión de la enseñanza que realiza a una asignatura, de una carrera informática universitaria.

Consideramos que, además de todas las herramientas disponibles, la aplicación de la familia de normas ISO 9001:2015 junto con el apoyo de un modelo y prototipo adecuado de software y con la especificación de requerimientos adecuada, se constituye en una estrategia importante para alcanzar la satisfacción de los estudiantes, y de cualquier interesado en la educación universitaria.

Actualmente se está considerando la posibilidad de alcanzar la certificación del proceso metodológico llevado adelante por la asignatura, por algunas de las entidades certificadoras, tales como Bureau Veritas Quality International e IRAM, las cuales son las más conocidas en nuestro país.

También se puso a disposición el prototipo del SGC a otras asignaturas de la universidad que estén interesadas en la calidad y la Familia de Normas ISO 9000.

4. FORMACION DE RECURSOS HUMANOS

La estructura del equipo de investigación es de 4(cuatro) miembros incluidos el Director de la línea de investigación en calidad.

Un miembro obtuvo la Especialidad en Ingeniería de Software de la Universidad Nacional de La Plata, con el trabajo "Sistema de Gestión de Calidad bajo Normas ISO".

Otro alcanzó el título de Licenciado en Análisis de Sistemas, otorgado por de la Universidad Nacional de Salta, con la tesis

"Análisis y Diseño de un Sistema de Gestión de Calidad basado en la Norma ISO 9.001".

Por último, se continúa con la dirección de tesis de grado de alumnos de la carrera Licenciatura en Análisis de Sistemas y Tecnicatura en Programación Universitaria de la Universidad Nacional de Salta.

5. BIBLIOGRAFIA

[1]Castellón Murcia, Roberto Aristides, "Aplicación de la norma ISO 9001 en el proceso enseñanza aprendizaje en la educación superior", UFG Editores, Agosto 2.009, pp. 91-121.

[2]Arias Figueroa, Daniel y otros, "Normas ISO y su Aplicación en Centros Educativos", XIII Workshop de Investigadores en Ciencias de la Computación, 2.011.

[3]Barberis, AngelR. y otros, "Modelo de sistema de gestión de calidad para la cátedra de Programación Numérica", XXII Workshop de Investigadores en Ciencias de la Computación, 2.010.

[4]Norma ISO 9.001 (2.015) Elaborada por el Comité Técnico ISO/TC176 de ISO.

[5]Guía de Interpretación de la IRAM-ISO 9.001:2008 para la educación. IRAM 30.000:2.001.

[6]GeneXus, sitio oficial de la herramienta <http://www.genexus.com>

[7]Sitio oficial de capacitación de la GeneXus <http://www.gxtraining.com>

[8]C.I.D.I.A. <http://cidia.unsa.edu.ar>

Integración de herramientas de prueba automatizadas para evaluar la accesibilidad en aplicaciones web

Marcela Daniele, Marcelo Uva, Nicolás Streri

Departamento de Computación, Facultad de Ciencias Exactas, Físico-Químicas y Naturales, Universidad Nacional de Río Cuarto
Ruta 36 Km. 601 –CP 5800 - Río Cuarto – Córdoba - Argentina Tel. (0358) 4676235
{marcela, uva}@dc.exa.unrc.edu.ar, nicolasstreri@gmail.com

RESUMEN

Las aplicaciones web constituyen una eficaz alternativa para satisfacer los requerimientos de software actuales del mercado, por lo que son cada vez más utilizadas en múltiples y diversos dominios. La importancia creciente de garantizar la accesibilidad de las aplicaciones web, y en particular brindar condiciones adecuadas y seguras de acceso a todas las personas, aún con capacidades diferentes, llevaron a que en la actualidad existan numerosas herramientas automáticas de prueba de accesibilidad web, que permiten evaluar el cumplimiento de las pautas definidas y legislaciones vigentes. En este estudio, se ahonda en la investigación de este tipo de herramientas y los resultados que brindan. Se realiza un análisis exploratorio de herramientas, open source y de licencia de software libre, utilizadas para evaluar la accesibilidad en aplicaciones web. Se procede a analizar su funcionamiento y a recolectar las técnicas aplicadas para medir la accesibilidad en aplicaciones web.

Con esta propuesta, se pretende definir una arquitectura modular con una interfaz de comunicación común que permita la integración, interacción y evaluación de aplicaciones web. Se plantea el diseño y construcción de una plataforma que

integre herramientas automáticas de prueba utilizadas en la evaluación de aplicaciones web, y en particular, la verificación del cumplimiento de los criterios establecidos para garantizar la accesibilidad a sitios web de todas las personas,

Palabras clave: prueba, herramientas automáticas, aplicaciones web, accesibilidad, integración

CONTEXTO

La línea de investigación presentada en este trabajo se desarrolla en el marco del programa de investigación titulado “Estudios Interdisciplinarios en Evaluación de Procesos de Software y sus aportes al Desarrollo del Pensamiento Computacional en Prácticas Educativas de Ciencias Naturales”, que incluye los proyectos de investigación “Título del Proyecto: Ingeniería de Software: Evaluación de la calidad de sistemas de software y la mejora continua de los procesos de desarrollo” y “El pensamiento computacional y las prácticas docentes en ciencias”. Corresponden a la convocatoria Proyectos y Programas de Investigación (PPI 2020-2023), aprobados y financiados por la Secretaría de Ciencia y Técnica de la Universidad Nacional de Río Cuarto. En

particular, la investigación presentada en este trabajo se desarrolla en una beca de investigación CIN y en un trabajo de tesis de grado de la Licenciatura en Ciencias de la Computación.

1. INTRODUCCIÓN

La vertiginosa expansión y crecimiento de la industria del software, demanda la revisión constante de metodologías de desarrollo y de trabajo, y la adopción de herramientas para automatizar y mejorar las actividades de desarrollo y gestión, como así también la automatización de las pruebas que mejoran la calidad en la producción del software. Con el objetivo de mejorar la confiabilidad y calidad del software, esta industria refuerza la adopción de procesos sistemáticos de desarrollo de software basados en estándares de calidad.

Es loable destacar que el crecimiento de la industria del software viene acompañado de numerosas y sólidas investigaciones y estudios en la disciplina; no obstante, existen aún muchos casos en los que no se ha podido evitar que se produzcan errores en el software y lleguen a etapas de producción. Estos estudios han permitido conocer las principales causas que afectan el desarrollo de software *libre de fallas*, tales como, estudios y análisis del problema deficientes con consecuencias en la definición de requerimientos, la complejidad intrínseca del software, la constante adaptación y extensión de los sistemas, la creciente presión por disminuir tiempos de producción y comercialización, entre otras.

Las pruebas del software, se vinculan directamente al incremento en la calidad del producto de software construido. El diseño y ejecución de las pruebas, en particular las pruebas de regresión, son

costosas e insumen gran cantidad de recursos y tiempo, y representan un desafío para la toma de decisiones de la gerencia de proyectos. En este marco, la automatización de las pruebas toma gran relevancia, ya que facilitan la labor tanto del ingeniero de pruebas como de la gerencia del proyecto, y permite obtener resultados de calidad, más confiables y en menor tiempo.

Las pruebas automáticas facilitan la comprobación de diferentes secuencias de entrada cubriendo una gran cantidad de casos, para buscar comportamiento indebido, encontrar defectos o asegurar la correctitud del programa. Además de ser usado en pruebas de regresión, el testeo automatizado es también empleado de manera habitual en la simulación de interfaces de usuario. Las aplicaciones de testeo automatizado son capaces de replicar funciones del usuario, como entradas de teclado y clicks del mouse, entre otros, y las respuestas de la interfaz gráfica de la simulación son guardadas y examinadas.

Las aplicaciones web son un tipo de software que constituyen una eficaz alternativa para satisfacer los requerimientos actuales del mercado. Las aplicaciones web son cada vez más utilizadas en múltiples y diversos dominios. Desde el surgimiento de la World Wide Web (WWW) en 1991, ha existido un continuo y acelerado crecimiento en el desarrollo de sitios web, llegando éstos a convertirse en un elemento frecuente y necesario en la vida cotidiana. Existen millones de usuarios activos en Internet que acceden diariamente a sitios web con diferentes propósitos como: trabajo, trámites, entretenimiento, comunicación, finanzas e inversión, educación, diversión, entre otros.

Actualmente, la clara transformación de las actividades cotidianas, que

incluyen procesos de digitalización, demandan al desarrollo de aplicaciones web prestar especial atención en el cumplimiento de *estándares de accesibilidad* establecidos, garantizando que cualquier usuario pueda acceder, navegar y utilizar dicho sitio. Las normas que regulan la accesibilidad en el software, ISO 9241-171 [1], definen a la *Accesibilidad* como “*la usabilidad de un producto, servicio, entorno o instalación por parte de personas con la gama más amplia de capacidades*” y a la *Usabilidad* como “*el grado en el que un sistema, producto o servicio puede ser utilizado por usuarios específicos para lograr objetivos específicos con eficacia, eficiencia y satisfacción en un contexto de uso específico*”.

El comité internacional (W3C-World Wide Web Consortium) que controla el uso y crecimiento de internet, definió las pautas de accesibilidad (WCAG) [2] dejando establecidos los requisitos funcionales y no funcionales que los sitios web debe cumplir, minimizando problemas de accesibilidad y permitiendo el acceso a todas las personas. Incluso, algunos países, como Estados Unidos y la Unión Europea, han avanzado en la definición de leyes que regulan el desarrollo de sitios web y cumplan con los criterios o pautas de accesibilidad.

La importancia creciente de garantizar la accesibilidad de las aplicaciones web, y en particular brindar condiciones adecuadas y seguras de acceso a todas las personas, aún con capacidades diferentes, llevaron a que en la actualidad existen numerosas herramientas automáticas de prueba de accesibilidad web, que permiten evaluar el cumplimiento de las pautas y legislaciones mencionadas, como: Axe, HTML CodeSniffer, Total Validator, Accessibility Checker, solo por nombrar algunas.

Por otra parte, la accesibilidad en aplicaciones web ha sido abordada en numerosos estudios, y a partir de la evaluación de criterios de accesibilidad en sitios web en casos seleccionados, aseguran la existencia de un importante número de ellos que no cumple con los estándares y legislaciones establecidas en torno a la accesibilidad web.

No obstante, la existencia de numerosas herramientas que automatizan las pruebas de software, en particular las pruebas de accesibilidad web, no garantiza ni facilita la selección adecuada. El principal problema que enfrentan los ingenieros de pruebas, es la selección y elección de las herramientas más adecuadas y beneficiosas para el análisis y evaluación que se requiere. Algunas investigaciones como las de Mehmet [3] y Al-Ahmad [4] compararon distintas herramientas y obtuvieron resultados similares: diferentes herramientas encuentran diferentes fallas, ninguna las encuentra a todas y tienen un índice de cobertura de fallas diferente. En conclusión, si el desarrollador quiere automatizar el proceso de encontrar dónde ocurren los problemas de accesibilidad, deberá usar más de una herramienta para obtener más cobertura. Algunos trabajos analizados que abordan el grado de accesibilidad que actualmente presentan los sitios web, y que resultan de interés para esta investigación son: [5] realiza una exhaustiva comparación y diferenciación entre el concepto de usabilidad y el de accesibilidad, en [6] los autores hacen un análisis completo de la evolución las pautas de accesibilidad WCAG 2.0. Y en los trabajos [7], [8] y [9] analizan con herramientas automáticas de testing la accesibilidad de sitios web de universidades de Indonesia, de municipalidades de Polonia y de universidades argentinas respectivamente.

En este estudio, se ahonda en la investigación de herramientas de automatización de pruebas existentes y los resultados que brindan. Se realiza un análisis exploratorio de herramientas, open source y de licencia de software libre, utilizadas para evaluar la accesibilidad en aplicaciones web. Se procede a analizar su funcionamiento y a recolectar las técnicas aplicadas para medir la accesibilidad en aplicaciones web.

Con esta propuesta, se pretende definir una arquitectura modular con una interfaz de comunicación común que permita la integración, interacción y evaluación de aplicaciones web. Se plantea el diseño y construcción de una plataforma que integre herramientas automáticas de prueba utilizadas en la evaluación de aplicaciones web, y en particular, la verificación del cumplimiento de los criterios establecidos para garantizar la accesibilidad a sitios web a todas las personas, aún con capacidades diferentes. Esto permitirá enriquecer la actividad de prueba del software mediante la aplicación de diferentes tipos de prueba, y evaluar distintos aspectos de una misma aplicación web a partir de cada una de las herramientas integradas. Se pretende colaborar mayoritariamente con la labor de los ingenieros de pruebas, como así también, brindar un valioso aporte al cumplimiento y control de las pautas y legislaciones de accesibilidad establecidas para los sitios web.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

El proyecto de investigación en el que se enmarca esta propuesta, persigue la definición y redefinición de modelos que aporten a la mejora de procesos de desarrollo de software, siendo uno de sus

objetivos específicos, el desarrollo de herramientas que soporten la automatización de procesos de desarrollo de software. El abordaje que se propone considerar con éste trabajo de investigación promete un importante aporte a la mejora de la etapa de prueba del proceso de desarrollo de software, y la aplicación propuesta como resultado de la integración de herramientas automáticas de testing para aplicaciones web, permitirá aprovechar las ventajas que ofrece cada una de las herramientas incluidas, combinarlas, y corregir las dificultades que las mismas presenten.

3. RESULTADOS ESPERADOS

Con este trabajo se busca promover el uso de herramientas automáticas de prueba en proyectos de desarrollo de software de mediana y alta complejidad, que aborden el desarrollo de aplicaciones web, ya que permitirán optimizar los procesos, probar un gran número de casos de prueba, y obtener resultados más confiables. En particular, se espera definir una arquitectura que posibilite la integración modular y extensible de herramientas automáticas de testing aplicadas a comprobar los criterios de accesibilidad incluidos en el desarrollo de aplicaciones web. También se propone construir una aplicación, que implemente la arquitectura definida, con una interfaz que integre y unifique la ejecución de estas herramientas de prueba automáticas y, facilite a ingenieros de prueba y desarrolladores, y proporcione mejores resultados en la comprobación de pautas de accesibilidad de sitios web.

4. FORMACION DE RECURSOS HUMANOS

Los integrantes del equipo de investigación que trabajan en las líneas del programa y proyectos en los que se enmarca este trabajo, conforman un equipo interdisciplinario integrado por informáticos, químicos, físicos y matemáticos, con trayectoria en investigación en ingeniería de software y en didáctica de las ciencias. Además, integran este equipo, estudiantes de las carreras de computación donde se forman científicamente y también desarrollan sus tesis de grado. También integra este equipo, un técnico informático y una asesora pedagógica. Actualmente, se están desarrollando dos tesis de Licenciatura en ciencias de la computación, una de Magister en Ingeniería de Software, una Magister en Ciencias, y una Tesis de Doctorado en informática. Particularmente, en la línea de investigación presentada este trabajo participa un becario CIN y becario en investigación SECYT-UNRC, y le permitirá concluir su tesis de grado, con proyecciones a continuar con estudios de posgrado.

5. BIBIOGRAFÍA

- [1] ISO 9241-171:2008. Ergonomics of human-system interaction — Part 171: Guidance on software accessibility. <https://www.iso.org/standard/39080.html>
- [2] Web Content Accessibility Guidelines (WCAG) 2.1. W3C 05 June 2018. <https://www.w3.org/TR/WCAG21/>
- [3] D. Mehmet, "What we found when we tested tools on the world's least-accessible webpage", *Accessibility in government*, 2017.
- [4] A. Al-Ahmad, I. Ahmaro and M. Mustafa, "Comparison between web accessibility Evaluation tools", 2013.
- [5] D. Alexander, "Usability and Accessibility: Best friends or Worst enemies", Monash University, 2006.
- [6] M. Ribera, M. Porras, M. Boldu, M. Termens, A. Sule and P. Paris, "Web Content Accessibility Guidelines 2.0", *Program*, vol. 43, no. 4, pp. 392-406, 2009.
- [5] M. Vigo, J. Brown and V. Conway, "Benchmarking web accessibility evaluation tools", *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility - W4A '13*, 2013.
- [7] W. Arasid et al., "An Analysis of Website Accessibility in Higher Education in Indonesia Based on WCAG 2.0 Guidelines", *IOP Conference Series: Materials Science and Engineering*, vol. 306, p. 012130, 2018.
- [8] K. Król and D. Zdonek, "Local Government Website Accessibility - Evidence from Poland", *Administrative Sciences*, vol. 10, no. 2, p. 22, 2020.
- [9] Laitano, M. I, "Accesibilidad web en el espacio universitario público argentino.", *Revista Española de Documentación Científica*, 38(1): e079, 2015.

La Sostenibilidad como Característica Transversal a la Calidad del Software

Leo Rosana*, Salgado Carlos⁺, Sanchez Alberto⁺, Peralta Mario⁺

* Departamento Académico de Ciencias Exactas, Físicas y Naturales. Universidad Nacional de la Rioja. La Rioja, Argentina
leorosana@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
e-mail: {csalgado, mperalta, [asanchez](mailto:asanchez@unsl.edu.ar)}@unsl.edu.ar

RESUMEN

En la actualidad, es difícil pensar algún proceso de gestión que no aproveche las ventajas que brinda un software, en ese sentido la calidad del mismo se constituye en un punto importante a tener en cuenta en vista de la satisfacción de necesidades expresas o implícitas. Para tal fin es necesario utilizar modelos o estándares.

Por otro lado, se ha generalizado la idea de *sostenibilidad*, el color verde, la concientización por la preservación de los recursos del planeta, etc. incluso en el área de las Tecnologías de la Información y las Comunicaciones (TIC) y sobre todo en aspectos de infraestructura (centros de cómputos, servidores, dispositivos que consumen energía).

La sostenibilidad se analiza desde tres dimensiones: ambiental (para preservar la biodiversidad sin renunciar al progreso económico y social), económica (para lograr rentabilidad) y social (para lograr consenso social).

El presente trabajo pretende demostrar la importancia de incorporar como característica de calidad del software a la sostenibilidad, basado en estándares de calidad, como por ejemplo el modelo de calidad definido por la norma ISO/IEC 25010, pero teniendo en cuenta las guías y buenas prácticas que se correspondan con la sostenibilidad.

Palabras Clave: *Calidad de Software, Sostenibilidad, ISO 25010, Modelos de Calidad.*

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 N° 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 calidad de los sistemas informáticos se ha convertido hoy en día en uno de los principales objetivos estratégicos de las organizaciones debido a que, cada vez más, sus procesos más importantes y, por lo tanto, la propia supervivencia de las organizaciones depende de los sistemas informáticos según se menciona en [1]. Cuando hablamos de sistemas informáticos, debemos mencionar que los

productos de software son un componente de importancia dentro del sistema que los contiene, y la calidad de éste afectará a la calidad del sistema en su conjunto. Aquí, entonces, sería importante dar alguna definición sobre calidad de software, si bien en la literatura hay varias definiciones, podemos mencionar la de [2] donde se define a la calidad de software como el cumplimiento de los requisitos de funcionalidad y desempeño explícitamente establecidos, de los estándares de desarrollo explícitamente documentados, y de las características implícitas que se espera de todo software desarrollado profesionalmente. También en [3] se la define como: “Grado en que el producto software satisface las necesidades expresadas o implícitas, cuando es usado bajo condiciones determinadas”.

La especificación y evaluación de la calidad del software es factor clave para garantizar el valor a las partes interesadas. Esto se puede lograr mediante la definición de las características de calidad necesarias y deseadas asociadas con las metas y objetivos del sistema. Es importante que dichas características se especifiquen, midan y evalúen siempre que sea posible utilizando medidas y métodos de medición validados o ampliamente aceptados [4].

Para garantizar la calidad de software es importante implementar algún modelo o estándar que permita la gestión de atributos en el proceso de construcción de software, teniendo en cuenta que la concordancia de los requisitos y su construcción son la base de las medidas de calidad establecidas [5].

La ventaja de estos estándares es que la calidad se convierte en algo concreto, que se puede definir, medir y, sobre todo, planificar. Ayudan también a comprender las relaciones que existen entre las diferentes características de un producto

de software.

Si bien, la norma ISO/IEC 14001:2015 [6] proporciona a las organizaciones un marco con el que proteger el medio ambiente y responder a las condiciones ambientales cambiantes, manteniendo el equilibrio con las necesidades socioeconómicas, este estándar no incluye requisitos específicos para otros sistemas como lo son los de gestión de la calidad.

Por otro lado, si se habla de satisfacer necesidades, es importante tener en cuenta un término que se escucha cada vez con más frecuencia, como lo es *sostenibilidad*. Surge en el año 1987 en el Informe Brundtland [7], titulado “Nuestro Futuro Común”, y por la necesidad de estudiar y delimitar el impacto de las actividades humanas sobre el medio ambiente. Allí se define la sostenibilidad como “el desarrollo que satisface las necesidades del presente sin comprometer la habilidad de generaciones futuras de satisfacer sus propias necesidades”, todo ello garantizando el equilibrio entre el crecimiento económico, el medio ambiente y el bienestar social, pilares del término que se analiza.

La autora en [8] presenta revisiones sistemáticas a la literatura referida a la sostenibilidad en el ámbito del software. Realiza una cronología desde 1987 cuando la Comisión Mundial de Medio Ambiente y Desarrollo (CMMAD) expone [7]. Presenta los términos Green IT (Tecnologías de la información verdes) y Green IS (Sistemas de información verdes) con diversas definiciones y en el caso de Green IT, varias subdivisiones.

Afirma que “El proceso de desarrollo del software para ser sostenible tiene que hacer uso responsable de los recursos que consumen energía y usan materias primas no renovables. Y el software debe ser fácil de mantener”.

El involucrar a los equipos de desarrollo permitirá evaluar cuán sostenible es el proceso de construcción y el producto de software, como así también, tomar conciencia de las ventajas que implica dicho producto en la toma de decisiones todo vinculado a la reducción del impacto medio ambiental en relación al uso de las TIC.

En [9] se explica que la Sostenibilidad actualmente, es un aspecto fundamental en las TIC, apunta al cuidado de los recursos de nuestro planeta, aunque tiene más alcance que el factor ecológico si nos basamos en sus tres pilares fundamentales: Social, Económico y Ecológico (ambiental).

Al aplicar y evaluar criterios de sostenibilidad en cada elemento de un proyecto, se podrá establecer el grado o nivel de sostenibilidad (no concienciado, equitativo, viable, soportable, sostenible)

Valorar la sostenibilidad mediante una sencilla puntuación o medición en un rango acotado, permitirá tener una idea cuantitativa de la misma. Estableciendo un umbral se podrá determinar si el objeto de estudio es o no sostenible.

La Sostenibilidad será un criterio esencial en las organizaciones a fin de alinearse a los compromisos de preservación de los recursos naturales, contribuyendo a la toma de decisiones estratégicas, tácticas y operativas.

La Norma Internacional ISO/IEC 25010:2011 [4], una de las divisiones de la serie SQUARE, describe el modelo de calidad para el producto software, presentando características y subcaracterísticas de calidad, criterios a tener en cuenta al momento de la evaluación.

En el 2015 un manifiesto expone principios y compromisos vinculados con

el diseño sostenible [10]. Infiriendo que la sostenibilidad es un concepto sistémico, tiene múltiples dimensiones (social, medioambiental, económica, técnica y humana) y todas deben analizarse.

Al hablar de tecnologías sostenibles, se piensa en un menor consumo de energía, empleo de menor cantidad de recursos, la no contaminación, el reciclado o reutilización, siempre enfocados en satisfacer las necesidades de la sociedad. De allí se define un producto sostenible como aquel que aporta beneficios ambientales, sociales y económicos resguardando la salud pública, el bienestar y el medio ambiente en todo su ciclo de vida.

Bajo estos conceptos, se puede decir que el software también puede ser sostenible, cuando su desarrollo se basa en el uso adecuado de recursos y cuando su impacto negativo en la economía, la sociedad y el medio ambiente, es mínimo o, en el mejor de los casos resulta positivo respecto del desarrollo sostenible [11]. Pero, ¿cómo lo medimos?

Una medida es un valor asignado a algún atributo de dicho producto. Una métrica interpreta esos valores e incluso puede relacionarlos. El IEEE Standard Glossary of Software Engineering Terminology define métrica como “una medida cuantitativa del grado en el que un sistema, componente o proceso posee un atributo determinado”. Al evaluar las métricas, surgen los indicadores. Un indicador es una métrica o combinación de ellas que proporciona comprensión sobre el proceso, proyecto o producto software [2].

Los Modelos de Calidad del Software (MC), son artefactos específicamente diseñados y construidos para apoyar en estos procesos.

Según Calero et al. en [12], y el estándar ISO 8402 [13] un modelo de calidad puede definirse como el conjunto

de factores de calidad, y de relaciones entre ellos, que proporciona una base para la especificación de requisitos de calidad y para la evaluación de la calidad de los componentes software. Los modelos de calidad se estructuran generalmente como una jerarquía (ya sea un árbol, ya sea un grafo dirigido), donde factores de calidad más genéricos, como eficiencia o usabilidad, se descomponen en otros más particulares, como tiempo de respuesta o facilidad de aprendizaje, probablemente en diversos niveles de descomposición. También menciona que 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. (1997) [14], Boehm et al. (1978) [15], Keller et al. (1990) [16] y FURPS Grady y Caswell (1987) [17]. Para los modelos de calidad a medida existen diversas propuestas de métodos para crearlos entre las que podemos destacar a GQM (Goal-Question-Metric) de Basili [18] y la del estándar IEEE 1061 [19].

Para el caso de los modelos de calidad mixtos se pueden destacar el ADEQUATE Horgan [20], el modelo de Gilb [21] y el modelo propuesto en el estándar ISO/IEC 9126-1 [22], este último es actualizado y reemplazado por el estándar ISO/IEC 25010 [4].

Coincidiendo con lo expresado en [10], 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.

Por lo expresado anteriormente, el objetivo de este trabajo es desarrollar un modelo que favorezca la evaluación de la calidad del software en una organización, con el fin de detectar los requerimientos de los usuarios y/o necesidades del dominio no cubiertas, en pos de mejorar la eficiencia organizacional a través de la mejora en la calidad del software que da soporte a los procesos de la organización teniendo en cuenta criterios de sostenibilidad.

El modelo que se propone se centrará en la calidad del producto de software, tomando como punto de partida el modelo conceptual de los procesos de la organización.

Es de gran importancia que las organizaciones puedan contar con modelos de calidad que partan del modelo conceptual de sus procesos y que les permita la evaluación objetiva y de manera sistémica de la calidad del software que da soporte a los procesos organizacionales, respetando y cumpliendo con los requisitos de sostenibilidad que las sociedades actuales exigen.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de trabajo de esta línea de investigación están asociados a:

- Evaluación de la calidad de productos de software.
- Estudio de modelos conceptuales con base en la sostenibilidad aplicados a la calidad de productos de software.
- Estudio de estándares y metodologías aplicadas a la construcción de MC.

3. RESULTADOS OBTENIDOS/ESPERADOS

La línea de I+D presentada en este trabajo se está desarrollando, tomando en cuenta los objetivos del Proyecto de I+D que la contiene.

De manera específica para esta línea se obtuvieron hasta la fecha los siguientes resultados:

- Lograr, a través de un MC, una evaluación objetiva de un determinado software para saber si este cubre las demandas organizacionales y las necesidades de los usuarios, y en función de ello proponer las mejoras necesarias para cubrir las necesidades no cubiertas por el mismo.
- Estudio de los modelos y normas de calidad aplicados a productos de software.
- Estudio de metodologías o métodos que guíen la construcción de modelos de calidad.

Los resultados esperados son:

- Concluir con la definición del modelo conceptual de calidad aplicado a productos software a través de la metodología y norma de calidad seleccionada, sin perder de vista los criterios de sostenibilidad.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se trabaja en lo referente a distintos modelos y métodos de evaluación de calidad. Se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) de la Lic. Rosana Leo, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

Como así también se están llevando a

cabo algunas tesinas de grado para la Licenciatura en Computación. Y trabajos finales de carrera de la Ingeniería en Informática e Ingeniería en Computación de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

5. BIBLIOGRAFÍA

- [1]. Piattini Velthuis et al., Calidad de sistemas de información. 2ª edición: RaMa, 2011.
- [2]. R. Pressman, Ingeniería del Software. 7^{ma} Ed: Mcgraw-Hill, 2010.
- [3]. ISO/IEC 25000 Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE)-System and software quality models.
- [4]. ISO/IEC 25010:2011 Systems and software Engineering-Systems and Software Quality Requirements and Evaluation (SQUARE) System and software quality models
- [5]. Callejas-Cuervo, Mauro; Alarcon-Aldana, Andrea Catherine; Álvarez-Carreño, Ana María Modelos de calidad del software, un estado del arte. En: Entramado. Enero - Junio, 2017. vol. 13, no. 1, p. 236-250.
- [6]. ISO/IEC 14001:2015 Environmental Management Systems. Requirements with guidance for use.
- [7]. CMMAD. Comisión Mundial de Medio Ambiente y Desarrollo (1987). Informe Brundtland: "Nuestro futuro común".
- [8]. Dra. Anaisa Hernandez Gonzalez. "La sostenibilidad y el software"
- [9]. Montoya Guillermo - CEO Deiser S.L "Sostenibilidad del software: más allá de Green IT"
- [10]. Christoph Becker. Manifiesto Karlskrona. Sustainability design

- and software. 2015.
- [11]. Naumann - Dick - Kern – Johan. El modelo GREENSOFT: un modelo de referencia para el software verde y sostenible y su ingeniería
- [12]. C. Calero, M. Piattini, M. Moraga, Calidad del producto y proceso software: Ra-Ma, 2010.
- [13]. International Standards Organization (1986). ISO International Standard 8402: Quality Management and Quality Assurance-Vocabulary
- [14]. McCall, J.A., Richards, P.K. Y Walters, G.F. (1977) “Factors in Software Quality”. RADC TR-77-369, Vols I, II, III, US Rome Air Development Center Reports NTIS AD/A-049.
- [15]. Boehm.W., Brown, J.R., Kaspar, H., Lipow, M., Macleod, G.J. y Merritt, M.J. (1978). “Characteristics of Software Quality”. North Holland Publishing Company.
- [16]. Keller, S., Kahn, L. y Panara, R. (1990) “Specifying Software Quality Requirements with Metrics”. Systems and Software Requirements Engineering - IEEE Computer Society Press – Tutorial
- [17]. Grady R.B. y Caswell, D.L. (1987). “Software Metrics: Establishing a Company-Wide Program”. Prentice-Hall.
- [18]. Basili, V.R., Caldiera, G. y Rombach, H.D. (1994). “Goal Question Metric Paradigm”. En: Encyclopedia of Software Engineering 1, John Wiley & Sons.
- [19]. Institute Of Electrical Electronic Engineering (1998). IEEE Std 1061-1998 IEEE Standard for a Software Quality Metrics Methodology – Description.
- [20]. Horgan, G., Khaddaj, S. Y Forte, P. (1999) “Anessential Views Modelfor Software Quality Assurance”. En: Project Control for Software Quality, Shaker Publishing.
- [21]. Gilb, T. (1988). “Principles of Software Engineering Management”. Addison Wesley.
- [22]. ISO/IEC 9126-1. (2001). Software engineering — Product quality — Part 1: Quality model.

MARCO DE TRABAJO PARA LA GESTIÓN DEL CONOCIMIENTO EN LA ADMINISTRACIÓN DE PROYECTOS DE DESARROLLO DEL SOFTWARE

Gerlero, Patricia¹; Straccia, Luciano¹; Pollo-Cattaneo, Ma. Florencia¹

¹Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS) –
Programa de Maestría en Ingeniería en Sistemas de Información
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires Tel +54 11 4867-7511

pagerlero@yahoo.com.ar, lstraccia@frba.utn.edu.ar, flo.pollo@gmail.com

Resumen

La sistematización de las lecciones aprendidas es utilizada por la industria del software como herramienta para ayudar a los gerentes a determinar qué información es la más relevante con el fin de alcanzar con éxito sus proyectos complejos. Sin embargo, los proyectos se ven atravesados por la ambigüedad de las relaciones humanas, la dinámica del entorno que influye de manera compleja y la limitación en la disponibilidad de los recursos. Es necesario ampliar la comprensión de la complejidad del proyecto como una noción subjetiva, que refleja las experiencias vividas de las personas involucradas. La clave para lograr la institucionalización de ese conocimiento es cambiar el paradigma de Gestión de proyecto y acoplarlo al sistema de Gestión de conocimiento.

La Gestión del conocimiento (GC) es un proceso a través del cual las organizaciones logran descubrir, utilizar y mantener el conocimiento, con la idea de alinearlos con las estrategias de negocio para la obtención de ventajas competitivas, e implica su análisis desde diversas perspectivas.

En este contexto, el objetivo del proyecto es definir un marco de trabajo

para la Gestión del conocimiento en la administración de proyectos de desarrollo de software.

Palabras clave: gestión del conocimiento, administración de proyectos, desarrollo de software, marco de trabajo.

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 se ha conformado en el año 2009 el Grupo GEMIS, con dependencia del Departamento de Ingeniería en Sistemas de Información, integrado por un equipo

de docentes, alumnos y graduados con interés en la sistematización de conocimientos y su promoción sobre el campo de la Ingeniería en Sistemas de Información y la Ingeniería de Software incluyendo sus aplicaciones y abordajes metodológicos en todo tipo de escenarios.

Introducción

Administración de proyectos

En 1994 el Standish Group publica el primer reporte donde analiza los principales factores de fracaso de proyectos de desarrollo de software, identificando a los requerimientos incompletos, falta de participación del usuario, falta de recurso y la falta de planificación. En contraposición, la participación del usuario y el apoyo de la dirección son considerados como principales factores de éxito [2]. En 2015 incluye como relevante la madurez emocional y los recursos calificados [3].

Por su parte, Steve McConnell encuesta durante el 2007 a más de 500 profesionales de software [4] para determinar la frecuencia y gravedad de los errores más comunes en los proyectos de desarrollo de software: la planificación, la gestión de riesgo, la estimación insuficiente, la participación del usuario se registra entre los primeros de la lista. El foco se centra en los procesos evidenciando la importancia de la intervención de los usuarios y los conflictos asociados a la misma participación.

Revisiones literarias como las realizadas por Rezvani y Khosravi [5], basadas en la identificación y agrupamiento de factores en varios ejes dimensionales, conforman el punto de partida de muchos otros investigadores. Si bien dejan de lado la complejidad y las múltiples interrelaciones, presentan

diversos indicadores de las lecciones aprendidas, adaptadas y utilizadas por la industria del software como método para ayudar a los gerentes a determinar qué información es la más relevante con el fin de alcanzar con éxito sus proyectos complejos [6].

El conocimiento explícito, en muchos casos institucionalizado mediante las lecciones aprendidas, presentan una clara limitación: la realidad única y cambiante de los proyectos de desarrollo de software donde el proceso social, la creación de valor y la importancia de comprender las experiencias vividas conforman en sí mismo sistemas complejos.

Los proyectos, independientemente de su tamaño, naturaleza o misión se ven atravesados por la ambigüedad de las relaciones humanas, la dinámica del entorno que influye de manera compleja y la limitación en la disponibilidad de los recursos [7]. Es necesario ampliar la comprensión de la complejidad del proyecto como una noción subjetiva, que refleja las experiencias vividas de las personas involucradas [8]. Para crear conocimiento, las habilidades compartidas con otros necesitan ser internalizadas, reformuladas, enriquecidas y traducidas para adaptarse a la nueva identidad [9]. La percepción, intuición y corazonadas como parte subjetiva tiene que ser incorporado dentro de la jerarquía y como un eslabón fundamental a la hora de predecir el posible desenlace al momento del intercambio de ideas sobre la problemática planteada.

La clave para lograr la institucionalización de ese conocimiento es cambiar el paradigma de Gestión de proyecto y acoplarlo al sistema de Gestión de conocimiento. Crear nuevo conocimiento significa literalmente recrear la organización [10] o crearla en el caso de organizaciones temporales.

Sistema de Gestión del Conocimiento

La Gestión del conocimiento (GC) es un proceso a través del cual las organizaciones logran descubrir, utilizar y mantener el conocimiento, con la idea de alinearlo con las estrategias de negocio para la obtención de ventajas competitivas [11].

Según Sanchez-Sanchez [12], “la Gestión del conocimiento desde la psicología social de las organizaciones radica en un proceso de influencia social de grupos colaborativos en torno a la transferencia de un conocimiento implícito hacia un resguardo del conocimiento tácito a fin de poder aprovechar la experiencia y habilidades de talentos y líderes ante las contingencias del entorno, o bien, los riesgos y las amenazas del contexto”.

Avendaño Perez y Florez Urbáez [13] definen la GC como “un enfoque gerencial o disciplina emergente que busca de manera estructurada y sistemática aprovechar el conocimiento generado para alcanzar los objetivos de la organización y optimizar el proceso de toma de decisiones.”

A medida que las organizaciones y la sociedad han ido notando importancia de la GC, diversos investigadores han propuesto una serie de pasos o guías para identificar cómo debe ser llevado a cabo el proceso de GC para que sea exitoso y genere los beneficios esperados para la organización. Estas propuestas las han hecho mediante modelos de GC. En relación con esto, Pérez y Urbáez [13] indican que los investigadores en las últimas décadas, a través de sus propuestas teóricas y de sus investigaciones de experiencias en empresas, han intentado aproximarse a explicar cómo debe organizarse, gestionarse y fluir el conocimiento en la organización.

Rodriguez Gómez [14] indica que “la multidisciplinariedad inherente al estudio de la gestión del conocimiento supone la existencia de diferentes perspectivas para el desarrollo y el estudio de los sistemas y modelos de gestión del conocimiento”. La definición de un modelo de GC requiere un abordaje multidisciplinar y desde diversas perspectivas, presentadas en [15]: roles y responsabilidades, proceso, gobernanza y tecnología (basadas en [16]), aspectos organizaciones (en los términos de [17]) y representación del conocimiento [18].

Wigg [19], propuso uno de los primeros modelos o guías para la GC en las organizaciones, haciendo énfasis en la falta de este tipo de guías en su época, y en la importancia de gestionar el conocimiento para obtener ventajas competitivas para la organización.

Posteriormente, diversos autores propusieron sus modelos de GC, entre los cuales puede hallarse Nonaka y Takeuchi [20], Sveiby [21], Earl [22], Kerschberg [23], Mc. Elroy [24], CEN [25], entre otros. Sin embargo estos modelos no son disciplinares, específicos, sino generales. En [26,27] se presentan modelos vinculados a la industria del software y Maulini [28,29] propone un nuevo modelo de gestión de conocimiento específicamente orientado a las pequeñas y medianas fábricas de software, analizando las dificultades que enfrenta la industria y tipología de organizaciones específica y las características de los diferentes modelos trabajados por diferentes autores.

Líneas de Investigación, Desarrollo e Innovación

El proyecto presentado en este trabajo propone avanzar en el estado del conocimiento para la Gestión del

conocimiento, adentrándose en las particularidades de la actividad de Administración de Proyectos para el ámbito de la industria del software.

En el período 2018-2021 se llevó adelante un Proyecto de Investigación y Desarrollo (PID) denominado “La gestión del conocimiento en pequeñas y medianas fábricas de software en el Área Metropolitana de Buenos Aires” que indaga en los modelos de gestión del conocimiento.

Para el período 2021-2023 se lleva adelante el PID “Arquitectura tecnológica para la implementación de Gestión del Conocimiento”.

Como una nueva línea en el marco de los proyectos desarrollados, se propone ahondar en las particularidades de tareas o áreas específicas de la industria del software y la gestión del conocimiento.

De esta manera, se articula dentro de los objetivos de GEMIS en el campo de la Informática, la generación de nuevos conocimientos en el área de la Ingeniería de Software y la Gestión del Conocimiento.

Resultados esperados

Objetivo General

Definir un marco de trabajo para la gestión del conocimiento en la administración de proyectos de desarrollo de software

Objetivos Específicos

Los objetivos específicos son: a) identificar los criterios asumidos como indicadores de éxito y fracaso y sus factores condicionantes en los proyectos de desarrollo de software; b) relevar marcos de trabajo para la Gestión del conocimiento disciplinar; c) proponer un marco de trabajo para la GC en la administración de proyectos de desarrollo de software.

Resultados esperados

Como resultado de las tareas a desarrollar se espera:

- contar con un marco de trabajo para la GC en la administración de proyectos de desarrollo de software;
- validar en una organización real el marco de trabajo propuesto;
- vincular con otros PID del Grupo, relacionados con Sistemas Expertos y Sistemas Inteligentes.

Formación de Recursos Humanos

El equipo se encuentra conformado por investigadores formados, investigadores de apoyo, graduados de grado, posgrado y alumnos de la carrera de Ingeniería en Sistemas de Información y de la Maestría correspondiente. Esta nueva línea de trabajo permitirá:

- la generación nuevas oportunidades y experiencias para la formación de los investigadores;
- la incorporación de nuevos becarios graduados a través de las becas BINID;
- el desarrollo de por lo menos 2 tesis de la Maestría en Ingeniería en Sistemas de Información. Además, se prevé el inicio de por lo menos 2 trabajos finales de la Especialidad en Ingeniería en Sistemas de Información

Referencias

- [1] Universidad Tecnológica Nacional (2016). Resolución Nro. 2508/16
- [2] Standish Group. (1994). Chaos Report. Recuperado el 10 de 05 de 2020, de https://www.standishgroup.com/sample_research_files/chaos_report_1994.pdf
- [3] Standish Group. (2015). Chaos Report. Recuperado el 15 de 02 de 2020, de

https://www.standishgroup.com/sample_research_files/CHAOSReport2015-Final.pdf

[4] McConnell, S. (2008). Software Development's Classic Mistakes 2008. Recuperado el 01 de 07 de 2020, de <http://www.construx.com/Page.aspx>

[5] Rezvani, A., & Khosravi, P. (2019). Identification of failure factors in large scale complex projects: an integrative framework and review of emerging themes. *International Journal of Project Organisation and Management*, 11(1), 1-21

[6] Garousi, V., Tarhan, A., Pfahl, D., Coşkunçay, A., & Demirörs, O. (2019). Correlation of critical success factors with success of software projects: an empirical investigation. *Software Quality Journal*, 27(1), 429-493.

[7] Cicmil, S., Cooke-Davies, T., Crawford, L., & Richardson, K. (2017). Exploring the complexity of projects: Implications of complexity theory for project management practice. Project Management Institute.

[8] Mikkelsen, M. F. (2020). Perceived project complexity: a survey among practitioners of project management. *International Journal of Managing Projects in Business*.

[9] Nonaka, L., Takeuchi, H., & Umemoto, K. (1996). A theory of organizational knowledge creation. *International Journal of Technology Management*, 11(7-8), 833-845

[10] Bueno, E. (2000). Dirección del Conocimiento y Aprendizaje: Creación, distribución y mediación de Intangibles.

[11] Sánchez-Sánchez, Arturo, Valés-Ambrosio, Oscar, García-Lirios, Cruz, & Amemiya-Ramirez, Michiko. (2020). Confiabilidad y validez de un instrumento que mide la gestión del conocimiento. *Espacios en blanco. Serie indagaciones*, 30(1), 1-10

[12] Cuesta, A. (2012). Modelo integrado de gestión humana y del conocimiento: una tecnología de aplicación. *Revista Venezolana de Gerencia*, 57, 86-98.

[13] Avendaño Perez, V.; Flores Urbáez, M. (2016). Modelos teóricos de gestión del conocimiento: descriptores, conceptualizaciones y enfoques. En *Entreciencias: diálogos en la Sociedad del Conocimiento*, 4(10), agosto-noviembre, 2016. Universidad Nacional Autónoma de México. León, México.

[14] Rodríguez Gómez, David (2006). Modelos para la creación y gestión del conocimiento: una aproximación teórica. *Educación* 37, 25-39.

[15] Straccia, L.; Ramacciotti, C.; Pollo-Cattáneo, M.F. (2020). Una visión de la tecnología para la Gestión del Conocimiento. Resultados en la literatura latinoamericana. En "Desarrollo e Innovación en Ingeniería - Quinta Edición" (Ed. Prof. Edgar Serna M). Pág. 135-142. Editorial Instituto Antioqueño de Investigación. ISBN: 978-958-52333-4-8.

[16] Milton, Nick (2014). The 4 legs on the Knowledge Management table

[17] Zytnewski, Marius (2016). Integration of knowledge management systems and business processes using multi-agent systems. *Int. J. Computational Intelligence Studies*, Vol. 5, No. 2.

[18] Meneses-Ortegón, Juan; Tové, Teodor; Puiggali, Joan; Fabregat, Ramón (2020). Representación del conocimiento de un proceso de co-creación de material educativo. En *TecnoLógicas*, vol 23, nro 47. Medellín, Colombia.

[19] Wiig, K. (1993). *Knowledge Management Foundations: Thinking about thinking – How people and organizations create, represent, and use knowledge*. Arlington, TX: Schema.

[20] Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford university press

[21] Sveiby, K. E. (2001). A knowledge-based theory of the firm to guide in strategy formulation. *Journal of intellectual capital*, 2(4), 344-358.

[22] Earl, M. (2001). Knowledge management strategies: Toward a taxonomy. *Journal of management information systems*, 18(1), 215-233.

[23] Kerschberg, L. (2001, September). Knowledge management in heterogeneous data warehouse environments. In *International Conference on Data Warehousing and Knowledge Discovery* (pp. 1-10). Springer Berlin Heidelberg.

[24] McElroy, M. (2003). *The new knowledge management: complexity, learning and sustainable innovation*, Butterworth-Heinemann, Boston.

[25] CEN (2004) CEN CWA 14924 European Guide to good Practice in Knowledge Management - Part 1 to 5, Brussels.

[26] Bovea, T. L. C., & García, V. H. M. (2011). Modelo de gestión del conocimiento aplicado a un sistema complejo: Desarrollo de fábricas de software. Ninth LACCEI Latin American and Caribbean Conference. Colombia.

[27] Pons, N. L., Pérez, Y. P., Stiven, E. R., & Quintero, L. P. (2014). Diseño de un modelo de Gestión del Conocimiento para mejorar el desarrollo de equipos de proyectos informáticos. *Revista Española de Documentación Científica*, 37(2), 044.

[28] Maulini, A.; Straccia, L; Pollo-Cattaneo, M.F. (2018). Una aproximación a un modelo de gestión de conocimiento aplicable a las pequeñas y medianas fábricas de software. En Serna, E. (ed) (2018). *Desarrollo e Innovación en Ingeniería*. Editorial Instituto Antioqueño de Investigación, Medellín, Colombia. ISBN 978-958-59127-9-3.

[29] Maulini, A.; Straccia, L; Pollo-Cattaneo, M.F. (2019). Un modelo de gestión de conocimiento aplicable a las pequeñas y medianas fábricas de software. En "Desarrollo e Innovación en Ingeniería - Cuarta Edición" (Ed. Prof. Edgar Serna M). Capítulo 1, Pág. 5-16. Editorial Instituto Antioqueño de Investigación.

MEDICIÓN AUTOMÁTICA DE ASPECTOS HUMANOS EN EQUIPOS VIRTUALES DE DESARROLLO DE SOFTWARE

Gustavo Sevilla, Sergio Zapata, Facundo Gallardo, Estela Torres

Instituto de Informática / Facultad de Ciencias Exactas, Físicas y Naturales / Universidad Nacional de San Juan; {gsevilla,szapata,etorres,fgallardo}@iinfo.unsj.edu.ar

RESUMEN

En los últimos años la comunidad científica ha puesto mayor atención en el importante rol que juegan los aspectos humanos y sociales en el proceso de desarrollo de software. Hay un fuerte consenso que estos aspectos impactan significativamente tanto en el proceso de software como en el producto resultante.

Paralelamente, el progreso de las tecnologías de información y comunicación (TICs) y el crecimiento de prácticas de globalización han fomentado la creación de los denominados equipos virtuales de desarrollo de software. Equipos en donde los miembros interactúan en forma remota y mediante TICs. En este contexto aparece el uso extendido de las plataformas de colaboración para el desarrollo de software (software forges) como los repositorios Github, Jira, Stackoverflow, etc.

Este trabajo pretende sacar provecho de los datos que quedan registrados en estos repositorios de software y que de alguna manera reflejan las características de las interacciones entre los miembros de equipos virtuales de software. La medición y análisis de estos datos puede arrojar evidencia de la calidad de las relaciones humanas del equipo virtual.

Palabras clave: Equipos Virtuales, Software forges, Aspectos humanos y sociales en el desarrollo de software

CONTEXTO

Este trabajo está enmarcado en el proyecto “Aspectos Humanos del Desarrollo Global de Software” financiado por la Universidad

Nacional de San Juan, siendo la unidad ejecutora el Instituto de Informática (IdeI) de la misma Universidad.

El proyecto es de ejecución bianual y su comienzo fue en enero de 2019. El mismo está inserto dentro del Programa de Incentivos a docentes-investigadores de la Secretaría de Políticas Universitarias del Gobierno Nacional.

1. INTRODUCCION

En los proyectos de desarrollo global o distribuido de software los integrantes de los equipos virtuales no tienen interacción personal cara a cara, excepto esporádicas y costosas reuniones localizadas. Aquí los aspectos humanos se podrían ver perjudicados por la distancia física, cultural y horaria. En tales entornos, conocer el nivel de relaciones humanas es esencial para que los líderes de proyectos adviertan posibles problemas que pudieran afectar el desarrollo de software. El desarrollador más talentoso podría tener un rendimiento inferior y, finalmente, abandonar el proyecto solo porque no está contento con su entorno o sus colegas [1].

Un equipo de desarrollo global de software es aquel que está separado geográficamente mientras colabora activamente en un proyecto de software común [2].

Los Equipos Virtuales (EV) o globales, permiten, entre otras ventajas, integrar personal calificado de distinta localización geográfica y muchas veces a un menor costo laboral. A su vez, los EV plantean problemas o desafíos tal como la gestión remota de proyectos, donde uno de los principales

factores que la afectan son las relaciones sociales entre los miembros. Las buenas relaciones sociales interpersonales en EV son importantes para la comunicación, la motivación y la creación de un buen ambiente de trabajo que contribuya a la mejora de la productividad.

La cohesión del equipo es el grado en que sus miembros se identifican entre sí y con el grupo en general [3, 4]. Los equipos globales tienden a menores niveles de cohesión que los equipos colocalizados, en gran parte debido a que colaboran remotamente a través de tecnología en lugar de hacerlo cara a cara. Por otro lado, investigaciones previas sugieren que la falta de cohesión grupal se puede superar mediante el intercambio de más comunicaciones sociales entre los miembros del grupo [5]. Las interacciones sociales permiten a los miembros del equipo crear vínculos más fuertes, lo cual permitirá aumentar los niveles de cohesión del grupo.

Por otra parte, una de las repercusiones del uso de las TIC en las actuales relaciones sociales tiene que ver con el registro digital de ellas en repositorios de datos llamadas plataformas de software forges, las cuales ofrecen una oportunidad de analizar esos registros para medir y evaluar estas relaciones. Las software forges proveen varias herramientas y facilidades a los equipos de desarrollo distribuido, tales como sistemas de control de código fuente, listas de correos electrónicos y foros de comunicación, sistemas de seguimientos de fallas, espacio para alojamiento web y similares [6].

Así, el presente trabajo ofrece un aporte en este contexto, proponiendo una aplicación que explote los datos sociales almacenados en las software forges.

Github es un repositorio, extensamente usado, para alojar proyectos de software que utiliza el sistema de control de versiones de Git. La plataforma integra una serie de características

y funcionalidades que hacen única su información sobre las actividades de los desarrolladores de software a código abierto [7].

Github utiliza una forma de trabajo basada en Pull Request (PR), un nuevo método colaborativo de desarrollo distribuido de software. Un PR es un mecanismo de solicitud de cambio de código de software que facilita la colaboración entre desarrolladores, ese cambio puede ser para implementar una mejora, adicionar una nueva funcionalidad o para remover un defecto. El PR es la unidad de trabajo de GitHub y representa de alguna manera el trabajo colaborativo realizado por sus miembros. Un PR es una solicitud de cambio al software.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Comprender el impacto de los aspectos humanos de los equipos virtuales de desarrollo de software en el proceso de desarrollo de software.

Objetivos específicos: a) Medir y evaluar aspectos humanos del desarrollo de software basados en las interacciones registradas en las plataformas software forges. b) Medir y evaluar indicadores de calidad y éxito de los procesos de desarrollo de software basados en las interacciones registradas en las plataformas software forges. c) Relacionar ambas mediciones realizadas.

3. RESULTADOS OBTENIDOS

Se desarrolló una aplicación web que, en base a datos extraídos automáticamente desde un repositorio de software, respecto de un proyecto de desarrollo de software, brinda información sobre los aspectos humanos del equipo de desarrollo.

Específicamente la aplicación genera información analítica, en la forma de gráficos, en base a datos extraídos de la plataforma Github con el fin de poder medir, visualizar y analizar las interacciones, y de allí relaciones humanas, que se producen entre los participantes de equipos de trabajo distribuidos.

3.1. Arquitectura de la aplicación

La aplicación se implementa como servicio web. En la Figura 1 se puede observar la arquitectura de software diseñada. Para su desarrollo se utilizó el framework Vue.js desplegado en la plataforma Netlify. Vue es un framework de javascript progresivo que permite desarrollar de manera rápida e incremental. Netlify es una plataforma que ofrece servicios de alojamiento y backend sin servidor (serverless) para sitios web estáticos.

Para hacer el relevamiento de información sobre los usuarios y sus actividades dentro del proyecto de software alojado en el repositorio de Github se utilizó la API v4 del mismo. Esta versión de la API utiliza GraphQL, el cual es un lenguaje de consulta para API y un motor en tiempo de ejecución para cubrir esas consultas con los datos existentes. Tanto las consultas como la respuesta de parte de la API usan el formato JSON.



Figura 1: Arquitectura de la aplicación

A los datos obtenidos se les aplicó un algoritmo propio de conteo que genera una matriz de interacciones entre los miembros de un PR. Se considera interacción de un miembro p1 hacia otro p2 cuando hay un

comentario o reacción (emoji en respuesta a un comentario) dirigido de p1 a p2 o de p1 a todos.

3.2. Métricas para la Medición de Interacciones

Con la matriz de interacciones generada es posible aplicar técnicas de extracción de evidencia. Utilizamos las fórmulas expuestas en la Tabla 1 para el cálculo de aspectos sociales.

Tabla 1. Métricas de Aspectos Sociales

Métrica	Fórmula
Cohesión Interpersonal entre el miembro i del equipo y el miembro j . (M_{ij} mensajes del miembro i al miembro j) [8].	$CINT_{ij} = 1 - \frac{ M_{ij} - M_{ji} }{M_{ij} + M_{ji}}$
Cohesión Individual del miembro i del equipo. (G conjunto de miembros del grupo) [8].	$CIND_i = \frac{\sum_{j=1}^{ G -1} CINT_{ij}}{ G - 1}$
Cohesión grupal del equipo e . (G conjunto de miembros del grupo) [8].	$CGR_e = \frac{\sum_{i=1}^{ G } CIND_i}{ G }$
Mímica o Similitud de vocabulario entre los miembros i y j del equipo. SC es la función coseno de similitud, W_i conjunto de palabras usadas por miembro i y W_j conjunto de palabras usadas por miembro j [9,10].	$CMM_{ij} = SC(W_i, W_j)$
Polaridad de los comentarios entre dos miembros. Donde C_{+ij} es la cantidad de comentarios con polaridad positiva de i a j , y C_{ij} es el total de comentarios de i a j (comentarios que i escribió en PR de j o comentarios donde i mencionó a j) [10].	$P_{ij} = \frac{C_{+ij}}{C_{ij}}$
Capacidad de un miembro donde $PRa(i)$ es la cantidad de PRs creados por i que hicieron merge, y $PR(i)$ es la cantidad de PRs creados por i [10].	$M_i = \frac{PRa(i)}{PR(i)}$
Colaboración entre los miembros donde $I(mi, mj)$ es el número de interacciones entre i y j , y $I(mi)$ es el total de interacciones de i y j con los miembros del PR [10].	$CL_{ij} = \frac{I(mi, mj)}{I(mi)}$
Colaboración Individual del miembro i del equipo. (G conjunto de miembros del grupo) [10].	$CL_i = \frac{\sum_{j=1}^{ G -1} CL_{ij}}{ G - 1}$
Colaboración grupal del equipo e . (G conjunto de miembros del grupo) [10].	$CLG_e = \frac{\sum_{i=1}^{ G } CL_i}{ G }$

3.3. Funcionalidades principales

La aplicación emplea un conjunto de cálculos que implementan las métricas mencionadas en la sección anterior, ofrece dos módulos

funcionales que reportan datos estadísticos, uno a nivel de PRs y el otro a nivel de Proyecto.

3.3.1. Módulo de métricas de Pull Requests

Este módulo muestra información sobre las interacciones sociales entre los miembros de un determinado PR. Por una parte se muestra un resumen de las estadísticas grupales, ver Figura 2. En ella se informa, en una escala porcentual, el grado de cohesión, de colaboración, de mímica y de polaridad alcanzado por el grupo de miembros del PR. Estas estadísticas grupales se calculan promediando los resultados de las métricas individuales de los miembros del PR.

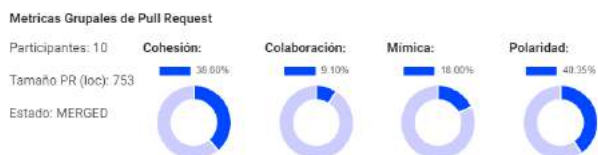


Figura 2: Gráficos Métricas Grupales en un PR

En el ejemplo mostrado en la Figura 2 se puede observar un bajo nivel de colaboración en el grupo lo cual puede ser una oportunidad para tomar decisiones al respecto por parte del líder del grupo.

También se informan, en una escala porcentual, métricas individuales como nivel de cohesión y colaboración de cada miembro del PR, ver Figura 3. La cohesión individual nos indica cuán cohesionado está cada miembro respecto del grupo de trabajo, mientras que la colaboración nos señala el nivel de actividad colaborativa individual.

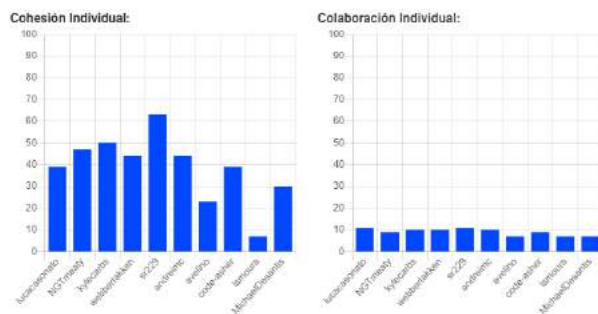


Figura 3: Gráficos de Cohesión y Colaboración Individual en un PR.

En el ejemplo presentado en la Figura 3 se puede notar que el miembro ismoura tiene bajos niveles de cohesión y colaboración, ésta es otra posible oportunidad de toma de decisiones para los administradores del proyecto.

Por último, se muestra una tabla con la cantidad de mensajes enviados, mensajes recibidos, mímica y polaridad individual de cada miembro, ver Figura 4.

Participante	Mjg. Enviados	Mjg. Recibidos	Mimica	Polaridad
lucacasonato	71	24	0.30	0.83
NGTinsaty	20	19	0.20	0.00
lyfeyarbo	22	23	0.19	0.50
weberlakken	12	22	0.26	1.00
sr229	18	21	0.12	1.00
andremc	39	18	0.25	0.27

Figura 4: Tabla de mensajes enviados y recibidos, Mímica y Polaridad en comentarios de un PR.

3.3.2. Módulo de métricas de proyecto

Este módulo reporta datos de todos los PR de un proyecto de software determinado. En primer lugar muestra una tabla donde cada fila es un PR del proyecto, y las columnas indican los resultados de las métricas de interacción grupales (cohesión, colaboración, mímica, polaridad), junto con datos extraídos de Github sobre el PR. En este caso los líderes de proyecto pueden observar los aspectos sociales en una forma comparativa entre los distintos PRs. De esa manera podrían detectar la evolución de dichos aspectos en el tiempo.

También se puede ver una tabla donde cada fila es un miembro que participó en algún PR del proyecto y las columnas indican los resultados de las métricas de interacción individuales (cohesión individual, colaboración individual, polaridad, habilidad), junto con datos extraídos de Github sobre el miembro. En la Figura 5 se puede observar un ejemplo de los datos mencionados.

Miembro	Cohesión Individual	Colaboración Individual	Capacidad	Cant. PE Autor	Polaridad	Cant. PE Participa	Rol
Noel2	0.23	0.55	1.00	2	0.32	5	contribuyente
Victorescco	0.13	0.22	0.60	34	0.34	59	contribuyente
andrei	0.60	0.51	0.50	6	0.45	78	miembro
NOTimeaty	0.42	0.60	0.70	12	0.25	16	contribuyente
vapur	0.30	0.28	1.00	1	0.20	1	contribuyente
CodeMar	0.66	0.37	0.88	87	0.68	134	miembro

Figura 5: Estadísticas de Miembros en un Proyecto de Software.

El módulo también muestra un resumen de las estadísticas grupales del proyecto, cada una de las métricas mostradas en el resumen se calcula promediando los valores obtenidos de la tabla anterior. En la Figura 6 se puede observar un ejemplo de los datos mencionados.



Figura 6: Gráficos Métricas Grupales en un Proyecto de Software.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo está formado por los autores del trabajo y profesores colaboradores de las universidades del Quindío (Colombia), ORT (Uruguay) y de la Frontera (Chile).

Con la presente línea de Investigación y Desarrollo se relacionan dos tesis de Doctorado, una en desarrollo y la otra en sus inicios, más una tesina de grado muy próxima a ser presentada.

5. BIBLIOGRAFIA

[1] Murgia, A., Tourani, P., Adams, B., & Ortu, M. (2014, May). Do developers feel emotions? an exploratory analysis of emotions in software artifacts. In Proceedings of the 11th working conference on mining software repositories (pp. 262-271). ACM

[2] Carmel, E. (1999). Global software teams: collaborating across borders and time zones. Prentice Hall PTR.

[3] Carron, A. V., Widmeyer, W. N., and Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire, *Journal of Sport Psychology* 7 244-266.

[4] Chidambaram, L. (1996). Relational development in computer-supported groups, *MIS Quarterly* 20(2) 143-165.

[5] Carless, S. A., & De Paola, C. (2000). The measurement of cohesion in work teams. *Small group research*, 31(1), 71-88.

[6] Squire, M., & Williams, D. (2012, January). Describing the software forge ecosystem. In 2012 45th Hawaii International Conference on System Sciences (pp. 3416-3425). IEEE.

[7] Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012, February). Social coding in GitHub: transparency and collaboration in an open software repository. In Proceedings of the ACM 2012 conference on computer supported cooperative work (pp. 1277-1286).

[8] Gallardo, F., Sevilla, G., Torres, E. y Zapata, S. (2019). Análisis de Datos de GitHub para Evaluar Cohesión en Equipos Virtuales de Desarrollo de Software. Instituto de Informática – Universidad Nacional de San Juan.

[9] Vincent (MSVCode). (2018). <https://github.com/MSVCode/doc-similarity>.

[10] Da Cruz, G. A. M., Moriya-Huzita, E. H., & Feltrim, V. D. (2018). ARSENAL-GSD: A framework for trust estimation in virtual teams based on sentiment analysis. *Information and Software Technology*, 95, 46-61.

Modelo de Diseño Instruccional en e-Learning

Carlos E. Vera¹, Iván Balmaceda Castro¹, Mariela Fernandez¹, Sandra Rodriguez¹

¹Centro de Investigación y Desarrollo Informático
Departamento de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de La Rioja, La Rioja, Argentina.
carloseduardonvn@gmail.com, {ibalmaceda, mfernandez, sirodriguez}@unlar.edu.ar

RESUMEN

En la actualidad, la exigencia en materia educativa, en la necesidad de una formación de calidad en respuesta a los requerimientos de los estudiantes, y con la mediación tecnológica, que buscan que sean flexible en tiempo y espacio, es pertinente hablar de entorno virtual de aprendizaje que integre la totalidad de recursos, contenidos y herramientas que contribuyan al proceso de enseñanza aprendizaje y que permitan el desarrollo de competencias. En ello deviene, el diseño y desarrollo de estos entornos que se centra en potenciar la autogestión del proceso de aprendizaje de los alumnos y promover el aprendizaje autónomo, autorregulado y colectivo [1-2]. De esta manera, el Diseño por Instrucciones, que conjuga la dimensión tecnológica y pedagógica, permite implementar estrategias didácticas orientadas a aprovechar las dos dimensiones, y de esta manera lograr una buena experiencia del usuario en el proceso.

Palabras clave: Diseño Instruccional, Experiencia de Usuario, UX, Educación

CONTEXTO

El trabajo se enmarca en el proyecto de investigación “*Modelo de Proceso y Evaluación Centrado en el Usuario incorporando requisitos de Usabilidad y Experiencia de Usuarios*” - Centro de Investigación y Desarrollo Informático, Departamento Académico de Ciencias Exactas, Físicas y Naturales - Sede Regional Chamental y Delegación Académica Tama de la

Universidad Nacional de La Rioja. Cod. Proy. 27PIN/C0003, financiado por Consejo de Investigaciones Científicas y Tecnológicas UNLaR Res N°052/18.

1. INTRODUCCIÓN

Muchas de las instituciones, no solo educativas, sino también consultores, pymes, startups, corporativos, en los últimos años, han ido adoptando una modalidad 100% en línea, conocida como e-learning, lo que ha abierto grandes retos y oportunidades, no tan sólo pedagógicas, sino también tecnológicas en el sector educativo. Este modelo de enseñanza genera interacción entre el estudiante-profesor con el material, a través de la utilización de diversas herramientas informáticas.

El e-Learning, con el transcurso de los años, ha presentado nuevos desafíos, los diseños han debido adaptarse a nuevos requerimientos sociales, a nuevas necesidades educativas y a soportes tecnológicos que han terminado por romper el tiempo y la distancia para la formación. [3]

La experiencia de usuario (UX) en e-learning ha tomado un gran protagonismo para conseguir el éxito de cualquier contenido online. Entonces, UX en e-learning es un proceso que organiza y estructura los recursos adaptando a necesidades del estudiante (usuario) y, por sobre todo, facilitando el proceso de enseñanza-aprendizaje. [4] En ello, plantea conocer a los usuarios, pero por sobre todo entenderlos, es decir, saber: quiénes son, qué necesitan, cuál es su contexto.

El Diseño Instruccional o ID, sirve de referencia para producir una variedad de materiales educativos, a través de una metodología de planificación pedagógica [5]. Asimismo, genera “experiencias de instrucción que hacen la adquisición de conocimientos y habilidades más eficiente, eficaz y atractiva.”[6]

La forma en que el estudiante, va interaccionando con el contenido y la vivencia que se produce en este proceso conforman la experiencia de usuario. Entonces, el ID es la “planificación instruccional sistemática que incluye la valoración de necesidades, el desarrollo, la evaluación, la implementación y el mantenimiento de materiales y programas” [7].

La ISO 9241 define a la usabilidad de una interfaz, como “la medida de la efectividad, eficiencia y satisfacción con la cual determinados usuarios pueden alcanzar determinados objetivos en un entorno particular con dicha interfaz” [8]. También, se considera de gran importancia la usabilidad pedagógica de la plataforma de aprendizaje, logrando una adecuada interrelación entre la tecnología y el aprendizaje, esta es una articulación entre la didáctica y la tecnología, que está referida a la infraestructura tecnológica, los materiales de estudio, los recursos y herramientas que se utilicen [9], es por ello que la interfaz de usuario y la usabilidad pedagógica, es de gran importancia para que la misma sea fácil de usar en la interacción entre profesores y estudiantes.

El Diseño Instruccional es de suma importancia para el e-learning, ya que de manera principal participa en el desarrollo, selección o adaptación de contenidos que tengan una calidad académica y que resulten adecuados para el desempeño de la formación en cuestión.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Acorde a los trabajos presentados en WICC 2019 [10], CICCSI 2019 [11], WICC 2020 [12], Jornada de Ciencia y Tecnología UTN [13] se continúa trabajando en la definición de distintos modelos y métodos de evaluación de calidad concernientes a la usabilidad, UX y accesibilidad.

El objetivo planteado en el proyecto "Modelo de Proceso y Evaluación Centrado en el Usuario incorporando requisitos de Usabilidad y Experiencia de Usuario" perteneciente a la Universidad Nacional de La Rioja, donde los ejes principales de la investigación se centran en el análisis de las pautas de Ingeniería de Usabilidad e Ingeniería de Software, análisis y definición de requerimientos funcionales y no funcionales basados en dichas pautas, definición e integración de un modelo que incluya los requisitos de experiencia de usuario, accesibilidad y usabilidad, como la definición de métodos que permitan evaluar dicho modelo en el dominio de educación.

3. RESULTADOS ESPERADOS/OBTENIDOS

El objetivo general de la presente línea de investigación consiste en proponer un modelo de diseño instruccional para e-learning, basado en los sistemas de desarrollos responsivos con una funcionalidad efectiva, para facilitar el aprendizaje del alumno mejorando la experiencia de usuario. En este contexto, se plantean los siguientes objetivos particulares:

- Definir el estado del arte.
- Identificar modelos del proceso de diseño que describen los diferentes pasos que deben darse para crear un ambiente de aprendizaje.
- Reconocer fortalezas y debilidades de los modelos analizados.
- Proponer un modelo que cumpla con los estándares de diseño internacional para plataformas digitales.
- Validación del modelo propuesto aplicado a casos de estudio de sistemas

e-Learning, como, por ejemplo, el Campus Virtual de la UNLAR.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación se desarrolla en el Centro de Investigación y Desarrollo Informático de la Universidad Nacional de La Rioja con docentes, graduados y estudiantes de la Sede Regional Chamical y Delegación Académica Tama.

En dicha línea, se está trabajando en el Proyecto de investigación en el marco de adscripción docente del Lic. Carlos E. Vera.

Como así también se están llevando a cabo algunos trabajos de grado para la Licenciatura en Sistemas de Información y la Tecnicatura en Informática carreras pertenecientes al Departamento Académico de Ciencias Exactas, Físicas y Naturales Universidad Nacional de La Rioja.

5. BIBLIOGRAFÍA

1. Trends, R. E. (2014). Aprendizaje invertido. Observatorio de Innovación Educativa. Tecnológico de Monterrey, 1-29.
2. Raposo, M. (2013). Orientaciones pedagógicas para los MOOC. Recuperado de http://gtea.uma.es/congresos/wp-content/uploads/2013/12/Texto_Congreso MRaposo-def.pdf
3. Vélez Morales, B. (2017). Diseño estratégico de plataformas digitales E-Learning (Master's thesis, Universidad Autónoma del Estado de México).
4. Balmaceda Castro, I., Rusu, C., & Aciar, S. (2020, July). Customer eXperience in e-Learning: A Systematic Mapping Study. In *International Conference on Human-Computer Interaction* (pp. 158-170). Springer, Cham.
5. Reigeluth, C. (1983) *Intructional-design Theoris and Model: An overview of their Current status. Hillsdales,N.J: Lawrence Erlbaum.*
6. Merrill, M. D.; Drake, L.; Lacy, M. J.; Pratt, J. (1996). «Reclaiming instructional design». *Educational Technology* 36 (5): 5-7.
7. Richey, R. C., Fields, D. C., & Foxon, M. (2001). *Instructional design competencies: The standards.* ERIC Clearinghouse on Information & Technology, Syracuse University, 621 Skytop Rd., Suite 160, Syracuse, NY 13244-5290.
8. ISO 9241, Ergonomic requirements for office work with visual finishes part 11-1998
9. Aguilar, B. L. C., Escobar, M. A. D. O., & Uscanga, M. E. E. S. La usabilidad pedagógica en los entornos virtuales de aprendizaje.
10. Balmaceda Castro, I., Salgado, C. H., Peralta, M., Sánchez, A., Fernández, M., Magaquian, J., & Fuentes, N. “Experiencia de usuario en plataforma virtual de aprendizaje.” XXI Workshop de Investigadores en Ciencias de la Computación, San Juan 2019.
11. Balmaceda Castro, I., Fuentes, L., Magaquian, J., Fernandez, M., Salgado, C., Rodriguez, S., & Gimenez, M. (2019, November). Modelo de Proceso incorporando Requisitos de Usabilidad y Experiencia de Usuario en el Desarrollo de Software. In *CICCSI* 2019.
12. Balmaceda Castro, I., Salgado, C. H., Peralta, M., & Sanchez, A. (2020). Heurística de Evaluación de la Experiencia de Usuario en Sistemas e-Learning. In *XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz).*
13. Vera, C., Fernandez, M., Rodriguez, S., & Castro, I. B. DISEÑO CENTRADO EN USUARIO: INCORPORANDO REQUISITOS Y ASPECTOS DE USABILIDAD. *Jornadas de Ciencia y Tecnología 2020 “50 aniversario”*, 169.

MODELO PARA CURADURÍA DE PROYECTOS SOFTWARE DE FUENTE ABIERTA PARA ESTUDIOS EMPÍRICOS EN INGENIERÍA DE SOFTWARE

Emanuel Irrazabal; Juan Andrés Carruthers; Juan Alberto Pinto Oppido

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
{eirrazabal, jacarruthers}@exa.unne.edu.ar, juan_al_pinto@hotmail.com

RESUMEN

Este es el cuarto año del proyecto F018-2017; una continuación de los proyectos F07-2009 y F10-2013, ambos enfocados en modelos, métodos y herramientas para la calidad del software. Este proyecto se enfoca en el estudio de la calidad del Software desde el punto de vista del código fuente por medio de métricas e indicadores cuantificables en este. Esta línea de investigación financiada por CONICET a través de una beca interna doctoral, busca generar métodos empíricos que contribuyan a identificar aquellos factores o atributos vinculados directamente con la calidad del producto software.

En particular, se está trabajando la construcción de un catálogo de proyectos Software de calidad. Esto ha incluido el desarrollo de un mapeo sistemático para reconocer las características de los proyectos utilizados en estudios. Se atiende la necesidad de los grupos de investigación de obtener muestras curadas de proyectos imprescindibles para la generación de resultados confiables y generalizables en la experimentación de estudios empíricos de la calidad de Software, proporcionando los insumos y procedimientos necesarios para conseguirlo de manera efectiva.

Palabras clave: calidad de software, proyecto fuente abierta, catálogo de proyectos.

CONTEXTO

Las líneas de Investigación y Desarrollo presentada en este trabajo corresponden al proyecto PI-17F018 “Metodologías y herramientas emergentes para contribuir con la calidad del software”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021, y a la beca interna doctoral de CONICET otorgada por RESOL-2021-154-APN-DIR#CONICET para el período 2021-2025.

1. INTRODUCCIÓN

En la Ingeniería de Software se trabaja mayoritariamente con la construcción de aplicaciones software multi-versión [1]. Por lo tanto, muchas de las actividades asociadas con una aplicación software provocan revisiones para mejorar la funcionalidad o para corregir errores, especialmente en las metodologías ágiles [2].

En el desarrollo software, la calidad puede estudiarse desde el punto de vista de: i) la calidad del proceso de desarrollo software, y ii) la calidad del código fuente [3]. En este último caso es necesario obtener métodos empíricos para demostrar la calidad del software [4] y utilizar evidencia directamente relacionada con el producto software resultante a partir de métricas e indicadores que se vinculen directamente con la calidad [5]. Sin embargo, esto no siempre es sencillo de realizar, ya que las conclusiones generales de los estudios

empíricos en Ingeniería de Software a menudo dependen de una gran cantidad de variables [6].

El uso masivo de repositorios para el código fuente (por ej., SourceForge, GitHub o Maven) le ha otorgado a los investigadores e ingenieros de software el acceso a millones de proyectos y, por lo tanto, datos para el desarrollo de estudios empíricos [7] [8] [9]. No obstante, la proporción de ruido en una muestra aleatoria tomada de repositorios podría sesgar el estudio, y puede llevar a los investigadores a conclusiones poco realistas, potencialmente inexactas [10].

Esto se contrapone con la condición de reproductibilidad y generalidad de los resultados empíricos, tal y como lo indica el énfasis actual en la Ingeniería de Software basada en evidencia. En particular, la reproductibilidad es una condición necesaria, no solo en publicaciones en revistas o conferencias de prestigio de la disciplina, sino también para las empresas de desarrollo de software que a menudo desean analizar la evolución de sus propios proyectos o como patrón comparativo en auditorías de software. En este contexto, una práctica para demostrar la efectividad de las métricas como predictores de las características de calidad del software es la construcción de los denominados corpus o catálogos de proyectos [11].

Los catálogos de proyectos son un insumo para los grupos de trabajo y sirven como mecanismo de comparación para distintos tipos de experimentos. Existen varios ejemplos en la literatura, como los realizados por Barone y R. Sennrich [12], Allamanis y Sutton [13] o Keivanloo [14] y se diferencian por la cantidad, calidad de proyectos que lo componen; como también los criterios y métodos para agruparlos.

Un catálogo popular en Ingeniería de Software es el Qualitas corpus [11], cuya última versión es de 2013. En general, se observa que no existe información actualizada en la mayoría de los catálogos, lo que hace necesario revisar los proyectos

y sus versiones, y generar nuevas métricas sobre ellos. Una alternativa interesante para ello es la utilización de herramientas de código abierto, que han demostrado ser de gran utilidad en entornos de trabajo profesionales, tanto tradicionales [15] como aplicadas a nuevas técnicas de desarrollo [16].

Por lo tanto, la línea de trabajo del proyecto tiene que ver con las características de calidad del código fuente del software.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se describe la línea de investigación y desarrollo:

Curaduría de proyectos

Se están estudiando los aspectos metodológicos y criterios considerados por la comunidad científica para conformar los catálogos de proyectos. Para esta etapa de documentación se utilizarán dos métodos de investigación. Por un lado, el método de revisiones sistemáticas [17] para identificar, evaluar, e interpretar toda la información relativa a un tema de investigación en particular, de un modo sistemático y replicable. Como segundo método de documentación se utilizarán las encuestas [18], para recopilar información de los grupos de investigación y los equipos de trabajo de las empresas privadas.

El siguiente paso consiste en la creación de un modelo de procedimientos para la construcción, mantenimiento y curaduría de un cuerpo de proyectos software y sus métricas de calidad de producto. Teniendo como fin proveer una estrategia para la construcción de estudios empíricos en Ingeniería del Software que brinde una mejor reproducibilidad, consistencia experimental, y flexibilidad para una evolución gestionada del cuerpo de proyectos en el tiempo. Y, finalmente, se

propone implementar el modelo en un ambiente real de trabajo.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el marco de este proyecto y respecto de la línea de curaduría de proyectos se lograron los siguientes objetivos:

1. Relevamiento de información a considerar en la selección de proyectos Software y metadatos utilizados en el campo de estudio.
En primer lugar, se realizó un mapeo sistemático (SMS) con el objetivo de identificar las características de los proyectos Software, tipos de metadatos consumidos, distintas herramientas usadas para recolectar los metadatos y análisis posteriores realizados a los mismos tenidos en cuenta por la literatura de la ingeniería del Software.

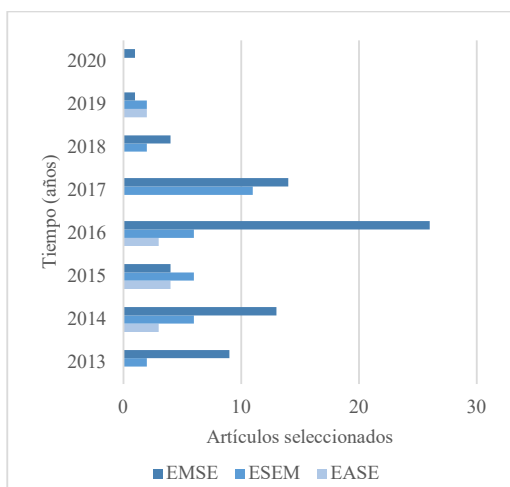


Fig. 1: Distribución de los artículos seleccionados de cada revista por año.

Las fuentes seleccionadas fueron la revista Empirical Software Engineering (ESE) y las conferencias Empirical Software Engineering and Measurement (ESEM) e International Conference on Evaluation and Assessment in Software Engineering (EASE) por ser representativas del área de investigación y haber sido

utilizadas en otros estudios. En Fig. I se puede observar la distribución de artículos recolectados por revista. El estudio fue realizado en el marco de una beca de pregrado de la UNNE.

2. Tesis de grado y desarrollo de herramienta para extracción de metadatos de plataforma de análisis estático de código fuente.



Fig. II: Página de Inicio SET.

Se construyó la aplicación web “Sonar Exporting Tool” (SET) para extraer las métricas de la plataforma Sonar Cloud (SC) en formatos de datos consumibles y difundir el código fuente de los proyectos Software analizados. Todos los proyectos seleccionados forman parte del Qualitas Corpus [11]. Las Fig. II y Fig. III son capturas de la aplicación.

SET permite exportar medidas de métricas de SC en los formatos de datos csv, json y xml; la actualización automática y manual de la base de datos. También se creó un procedimiento para el análisis automatizado de colecciones de proyectos por medio del cliente SonarScanner.

Además, se busca mantener actualizado cada uno de los proyectos dentro de la aplicación en su versión estable más reciente.

Projects Hosted in Sonar Cloud				
Key	Name	Last Analysis	Version	
unine:sonar:spotbugs	Spotbugs	12/8/2020, 7:30:99 PM	4.2.0	<input type="checkbox"/>
unine:sonar:apache-ant	Apache Ant	12/8/2020, 7:26:41 PM	1.10.9	<input type="checkbox"/>
unine:sonar:antigo	Xmogo	9/4/2019, 5:38:41 AM	5.0.0	<input type="checkbox"/>
unine:sonar:xerces	Xerces	9/4/2019, 5:28:34 AM	2.10.0	<input type="checkbox"/>
unine:sonar:xaian	Xaian	9/4/2019, 5:17:27 AM	2.7.1	<input type="checkbox"/>
unine:sonar:weka	Weka	9/4/2019, 4:56:27 AM	3.7.9	<input type="checkbox"/>
unine:sonar:webmail	Webmail	9/4/2019, 4:51:32 AM	0.7.10	<input type="checkbox"/>
unine:sonar:wict	WICT	9/4/2019, 4:36:02 AM	1.5.2	<input type="checkbox"/>
unine:sonar:velocity	Velocity	9/4/2019, 4:29:48 AM	1.6.4	<input type="checkbox"/>
unine:sonar:trove	Trove	9/4/2019, 4:24:48 AM	2.1.0	<input type="checkbox"/>
unine:sonar:tomcat	Tomcat	9/4/2019, 4:09:43 AM	7.0.2	<input type="checkbox"/>
unine:sonar:tapestry	Tapestry	9/4/2019, 3:55:56 AM	5.1.0.5	<input type="checkbox"/>
unine:sonar:sunflow	Sunflow	9/4/2019, 3:46:10 AM	0.07.2	<input type="checkbox"/>

Fig. III: Listado de proyectos en SET.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo del Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 3 docentes investigadores, un becario interno doctoral de CONICET y un becario de investigación de pregrado.

5. REFERENCIAS

- [1] Parnas, D. L. (2001). Some software engineering principles. Hoffman, D. M. y Weiss, D. M. (Ed), Software fundamentals: collected papers by David L. Parnas (pp. 257–266). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
- [2] Irrazabal, E., Vásquez, F., Díaz, R., & Garzás, J. (2001) Applying ISO/IEC 12207: 2008 with SCRUM and Agile methods. In International Conference on Software Process Improvement and Capability Determination (pp. 169-180). Springer, Berlin, Heidelberg.
- [3] Lehman, M. M. (1996). Laws of software evolution revisited. En Montangero C. (Ed), Laws of Software Evolution Revisited (pp. 108-124). Berlin, Germany: Springer-Verlag.
- [4] Kitchenham, B. y Pfleeger, S. L. (1996). Software quality: the elusive target. IEEE Software, 13(1), 12-21.
- [5] Garvin, D. A. (1984). What Does “Product Quality” Really Mean?. MIT Sloan Management Review, 25-43.
- [6] Basili, V.R., Shull, F., and Lanubile, F. Building knowledge through families of experiments. Software Engineering, IEEE Transactions on, 25 (1999), 456--473.
- [7] Vidal, S, Bergel, A., Marcos, C., Díaz Pace, J. A.: Understanding and addressing exhibitionism in Java empirical research about method accessibility. Empirical Software Engineering 21(2): 483-516 (2016)
- [8] Vidal, S., Bergel, A., Díaz Pace, J. A., Marcos, C.: Over-exposed classes in Java: An empirical study. Comput. Lang. Syst. Struct. 46: 1-19 (2016)
- [9] Vázquez, H., Bergel, A., Vidal, S., Díaz Pace, J. A., Marcos, C.: Slimming javascript applications: An approach for removing unused functions from javascript libraries. Inf. Softw. Technol. 107: 18-29 (2019)
- [10] Munaiah, N.. Curating github for engineered software projects. Empirical Software Engineering, 2017, vol. 22, no 6, p. 3219-3253.
- [11] Tempero, E., Anslow, C., Dietrich, J., Han, T., Li, J., Lumpe, M., Melton, H y Noble, J. (2010). The Qualitas Corpus: A curated collection of Java code for empirical studies. IEEE, 336-345.
- [12] Barone, A. V. M., y Sennrich, R. (2017). A parallel corpus of Python functions and documentation strings for automated code documentation

- and code generation. arXiv preprint arXiv:1707.02275.
- [13] Allamanis, M., y Sutton, C. (2013). Mining source code repositories at massive scale using language modeling. IEEE Press, 207-216.
 - [14] Keivanloo, I., Rilling, J., y Zou, Y. (2014). Spotting working code examples. ACM, 664-675.
 - [15] Irrazábal, E., Garzás, J., & Marcos, E. (2011). Alignment of Open Source Tools with the New ISO 25010 Standard-Focus on Maintainability. In International Conference on Software and Data Technologies (Vol. 2, pp. 111-116).
 - [16] Mascheroni, M. A., & Irrazábal, E. (2018). Continuous testing and solutions for testing problems in continuous delivery: A systematic literature review. *Computación y Sistemas*, 22(3), 1009-1038.
 - [17] Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele University*, 33(2004), 1-26.
 - [18] Kitchenham, Barbara, and Shari Lawrence Pfleeger. "Principles of survey research." *ACM SIGSOFT Software Engineering Notes* 27.5 (2002): 17-20.

Modelo Semántico de Repositorio de Procesos de Negocio para la Gestión de Procesos de Negocio Colaborativos

Juan Pablo Ferreyra*, Marisa Pérez*, Diego Cocconi*, Claudia Verino*

Departamento de Ingeniería en Sistemas de Información / Facultad Regional San Francisco / Universidad Tecnológica Nacional (UTN)
Av. de la Universidad 501, San Francisco (2400), Córdoba, Argentina, (03564) 431019 / 435402
*{jpferreyra, mperez, dcocconi, cverino}@sanfrancisco.utn.edu.ar

RESUMEN

La colaboración cada vez más frecuente entre organizaciones para combinar fuerzas en la generación de productos y servicios en escenarios tan dinámicos y cambiantes como los actuales genera nuevas necesidades de flexibilidad y tiempos de respuesta. Las nuevas tendencias organizacionales y las tecnologías de información pueden ser aprovechadas para establecer relaciones de colaboración entre diferentes organizaciones para obtener beneficios comunes. De este modo, es posible conformar redes colaborativas que posibiliten llevar a cabo procesos de negocio inter-organizacionales. La colaboración es posible a través de la ejecución de procesos de negocio colaborativos (CBP, del inglés Collaborative Business Processes). La gestión de repositorios de modelos de procesos en colaboraciones inter-organizacionales es una funcionalidad básica requerida en redes colaborativas de organizaciones. Si bien existen propuestas que permiten gestionar repositorios de modelos de procesos de negocio, las mismas presentan distintas deficiencias, como la ausencia de soporte para CBP o la imposibilidad de garantizar la interoperabilidad. El problema de reunir sistemas de información heterogéneos y distribuidos se conoce como problema de interoperabilidad. Para garantizar la interoperabilidad en el intercambio de información en una colaboración inter-

organizacional (heterogénea) es necesario el uso de modelos semánticos basados en ontologías. En el presente proyecto se propone un modelo basado en ontologías y tecnologías semánticas para la gestión de procesos de negocio inter-organizacionales, que garantice interoperabilidad, además de ofrecer soporte para búsquedas enriquecidas y gestionar colaboraciones mediante un protocolo de negociación.

Palabras clave: gestión de procesos de negocio, procesos de negocio inter-organizacionales, redes colaborativas, repositorios de procesos de negocio, sistemas de información.

CONTEXTO

El presente trabajo se encuadra dentro del proyecto de investigación I+D UTN 8218 “Modelo Semántico de Repositorio de Procesos de Negocio para la Gestión de Procesos de Negocios Colaborativos”. El mismo se encuentra homologado como proyecto de investigación y desarrollo por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional (UTN).

El objetivo general de dicho proyecto es proponer un repositorio de modelos de procesos de negocio semántico, basado en ontologías, que permita contribuir a la gestión de colaboraciones en entornos inter-

organizacionales, dando solución a los problemas de interoperabilidad que se originan en el intercambio de información en entornos heterogéneos.

1. INTRODUCCIÓN

La colaboración cada vez más frecuente entre organizaciones para combinar fuerzas en la generación de productos y servicios en escenarios tan dinámicos y cambiantes como los actuales genera nuevas necesidades de flexibilidad y tiempos de respuesta [1].

Las organizaciones tienden a establecer relaciones de integración, cooperación y colaboración, lo que resulta en nuevas formas de redes colaborativas [2][3]. Una red colaborativa consiste en organizaciones autónomas, geográficamente distribuidas y heterogéneas que colaboran para lograr objetivos comunes [4] [3].

La colaboración es posible a través de la ejecución de *procesos de negocio colaborativos* (CBP, del inglés *Collaborative Business Processes*) o *coreografías de proceso* [5], que abarcan a todas las organizaciones de una cadena de valor colaborativa [6]. Un CBP es un proceso abstracto, no ejecutable directamente [7], que permite definir el comportamiento de las interacciones entre las organizaciones, esto es, cómo coordinan sus acciones e intercambian documentos de negocio, con el propósito de tomar decisiones en forma conjunta para alcanzar metas en común [8] [9].

Para implementar y ejecutar un proceso colaborativo en forma descentralizada, se requiere que cada organización defina y gestione sus procesos de negocio internos, denominados *procesos de interfaz* (públicos) y *procesos de integración* (privados) [10]. Un proceso de interfaz define el comportamiento público y externamente visible de una organización, el cual es expresado en términos de las actividades que soportan el envío y la recepción de mensajes con otras organizaciones [10]. Un proceso de integración [11], también llamado *proceso de orquestación* [9] o *proceso público* [5], define y combina el comportamiento y las

actividades públicas (derivadas de un proceso de interfaz) con las actividades privadas que una organización debe ejecutar para dar soporte al rol que ésta desempeña en un proceso colaborativo. Un proceso de integración incorpora las actividades privadas, las cuales permiten generar y procesar la información intercambiada entre las organizaciones, realizar la transformación de datos e invocar a sistemas de información internos [10].

La gestión de repositorios de modelos de procesos en colaboraciones inter-organizacionales es una funcionalidad básica requerida en redes colaborativas de organizaciones [10]. Un repositorio es una base de datos compartida con información sobre artefactos creados o utilizados por una organización [12]. Un repositorio de modelos de procesos de negocio es un repositorio especializado, específico para almacenar y administrar modelos de procesos de negocio [12]. En este tipo de repositorios surgen problemas típicos, como ser, la capacidad de encontrar un proceso particular en una colección, administrar diferentes versiones de los procesos y mantener la coherencia cuando varias personas editan el mismo proceso al mismo tiempo [12]. La disponibilidad de una gran colección de procesos abre nuevas posibilidades, como: extraer el conocimiento sobre las operaciones de la organización de la colección o reutilizar fragmentos de procesos (mejores prácticas) de la colección para diseñar nuevos procesos [12].

El intercambio de información en redes colaborativas presenta una serie de desafíos, uno de ellos está ligado al problema de interoperabilidad. El problema de reunir sistemas de información heterogéneos y distribuidos se conoce como problema de interoperabilidad [13]. Los problemas que pueden surgir debido a la heterogeneidad de los datos son: *heterogeneidad estructural* (heterogeneidad esquemática) y *heterogeneidad semántica* (heterogeneidad de datos) [13]. La heterogeneidad estructural significa que diferentes sistemas de información almacenan sus datos en diferentes estructuras. La heterogeneidad

semántica considera el contenido de un elemento de información y su significado previsto. Para poder lograr la interoperabilidad semántica en una red colaborativa, el significado de la información que se intercambia debe entenderse en todos los sistemas. Los conflictos semánticos ocurren cuando dos contextos no usan la misma interpretación de la información [13]. La interoperabilidad semántica y el intercambio de integración de información son uno de los propósitos principales de la aplicación de las ontologías [14].

Si bien es muy amplia su aplicación en el ámbito de las Ciencias de la Computación, el término ontología se puede utilizar para referir a un artefacto que representa la semántica de un dominio dado [14]. Por lo tanto, para garantizar la interoperabilidad en el intercambio de información en una colaboración inter-organizacional (heterogénea) es necesario el uso de modelos semánticos basados en ontologías.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo se encuadra dentro de una línea de investigación que abarca las siguientes áreas temáticas:

- Gestión de procesos de negocio.
- Procesos de negocio inter-organizacionales.
- Redes colaborativas.
- Repositorios de procesos de negocio.
- Interoperabilidad semántica.

3. RESULTADOS OBTENIDOS Y ESPERADOS

El objetivo general del presente proyecto es proponer un repositorio de modelos de procesos de negocio semántico, basado en ontologías, que permita contribuir a la gestión de colaboraciones en entornos inter-organizacionales, dando solución a los problemas de interoperabilidad que se

originan en el intercambio de información en entornos heterogéneos.

Para poder llegar al objetivo general se proponen los siguientes objetivos específicos:

- Definir un modelo basado en ontologías y tecnologías semánticas para la gestión de procesos de negocio colaborativos, que garantice interoperabilidad.
- Ofrecer una solución que, por medio de búsquedas enriquecidas, permita extraer conocimiento sobre las operaciones llevadas a cabo por las organizaciones en la colaboración.
- Ofrecer una solución que, permita el descubrimiento de procesos, en base a las necesidades de la colaboración, por medio de búsquedas enriquecidas, para favorecer la reutilización de procesos de negocio (mejores prácticas).
- Ofrecer una solución para gestionar la negociación necesaria para llevar a cabo un CBP, donde una organización pueda encontrar (descubrimiento de organizaciones) potenciales candidatos para establecer una colaboración inter-organizacional, y por medio de un protocolo de negociación, permita gestionar acuerdos, ofertas, contraofertas, rechazos y demás términos necesarios para establecerla.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está conformado por docentes, graduados y alumnos de la carrera de Ingeniería en Sistemas de Información. Entre los docentes, tres de ellos se encuentran en la etapa de desarrollo de sus tesis de maestría y otro está preparando su tesis de doctorado (mención Ingeniería en Sistemas de Información), todos ellos con temas altamente vinculados al área de estudio del proyecto.

Como iniciativa del grupo, se pretende generar las siguientes actividades:

- Capacitación y formación de recursos humanos: cursos y seminarios enfocados a estudiantes y docentes de otras materias de la carrera.
- Dirección y asesoramiento sobre el tema a interesados de la industria local (por medio de talleres, cursos, charlas y transferencias).
- Transferencia de conocimiento obtenido a las cátedras del tronco integrador de la carrera Ingeniería en Sistemas de Información.
- Involucrar a alumnos y graduados de la carrera Ingeniería en Sistemas de Información en la realización de actividades del proyecto, incentivándolos a acercarse a propuestas de becas (actualmente el grupo cuenta con dos alumnos en estas condiciones).

5. BIBLIOGRAFÍA

- [1] Ferreyra, J. P., Roa, J., Cocconi, D., Perez, M., Verino, C., Villarreal, P. D. (2017). "Estado actual de la Gestión de Procesos de Negocio basada en Computación en la Nube". *5to. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI), 2017. Santa Fe, Argentina.*
- [2] Chituc, C. M., Azevedo, A., Toscano, C. (2009). "A framework proposal for seamless interoperability in a collaborative networked environment". *Computers in industry*, 60(5), 317-338.
- [3] Cocconi, D., Roa, J., Villarreal, P. (2017, Septiembre). "Cloud-based platform for collaborative business process management". *2017 XLIII Latin American Computer Conference (CLEI)* (pp. 1-10). IEEE.
- [4] Camarinha-Matos, L. M., Afsarmanesh, H., Galeano, N., & Molina, A. (2009). "Collaborative networked organizations—Concepts and practice in manufacturing enterprises". *Computers & Industrial Engineering*, 57(1), 46-60.
- [5] Object Management Group. Business Process Modeling Notation V2.0, OMG, <http://www.omg.org/spec/BPMN/2.0.2/>, 2013.
- [6] Andres, B., Macedo, P., Camarinha-Matos, L. M., & Poler, R. (2014, October). "Achieving coherence between strategies and value systems in collaborative networks". In *Working Conference on Virtual Enterprises* (pp. 261-272). Springer Berlin Heidelberg.
- [7] Lazarte, I.M. Diseño y Gestión de Modelos de Procesos de Negocio en Colaboraciones Interorganizacionales. Tesis Doctoral. Director: Villarreal. Doctorado en Ingeniería en Sistemas de Información (CONEAU A) UTN-FRSF.
- [8] Villarreal, P, Lazarte, I., Roa, J., Chiotti, O. A Modeling Approach for Collaborative Business Processes based on the UP-ColBPIP Language. *Lecture Notes in Business Information Processing (LNBIP)*, vol. 43, pp. 318-329, Springer-Verlag Berlin Heidelberg, 2010.
- [9] Weske, M. (2012). "*Business Process Management. Concepts, Languages, Architectures*". 2nd Edition. Springer-Verlag Berlin Heidelberg.
- [10] Lazarte, I. M., Thom, L. H., Iochpe, C., Chiotti, O., Villarreal, P. D. (2013). "A distributed repository for managing business process models in cross-organizational collaborations". *Computers in Industry*, 64(3), 252-267.
- [11] Lazarte, I.M., Villarreal, P.D., Chiotti, O., Thom, L.M., and Iochpe, C., An MDA-based Method for Designing Integration Process Models in B2B Collaborations. *Proceedings of the 13th Int. Conference on Enterprise Information Systems (ICEIS 2011)*, Beijin, China, pp. 55–65, 2011.
- [12] Z. Yan, R. Dijkman, P. Grefen, Business process model repositories Framework and survey, *Information and Software Technology* 54 (2012) 380–395.
- [13] Wache, Holger & Vögele, Thomas & Visser, Ubbo & Stuckenschmidt, Heiner & Schuster, Gerhard & Neumann, H & Hübner, Sebastian. (2001). *Ontology-based integration*

of information a survey of existing approaches.

[14] Rico, M.G. Soporte para enriquecer la representación de entidades en una ontología. Tesis Doctoral. Santa Fe: UTN - FRSF, 2011.

Reusabilidad en el contexto de Desarrollo de Sistemas para Big Data

Agustina Buccella, Juan Luzuriaga, Alejandra Cechich,
Líam Osycka, Facundo Paterno, Matias Pol'la, Marcos Cruz
Rodolfo Martínez, Rafaela Mazalu, Marcelo Moyano
GIISCO Research Group
Departamento de Ingeniería de Sistemas
Universidad Nacional del Comahue
Neuquen, Argentina
agustina.buccella@fi.uncoma.edu.ar

1. Resumen

Actualmente, el crecimiento de sistemas de Big Data (SBD) está obligando a la comunidad de Ingeniería de Software a replantearse un cambio de paradigma en el desarrollo de estos sistemas. Consecuentemente, en este contexto existen varios desafíos para arquitectos y diseñadores, particularmente sobre los requerimientos que una arquitectura para SBDs debe cumplir; entre ellos, considerar las cinco “Vs” (Volumen, Velocidad, Variedad, Variabilidad y Veracidad). En particular, en nuestra investigación nos centramos en una de estas características, *Variabilidad*, que se refiere a la naturaleza evolutiva de los datos.

Nuestro trabajo se enfoca en incorporar Variabilidad en SBDs a través del modelado de elementos reusables de un dominio – sea este de negocios o tecnológico. Esto nos lleva a incorporar información (y sus posibles usos), a modo de línea de productos software. El presente proyecto tiene como fin desarrollar técnicas y herramientas que mejoren la explotación de grandes volúmenes de datos, favoreciendo el desarrollo de ambientes inteligentes que permitan reusabilidad.

Palabras Clave: Reusabilidad - Líneas de Producto de Software - Big Data

2. Contexto

La línea presentada se inserta en el contexto del *Programa: Desarrollo de Software Basado en Reuso - Parte II (04/F009)*. Directora: Dra. Alejandra Cechich, y *SubProyecto: Reuso Orientado a Dominios - Parte II*. Incluido dentro del Programa. Directora: Dra. Agustina Buccella, Codirector: Mg. Juan Manuel Luzuriaga.

3. Introducción

A pesar de que existen mecanismos para manejar el cambio de esquemas y datos en un modelo relacional [2], alcanzar un nivel similar en Big Data es todavía un desafío importante debido a la naturaleza flexible de los esquemas de almacenes NoSQL [6, 7].

En [3], se presentan seis arquitecturas de referencia¹ propuestas actualmente en la literatura para SBDs y se analiza el cumplimiento de las cinco Vs como requerimientos tradicionales. El análisis se hace de manera transversal; es decir, se analiza cada requerimiento por separado y luego sus interacciones con otros requerimientos. Por ejemplo, en el contexto

¹Una arquitectura de referencia combina el conocimiento sobre arquitecturas en general con la experiencia en requerimientos específicos de una solución en un dominio de problema.

de SBDs se explotan almacenes NoSQL que guardan datos con bajo acoplamiento y flexibles; sin embargo, no satisfacen requerimientos de Variabilidad y Veracidad, por lo que, en la práctica, las aplicaciones resultan en modelos ad hoc. De ahí la necesidad de contar con arquitecturas de referencia que faciliten el desarrollo concreto, conociendo los componentes y sus relaciones previamente para permitir el análisis de propiedades. Del análisis realizado en [3], se desprende que sólo una propuesta (Bolster, [8]) satisface los requerimientos de Variabilidad por medio de (1) el almacenamiento del esquema de información de los elementos incorporados; (2) la existencia de estadísticas descriptivas para acceder a la evolución de los datos; y (3) el almacenamiento de la información sobre fuentes de datos usando un repositorio de metadatos (MetaData Management system).

Por otra parte, en [5] se presenta el desarrollo de una arquitectura para un SBD en un caso específico, haciendo uso de una arquitectura de referencia que define una familia de sistemas relacionados. En particular, como el dominio es demasiado amplio, la arquitectura de referencia se ve acotada por medio de casos de uso (ej. visualización y análisis de información geoespacial estratégica, análisis inteligente de señales, etc.). Esta arquitectura de referencia sirve como mecanismo para capturar y compartir conocimiento, conteniendo tanto conocimiento del dominio (casos de uso) como conocimiento de la solución (la correspondencia a tecnologías concretas).

Considerando estas propuestas, nuestro trabajo se enfoca en incorporar Variabilidad en SBDs a través del modelado de elementos reusables de un dominio – sea este de negocios o tecnológico. Esto nos lleva a incorporar información (y sus posibles usos), a modo de línea de productos software.

En particular, los desarrollos en las líneas de productos de software (LPS) [10] se centran en identificar similitudes y variabilidades dentro de dominios particulares para ser reutilizados cuando se desarrollan nuevos productos. A su vez, este reuso de dominios puede ser extendido a subdominios, en especial cuando entre ellos existen relaciones o aspectos similares.

En la Figura 1 se muestra una primera aproximación de los elementos que componen nuestra propuesta de una arquitectura de referencia para SBDs basada en reuso. Como puede verse, los aspectos de negocios (dominio), aplicación (software y análisis) y tecnológicos se abordan en niveles separados; siendo transversales aspectos como el uso/reuso de estándares, taxonomías y conocimiento.

Los componentes principales de la arquitectura son:

- *Taxonomía de Dominio y Estándares:* Las taxonomías específicas de dominio permiten clasificar elementos de acuerdo a determinados criterios, relaciones y propiedades [4]. Su principal objetivo es capturar el conocimiento del dominio basándose en divisiones de las entidades acordes a lo que se intenta especificar. Estas entidades pueden ser objetos del dominio, servicios e incluso cualquier otro elemento o conjunto de elementos que se desee clasificar. Por lo tanto, la creación de taxonomías debe contribuir a la definición de un vocabulario común y controlado para todos los participantes. Al mismo tiempo, para garantizar interoperabilidad y luego reuso, las taxonomías deben construirse en base a los estándares existentes. Así en este componente se deben considerar los estándares creados para la ingeniería de software, los definidos para Big Data y aquellos definidos para el dominio en que se este trabajando. Por ejemplo, en el caso de Big Data en los últimos años han surgido una serie de estándares respecto a su arquitectura de referencia, interoperabilidad, terminología, etc. Estos esfuerzos de estandarización han sido llevados a cabo por el comité ISO/IEC JTC 1/SC 42 (Artificial intelligence)² y por el grupo de trabajo del NIST (Big Data Public Working Group - NBD-PWG)³. A su vez, la información estandarizada del dominio se refiere a los estándares existentes en el dominio que se este analizando.

²<https://www.iso.org/committee/6794475.html>

³<https://bigdatawg.nist.gov/>



Figura 1. Arquitectura de referencia para SBDs basada en reuso

- *Activos orientados al dominio:* Estos activos agrupan todos aquellos artefactos de software que son creados para el dominio en el que se está trabajando. Así, además de incluir a los participantes del desarrollo del SBD, como los ingenieros de software, científicos de datos, desarrolladores, usuarios expertos, etc., involucra los requerimientos del proyecto y del dominio, restricciones, modelos (artefactos de análisis y diseño generados) y casos de uso. Es importante resaltar que estos activos deben generarse a partir de las *taxonomías de dominio y estándares* y de los *activos basados en conocimiento*. De esta forma, se deben crear artefactos enfocados en que puedan ser reusados en el mismo dominio e incluso en otros dominios relacionados (artefactos para reuso), y/o que puedan desarrollarse en base a otros artefactos ya creados (artefactos con reuso). A su vez aquí es importante contemplar la variabilidad. Es decir, estos activos deben considerar los aspectos comunes y aquellos variables dentro del dominio. Por ejemplo, como describimos previamente, es importante crear activos flexibles que puedan adaptarse a la evolución de los esquemas y datos fuentes para que puedan seguir siendo útiles en los análisis realizados.
 - *Software/Analítica Reusable:* La analítica de datos (data analytics) es un término abarcativo que se encarga de gestionar los datos en todo su ciclo de vida. Como es sabido, los datos sin procesar (raw data) por sí mismos no tienen un significado útil, por lo que la analítica de datos se dedica al proceso de extraer y crear información desde estos datos por medio de la recolección, limpieza, organización, almacenamiento, procesamiento, contextualización, análisis y gobernanza de datos. Existen 4 tipos de categorías para la analítica de datos que dependen de los resultados que producen: descriptiva, diagnóstica, predictiva y prescriptiva. Para cada una de estas categorías existen también diversas técnicas o algoritmos de análisis de datos como clasificación, regresión, clustering, visualización, etc.
- En este componente de la arquitectura, se deben definir y diseñar los tipos de análisis necesarios en el dominio junto con la forma de realizarlos, es decir el diseño de

los algoritmos. Esto no es una tarea sencilla ya que cada uno de ellos requiere un conjunto de datos de entrada específico, es decir formatos y estructuras específicas requeridas por cada uno. Al mismo tiempo, se deben documentar los procesos de diseño de estos algoritmos de forma que puedan ser reusados en diferentes dominios.

- **Tecnología:** La tecnología disponible para crear sistemas de Big Data es muy variada. Considerando el cumplimiento de las 5 V's que se describieron previamente, existen diferentes tecnologías que se abocan en una o varias de ellas. Por ejemplo, para lidiar con los problemas de la Variabilidad, en cuanto a las diferentes fuentes de datos disponibles en sus muy diversos formatos, tenemos un conjunto de herramientas independientes como OpenRefine⁴, Optimus⁵, o lenguajes con librerías específicas como R⁶ o python⁷. También existen varios frameworks especializados en la analítica de los datos como Spark MLib⁸ el cual también forma parte del ecosistema Hadoop. Al mismo tiempo conviven los diferentes tipos de repositorios que permiten almacenar la información extraída desde las fuentes de datos y hacerla disponible para su análisis y visualización. Estos repositorios son muy variados, desde aquellos basados en tecnologías NoSQL como MongoDB⁹ o CouchDB¹⁰, sistemas de archivos distribuidos como HDFS¹¹ (también creado como parte de Hadoop) etc. Finalmente también podemos citar aquellas plataformas y frameworks provistos en la nube que surgen principalmente para lidiar con la escalabilidad. Entre los mas conocidos podemos citar Google Cloud¹², AWS¹³ y

Azure¹⁴.

- **Activos basados en Conocimiento:** Los repositorios para reuso ponen a disposición un amplio rango de activos que los ingenieros de software pueden usar para desarrollar sistemas y así reducir la necesidad de crear nuevamente componentes que provean la misma funcionalidad. Existen diversas alternativas en la elaboración de estos repositorios (basados en componentes, en modelos, etc.). Las funcionalidades asociadas a SBDs son activos potencialmente reusables en el mismo sentido, agregando además la capacidad de reutilización de los datos en sí mismos; lo que hace que los repositorios de experiencias se conviertan en un elemento clave para alcanzar el reuso de activos de dominio. Respondiendo a las cuestiones abiertas planteadas en [9] sobre cómo alcanzar beneficios efectivos al compartir datos, nuestro enfoque intenta identificar objetivos de los posibles usos (ej. interoperabilidad, integración, etc.) así como determinar en qué medida el uso de formatos o estándares facilita el reuso, o cómo distinguir entre datos/usos potencialmente reusables en contexto.

4. Líneas de Investigación y Desarrollo

En investigaciones previas, hemos realizado amplios avances en lo que respecta al área de LPSs definiendo y refinando una metodología de desarrollo a nivel de subdominios. Dentro de la metodología, hemos presentado sus bases y diseñado artefactos que se utilizan en el análisis de dominios y en el análisis organizacional de una LPS [1] y tienen la particularidad de favorecer el reuso basado en una taxonomía de servicios. Es precisamente esta ventaja la que nos permitió luego realizar extensiones hacia otros subdominios. De esta forma hemos podido así avanzar en el desarrollo de múltiples LPSs basadas en la jerarquía de dominios definida.

⁴<https://openrefine.org/>

⁵<https://hi-optimus.com/>

⁶<https://www.r-project.org/>

⁷<https://www.python.org/>

⁸<https://spark.apache.org/mllib/>

⁹<https://www.mongodb.com/>

¹⁰<https://couchdb.apache.org/>

¹¹<https://hadoop.apache.org/>

¹²<https://cloud.google.com/>

¹³<https://aws.amazon.com/>

¹⁴<https://azure.microsoft.com/>

5. Resultados Obtenidos/Esperados

El objetivo principal de la línea de investigación es *desarrollar técnicas y herramientas que mejoren los procesos y técnicas aplicadas a la explotación de grandes volúmenes de datos, favoreciendo el desarrollo de ambientes inteligentes que permitan reusabilidad.*

Previamente, hemos trabajado en el área de Líneas de Productos, donde hemos definido y aplicado nuevos métodos y técnicas para la creación de LPSs con soportes inteligentes dentro del dominio geográfico que contemplan las particularidades de los subdominios incluidos. A su vez, hemos realizado formalizaciones de reglas y patrones para soportar el desarrollo asistido, de manera que sean lo suficientemente generales para ser aplicados en otros subdominios geográficos.

En la presente investigación, se esperan aplicar estos resultados previos para extender e instanciar el modelo presentado en este artículo. Éste se validará en casos de predicción de calidad del agua en golfos de la Patagonia (San Jorge, San Matías) y en ríos (Río Limay, Río Negro), midiendo el grado de extensibilidad y reusabilidad de los modelos.

6. Formación de Recursos Humanos

El proyecto reúne a 13 investigadores, entre los que se cuentan docentes y alumnos de UNComa, y colaboradores. A su vez, cuenta con dos doctores y un magister. Varios de los docentes-investigadores de GIISCo-UNComa han terminado o se encuentran próximos a terminar carreras de postgrado. Uno de ellos se encuentra finalizando su doctorado en el transcurso de este año, en el área Gestión de Variabilidad. A su vez varios de los integrantes se encuentran finalizando sus tesis de grado. Por último, este año seguiremos con la supervisión del trabajo de 2 becarios EVC-CIN.

Referencias

- [1] A. Buccella, A. Cechich, M. Arias, M. Pol'la, S. Doldan, and E. Morsan. Towards systema-

tic software reuse of gis: Insights from a case study. *Computers & Geosciences*, 54(0):9 – 20, 2013.

- [2] Carlo Curino, Hyun Jin Moon, Alin Deutsch, and Carlo Zaniolo. Automating the database schema evolution process. *International Journal on Very Large Data Bases*, 22(1):73–98, 2013.
- [3] Ali Davoudian and Mengchi Liu. Big data systems: A software engineering perspective. *ACM Computing Surveys*, 53(5), 2020.
- [4] I. Hunink, E. Rene, S. Jansen, and S. Brinkkemper. Industry taxonomy engineering: the case of the european software ecosystem. In *Proceedings of the Fourth European Conference on Software Architecture: Companion Volume*, ECSA '10, pages 111–118, New York, NY, USA, 2010. ACM.
- [5] John Klein, Ross Buglak, David Blockhow, Troy Wuttke, and Brenton Cooper. A reference architecture for big data systems in the national security domain. In *Proceedings of the 2nd International Workshop on BIG Data Software Engineering*. ACM/IEEE, 2016.
- [6] Jiaheng Lu and Irena Holubová. Multi-model databases: A new journey to handle the variety of data. *ACM Computing Surveys*, 52(3), 2019.
- [7] Loup Meurice and Anthony Cleve. Supporting schema evolution in schema-less nosql data stores. In *24th International Conference on Software Analysis, Evolution and Reengineering (SANER'17)*, pages 457–461. IEEE, 2017.
- [8] Sergi Nadal and et al. A software reference architecture for semantic-aware big data systems. *Information and Software Technology*, 90:75–92, 2017.
- [9] B.M. Pasquetto, I.V. and Randles and C.L. Borgman. On the reuse of scientific data. *Data Science Journal*, 16, 2017.
- [10] Klaus Pohl, Günter Böckle, and Frank J. van der Linden. *Software Product Line Engineering: Foundations, Principles and Techniques*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

Reuso de Servicios Heterogéneos basado en CBR

Andres Flores, Martin Garriga, Alejandra Cechich,
Marcelo Moyano, Alan De Renzis, Diego Anabalón, Franco Corgatelli
Grupo de Investigación en Ingeniería de Software del Comahue (GIISCO)
<http://giisco.uncoma.edu.ar>

Departamento de Ingeniería de Sistemas – Facultad de Informática
Universidad Nacional del Comahue
Buenos Aires 1400, (8300) Neuquén

Contacto: [andres.flores, martin.garriga]@fi.uncoma.edu.ar

Resumen

El paradigma de Computación Orientada a Servicios (SOC), promueve el desarrollo de aplicaciones distribuidas en ambientes heterogéneos, que son construidas ensamblando o componiendo servicios reusables, que se publican a través de una red y se acceden mediante protocolos específicos. SOC ha sido ampliamente adoptado con la tecnología de Servicios Web. Existen diferentes estilos de Servicios Web que amplían las oportunidades de reuso, pero generan un desafío de recuperabilidad en torno a la evaluación de servicios heterogéneos, considerando sus tecnologías subyacentes distintivas. Entre los estilos se encuentran los servicios SOAP (con descripciones WSDL) y los servicios RESTful (con múltiples lenguajes de descripción tal como WADL, OpenAPI, etc.). Para afrontar estos desafíos se definió un mecanismo de reuso aplicando Razonamiento basado en Casos (CBR) sobre un Metamodelo de Servicios Heterogéneos que permite la evaluación y composición de servicios. Se ha establecido un mapeo de las descripciones de servicios de los diferentes estilos hacia el Metamodelo, para extender en forma transparente la recuperabilidad de los

servicios de terceras partes de distintos estilos y tecnologías.

Palabras Clave: Ingeniería de Software basada en Reuso – Software Orientado a Servicios – Servicios Web – WSDL – REST.

Contexto

La línea presentada se inserta en el contexto de los siguientes proyectos y acuerdos de cooperación:

- 04/F009-2: “Reuso Orientado a Servicios – Parte II”. Financiado por UNCo. (2017-2021).
- PIP GI 11220170100951CO: “Construcción de Líneas de Productos Software guiada por Estándares de Dominio”. Financiado por CONICET. (2018-2021).
- PICT-2017-1725: “Extensión de Metamodelos de Aplicaciones Orientadas a Servicios para Descripción, evaluación y despliegue de Servicios en la Nube”. Financiado por ANPCyT. (12/2018-12/2021).
 - Investigaciones conjuntas con ISISTAN-UNICEN, Tandil.
- Acuerdo de Cooperación con el Grupo Alarcos, Escuela Superior de Informática, Universidad de Castilla-La Mancha, España.

Introducción

Actualmente la industria de software observa cómo mediante el reuso de software se puede alcanzar un proceso de desarrollo de software acelerado y confiable al basarse en artefactos software que ya han sido probados en diferentes contextos de aplicación. Para ello se adopta el concepto denominado “tercerización”, por medio del cual se acuerdan contratos para adquisición y provisión de artefactos software reusables y se establecen relaciones comerciales entre vendedores y clientes. Por lo tanto, desde el punto de vista de un cliente implica la posibilidad de acelerar el desarrollo de un producto software para reducir el lanzamiento al mercado, y desde el punto de vista de un proveedor implica la posibilidad de observar sus productos con una perspectiva nueva que los coloque dentro del mercado de artefactos reusables.

Un paradigma que promueve fuertemente el reuso de software se denomina Computación Orientada a Servicios (SOC), donde 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 [SH05,PTDL07]. El paradigma SOC encontró una plataforma potencial de aprovechamiento mediante la Web, desde donde se desarrolló la tecnología de Servicios Web [NSS03, Wetal05], con notaciones formales específicas para la descripción las interfaces de servicios. Así el paradigma SOC bajo la implementación con Servicios Web ha logrado su amplia adopción en la industria, principalmente bajo la flexibilidad de ejecución remota que permite descentralizar aún más los procesos de negocios de las compañías, con la ventaja de que las plataformas

específicas de ejecución se encuentran ocultas, por lo cual no se requiere de inversiones adicionales en tecnología (incluyendo costos y esfuerzo de aprendizaje), al adquirir funcionalidad de terceras partes. El beneficio que la tecnología de Servicios Web provee al paradigma SOC se ha denominado “relación sin responsabilidad”, donde una aplicación cliente no requiere asumir cómo se ha implementado el servicio con el que se comunica. Sin embargo, los proveedores de servicios tienen la responsabilidad de evaluar la calidad de los productos ofrecidos como servicios y los consumidores de servicios a su vez deben ser capaces de identificar tal calidad que influirá sobre las aplicaciones en desarrollo.

El funcionamiento concreto del paradigma SOC se basa en la Arquitectura orientada a Servicios (SOA) [SH05] que se encuentra compuesta por tres actores principales: un proveedor, un consumidor y un registro de servicios; donde el proveedor desarrolla y publica servicios en el registro, para que luego el consumidor busque servicios y establezca una comunicación con el proveedor. Sin embargo, la búsqueda de servicios publicados en un registro UDDI (según la tecnología de servicios Web) [OASIS04], en general requiere invertir un esfuerzo considerable para distinguir servicios candidatos que satisfagan los requerimientos de la aplicación cliente [NSS03, Wetal05]. En particular, cuando varios candidatos ofrecen funcionalidades similares se requieren métodos eficientes de selección de servicios que discriminen tanto aspectos funcionales como no-funcionales, considerando además las interacciones válidas para un servicio candidato en función de los procesos de negocio que implementará la aplicación cliente. En particular el ensamblaje de servicios considerando procesos de negocio e interoperabilidad de servicios plantea el uso de dos conceptos de reciente investigación: Orquestación y

Coreografía de servicios [P03, Wetal05]. El primero relacionado a una aplicación particular que describe un proceso de negocios específico, y el segundo relacionado a las interacciones válidas que pueden ocurrir entre distintos servicios predestinados a intervenir en una colaboración.

Los Servicios Web pueden ser desarrollados mediante diferentes estilos distintivos, entre los cuales se encuentran los servicios SOAP y los servicios RESTful. Los servicios SOAP se describen mediante el lenguaje WSDL (Web Service Description Language) que permite especificar las operaciones e intercambio de mensajes para un Servicio Web. Existen diferentes versiones del lenguaje WSDL (1.0, 1.1, 2.0), siendo la versión WSDL 2.0 definido como estándar OMG. Los servicios SOAP se consideran servicios “grandes”, debido a que el intercambio de mensajes se realiza por medio de documentos XML, que también es la base para el lenguaje WSDL. Esto hace que se deba afrontar una sobrecarga de transferencia de archivos. Los servicios RESTful se pueden describir mediante diferentes lenguajes que intentan solventar la sobrecarga mencionada, mediante el intercambio de mensajes livianos, generalmente usando JSON o YAML. El primer intento de notación para REST fue WADL (Web Application Description Language) como analogía de WSDL, y también basado en XML. La industria ha propuesto diferentes lenguajes, entre los que se encuentra OpenAPI/Swagger, REST API Modeling Language (RAML), o MIT’s API Blueprint, entre otros.

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 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 la evaluación y selección de servicios. 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 (Liao et al., 1998). basados en servicios, evaluando las descripciones de la interfaz WSDL. El Metamodelo de Servicios permite representar información de 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,

Desarrollo e Innovación

El perfil de esta línea puede definirse en base a las actividades de investigación y transferencia, a las que da soporte el grupo GIISCo. Los temas específicos consideran los desafíos diferentes asociados al crecimiento de la Tecnología de la Información y las Comunicaciones.

Actualmente, abordamos los siguientes aspectos:

- Compatibilidad y selección de servicios.
- Adaptación y Composición de servicios.
- Complejidad y legibilidad de servicios
- Refactorización de servicios
- Testing de servicios.
- Herramientas para evaluación, selección, composición y testing de servicios.
- Definición de aplicaciones en dominios específicos.

Resultados y Objetivos

En [FCGMRAC20] hemos enumerado una serie de contribuciones anteriores. Durante el año 2020, hemos profundizado la investigación en aspectos de *evaluación de compatibilidad y complejidad de servicios*, generando métodos y herramientas enfocados en las interfaces y comportamiento dinámico de los servicios [GF19].

Las líneas de investigación convergen en el tratamiento del desarrollo de software basado en el reuso de servicios desde la perspectiva de las aplicaciones orientadas a servicios. Una aplicación orientada a servicios implica una solución de negocio que consume servicios de uno o más proveedores y los integra en un proceso de negocio [SW04]. Además puede verse como una aplicación basada en componentes que integra dos tipos de componentes: internos localmente empotrados en la aplicación, y externos estática o dinámicamente enlazados a algún servicio [CMZC14]. El Metamodelo de Servicios Heterogéneos habilita la posibilidad de reusar Servicios Web de diferentes estilos y tecnologías, extendiendo las oportunidades para la solución de aplicaciones cliente. Las capacidades que están desarrollando para

la composición de servicios heterogéneos abren otro abanico de posibilidades cuando los servicios candidatos recuperados no proveen las funcionalidades requeridas. El Metamodelo de Servicios ha sido complementado con nuestro framework para evaluación y selección de servicios desarrollado en trabajos previos [GRLFMCZ18]. Por otro lado la refactorización de servicios basado en métricas de complejidad permite no solamente la adecuación de Servicios Web que un proveedor ofrece, sino también la posibilidad de estudiar escenarios de composición de servicios, para validar nuestra propuesta. Se prevee la aplicación de estos modelos y las herramientas de soporte a dominios específicos, con particular énfasis en aquellos que requieran rigurosidad como aporte de validación efectiva. La visión de esta línea de investigación se resume en:

“Definir técnicas y herramientas para la mejora del desarrollo de software, en función del reuso de servicios web. La definición de modelos de evaluación, selección y refactorización de servicios, y la posibilidad de composición de servicios”.

Formación de Recursos

Humanos

Este proyecto se compone de 11 investigadores, entre los que se cuentan docentes y estudiantes del Grupo GIISCO de UNComa y asesores externos. Algunos de los docentes–investigadores se encuentran realizando carreras de postgrado. Se cuenta actualmente con 3 doctores (1 investigador adjunto y 1 investigador asistente CONICET), 2 doctorandos y 1 maestrando entre los miembros del proyecto. Dirección de Tesis durante 2020: Postgrado (4 tesis), Grado (4 tesis).

Referencias

- [CMZC14] Crasso, M., Mateos, C., Zunino, A., Campo, M. (2014). *EasySOC: Making Web Service Outsourcing Easier*. International Journal on Information Sciences, 259: 452–473.
- [GRLFMCZ18] Garriga, M., De Renzis, A., Lizarralde, I., Flores, A., Mateos, C., Cechich, A., Zunino, A. (2018). *A Structural-Semantic Web Service Selection Approach to Improve Retrievability of Web Services*. Information Systems Frontiers, 20(6): 1319-1344 (**Indexed SCI IF JCR2017: 3.232**).
- [GF19] Garriga, M., Flores, A. (2019). *Standards-driven Metamodel to Increase Retrievability of Heterogeneous Services*. 34th ACM/SIGAPP Symposium on Applied Computing, pp. 2507-2514. (**Qualis A1**).
- [FCGMRAC20] Flores, A., Garriga, M., Cechich, A., Moyano, M., De Renzis, A., Anabalon, D., Corgatelli, F. (2020). *Reuso, Composición y Refactorización de Servicios Heterogéneos*. WICC'20, XXII Workshop de Investigadores en Ciencias de la Computación. Abril, Santa Cruz.
- [LZM98] Liao, T., Zhang, Z., Mount, C. (1998). *Similarity Measures for Retrieval in Case-based Reasoning systems*. Applied Artificial Intelligence, 12, 267–288.
- [NSS03] Nagappan, R.; Skoczylas, R.; Sriganesh, R. (2003). *Developing Java™ Web Services: Architecting and Developing Secure Web Services Using Java*. Wiley Publishing Inc.
- [OASIS04] OASIS Consortium (2004). *UDDI Version 3.0.2*. UDDI Spec Technical Committee Draft, October.
- [P03] Peltz, C. (2003). *Web Services Orchestration and Choreography*. IEEE Computer, 36(10): 46–52.
- [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.
- [SH05] Singh M., Huhns, M. (2005). *Service-oriented computing: Key concepts and principles*. IEEE Internet Computing, 9(1): 75–81.
- [SW04] Sprott, D., Wilkes, L. (2004). *Understanding Service-Oriented Architecture*. The Architecture Journal. MSDN Library. Microsoft Corporation, <http://msdn.microsoft.com/en-us/library/aa480021.aspx>
- [W3C05] W3C Candidate Recommendation (2005) *Web Services Choreography Description Language Version 1.0*. <http://www.w3.org/TR/ws-cdl-10/>
- [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.

SISTEMAS DE INFORMACIÓN ORGANIZACIONALES ASISTIDOS POR ANALÍTICA E INTELIGENCIA DE NEGOCIOS

Romagnano, M.^{1,2}, Ganga, L.¹, Pantano, J. C.^{1,2}, Herrera, M.^{1,2}, López, H.¹, Becerra, M.¹, Aballay, A.¹, Sarmiento, A.¹, Gordillo, M. L.¹, Aguilera, C.², García, A.¹

¹Departamento de Informática, FCFN, Universidad Nacional de San Juan

²Instituto de Informática, FCFN, Universidad Nacional de San Juan

maritaroma@iinfo.unsj.edu.ar, juancruz871@hotmail.com, leonelganga@iinfo.unsj.edu.ar, mcbecerra2008@gmail.com, prof.alicia.gordillo13@gmail.com, cynaguilera95@gmail.com, garcia.exe@gmail.com, adriva2005@yahoo.com.ar, lepezhr@yahoo.com.ar

RESUMEN

Hace casi dos décadas las organizaciones trabajan y realizan un enorme esfuerzo para mantenerse en un entorno global, altamente competitivo y cambiante. Continuamente, se les exige tomar decisiones estratégicas para poder permanecer, de manera exitosa y rentable, en este turbulento mercado. La toma de decisiones conlleva a la conversión de datos en información y de información en conocimiento. Requiere la sincronía de muchos especialistas y de especialidades dentro de una organización. Esta transformación representa un desafío a diario, ya que tienen que lidiar con grandes cantidades de datos que a menudo se generan en las operaciones cotidianas.

La analítica e inteligencia de negocios se han convertido en una apuesta tecnológica que las organizaciones y empresas líderes deberían adoptar. A través de sus técnicas y herramientas ofrecen la gestión del conocimiento, permitiendo a las organizaciones responder dinámica y rápidamente al vertiginoso mercado en el cual se desenvuelven. Las nuevas tecnologías y capacidades informáticas que se han ido desarrollando al pasar los años, permiten que la inteligencia y analítica de negocios, comúnmente conocidas como Business Intelligence (BI) y Business Analytics (BA), respectivamente, puedan aplicarse satisfactoria y exitosamente en tareas cotidianas de una organización, convirtiéndose en las principales causas de grandes ventajas competitivas.

Por lo tanto, atendiendo a estas necesidades de gestión de la información por la cual transitan actualmente las organizaciones regionales, y observando la debilidad en la actual currícula académica, este trabajo propone determinar cómo los Sistemas de Inteligencia de Negocios (SIN) aportan a los Sistemas de Información Organizacionales (SIO), para la toma de decisiones.

Palabras Claves: Sistemas de Información, Toma de Decisiones, Analítica de Negocios, Inteligencia de Negocios, Organización.

CONTEXTO

El presente trabajo se encuentra enmarcado en el proyecto “Aporte de los Sistemas de Inteligencia de Negocios a los Sistemas de Información Organizacionales para la Toma de Decisiones”, presentado en la convocatoria del Consejo de Investigaciones Científicas y Técnicas y de Creación Artística (CICITCA) de la Universidad Nacional de San Juan, para ser desarrollado durante el período comprendido entre 01/01/2020 al 31/12/2022.

Las tareas de investigación se desarrollan en el Laboratorio de Sistemas de Información, en el ámbito del Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.

El grupo de investigación se encuentra conformado por integrantes de distintas disciplinas (Informática, Matemática,

Estadística, Administración de Empresas y Abogacía) que, en su mayoría, cuentan con una experiencia de más de 20 años en la disciplina que les compete.

Desde el año 1995 hasta el 2010 la línea de investigación estuvo orientada a la gestión de los Sistemas de Información (SI). Luego, en el año 2011 se puso énfasis en los fundamentos conceptuales de los SI, pudiendo identificar las restricciones que existen en los SI. Se continuó estudiando la crisis por la que transitaba la disciplina al confundir los SI con Tecnologías de Información (TI). Posteriormente se trabajó en la construcción de un modelo disciplinar, basado en constructos. Desde el año 2016 se vienen identificando y caracterizando los modelos conceptuales que aportarán conocimiento a los SI.

Si bien el eje central del equipo de trabajo sigue siendo los SI, actualmente se ha redireccionado la investigación a estudiar el aporte de la Inteligencia Artificial, específicamente la adquisición de conocimiento a través de los Sistemas de Inteligencia de Negocio, a los Sistemas de Información Organizacionales, lo cual permite a la organización tomar decisiones acertadas y oportunas.

1. INTRODUCCIÓN

En los tiempos que corren, la necesidad de las organizaciones por lograr una mejora continua en sus procesos de negocio, juega un rol protagónico y decisivo. A su vez, aquellas organizaciones categorizadas como empresas tienen que enfrentarse a nuevos modelos económicos, obligadas a realizar frecuentes cambios, tratando de acercarse todo lo posible a sus consumidores. Deben dedicar gran parte de su tiempo y de sus recursos, económicos y humanos, a la obtención, procesamiento, aplicación y proyección de la información.

Inmersas dentro de la Ciencia de los Datos, BI y BA son dos de las tendencias actualmente consideradas como muy beneficiosas para una organización. Esto se debe a que, utilizadas adecuadamente, pueden presentar ventajas

competitivas, permitiéndole conocer, con alta precisión, su estado actual y, en base a la información presentada, ser capaces de pronosticar futuros comportamientos del mercado y llevar a cabo acciones proactivas en base a análisis predictivos y prescriptivos. En la actualidad, estos conceptos, están ganando fuerte reconocimiento y popularidad, aunque no son nuevos ni de reciente aparición, especialmente BI.

BI hace referencia al manejo optimizado de los datos que almacena, recopila y analiza una organización, siendo capaz de transformarlos en decisiones estratégicas que permitan el diseño de acciones orientadas a alcanzar el éxito empresarial (López Benítez, 2018). Los elementos en que se sustenta la conceptualización de Inteligencia de Negocios son los Sistemas de Información, los mecanismos de innovación y los procesos de toma de decisiones. En cada uno de ellos se implementan estrategias que pueden llevar a la organización a adquirir conocimiento y a mejorar la manera en que este incrementa el valor de los productos y servicios que se ofrecen (Ahumada Tello y Perusquia Velasco, 2015).

La industria global gasta en software BI un promedio anual de 14 billones de dólares. Gartner Group pronosticó que en el 2018 más de la mitad de las grandes organizaciones de todo el mundo competirán utilizando Advanced Analytics (AA) y algoritmos propietarios, causando volúmenes de análisis de datos a gran escala. Además, se pronosticó que para el 2020 el segmento de más rápido crecimiento en el mercado analítico será el de BI, representando más del 40% de las nuevas inversiones en una empresa (Gartner Group, 2016). Asimismo, se había previsto que las empresas inviertan un 30% más en Inteligencia Artificial (AI) en 2017 que en 2016 a fin de que el aprendizaje automático pudiera potenciar el análisis de datos a una escala superior a la humana. Con ello se busca impulsar decisiones más rápidas y acertadas en marketing, comercio electrónico, gestión de productos, entre otros, ayudando a cerrar la brecha entre los supuestos y la acción (Evelson y Bennett,

2017). En el 2021, el mercado de BI moverá alrededor de los 23.100 millones de dólares y crecerá a ritmos anuales del 7,6% entre 2020 y 2025 (Dir&Ge, 2020).

También, BA se ha convertido en una frase de moda en la era actual de la economía. Principalmente, la proliferación del uso de Internet y las tecnologías de la información han convertido el BA en un área de aplicación sólida. Es imposible negar su importancia y su impacto en las TI, en los métodos cuantitativos y en las Ciencias de las Decisiones (Cegielski y Jones-Farmer, 2016). Tanto en las industrias como en el ámbito académico se busca contratar personas con talento en estas áreas, con la esperanza de desarrollar competencias organizacionales que les permitan obtener y mantener ventajas competitivas frente al mercado (Chahal, J. Jyoti y J. Wirtz (ed.), 2019).

Hawley plantea a la BA desde otro punto de vista. Se centra más en comprender la cultura organizacional que en la mera tecnología. Propone que, para una implementación y aprovechamiento exitoso de los beneficios de la analítica de negocios, se debe tener conocimiento de la motivación de una organización, sus fortalezas y debilidades (Hawley, 2016).

En cuanto a SI, Peña los define como un conjunto de elementos interrelacionados con el propósito de prestar atención a las demandas de información de una organización, para elevar el nivel de conocimientos que permitan un mejor apoyo a la toma de decisiones y desarrollo de acciones (Peña, 2006). Otros autores lo definieron como un conjunto de elementos que interactúan entre sí con el fin de apoyar las actividades de una empresa o negocio. En un sentido amplio, un SI no necesariamente incluye equipos electrónicos (hardware). Sin embargo, en la práctica se utiliza como analítica de sinónimo de “Sistema de Información Computarizado” (Cohen y Asín Lares, 2005).

Por último, como estos sistemas se encuentran inmersos en la organización, cabe mencionar su definición. Así Chiavenato (2000) lo define como la coordinación de diferentes actividades

de contribuyentes individuales, con la finalidad de efectuar intercambios planteados con el ambiente.

Por lo tanto, BI considerada como una especificidad de la Ciencia de los Datos, puntualmente como una especificación de la IA para el Negocio, viene a ofrecer esta gestión del conocimiento. Se presenta como un conjunto de procesos, aplicaciones y tecnologías que facilitan la obtención rápida y sencilla de datos provenientes de distintos Sistemas de Información Organizacionales (SIO). Los datos analizados e interpretados, se transforman en conocimiento apropiado para la toma de decisiones organizacionales.

Además, uno de los objetivos de los SI es brindar información que provea soporte a la toma de decisiones en la organización. A su vez, uno de los objetivos de BI es la recolección y procesamiento de datos, de tal manera que puedan ser utilizados para generar y mantener estrategias que brinden ventajas competitivas. Se puede decir que los SI utilizan BI para procesar la información y luego recibir una respuesta concreta y acertada del mismo. Al mismo tiempo, los SIN utilizan los SI como una herramienta para poder, justamente, procesar estos datos que recolecta (Dedić N., y Stanier C., 2016). Entonces, los SIN son, a fin de cuentas, SI con un objetivo principal de apoyar la toma de decisiones.

Leslie Bell-Friedel menciona: “la tecnología está y los datos también, el problema radica en que las organizaciones desconocen cómo explotarlos de la mejor manera posible o ignoran el potencial beneficioso de la aplicación de estos conceptos” (Bell-Friedel, 2017).

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Para enfrentar el mercado fluctuante y dinámico, las organizaciones deben actuar con prudencia al optimizar sus recursos y al tomar decisiones estratégicas. El conocimiento logrado, y su posterior gestión puede ser su principal ventaja competitiva. Heavin, Daly y

Adam planteaban que la habilidad de una organización para administrar el conocimiento es esencial en términos de su desarrollo y como un activo estratégico (Heavin, Daly y Adam, 2014).

Consecuentemente, el equipo de investigación, considera la temática propuesta de relevancia social, organizacional y que además posee implicancias prácticas, altamente propicia debido a que actualmente Argentina cuenta con políticas de estado tales como Argentina Innovadora 2020 y 2030, Ley de Promoción de la Economía del Conocimiento, Industria 4.0, Plan de Inteligencia Artificial, entre otras. Precisamente, San Juan posee polos tecnológicos como Casetic San Juan, Servicios Mineros y San Juan TEC, lo cual puede tomarse como puntapié inicial para trabajar conjuntamente con el estado en la sustentabilidad local, aportando conocimiento y casos de estudio desde la academia.

Los autores de este artículo e integrantes del proyecto se han dividido en subgrupos, de acuerdo a su formación y especialización, y se encuentran trabajando en las siguientes líneas de investigación y desarrollo:

- Inteligencia y analítica del negocio.
- Análisis y procesamiento estadístico de los datos.
- Preservación y legislación de los datos.
- Sistemas de información organizacionales.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Durante el año 2020, primer año de desarrollo del proyecto y debido a las conocidas restricciones nacionales y mundiales que rigen por la pandemia, el grupo de investigación se enfrentó a la dificultad de contar con pocos datos provenientes de organizaciones y empresas de carácter público y privado. Sin embargo, se logró avanzar en la etapa de revisión y búsqueda bibliográfica de la temática, se pudo obtener un instrumento de caracterización de organizaciones locales, sin poder validar y refinar aún. Solo se pudieron

identificar los beneficios y los distintos tipos de conocimientos que aportan los SIN a los SIO en una empresa de transporte local y en una entidad bancaria. Por último, se realizaron publicaciones y difusiones en una revista internacional y en 6 congresos, nacionales e internacionales.

Para el presente año se espera avanzar en la adquisición, tratamiento y análisis de datos provenientes de empresas y organizaciones, públicas y privadas, y así poder generalizar cuáles son los aportes que los SIN hacen a los SIO.

Los resultados del proyecto tienen una inmediata transferencia al medio local, principalmente a la alta gerencia de organizaciones locales de cualquier tipo, y a la comunidad científica. Además, se trabajará en experimentación y difusión de resultados con investigadores de la Facultad de Ingeniería de la Universidad de Quindío, Colombia y de la Escuela de Ingeniería en Transporte, Pontificia Universidad Católica de Valparaíso, Chile.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del grupo de investigación se identifican integrantes formándose en:

- Licenciatura en Ciencias de la Computación, UNSJ.
- Licenciatura en Ciencias de la Información, UNSJ.
- Maestría en Estadística Aplicada, Escuela de graduados de la Facultad de Ciencias Económicas, UNC.
- Estancia de formación de docencia e investigación, Programa de Movilidad Internacional, UNSJ.
- Doctorado en Ingeniería, Facultad de Ingeniería, UNCuyo.
- Doctorado en Demografía, Escuela de graduados de la Facultad de Ciencias Económicas, UNC.

5. BIBLIOGRAFÍA

Ahumada Tello, E. y Perusquia Velasco, J. M. (2015). Business intelligence: Strategy for competitiveness development in technology-based firms. Disponible en www.sciencedirect.com, marzo de 2019.

Bell-Friedel, L. (2017). Marine Propulsion & Auxiliary Machinery. Pp.71-72. Disponible en: https://issuu.com/rivieramaritimemedia/docs/mp-aug17_text, noviembre 2020.

Cegielski, C. y Jones-Farmer, L. (2016). Knowledge, skills, and abilities for entry-level business analytics positions: A multi-method study. *Decision Sciences Journal of Innovative Education*, Vol. 4(1), enero de 2016, Pp.91–118. <https://doi.org/10.1111/dsji.12086>.

Chahal, H., Jyoti, J. y Wirtz, J. (ed.) (2019). *Understanding the Role of Business Analytics. Some Applications*. Springer Nature Singapore Pte. Ltd. 216 pag. ÑO 2019. ISBN: 978-981-13-1334-9.

Chiavenato I. (2000). *Administración de recursos humanos. El capital humano de las organizaciones*. Octava edición. ISBN 970-10-6104-7. McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V.

Cohen, K., y Asín Lares, E. (2005). *Sistemas de información para los negocios: un enfoque de toma de decisiones*.

Dedić N., y Stanier C. (2016). *Measuring the Success of Changes to Existing Business Intelligence Solutions to Improve Business Intelligence Reporting*. Lecture Notes in Business Information Processing. Springer International Publishing. Volume 268, pp. 225-236.

Dir&Ge (2020). El mercado de business intelligence crecerá a ritmos anuales del 7,6% entre 2020 y 2025. Disponible en: <https://directivosygerentes.es/innovacion/mercado-business-intelligence-crecimiento>, enero de 2021.

Evelson, B. y Bennett, M. (2017). *Las Plataformas Forrester Wave: Enterprise BI con la Mayoría de las Implementaciones en las Instalaciones*. Cambridge: Forrester Research. Disponible en: <https://www.forrester.com/report/The+Forrester+Wave+Enterprise+BI+Platforms+With+Majority+Cloud+Deployments+Q3+2017/-/E-RES137263>, diciembre de 2020.

Gartner Group (2016). *The 2016 Gartner Magic Quadrant for Business Intelligence and Analytics*. Stamford: Gartner Research. Disponible en: <https://www.gartner.com/>, noviembre de 2020.

Hawley, D. (2016). Implementing business analytics within the supply chain: success. *The Electronic Journal Information Systems Evaluation*, 19(2), 112–120.

Heavin, C., Daly, M., y Adam, F. (2014). Small data to Big Data: The Information Systems (IS) continuum. *KMIS 2014 - Proceedings of the International Conference on Knowledge Management and Information Sharing*. 289-297.

López Benítez, Y. (2018). *Business Intelligence*. ADGG102PO. IC Editorial. 1º Edición. Andalucía – España. Vol. 1, 151 pag. Año 2018. ISBN: 978-84-9198-467-2.

Peña Ayala, A. (2006). *Ingeniería de software: Una guía para crear sistemas de información*. México DF (México): Instituto Politécnico Nacional.

Sistemas móviles, accesibles e inteligentes para una sociedad inclusiva

Susana I. Herrera, Marta C. Fennema, Marilena Maldonado, María I. Morales, Rosa A. Palavecino, Federico Rosenzvaig, Adriana Macedo, Adriana Jiménez, Pablo Divi, Paula Barrera Nicholson

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero

{sherrera, marilena}@unse.edu.ar, {crisfen, mines_morales}@yahoo.com, {rosypgg, adrumacedo, paulabarreranicholson, fedvaig, ajfraguas}@gmail.com, pablodivi81@hotmail.com

Resumen

En este trabajo se presenta una investigación que constituye una continuación de los proyectos sobre Computación Móvil, llevados a cabo en la Universidad Nacional de Santiago del Estero (UNSE) de 2012 a 2018, 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, se 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, personas con discapacidad y personas del colectivo LGBT+.

La investigación se lleva a cabo por un equipo interdisciplinario integrado por informáticos, médicos, sociólogos, fonoaudiólogos; como también profesionales de la Matemática y de Idiomas. La investigación se realiza en colaboración con el Ministerio de Salud de la Provincia de Santiago del Estero y cuenta con 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, aplicaciones móviles para colectivo LGBT+.

1 Contexto

Esta investigación se lleva a cabo en el marco del proyecto denominado “Sistemas móviles,

accesibles e inteligentes para una sociedad inclusiva”, del Instituto de Investigaciones en Informática y Sistemas de Información (IISI) de la UNSE, financiado por el Consejo de Ciencia y Técnica de dicha universidad, durante el período enero 2021-diciembre 2024. Es una continuación de los proyectos de investigación que el grupo de investigadores de Computación Móvil de UNSE viene desarrollando desde el año 2012.

A través de este proyecto, se pretende continuar la investigación sobre métodos y tecnologías de Computación Móvil que permitan proveer soluciones innovadoras a problemáticas de personas vulnerables, así como también al medio ambiente.

Se propone llevar adelante una investigación flexible e interdisciplinaria, relacionando los siguientes dominios: Computación Móvil, Ingeniería del Software, Inteligencia Artificial, Educación y Salud. Teniendo como centro la Computación Móvil, se abordará desde ella el desarrollo de aplicaciones móviles (Delia, 2017; Delia et al. 2018) y, en menor medida, la problemática de las redes móviles.

Para el desarrollo de aplicaciones móviles, desde la Ingeniería del Software (IS), se estudiarán métodos ágiles (Scrum y MobileRA) y diversos frameworks para el desarrollo de aplicaciones móviles multiplataforma, nativas y web (Xamarin, Ionic, Angular, etc.).

También dentro de la IS, pero enfocado desde el área de Interacción Hombre Máquina, se investigarán tecnologías innovadoras de interacción, como Realidad Aumentada (RA) (Kipper & Rampolla, 2013; Azuma, 2001), sintetizadores de voz, modelos 3D animados. Además, se abordarán métodos y tecnologías para la evaluación de la accesibilidad (Mariño et al., 2018; Masri & Lujan, 2010; W3C, 2020)

web y móvil, tomando como referencia los estándares internacionales y nacionales.

Desde la Inteligencia Artificial, se investigará acerca de la integración de tecnologías de Machine Learning (Witen et al., 2017; Chollet, 2017) en las aplicaciones móviles, definiendo modelos de datos, métodos y herramientas que permitan un comportamiento inteligente.

Bajo un enfoque de investigación-acción, se desarrollarán prototipos de aplicaciones móviles para aprendizaje, entrenamiento auditivo y entrenamiento vocal. Las aplicaciones de m-learning estarán orientadas al aprendizaje de temas complejos (de Matemática y Programación), el aprendizaje de personas con discapacidad y personas que habitan en zonas de recursos limitados (Ej. zonas rurales) o personas en situación de pobreza. El entrenamiento auditivo estará dirigido a personas hipoacúsicas. Mientras que el entrenamiento vocal, estará dirigido a mujeres transgénero (Gomez Raya, 2018).

El grupo de investigadores de Computación Móvil (computacionmovil.unse.edu.ar) está compuesto por 25 personas provenientes de diferentes disciplinas. La propuesta cuenta con el asesoramiento del Laboratorio en Investigación en Informática LIDI (UNLP), de la Facultad de Ciencias Exactas de la UNSa y del Laboratorio en Cognición Humana y Artificial de la Universidad Paris 8 (Francia).

2 Introducción

La propuesta se realiza desde una perspectiva general sistémica, retroprospectiva y glocal (Morin, 2018; Herrera¹). En cuanto a **retroprospectiva**, implica que, ubicado el grupo de investigación en el presente, considerando sus antecedentes en investigación, se propone abordar problemáticas que contribuyan en el futuro a lograr una humanidad en equilibrio consigo misma y con su ambiente. En cuanto a **glocal**, significa que, ubicado el grupo en Santiago del Estero (Argentina), a través de la vinculación con redes científicas o

en asociación con otros grupos de investigación (locales, nacionales e internacionales), se contribuirá a resolver problemáticas de la comunidad local que también son comunes a otras comunidades del planeta.

Los antecedentes del grupo² se ubican en el dominio de la Ingeniería del Software, en el área de Computación Móvil (Talukder et al., 2010). La principal línea de trabajo ha sido el **desarrollo de aplicaciones móviles multiplataforma que utilizan técnicas de visualización innovadoras como la Realidad Aumentada** (Herrera et al, 2019; Fennema et al., 2018, Herrera et al., 2018; Fennema et al., 2016; Herrera et al., 2014; Fennema et al., 2014; Herrera et al., 2012). A su vez, a través de los avances en dicha línea, se ha intentado proveer soluciones innovadoras a problemáticas presentes en **grupos vulnerables, como aprendices en zonas rurales aisladas** (Rocabado et al., 2014) y **niños con discapacidad** (Herrera, 2020). También se han provisto soluciones para el aprendizaje de temas complejos de Matemática y Programación, y el aprendizaje colaborativo en general. Para obtener logros significativos en estas temáticas, el grupo se ha caracterizado por ser interdisciplinario (investigadores provenientes de la Informática, de la Salud y de la Educación), por contar con el asesoramiento y trabajar en colaboración con especialistas de otras universidad argentinas y europeas (Universidad Nacional de La Plata, Universidad Islas Baleares y Universidad Paris 8), por ser intergeneracional, por incorporar profesionales de la industria del software y por trabajar en colaboración con organizaciones del medio (se firmaron convenios con el Ministerio de Salud de Santiago del Estero, el Servicio de Detección Temprana de Hipoacusia de Santiago del Estero y el Ministerio de Ambiente y Desarrollo Sustentable de la Nación Argentina).

Hoy nuestro planeta se ve amenazado más que nunca por fenómenos complejos de diferentes matices (Chomsky, 2020; Morin, 2014): biológicos (Ej. Pandemia COVID-19), políticos

¹ La crisis de la Ciencia y la tecnología que ha generado el COVID-19. Una mirada desde Francia. Café Con-Ciencia - Dra. Susana Herrera - FCID UCSE - <https://youtu.be/RG600XKkrc>

² computacionmovil.unse.edu.ar

(Ej. amenazas de bombas nucleares algunos países), sociales (Ej. pobreza que produce desequilibrio educativo y nutricional), ambientales (Ej. calentamiento global). Algunos de estos problemas afectan a todos los seres vivos por igual, por ejemplo, el problema de la polución que está destruyendo el medio ambiente. Pero la mayoría de las problemáticas afectan principalmente a las personas más vulnerables: personas pobres, personas con discapacidad, personas que se encuentran en otros grupos que requieren atención especial. Considerando estas amenazas locales y globales, a través de este proyecto, se pretende continuar la investigación sobre métodos y tecnologías de Computación Móvil que permitan proveer soluciones innovadoras a problemáticas de personas vulnerables, así como también al medio ambiente (hoy vulnerable). Los grupos impactados serían: personas con discapacidad (principalmente discapacidad auditiva y visual), personas del grupo LGBTI+ y aprendices de zonas rurales aisladas. En cuanto a las fenoménicas a tratar, en general, se abordarán: el aprendizaje, el entrenamiento auditivo y el entrenamiento vocal. En cuanto al aprendizaje, principalmente se abordará el aprendizaje de disciplinas que involucran aprendizajes complejos, como Matemática y Programación; y también se abordará el aprendizaje de conceptos ambientales. Todo esto permitirá que la investigación en Computación Móvil contribuya a una educación inclusiva para un mundo sustentable.

Para atender dichas problemáticas, se propone llevar adelante una investigación amplia y flexible en los campos de Ingeniería del Software, Computación Móvil e Inteligencia Artificial. **Su objetivo general consiste en optimizar la accesibilidad e inteligencia de las aplicaciones móviles, mediante el estudio de métodos y tecnologías innovadoras de desarrollo y de interacción.** Por lo tanto, las principales variables a abordar serán la **inteligencia** y la **accesibilidad** de aplicaciones móviles. Además, se evaluará el impacto de estas aplicaciones en el **aprendizaje** y en el **entrenamiento auditivo** y en el **entrenamiento vocal** de **personas con discapacidad**

(principalmente auditiva y visual) y del colectivo LGBTI+. Dentro de las tecnologías de interacción innovadoras, se continuará con los estudios sobre **Realidad Aumentada** y con el uso de **sintetizadores de voz.**

3 Líneas de investigación y desarrollo

Para llevar adelante la propuesta, se identifican diferentes líneas de trabajo integradas por personas provenientes de diferentes disciplinas. Dichas líneas trabajarán interactuando entre sí, teniendo cada una sus objetivos bien definidos y sus responsables: 1) Ingeniería del Software, Computación Móvil e Inteligencia Artificial, 2) Investigación en Informática, Salud y Educación, 3) Accesibilidad de Aplicaciones Web y Móviles, 4) Aplicaciones móviles para el colectivo LGBTI+.

Las líneas 1 y 2 constituyen el soporte del resto de líneas del proyecto. Líneas 3 y 4, abordan temáticas más específicas. Podrán agregarse (durante los 4 años) líneas referidas a desarrollos específicos para los dominios salud o medioambiente, del tipo de la línea 4.

La línea 1 trata los métodos, técnicas y herramientas de la Ing. del Software, de la Inteligencia Artificial y de Redes Móviles que se aplican al desarrollo de aplicaciones móviles. Usando investigación acción, se investigan estos elementos mientras **se desarrollan aplicaciones** para entrenamiento/aprendizaje de personas con discapacidad y para m-learning en general. El grupo está integrado por especialistas en Informática.

La línea 2 da soporte a todos los **procesos de investigación** que se desarrollan en el proyecto, determinando en cada caso el enfoque de investigación a utilizar (cuantitativos, cualitativos y mixtos), hipótesis y diseños experimentales o trabajos de campo. La selección dependerá si la investigación impacta sobre la Informática o sobre la Salud o sobre la Educación. El grupo está integrado por especialistas en Investigación, en Educación y en Salud.

La línea 3 trata las normativas, métodos, técnicas y herramientas relacionadas con una de las variables principales del proyecto: la accesibilidad. Esta línea trabaja en general sobre las personas con discapacidad y la accesibilidad

web y la accesibilidad académica. El grupo está integrado por especialistas en Informática y en Educación.

La línea 4 trata del desarrollo de aplicaciones móviles para el colectivo LGBTI+, abordando principalmente el entrenamiento vocal de las mujeres transgénero (o trans). El grupo está integrado por especialistas en Informática y en Salud.

4 Resultados esperados

Se espera generar nuevo conocimiento científico-tecnológico en Informática (Computación Móvil, Ingeniería del Software, Interacción Hombre Máquina), relativos a:

- Ingeniería del Software: métodos y tecnologías blandas para el desarrollo de aplicaciones móviles accesibles e inteligentes (multiplataforma, nativas y web) con y sin RA.
- Accesibilidad Web y Móvil: pautas metodológicas para evaluación y construcción de software.
- M-learning: diseño de prácticas innovadoras usando las aplicaciones desarralladas (AlgeRA, ImaColab, Matemática y Sonidos, etc.).

En cuanto a **productos software**, se esperan los siguientes resultados:

- Se ampliarán y optimizarán las siguientes aplicaciones móviles con RA (agregando inteligencia y accesibilidad): **AlgeRA** (m-learning) y **Matemática y Sonidos** (m-learning y entrenamiento auditivo).
- Se optimizará la siguiente aplicación móvil nativa sin RA (agregando inteligencia): **Comunicador para niños con parálisis cerebral**.
- Se desarrollará una aplicación móvil multiplataforma inteligente sin RA para **entrenamiento vocal de mujeres transgénero**.

Se desarrollará una aplicación móvil multiplataforma sin RA para **aprendizaje de contenidos medioambientales** de jóvenes vulnerables (Programa Haciendo Lio del Ministerio de Ambiente de la Nación). M-learning: se diseñarán e implementarán experiencias, con y sin RA, y sus impactos en la

motivación intrínseca. Se optimizarán las aplicaciones de m-learning AlgeRA e ImaColab.

5 Formación de recursos humanos

La concreción de este proyecto contribuirá 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.

Esta investigación contribuye al desarrollo de dos tesis doctorales en Informática (UNLP), una tesis doctoral en Educación (UNSE), una tesis de Maestría en Informática Educativa (UNSE), y varios trabajos finales de grado de la Licenciatura en Sistemas de Información (UNSE).

Referencias

1. Azuma, R. (2001). *Augmented Reality: Approaches and Technical Challenges. Fundamentals of Wearable Computers and Augmented Reality*. Mahwah, New Jersey: W. Barfield, Th. Caudell.
2. Chollet, François. (2017). "Deep Learning with Python", Manning.
3. Chomsky, N. (2020). *Cooperación o extinción*. Traducción al español. B Editorial. ISBN 978-84-666-6792-0. Madrid.
4. Delía, L. (2017). *Desarrollo de aplicaciones móviles multiplataforma*. Trabajo Final de Especialista en Ingeniería Web. Facultad de Informática, Universidad Nacional de La Plata.
5. Delia, 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.
6. Fennema, M., Herrera, S., Palavecino, R., Budán, P., Rosenzvaig, F., Najjar 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.
7. Fennema, F., Herrera, S., Palavecino, R., Najar, P., Budán, P., Suarez, G., Cordoba, C. (2016). Aplicaciones Móviles: arquitecturas, visualización, realidad aumentada, herramientas de medición, desarrollo híbrido. Argentina. Concordia. Libro. Artículo Breve. Workshop. XVIII Workshop de Investigadores en Ciencias de la Computación. Red UNCI.
 8. Fennema, M., Palavecino, R., Herrera, S., Najar, P. (2014). Métodos y Técnicas de Desarrollo para Optimizar la Calidad de los Sistemas Móviles. XVI Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-750-605-570-7. Ushuaia.
 9. Gomez Raya, A. (2018). Intervención logopédica en la feminización de la voz en transexuales: revision bibliografica. Revista de Investigación en Logopedia, 21-42.
 10. Herrera, S., Fénema, M.C., Morales, M.I., Maldonado, M., Palavecino, R., Rosenzvaig, F., Macedo, A., Suárez, C., Pellicer, R., Villavicencia, R. (2019). Sistemas móviles multiplataforma con realidad aumentada. Aplicaciones en educación y salud. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan). ISBN: 978-987-3619-27-4.
 11. Herrera, S., Sanz, C., Morales, M., Palavecino, R., Maldonado, M., Irurzun, I. Carranza, A., Suarez, 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.
 12. Herrera, S., Morales, M., Fennema, M., Goñi, J. (2014). Diseño de estrategias de m-learning. Desarrollo de aplicaciones. XVI Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-750-605-570-7. Ushuaia.
 13. Herrera, S., Fennema, C., Rocabado, S., Goñi, J. Expositor: Rocabado, S. (2012). Optimización de la calidad de los Sistemas Móviles mediante la implementación de nuevas arquitecturas, realidad aumentada, técnicas de visualización y redes móviles Ad-Hoc. Workshop en Investigación en Ciencias de la Computación. Posadas, Abril 2012. ISBN: 978-950-766-082-5.
 14. Herrera, S. (2020). Mobile Learning with Augmented Reality to Improve Inclusion of Hearing-Impaired Children. Mobile Learning Week. UNESCO. Paris.
 15. Kipper, Gregory, Rampolla, Joseph. (2013). "Augmented Reality. An Emerging Technologies. Guide to AR". Elsevier. ISBN: 978-1-59749-733-6. USA.
 16. Mariño, Sonia I., Pedro L. Alfonso, Cintia Galain, Josue Maidana, Romina Alderete (2018). ACCESIBILIDAD WEB, APORTANDO A LA INCLUSIÓN. XX Workshop de Investigadores en Ciencias de la Computación, pp 694-697.
 17. Masri, F. and S. Lujan (2010). "Análisis de los métodos de evaluación de la accesibilidad Web," 7mo Congr. Int. Educ. Super. (Universidad 2010), no. 978-959-16-1164-2, pp. 1-10, 2010.
 18. Morin, E. (2014). La via: el futuro de la humanidad. Ed. Planeta. Traducción al español. 9788408126805. Barcelona.
 19. Morin, E. (2018). Me Philosophes. Editions Germina. Collection Pluriel. Paris.
 20. Rocabado, S., Herrera, S., Coronel, A., Campos, M. (2014). MANETs-Learn: modelo de MANETs para m-learning rural. XVI Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-750-605-570-7. Ushuaia.
 21. W3C, "WCAG 2.1, (2020). En: <https://www.w3.org/TR/WCAG21/>.
 22. Witten, Ian H., Eibe Frank, Mark A. Hall. (2017). Data Mining: Practical Machine Learning Tools and Techniques, (Fourth Edition). Morgan Kaufmann.

Técnicas y herramientas para la mejora de procesos, gobernanza digital y calidad de datos

Silvia Esponda , Ariel Pasini , Marcos Boracchia ,
Julieta Calabrese¹ , Rocío Muñoz¹ , Juan Santiago Preisegger¹ , Patricia Pesado 

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata

¹ Becario postgrado UNLP
50 y 120 La Plata Buenos Aires
Centro Asociado CIC-PBA

(sesponda, apasini, marcosb, jcalabrese, rmunoz, jspreisegger, ppesado) @lidi.info.unlp.edu.ar

RESUMEN

El III-LIDI (Instituto de Investigación en Informática LIDI) posee una línea de investigación orientada a la calidad de software, mejora de procesos y gobernanza digital.

Los temas relacionados con la calidad de software abarcan, la evaluación de calidad de producto y calidad de datos y la viabilidad de la aplicación de los estándares internacionales ISO/IEC en base de datos de organismos públicos y privados. Por otro lado, la línea abarca conceptos relacionados con gobernanza digital, implementando soluciones e-democracia y analizando técnicas de innovación en ciudades inteligentes sostenibles.

Palabras Claves

Ingeniería de Software – Calidad del proceso – Calidad del producto – Calidad de Datos - Normas de Calidad – ISO - Gobernanza Digital

CONTEXTO

Esta línea de investigación se enmarca en el proyecto “11/F023 Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” y en el subproyecto “Gobernanza Digital. Mejora de Procesos” (2018-2021), acreditado por el Ministerio de Educación de la Nación.

Además, se ha desarrollado el proyecto “Calidad de Datos” aprobado por la Facultad de Informática UNLP.

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 aplicando mejoras. Asimismo, participa en proyectos internacionales, entre ellos CAP4CITY y del Consorcio CC&BD&ET.

1. INTRODUCCION

Dentro del proyecto, se destacan los siguientes ejes principales:

1 – Ciudades Inteligentes Sostenibles

Generalmente, se relaciona el concepto de ciudades inteligentes con el uso de la tecnología en el desarrollo de las actividades de la ciudad, pero en realidad es mucho más que eso. Según [1] las ciudades inteligentes sostenibles (CIS) representan las últimas etapas de progresión a través de ciudades digitales y ciudades inteligentes, consideradas como un proceso transformador continuo, basado en la colaboración y el compromiso de diferentes actores, construyendo diferentes capacidades (humanas, técnicas e institucionales) de manera de mejorar la calidad de vida, proteger los recursos naturales y perseguir el desarrollo socioeconómico. La Unión

Internacional de Telecomunicaciones (UIT) de la ONU, estableció una de las definiciones pioneras de ciudad inteligente sostenible: "Una ciudad inteligente sostenible es una ciudad innovadora que utiliza tecnologías de información y comunicación (TIC), y otros medios, para mejorar la calidad de vida, la eficiencia de la operación y los servicios urbanos y la competitividad, al tiempo que se garantiza que satisfaga las necesidades de las generaciones, presentes y futuras, con respecto a los aspectos económicos, sociales, ambientales y culturales".

Las CIS se componen de cinco pilares: *Social, Económico, Ambiental, Gobernanza e Infraestructura Urbana*, para los que se requiere la prestación de servicios públicos que satisfagan las necesidades de sus ciudadanos [2]

En este caso desde el proyecto se busca generar técnicas y herramientas que permitan lograr una participación ciudadana en este tipo de procesos.

Los organismos ponen a disposición de su comunidad grandes volúmenes de datos abiertos con el fin de que aquellos interesados en la temática puedan procesarlo y generar aportes en el proceso de mejora de la ciudad. Pero al momento de acceder a los datos, aparecen diferencias técnicas como formatos de archivos, estructura del archivo, nombres de columnas, tipos de datos, magnitudes, etc., que dificultan y en algunos casos imposibilitan el análisis de la información. Por lo que se estudian técnicas y herramientas para la vinculación de los datos de los portales gubernamentales y se han realizado herramientas para de gestión de portales institucionales, enfatizando en las características que hacen a este concepto. Al mismo tiempo ha desarrollado un sistema que permite la generación de indicadores basados en datos abiertos, aplicado a datos públicos de medio ambiente.[3]

Otras de las líneas del proyecto relacionadas con las CIS son los servicios

de voto por internet, facilitando el proceso de elecciones a diferentes organismos públicos y privados.

La definición de CIS incluye el concepto de "innovación", en este sentido desde el proyecto se analizan diferentes técnicas y metodologías de innovación y co-creación que les permiten a los ciudadanos contribuir a la resolución de problemas públicos, como por ejemplo Ciencia de datos, Ciencia del comportamiento, Simulación y modelaje, Inteligencia colectiva y Design thinking. [4]

La innovación puede hacerse tanto en el sector privado como en el sector público. En cuanto a la prestación de servicios públicos, hay que tener en cuenta que muchas veces los receptores no tienen la opción de dejar de consumir los servicios, por lo que deben adaptarse. Por esas razones es importante que, al llevar a cabo una innovación en el sector público, los ciudadanos puedan participar de los procesos desarrollando ideas conjuntas que permitan satisfacer las expectativas de un mayor número de interesados. Innovar en el sector público implica dar impulso a mecanismos de participación ciudadana, colaboración y transparencia.[3]

Además, el grupo participa del proyecto CAP4CITY "Fortalecimiento de la capacidad de gobierno para ciudades inteligentes y sostenibles" del Programa ERASMUS. [5]

En el cual hemos creado una metodología que guía y orienta el diseño de MOOC, aplicada para el diseño de 31 cursos. En particular hemos desarrollado los cursos de "Aplicaciones para Ciudades Inteligentes" y "Diseño y Co-Creación de Servicios Orientados al Usuario" y estamos trabajando junto a la Universidad Nacional del Sur en la definición de una Maestría en Ciudades Inteligentes Sostenibles.

Asimismo, el III-LIDI participa del Consorcio de I+D+I en Cloud Computing, Big Data & Emerging Topics, en el tópic de ciudades inteligentes sostenibles.

2- Calidad de datos.

En los últimos años, es cada vez más notoria la importancia de los datos en una organización. El avance digital está influyendo en todos los sectores y ha convertido a los datos en el recurso más potente y en un aspecto clave para la toma de decisiones; sin embargo, estos sectores no suelen disponer de recursos accesibles que evalúen la calidad de sus datos. Los datos se pueden ver afectados por factores negativos: ruido, valores perdidos, inconsistencias, un tamaño demasiado grande en cualquier dimensión, entre otros. Está demostrado que una baja calidad de los datos conduce a una baja calidad del conocimiento. Es por ello que el no contar con datos de calidad puede generar grandes consecuencias en la organización al momento de prestar un mejor servicio que cumpla con las exigencias del cliente. [6]

No solo las organizaciones producen datos de forma masiva, en una CIS, los ciudadanos, constantemente, generan información que luego podrá ser utilizada en la toma de decisiones sobre el desarrollo de esa ciudad y, pasado un tiempo, influirá en la vida de esos ciudadanos. Lograr que los ciudadanos tengan acceso a los datos y se les permita participar en su análisis para construir el desarrollo de la ciudad, es un importante aporte del ciudadano para la ciudad inteligente.

En este sentido uno de los puntos del proyecto, apunta a estudiar los estándares de calidad de datos, propuestos por la familia ISO/IEC 25000 (SQuaRE) en particular las normas ISO/IEC 25012 - "Data Quality Model" [7], [8], la cual define un conjunto de características

destinadas a evaluar la calidad de los datos, e ISO/IEC 25040 - "Evaluation process" [9], que define el proceso de evaluación a llevar a cabo.

3- Mejora de los servicios de gobierno digital en organismos públicos de gobierno

Se continúa trabajando en la generación de portales institucionales y además se está desarrollando una herramienta para la vinculación de datos abierto.

Desde el año 2003, el III-LIDI trabaja en aplicaciones relacionadas con la gobernanza digital para la elección de autoridades mediante el voto electrónico. Se destacan la definición e implementación de tres modelos (presencial, semipresencial y remoto) en distintos tipos de votaciones.

4- Mejora en los procesos de gestión de la Facultad de Informática

El III-LIDI, en coordinación con el área de Gestión Digital y Calidad de la Facultad de Informática, trabaja en este proyecto, con el objeto de analizar, definir y establecer un plan a ser aplicado a distintos procesos de la Gestión Universitaria.

La Facultad de Informática posee un Portal de Gestión Administrativa que provee a sus Alumnos, Docentes, No Docentes y Graduados una guía para la realización de trámites de los principales servicios que presta. De cada uno de estos trámites se brinda información estandarizada sobre: Descripción, Forma de Acceso, Requisitos y Contacto [10]

Desde el año 2011, la Facultad de Informática certifica los procesos: "Curso de Nivelación a Distancia para el Pre-Ingreso a la Facultad de Informática" y "Concursos de Profesores y Auxiliares Docentes Ordinarios" bajo la norma IRAM-ISO 9001:2015.

En el marco de la obtención de la certificación EUROINF para las licenciaturas de la Facultad de Informática en el año 2020, el grupo se encuentra desarrollando el proceso de implementación un Sistema Interno de la Calidad (SIGCFI) abarcando los cuatro ejes principales: formación, investigación, extensión y transferencia.

5- Mejora de Procesos de gestión en el desarrollo de software

La aplicación de la ISO/IEC/IEC 29110 “Perfiles de ciclo de vida para pequeñas organizaciones”; es una norma creada para PyMEs [11], que aumenta la competitividad y optimización para desarrollar un proyecto de software, con mayor control en cada etapa e impacta directamente en la calidad de software mejorando la imagen de la empresa y reduciendo riesgos en el desarrollo.

Se trabaja en un relevamiento de la aplicación de la norma en empresas de la región.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Análisis y estudio de normas y modelos orientados a la Calidad de Datos.
- Aplicación de los conceptos de calidad de Datos a la base de datos abierta de los casos de COVID publicada por los organismos gubernamentales
- Asistencia en el proceso de certificación en organismos públicos y privados según los requisitos de IRAM - ISO 9001.
- Desarrollo de Voto electrónico presencial y Voto por Internet. Arquitecturas adaptadas a la legislación vigente.
- Análisis, discusión y estudio de herramientas para la aplicación de normas de calidad relacionadas con certificación de servicios gubernamentales.
- Análisis, discusión y estudio de técnicas y herramientas de mejoras de proceso en el desarrollo de software.

- Análisis de propuestas de innovación y co-creación de servicios públicos
- Análisis de aplicaciones móviles para ciudades inteligentes.

3. RESULTADOS OBTENIDOS/ ESPERADOS

- Se mantiene el Portal de Gestión Administrativa para la Facultad de Informática UNLP, donde se reúne información y enlaces sobre los diferentes servicios que brinda la facultad.
- Se continua en el proceso de re-certificación bajo la norma IRAM-ISO 9001:2015: “Pre-Ingreso a Distancia de la Facultad de Informática” y “Concursos Docentes de la Facultad de Informática”.
- Se avanza en la implantación del SIGCFI
- Se avanzó en la tesis de doctorado “Modelo de madurez de los servicios de gobierno electrónico en el ámbito universitario”.
- Se aplicó el Voto por Internet en distintos organismos.
- Se avanza en el desarrollo de una herramienta para asistir en la evaluación de calidad de datos utilizando GQM
- Se realizaron acciones de consultoría y asesoramiento en organismos públicos y privados.

4. FORMACION DE RECURSOS HUMANOS

- El proyecto cuenta con becarios de Maestría de la UNLP en los temas del área.
- Se desarrollan tesis de doctorado y tesinas de grado en el área.
- Los integrantes de esta línea de Investigación participan en el dictado de asignaturas/cursos de grado/postgrado en la Facultad de Informática de la UNLP y en otras universidades del país. En particular, en la UNLP, se dicta la asignatura “Calidad de Sistemas de Software”

5. BIBLIOGRAFIA

Aires, “Ambiente para la ayuda a la mejora de procesos en las PyMEs .”

- [1] E. Estevez, N. V. Lopes, and T. Janowski, “Smart Sustainable Cities - Reconnaissance Study. Operating Unit ON Policy-Driven. Electronic Governance,” *United Nations University, Canada*, 2017. .
- [2] S. Esponda *et al.*, “Fortalecimiento de la calidad en procesos de software y Gobernanza Digital,” *Sedici.Unlp.Edu.Ar*, pp. 574–578, [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/77106>.
- [3] 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, [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/113243>.
- [4] R. Muñoz, A. Pasini, and P. Pesado, “Innovación en el Sector Público para Ciudades Inteligentes Sostenibles.”
- [5] E. Estevez *et al.*, “Desarrollo de Recursos Humanos para la Gobernanza de Ciudades Inteligentes Sostenibles – Proyecto CAP4CITY,” 2020.
- [6] S. Esponda, A. Pasini, and P. Pesado, “Modelo de evaluación de datos utilizando el enfoque GQM,” vol. 25000, pp. 418–428.
- [7] ISO, “ISO/IEC 25000:2014 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Guide to SQuaREtle,” 2014.
- [8] ISO, “ISO/IEC 25012:2008 □ Software engineering -- Software product Quality Requirements and Evaluation (SQuaRE) -- Data quality model,” 2008.
- [9] ISO, “ISO/IEC 25040:2011 □ Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Evaluation process,” 2011.
- [10] A. Pasini, R. Muñoz, P. Pesado, and I. De Investigación, “Gestor de contenidos orientado a portales de organismos gubernamentales,” pp. 0–3.
- [11] S. Esponda, P. Pesado, L. Plata, and B.

Tendencia actual de la utilización de herramientas de Gestión de Proyectos Software en las instituciones de Educación de Superior de Jujuy

Melina Soledad Audisio, Marcela Rueda, Rodríguez, Rafael,
Fernanda Cabezas, Nancy Toconas

Universidad Católica de Santiago del Estero – Departamento Académico San Salvador
Lavalle 333 San Salvador de Jujuy – 0388 4236139
melinasoledad.audisio@ucse.edu.ar

RESUMEN

En la gestión de proyectos software el trabajo se encuentra automatizado realizado por grandes equipos multidisciplinarios, por lo que es necesario usar herramientas que permitan la administración del trabajo, mejorando los procesos de control y su seguimiento.

La provincia de Jujuy no se encuentra alejada de esta realidad ya que un gran número de equipos de desarrollo utilizan herramientas de gestión de proyectos en sus trabajos, esto se pudo observar de la investigación realizada por éstos investigadores en el periodo 2018/2019. De la investigación elaborada se propuso la realización de un proyecto de investigación el cual tiene como propósito identificar si durante el transcurso de la formación académica de carreras afines en Proyectos de software de las instituciones educativas superiores de la ciudad de San Salvador Jujuy, se enseñan y se utilizan como recurso en las prácticas docentes las herramientas de gestión de proyectos. De modo de poder identificar los factores por los cuales realizan su elección, tipo de herramienta y el impacto que tienen dentro del proceso de enseñanza-aprendizaje.

Palabras clave: Gestión de Proyectos Software, Ingeniería de software, Automatización de Proyectos.

CONTEXTO

La presente investigación se enmarca en el proyecto “*Tendencia actual de la utilización de aplicaciones de gestión de proyectos, como herramientas de apoyo, en espacios académicos de las instituciones de educación de superior*”.

Este proyecto, está aprobado por el Consejo del Departamento Académico San Salvador de la Universidad Católica de Santiago del Estero, mediante Disposición N°329-2020. El mismo tiene estimada una duración de nueve meses (2020-2021) y se lleva a cabo por docentes investigadores y estudiantes de cursos superiores de la carrera de Ingeniería Informática de esta casa de estudios.

INTRODUCCIÓN

La gestión de proyectos es una de las disciplinas de mayor crecimiento en cualquier organización y también en el campo de la informática. Este trabajo se centra en el uso y enseñanza de herramientas de gestión de proyectos software en carreras de Informática, las cuales permiten organizar las tareas necesarias para el desarrollo de un proyecto software logrando un control exhaustivo del proceso de desarrollo, ayudando a mejorar la eficiencia y productividad de todo el equipo de trabajo. Para algunos autores, es importante mencionar que no se deben tomar como una herramienta de control, sino como un espacio

de trabajo que permite trabajar de manera ordenada y lineal [1].

Existen diversas herramientas para la gestión de proyectos por lo que hay que identificar cuál de ellas se adapta a las necesidades de lo que se desea realizar. Algunas de las herramientas de software más utilizadas son:

- Trello[2]: es una herramienta sencilla e intuitiva la cual organiza en forma de tableros el trabajo de forma que todos los usuarios puedan visualizarse de la misma manera. Permite además añadir tareas, asignarlas a uno o varios usuarios y marcarlas una vez se han completado. Todos pueden comentar el progreso y realizar anotaciones en cualquiera de las tareas, o bien adjuntar cualquier tipo de archivos.

- Slack[3]: es una herramienta que ahorra comunicaciones entre los integrantes del equipo de trabajo permitiendo enviar una gran cantidad de emails. La comunicación la realiza en forma de un chat interno donde se pueden enviar todo tipo de archivos y evitar así cadenas de email infinitas, permitiendo descongestionar la bandeja de entrada del correo haciendo más ágil y eficaz la comunicación interna en el equipo.

- Asana[4]: permite una gestión completa del proyecto posibilitando dividir el trabajo en subproyectos de manera ordenada, los gráficos del progreso de los proyectos y la posibilidad de integrar herramientas externas.

- Facebook Workplace[5]: es una herramienta sencilla e intuitiva, un clon de su plataforma original. Permite crear grupos de trabajo, públicos o cerrados. Además, se pueden crear y agendar reuniones, dar acceso a documentos, y realizar comunicaciones a nivel interno.

- Remember the link[6]: esta herramienta permite gestionar online una lista de tareas (propias o compartidas), realizando notificaciones y comunicaciones por mail si es necesario.

- GloBoard[7]: es una plataforma para controlar las tareas y proyectos. Permite administrar los tiempos, recursos humanos y económicos de la organización. GloBoard se ha

constituido como una herramienta de alto valor debido a su capacidad de sincronizarse con un repositorio de código de GitKraken, lo cual agiliza la actividad de los desarrolladores, y también de otros miembros del equipo como los del área de Control de Calidad y Diseño.

- JIRA[8]: es una herramienta de múltiples usos que permite gestionar y organizar los equipos de una empresa, además del seguimiento en tiempo real de los avances del proyecto.

- Redmine[6]: es una herramienta de gestión de proyectos de código abierto. Se trata de una solución altamente configurable que se adapta a cualquier tipo de empresa o proyecto. Está orientada a la coordinación de tareas y la gestión de flujos de trabajos además de facilitar y fomentar el trabajo cooperativo y la comunicación de los participantes.

Como antecedentes se encuentra el trabajo de Castelli, Quiroz y Fernanda Rodríguez Aleua [09], los cuales presentan los resultados y beneficios obtenidos de experiencias realizadas con alumnos de 2º año de la cátedra de Análisis de Sistemas de la UTN-FRC para la gestión del conocimiento, cuyo soporte teórico está basado en tres ejes: framework de Scrum, herramientas colaborativas y plataforma Moodle.

Realizando un sondeo dentro de la provincia, no se encontraron estudios relevantes referidos al tema.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

De acuerdo con la naturaleza y características del objeto de estudio, esta investigación se enmarca en una investigación de tipo cuantitativa, se utilizarán técnicas de recolección de datos tales como encuestas y/o entrevistas.

Actualmente se trabaja en la definición de las hipótesis para la confección del cuadro de operacionalización de variables realizándose un análisis de las mismas, sus dimensiones,

indicadores y categorías, para la confección y aplicación de los primeros cuestionarios en los cursos en donde se enseñe gestión de proyectos en carreras de informática en instituciones de educación superior.

OBJETIVOS

Este proyecto tiene estipulados los siguientes objetivos:

- *General:*
Identificar si utilizan herramientas para la gestión de proyectos en Institutos de Educación Superior, que dictan carreras afinesal Desarrollo de Software, en la ciudad de San Salvador de Jujuy.
- *Específicos:*
Indagar si se enseñan y utilizan herramientas para la administración de proyectos.
Determinar qué tipo de herramientas utilizan
Realizar un análisis del resultado de las etapas anteriores para determinar cuáles herramientas se enseñan y se utilizan como recurso en las prácticas docentes e indagar entre los estudiantes que trabajan si las conocen y utilizan en las empresas en las cuáles de desempeñan.
Estos resultados podrían permitir determinar la importancia de estos temas en los programas de estudios, permitiendo actualizarlos y extender o acortar su estudio.

ESTRUCTURA DEL EQUIPO DE INVESTIGACIÓN

El proyecto está siendo desarrollado por un equipo conformado por docentes investigadores del Departamento Académico San Salvador perteneciente a la Universidad Católica de Santiago del Estero (UCSE – DASS) y por alumnos colaboradores de la carrera Ingeniería Informática.

- Directora: Ing. Melina Audisio
- Investigadores:

- Ing. Marcela Rueda.
- Ing. Rodríguez, Rafael.
- Alumnos:
 - Cabezas, Fernanda
 - Toconas, Nancy.

REFERENCIAS

- [1] Diaz, Fernando(2019). Las 5 mejores herramientas de gestión de proyectos y tareas. Disponible en: <https://nosinmiscookies.com/mejores-herramientas-gestion-proyectos-y-tareas/> consultado en mayo 2020.
- [2] Trello, disponible en <https://trello.com/>. Consultado en mayo 2020
- [3] Slack., disponible en <https://slack.com/intl/es-ar/>. Consultado en mayo 2020
- [4] Asana, disponible en <https://asana.com/es>. Consultado en mayo 2020
- [5] Facebook Workplace, disponible en <https://www.facebook.com/workplace>. Consultado en mayo 2020
- [6] Redmine, disponible en <https://www.easyredmine.com/>. Consultado en mayo 2020
- [7] GloBoard , disponible en <https://www.gitkraken.com/boards>. Consultado en mayo 2020
- [8] JIRA, disponible en <https://www.atlassian.com/es>. Consultado en mayo 2020
- [9] Castelli, Silvia Lanza ; Quiroz, Valeria Ortiz; Aleua, Fernanda Rodríguez . Gestión Ágil Del Conocimiento Y Competencias Generadas En Un Entorno De Aprendizaje Colaborativo En El Ámbito Universitario. Disponible en: https://cadi.org.ar/wp-content/uploads/2018/09/4_CADI_y_10_CAEDI_per_115.pdf, consultada mayo 2020.

TOMA DE DECISIONES CIENTÍFICA EN LA INGENIERÍA DE SOFTWARE MEDIANTE INTELIGENCIA COMPUTACIONAL Y ANÁLISIS DE DATOS

Carlos Casanova, Manuel Chichi, Leonardo Hoet, Fernando Pereyra Rausch, Lucas Prado, Giovanni Daián Rottoli, Esteban Schab, Anabella De Battista

Grupo de Investigación sobre Inteligencia Computacional e Ingeniería de Software, Departamento Ingeniería en Sistemas de Información, Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional
Entre Ríos, Argentina

{casanovac, chichim, hoetl@frcu.utm.edu.ar, pereyraf, pradol, rottolig, schabe, debattistaa}@frcu.utm.edu.ar

RESUMEN

La adopción de herramientas formales que complementen la experiencia y el buen juicio en las distintas actividades de un proceso de desarrollo de software todavía es un pendiente dentro la industria del software. La falta de conocimientos respecto de enfoques realistas para resolver problemas de la IS y la falta de herramientas software que auxilien a los tomadores de decisiones utilizando tales enfoques son dos carencias que pueden explicar las dificultades en esta adopción. Las líneas de investigación aquí propuestas tienden a suplir ambas. Para esta tarea se propone la utilización de tanto técnicas comprendidas en lo que se conoce como Inteligencia Computacional (IC), dentro de las cuales se encuentran la teoría de conjuntos difusos, las redes neuronales y la computación evolutiva, como también de herramientas de la Ciencia de Datos, incluyendo técnicas de aprendizaje automático, estadísticas y visualización de datos, entre otros. Estas técnicas son capaces de brindar la flexibilidad necesaria para crear métodos y modelos que sean tolerantes a la imprecisión, la falta de información y la aproximación, características que le son propias a los contextos de decisión en la IS.

Palabras clave: Ingeniería de Software, Optimización, Inteligencia Computacional, Ciencia de Datos, Preferencias.

CONTEXTO

El presente trabajo se desarrolla en el ámbito del Grupo de Investigación sobre Inteligencia Computacional e Ingeniería de Software (GIICIS), perteneciente al Departamento

Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Concepción del Uruguay.

1. INTRODUCCIÓN

El Vocabulario de la Ingeniería de Sistemas y Software de ISO/IEC/IEEE (SEVOCAB) define a la ingeniería de software como 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. El Software es un bien intangible que cumple un doble rol. Es un producto y al mismo tiempo es el vehículo para entregar el producto más importante de nuestro tiempo: la información.

El software se desarrolla o modifica con intelecto: no se manufactura en el sentido clásico. Los costos del software se concentran en la ingeniería. Esto significa que los proyectos de software no pueden administrarse como si fueran proyectos de manufactura.

En este contexto, un proceso de software es un conjunto de actividades relacionadas que conducen a la producción de un producto de software [1]. Sin embargo, no es una prescripción rígida de cómo elaborarlo. Por el contrario, debe ser ágil y adaptable (al problema, al proyecto, al equipo y a la cultura organizacional). Por tanto, un proceso adoptado para un proyecto puede ser significativamente distinto de otro adoptado para otro proyecto [1]. El proceso de software es complejo y, como todos los procesos intelectuales y creativos, confía en que las personas tomen decisiones con criterio.

1.1. Problemas que surgen en el contexto de la IS

La comunidad científica viene realizando esfuerzos en la definición realista de problemas propios de la ingeniería de software. El primer paso en cualquier metodología científica es la definición del problema. Sin una definición apropiada del problema, se corre el riesgo de encontrar una solución elegante para un problema que no existe. Características comunes que pueden identificarse son las siguientes:

- Hay una necesidad de **balancear objetivos que compiten** entre sí, es decir, se clasifican como problemas de optimización multi-objetivo.
- Existe 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. Por caso, puede darse que no se cuenta con suficiente información para estimar con un grado de certeza aceptable ciertos parámetros. También puede darse la falta de acuerdo entre varios expertos al realizar una estimación, o que la misma sea establecida de forma imprecisa, por ejemplo, puede requerirse que el acoplamiento entre dos módulos sea “bajo” o que ciertos *stakeholders* resulten “muy favorecidos” en la próxima versión.
- Existen muchas soluciones potenciales, usualmente caracterizadas por una **explosión combinatoria** de las variables de decisión.

1.2. 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 [2]. Como su nombre lo indica, el objetivo de esta disciplina implica “investigar sobre las operaciones”. En consecuencia, esta disciplina se aplica a la problemática relacionada con la conducción y la coordinación de actividades en una organización.

En cierto sentido, la IO involucra la investigación científica creativa de las propiedades fundamentales de las operaciones. Sin embargo, es más que esto. 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 [2].

1.3. 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 [3]. Estos múltiples métodos no son competitivos entre sí, sino que son complementarios y pueden ser utilizados juntos para resolver un problema dado.

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.

1.4. Ciencia de datos

La ciencia de datos, minería de datos, descubrimiento de conocimiento y otros términos similares, buscan agrupar un conjunto técnicas, herramientas y metodologías relacionadas con diferentes

aspectos del procesamiento de datos, tales como la recolección, la limpieza, el procesado, análisis y extracción de patrones implícitos a partir de los mismos [4].

Si bien una gran variedad de las técnicas utilizadas en el marco de este gran paraguas conceptual se pueden catalogar como parte de la inteligencia computacional, se nutre además de otras disciplinas como la estadística y la visualización de datos para extraer información valiosa y novedosa de repositorios de datos para asistir a la toma de decisiones [5].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1. Formulación realista de problemas de la IS como problemas de búsqueda/optimización

A nivel de industria de software mundial, ninguna técnica formal ha tenido un impacto significativo [6]. 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 [7]. En consecuencia, se busca formular modelos matemáticos utilizando mecanismos formales de captación de incertidumbre (como la teoría de conjuntos difusos, o la de posibilidades) que resulten adecuados para la realidad de un proceso de desarrollo de software. El primer paso, por tanto, consiste en formular el problema como uno de búsqueda/optimización. 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. Mecanismos de captación de preferencias

Existen múltiples enfoques de optimización multiobjetivo, entre ellos los basados en preferencias [8]. Estos suponen que puede incluirse información del tomador de decisiones para guiar la búsqueda hacia regiones prometedoras del espacio de soluciones. Idear mecanismos adecuados para captar esta información no es una tarea trivial y se trabaja en el estudio de los mecanismos disponibles, y el diseño e implementación de nuevos mecanismos para mejorar principalmente la usabilidad de los métodos, siendo los componentes analizables de estos mecanismos principalmente tres: la cantidad de interacciones con el usuario, qué información se le solicita, y en qué momento del proceso de optimización.

2.3. Ingeniería de Software basada en búsquedas conducida por los datos

Los datos producidos por los procesos de la ingeniería de software constituyen una fuente de conocimiento que puede ser explotada y utilizada para la resolución de los problemas, sirviendo de soporte a los algoritmos de optimización y actividades llevadas a cabo en el contexto de la Ingeniería de Software basada en Búsqueda (ISBB) [9], [10]. Son ejemplos de esta sinergia la comprensión de las preferencias de los tomadores de decisiones cuando estas no se encuentran explícitas a partir de las interacciones de los mismos con los algoritmos de búsqueda, o la posibilidad de explicar los resultados obtenidos por los algoritmos de optimización para su mejor comprensión y análisis.

En consecuencia, se propone la integración de algoritmos y técnicas del área del análisis de datos en los mecanismos utilizados para la exploración de las soluciones de la ingeniería de software basada en búsqueda para incrementar la usabilidad de los mismos.

2.4. 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.

2.5. 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 [15]. Para esto se desarrolló un prototipo en Octave que todavía debe ser mejorado.

Lograr un primer prototipo de interfaz usable para usuarios no expertos en este tipo de técnicas es un objetivo esperado. En esta línea un primer prototipo de una aplicación para la exploración de frentes de Pareto ha sido construido haciendo uso de la librería Dash para Python 3.7. Este prototipo 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. Todo este proceso se puede realizar de forma iterativa, facilitando la exploración de las soluciones y consolidando las preferencias del tomador de decisiones. Se busca continuar con el desarrollo y validación de este prototipo software. Además se proyecta poder construir *plugins* o mecanismos de interoperabilidad con *suites* existentes usadas en la industria.

Finalmente, se espera contar al final del proyecto con una colección de modelos altamente cohesivos que brinden información para facilitar la toma de decisiones relativas a distintos problemas de la Ingeniería de Software.

4. FORMACIÓN DE RECURSOS HUMANOS

La investigación presentada constituye las líneas fundacionales de un nuevo grupo de investigación dentro de la UTN-FRCU, el GIICIS. Dos investigadores se encuentran realizando su tesis de doctorado. Además participan en el proyecto dos becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación, dos alumnos realizan su práctica supervisada y uno su trabajo de fin de carrera en el contexto de esta investigación. En el marco de este proyecto ya se han defendido exitosamente dos prácticas profesionales supervisadas.

5. BIBLIOGRAFÍA

- [1] I. Sommerville, *Software Engineering*, 9.^a ed. Pearson Education, Inc., 2011.
- [2] F. S. Hillier y G. J. Lieberman, *Introducción a la Investigación de operaciones*. McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V., 2010.
- [3] J. M. Keller, D. Liu, y D. B. Fogel, *Fundamentals of computational intelligence: Neural networks, fuzzy systems, and evolutionary computation*. John Wiley & Sons, 2016.
- [4] C. C. Aggarwal, *Data Mining*. Cham: Springer International Publishing, 2015.
- [5] J. Han, M. Kamber, y J. Pei, *Data Mining: Concepts and Techniques*. Elsevier Inc., 2012.
- [6] 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.
- [7] 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, n.º c, pp. 51-52.
- [8] T. N. Ferreira, S. R. Vergilio, y J. T. de Souza, «Incorporating user preferences in search-based software engineering: A systematic mapping study», *Information and Software Technology*, vol. 90, pp. 55-69, 2017.
- [9] V. Nair *et al.*, «Data-Driven Search-based Software Engineering», en *2018 IEEE/ACM 15th International Conference on Mining Software Repositories (MSR)*, 2018.
- [10] T. Colanzi, W. Assunção, ... S. V.-I. and, y U. 2020, «The Symposium on Search-Based Software Engineering: Past, Present and Future», *Information and Software Technology*, vol. 127, 2020.
- [11] B. S. Mishra y S. Dehuri, «Parallel computing environments: A review», *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, vol. 28, n.º 3, pp. 240-247, may 2011.
- [12] A. Das y B. K. Chakrabarti, *Quantum annealing and related optimization methods*. Springer Science & Business Media, 2005.
- [13] C. Casanova *et al.*, «Aproximación del Frente Pareto-Óptimo de un Problema NRP Bi-Objetivo mediante un Algoritmo basado en Enjambres de Partículas», en *Anales VI Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación*, 2018.
- [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] M. L. Gabioud y C. Casanova, «Priorización en Ingeniería de Requerimientos con Preferencias Difusas», en *VII Congreso Nacional de Ingeniería Informática / Sistemas de Información*, 2019.

VERIFICACIÓN DE ENTREGABLES DEL PROCESO DE ELICITACIÓN EN PYMES ARGENTINAS DE LA REGIÓN CENTRO

Viviana Bourdetta, Carlos Casanova, Rafael Blanc, Karina Cedaro, Lourdes Pralong, Laura Elena Ríos, Rossana Sosa Zitto

Facultad de Ciencia y Tecnología, Universidad Autónoma de Entre Ríos
vivianab@gmail.com; carlos.casanova16@gmail.com; rafaellujanblanc@yahoo.com.ar;
karinacedaro@gmail.com; mdlpralong@gmail.com; elenrios2@gmail.com;
rossana.e.sosa.zitto@gmail.com

RESUMEN

La calidad de los entregables de la Elicitación de Requisitos es causa de gran parte de los desvíos significativos en los proyectos de desarrollo de software por el alto impacto de sus consecuencias sobre el producto final. Aun así las empresas no invierten los recursos suficientes, cuando ello les permitiría reducir esfuerzos, costos y obtener ventajas en un mercado altamente competitivo.

Si bien la Ingeniería de Requisitos intenta con sus aportes mejorar la situación, el alcance de sus propuestas no siempre aplica a las PyMES desarrolladoras de nuestra región, industrias que se enfrentan a diario con falta de recursos, habilidades y experiencia en su búsqueda por crear software de calidad y sobrevivir en el mercado.

En este marco el presente trabajo propone una herramienta basada en la aplicación de técnicas de mejora que permitirán complementar el rol de los ingenieros de requerimientos en lo que hace a la verificación de los resultados de este proceso. Se trabaja en el desarrollo y construcción de la misma para ser procesada por una aplicación de software a efectos de facilitar la comprensión del problema manifestado por el usuario y reducir la brecha conceptual entre los resultados del proceso de elicitación y su modelado.

Palabras clave: *Ingeniería de software, Ingeniería del Conocimiento, Ingeniería de Requisitos, Inteligencia Artificial. Calidad y Mejora Continua de Requisitos*

CONTEXTO

El presente trabajo se desarrolla en el ámbito del Grupo de Investigación de Ingeniería de Software (GIISW), perteneciente a la Sede Concepción del Uruguay de la Facultad de Ciencia y Tecnología dependiente de la Universidad Autónoma de Entre Ríos.

1. INTRODUCCIÓN

Actualmente, la industria nacional del sector SSI (Software y Servicios Informáticos) tiene como unos de sus retos fundamentales implementar aplicaciones que sean entregadas a tiempo, que no involucren presupuestos elevados y que satisfagan las necesidades del usuario, utilizando para ello metodologías y herramientas que guíen el proceso de desarrollo de Software [1].

El proceso de desarrollo de software es la aplicación de un conjunto de actividades, acciones y tareas que se ejecutan para crear algún producto de trabajo. Se obtiene un producto de alta calidad si estas actividades, acciones y tareas, se combinan con procedimientos y técnicas de la ingeniería de software. [2]. Por ello otra exigencia paralela que ocupa a la industria de software es la calidad de los recursos usados en sus procesos, debido a la directa implicancia sobre los resultados. Este tema día a día genera mayor interés, como lo demuestra el último informe presentado en 2019 por la OPSSI (Observatorio permanente de la industria del software y servicios informáticos) en el cual el

42% de las firmas realizaron mejoras en la calidad en el proceso de desarrollo [3].

Existen diversos estudios respecto de los fracasos en la industria del software, tal es el caso del reporte del estudio presentado por la consultora internacional *Standish Group*, en el que se cita que tan solo el 32% de los proyectos de desarrollo de software se pueden considerar exitosos; el 44% se entregaron fuera de plazo, excedieron su presupuesto y no cubrieron la totalidad de las características y funcionalidad pactada; y el 24% de los proyectos fueron cancelados. En este mismo estudio se puntualiza que el principal factor para el fracaso de un proyecto de desarrollo de software radica en la mala calidad de los requisitos y la definición poco clara de los mismos [4].

Por definición la Ingeniería de Requisitos es la ciencia y disciplina que tiene por objeto el análisis, documentación y validación de las necesidades y/o requerimientos de las partes interesadas para el desarrollo de un sistema. [5]. Su objetivo es desarrollar una especificación completa, consistente y no ambigua de los requisitos, la cual servirá como base para acuerdos comunes entre todas las partes involucradas y en dónde se describen las funciones que realizará el sistema. Tiene cuatro etapas fundamentales: la elicitación, el análisis, la especificación y la validación. Durante ellas constantemente se requiere colaborar, coordinar y comunicar con los distintos *stakeholders*, entre sí, pero manteniendo distintas perspectivas y puntos de vista lo que plantea grandes retos de comunicación.

A pesar de la diversidad de las empresas y el perfil de los ingenieros de software, se verifican problemas por el insuficiente y no correctamente elaborado proceso de elicitación de los requisitos. Los requisitos se describen mediante el uso de texto, que está orientado al cliente y no es apropiado para el desarrollador. Las descripciones ambiguas y los requisitos fragmentados en varios artefactos también socavan la comprensión. Los desarrolladores señalan que se requiere mucho esfuerzo para aclarar dudas cuando la especificación de requerimientos es insuficiente o se enfoca solo

en los requisitos funcionales aumentando el esfuerzo de codificación, prueba y mantenimiento. [6]

Actualmente se ha demostrado que aplicar un modelo de gestión de conocimiento elaborado permite mejorar la calidad del desarrollo de software reduciendo la cantidad de ciclos de prueba, el tiempo de atención del desarrollo y la cantidad de reversiones pos despliegue en producción [7]. La calidad de los productos de software se obtiene al implementar prácticas de calidad en el proceso de desarrollo [8][9][10]. La base de dicho proceso son los requisitos, ya que determinan la calidad del producto final, razón por la cual la Ingeniería de Requisitos (IR) es una etapa crucial dentro del proceso de desarrollo de software (SW) [11].

Implementar una gestión adecuada de los procesos permite a las empresas no solo reducir esfuerzos, sino además obtener ventajas competitivas frente a otras empresas. Esto no resulta sencillo especialmente cuando se trabaja en ambientes extremadamente dinámicos, los cuales conllevan un alto nivel de variabilidad. Es por ello que los constantes cambios que surgen en dichos entornos crean la necesidad de adaptar continuamente estos procesos. Adaptarse y soportar estos cambios resulta clave para el éxito de las organizaciones [12].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se realizan estudios de casos, utilizando diferentes métodos de recolección basados en encuestas con diferentes cuestionarios, lo que permite realizar un análisis descriptivo de la situación actual de las actividades realizadas en la etapa de elicitación de requerimientos en empresas pymes que desarrollan software en la zona centro de Argentina. El análisis de tal relevamiento nos permite obtener conclusiones respecto de los entregables de este proceso. Luego de analizar los resultados obtenidos se detecta, al igual que en estudios citados en este trabajo, la existencia de problemas para satisfacer la necesidad de estructurar y categorizar los datos provenientes del proceso de elicitación de requisitos a efectos de facilitar la comprensión del problema manifestado por el usuario. Esto, dicho en otros términos, consiste en formalizar los requisitos mediante técnicas de representación del conocimiento [13]. Éstas junto con el diseño de metodologías de gestión del conocimiento [14] son capaces de resolver problemáticas presentes en todas las etapas de la IR.

Por caso, en el proceso de Ingeniería de Conocimiento involucrando un experto (este es, el cliente, el experto en el negocio), un ingeniero de requisitos puede proyectar un software en una pantalla y trabajar junto al cliente en la elaboración del documento de requisitos, utilizando lenguaje visual basado en mapas conceptuales. En esta línea se propone el desarrollo de herramientas de software que soporten las tareas de los ingenieros de requisitos en cualquiera de sus etapas, incluyendo el uso de sistemas expertos y redes semánticas.

La aplicación de técnicas de procesamiento y análisis de lenguaje natural es otra línea de investigación, sobre todo para aquellas pequeñas empresas que no poseen un área de gestión de requisitos, o que son escépticos respecto de los beneficios que puedan aportar enfoques más sintácticos en el análisis de los mismos. En este sentido, se estudian técnicas provenientes de la inteligencia computacional, como las Redes Neuronales y la Lógica

Difusa, en el análisis de documentos de requerimientos, especialmente colecciones de historias de usuario, con el fin de establecer criterios de calidad (o ausencia de esta), como ambigüedad, conflictos intra documentales, duplicación y solapamiento, plausibilidad de la estimación, entre otros [15, 16, 17]. Este análisis es capaz de evidenciar falencias en los documentos de requerimientos y constituir un insumo para la posterior recomendación de acciones tendientes a la mejora, por caso, la aplicación de técnicas específicas de elicitación.

3. RESULTADOS OBTENIDOS Y ESPERADOS

En esta sección se detallan algunos resultados obtenidos a partir de un relevamiento por medio de cuestionarios diseñados con el objeto de relevar y describir las técnicas actuales de la toma de requisitos. Se contactaron 50 empresas de la región, las encuestas se realizaron a través de videoconferencia, enviándolas por correo electrónico, o de manera presencial. Los encuestados fueron, en mayor parte, directores y/o líderes de proyectos, con una antigüedad de más de cinco años en la empresa.

La región centro refiere a las provincias de Buenos Aires, Santa Fe, Córdoba y Entre Ríos. Respecto de la antigüedad de las empresas en el mercado, la más antigua fue fundada en el año 1982, y la más reciente en el año 2015. En promedio tienen 10 clientes, aunque se diferencian dos empresas con 60 y 200 clientes respectivamente. En cuanto al personal empleado, la empresa con mayor cantidad tiene 220 y la de menor cantidad 4 personas. Se relevaron tanto aspectos vinculados a los procesos de Ingeniería de requisitos, como la importancia e impacto que tiene esta etapa en los proyectos de desarrollo.

Del relevamiento surge que la etapa de ingeniería de requisitos se inicia a través del contacto con los clientes mediante reuniones formales, en todos los casos los requisitos son validados por el cliente y se modelan con lenguajes específicos, mayormente historias de usuario. Al consultarles por el tiempo que les

destinan a la etapa de análisis las respuestas son dispares van desde una empresa con mucha experiencia que le destina 10% hasta otras que le destinan un 40% de la duración total del proyecto.

El 92% de las firmas entrevistadas afirman no tener un área de requisitos estable como parte de su estructura formal, además un 15% de las mismas no cuenta con personal dedicado a esta función. Lo más frecuente es contar con entre 4 y 5 personas dedicadas a esta función, lo que se da en el 53% de los casos. Finalmente, el 31% de las firmas posee 1 o 2 personas asignadas a este rol. La formación de las personas que toman requisitos en las empresas es de carreras universitarias relacionadas a la informática ya sea que hayan terminado la misma (70%) o con la misma incompleta (23%) y finalmente en menor medida se da el caso de perfiles universitarios completos en carreras no relacionadas a la informática en un 15% de los casos. La decisión de no tener un área de requisitos no se relaciona con la falta de conocimiento sobre el impacto de los mismos para el éxito del proyecto, ya que el 80% de la muestra posee empleados con título universitario relacionado al área sistemas/informática que reconocen claramente la importancia de contar con esta área y con recursos asignados a tal fin.

El 77% de la muestra dijo que la metodología para manejar los requisitos varía dependiendo del proyecto y cliente. El 62% de las firmas tienen o tuvieron en algún proyecto problemas al momento de elicitar los requisitos de sus clientes. Todos utilizan sistemas de trazabilidad de los datos, y coinciden en que este ha mejorado con el tiempo y la experiencia adquirida.

En cuanto a la metodología utilizada para llevar adelante el proyecto se observa que *Agile Unified Process* es la considerada más flexible y utilizada en la mayoría de los proyectos.

Los datos obtenidos de las entrevistas permitieron detectar patrones comunes y distintivos en cuanto al tratamiento y seguimiento de los requisitos. Se observa la necesidad de implementar estrategias que garanticen el éxito de esta etapa, enfocando en

propuestas de mejora a la metodología, las técnicas y las herramientas utilizadas, así como también que incrementen el compromiso y la participación de todos los involucrados.

La propuesta, nace a partir del análisis de las necesidades y limitaciones de las PyMEs y se constituye en base a aquellas actividades y prácticas que han demostrado que aportarán valor en los casos analizados. En este sentido, se plantea un proceso ágil, sin demasiados requisitos para su implementación, con la mínima documentación posible y que genere productos de trabajo reutilizables para su despliegue en los diferentes proyectos de la mismas.

Se espera poder desarrollar una herramienta que a partir del establecimiento de niveles de calidad realice un análisis semántico de conjuntos de historias de usuario, midiendo el nivel de calidad en aspectos relevantes de los requisitos planteados al tiempo de sugerir acciones a llevar a cabo para la mejora de tal entregable. La aplicación continua de esta herramienta mejoraría la calidad de los requisitos relevados en la etapa de elicitación, disminuyendo los desvíos y aumentando la probabilidad de éxito en los proyectos de desarrollo.

4. FORMACIÓN DE RRHH

El impacto sobre las capacidades institucionales estará garantizado por el contacto con los grupos de investigación del país y el exterior más consolidados.

El proyecto permitirá dar continuidad a la línea de investigación en la temática calidad de software iniciada en el año 2013 y contribuirá a la formación como investigadoras de las docentes Ing. Elena Ríos, Ing. Viviana Bourdetta, Mg. Karina Cedaro y Lic. Lourdes Pralong. Participan en el proyecto alumnos becarios de la carrera Licenciatura en Sistemas de Información, y se prevé incorporar nuevos becarios para que inicien su formación en la investigación. Además, se hará transferencia de los resultados a la carrera licenciatura en sistemas de información en las materias relacionadas al proceso de toma de requisitos.

5. CONCLUSIONES Y PERSPECTIVAS

La clave para el éxito de un proyecto de desarrollo de software está vinculado directamente con la etapa de elicitación de requisitos. La obtención de requisitos correctos que lleven a productos acordes a las necesidades de los clientes de las empresas reside en la calidad del proceso de elicitación.

No obstante, la mayoría de las firmas relevadas no tienen personal dedicado en forma permanente a este proceso y es un rol de tipo flexible, y por otro lado gran parte de las empresas expresan tener desvíos al momento de trabajar en las actividades vinculadas con los requisitos. Este problema genera la necesidad de un proceso de verificación de requisitos que evite pérdidas económicas por no conformidades y retrabajos, ya que los desvíos se encuentran no solamente en la etapa de elicitación, sino también en etapas posteriores cuando se ha invertido más recursos en el proyecto de desarrollo.

Este trabajo plantea la construcción de una herramienta capaz de medir niveles de calidad en aspectos relevantes de los entregables al tiempo de sugerir acciones a llevar a cabo para la mejora de los mismos. Esta herramienta será el pilar fundamental de la metodología para la gestión y mejora continua de los requisitos.

6. BIBLIOGRAFÍA

- [1] F. Barletta, M. Pereira, and G. Yoguel, "Impacto de la Política de Apoyo a la Industria de Software y Servicios Informáticos," 2014.
- [2] Reporte anual sobre el Sector de Software y Servicios Informáticos de la República Argentina. OPSSI. 2019-2020.
- [3] A. Hernández Paez, Y. Fuentes Castillo, J. Anias Santos. Buenas prácticas para el desarrollo de los requisitos basado en componentes utilizando el modelo de capacidad y madurez integrada. 2020.
- [4] T. Clancy, "The standish group chaos report," 2014.
- [5] Kheirkhah, E., & Deraman, A. (2008, August). Important factors in selecting requirements engineering techniques. In *2008 International Symposium on Information Technology* (Vol. 4, pp. 1-5). IEEE.
- [6] Medeirosa, J.; Vasconcelos, A.; Silva, C. y Goulão, M. (2018). Quality of software requirements specification in agile projects: A cross-case analysis of six companies. *Journal of Systems and Software*. Volume 142, August, pp. 171-194.
- [7] Calderón Romero, M. E., Modelo de gestión del conocimiento para mejorar la calidad del desarrollo de software. 2020.
- [8] B. A. Nuseibeh and S. M. Easterbrook, "Requirements Engineering: A Roadmap," in *ICSE 2000 Proceedings of the Conference on The Future of Software Engineering*, 2010.
- [9] R. Sosa Zitto and R. Blanc, "Buenas prácticas de Scrum para alcanzar niveles de calidad en pymes de desarrollo de software," in *XVI Workshop de Investigadores en Ciencias de la Computación*, 2014, pp. 490–494.
- [10] R. Sosa Zitto, R. Blanc, L. Pralong, C. Álvarez, and S. Galáz, "El desafío de producir software de calidad aplicando prácticas de CMMI para las pymes de Concepción del Uruguay, Entre Ríos," in *CoNaIISI 2013*.
- [11] C. M. Zapata, G. L. Giraldo, and J. E. Mesa, "Una propuesta de metaontología para la educación de requisitos," *Ingeniare. Rev. Chil. Ing.*, vol. 18, no. 26–37, 2010.
- [12] V. A. Ledesma, G. D.S. Hadad, J. H. Doorn, J. P. Mighetti, N. A. Bedetti, M. C. Elizalde. "Evolución de los Factores Situacionales durante el Proceso de Requisitos". *XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste)*.
- [13] P. Pytel et al., "Propuesta de Aplicación de Técnicas de Representación de Conocimiento en el Análisis de Requisitos Software,". *Actas 1er Seminario Argentina-Brasil de TIC*, 2011.
- [14] J. Capote, C. J. Llanten Astaiza, C. J. Pardo Calvache, A. de J. González Ramírez, and C. A. Collazos, "Gestión del conocimiento como apoyo para la mejora de procesos software en las micro, pequeñas y medianas empresas," *Rev. Ing. e Investig.*, vol. 28. 2008.
- [15] Kenter, T., & De Rijke, M. (2015, October). Short text similarity with word

embeddings. In Proceedings of the 24th ACM international on conference on information and knowledge management (pp. 1411-1420).

[16] Azmi-Murad, M., & Martin, T. P. (2004, August). Using fuzzy sets in contextual word similarity. In International Conference on Intelligent Data Engineering and Automated Learning (pp. 517-522). Springer, Berlin, Heidelberg.

[17] Lucassen, Garm; Dalpiaz, Fabiano; Van der Werf, Jan Martijn; Brinkkemper, Sjaak. (2016). Improving agile requirements: the Quality User Story framework and tool. Requirements Engineering. 21.

Innovación en Sistemas de Software

Captura multimodal para la inferencia de estados emocionales aplicados a contextos de computación afectiva

Jorge Ierache , Iris Sattolo , Gabriela Chapperon, Rocío Ierache, Fernando Elkfury, Carlos Barrionuevo , Facundo Nervo

Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica (ISIER)

Escuela superior de Ingeniería, Informática y Ciencias Agroalimentaria

Secretaría de Ciencia y Tecnología Universidad de Morón Cabildo 134, Buenos Aires, Argentina

{jierache,isattolo,gchapperon}@unimoron.edu.ar

Resumen

En este trabajo se enuncian el contexto de desarrollo en materia de I+D aplicada a la captura multimodal de emociones en diversos ámbitos de aplicación (virtuales, simuladores, robots, gastronomía), los proyectos vinculantes y los resultados publicados, como así también las líneas de I+D+I, la formación de RRHH. Finalmente, se expresa el reconocimiento al Ing. Hugo Rene Padovani, quien apoyara la formación de docentes investigadores del área informática en la UM .

Palabras clave: Computación Afectiva, Estado emocional, captura multimodal.

Contexto

Esta investigación aplicada se desarrolla en el contexto del Proyecto de Investigación Científica Tecnológica Orientado (PICTO) aprobado por la Agencia Nacional de promoción de la investigación, el desarrollo tecnológico y la innovación (ANPCyT), denominado “Influencias del estado biométrico emocional de personas interactuando en contextos de entornos simulados, reales e interactivos con robots” PICTO UM 2019-00005. El mismo se desarrolla dentro del Instituto de Sistemas Inteligentes y Enseñanza

Experimental de la Robótica de la Universidad de Morón (ISIER-UM) y está auspiciado por la Secretaría de Ciencia y Tecnología con una duración de tres años. Este trabajo se sustenta sobre las bases iniciales de las investigaciones realizadas en el marco de los proyectos: “Influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales” Ping/17-03-JI-004,(2017), el proyecto denominado “Explotación de datos EEG y parámetros fisiológicos de usuarios interactuando en contextos virtuales” DC-019 80020190100007 UM (2019) , el proyecto “Valoración Emocional Multimodal aplicada en contextos gastronómicos” PIO 80020190300016 UM (2020).

Introducción

El estudio de los estados afectivos en las personas es un campo que se aplica en diferentes realidades tales como la medicina, educación, conducción de automóviles y otros, con el fin de solucionar ciertos problemas aparejados por el dominio. A lo largo de la historia, la intuición y los trabajos de muchos profesionales de diversas disciplinas han indicado la importancia de las emociones, así como la necesidad de conocerlas y aprender a relacionarnos con éstas de la forma más coherente posible. Como ejemplo se puede citar a Darwing en 1872, James y Lang en 1884 y otros pensadores, quienes ya en su época

argumentan sus teorías sobre cómo se producen las emociones. En el campo de la computación se habla de proveer a las máquinas de emociones, originando un nuevo campo denominado computación afectiva. S. Baldasarri [1] plantea que los sistemas “afectivos”, deben ser capaces de capturar y reconocer los estados emocionales del usuario a través de mediciones sobre señales generadas en la cara, la voz, el cuerpo, o cualquier otro reflejo del proceso emocional que se esté llevando a cabo. Albert Mehrabian en sus investigaciones afirma que en una comunicación la palabra hablada solo contiene el 7% del significado, mientras que el tono de voz un 38% y las expresiones faciales nada menos que el restante 55% [2]. Existen distintos enfoques para representar las emociones, el enfoque dimensional [3] y el enfoque categórico. El enfoque categórico, desarrollado inicialmente por el psicólogo Paul Ekman [4], afirma que existe un conjunto de seis emociones básicas y universales que no están determinadas por las culturas. Luego actualizó este conjunto a siete emociones, sumando el desprecio. Estas son: a) enojo, b) asco, c) felicidad, d) miedo, e) tristeza, f) sorpresa, g) desprecio. Por otro lado, el enfoque dimensional establece que los estados afectivos se distribuyen en un espacio continuo, cuyos ejes dimensionales indican una cuantificación de una característica. A pesar de que actualmente no existe un consenso acerca de lo que es la definición de emoción, si existe un acuerdo en que las emociones pueden clasificarse en tres dimensiones continuas: 1) Valencia (Valoración): un constructo bipolar, que va de agradable a desagradable, 2) Excitación (Arousal): que va de calmado hasta activado (o excitado), y 3) Dominancia (Control). Los métodos utilizados en la computación afectiva se basan en el procesamiento de los datos recopilados y etiquetados obtenidos mediante la ejecución de experiencias en laboratorios, donde se pueden

obtener etiquetas (definidas con categorías predefinidas o de forma dimensional) de diferentes maneras: durante la interacción con el sistema o retrospectivamente, o por el investigador durante o después de la interacción. De acuerdo con los tipos de datos que se utilizan en los sistemas propuestos pueden clasificarse en sistemas unimodales, los que exploran una sola fuente de datos, y multimodales, los que combinan dos o más fuentes de datos. Como trabajos unimodales se puede citar a: en [5] se captura el rostro a través de videos y en [6],[7] captura de rostro con imágenes. Como trabajos multimodales se puede citar a: en [8] agregan a la captura del rostro la posición de la cabeza, en [9] utilizan cuestionarios, teclado y micrófono, en [10] captura de teclado y Electroencefalograma (EEG). Para deducir el estado emocional de un individuo en un contexto de captura multimodal (de ahora en más MM) se tiene que registrar, simultáneamente, diversa información biométrica. En este contexto se está trabajando la construcción de un Framework MM en el cual se combinen distintas características, distintos sensores y distintas estrategias para una mayor fidelidad de los datos capturados. En trabajos previos del ISIER-UM se trabajó en: [11], [12], [13], [14] control de Robots, en [15] el control de artefactos aplicados a la domótica, y en [16] la lectura emocional del usuario, enfocando la lectura de la excitación y meditación. Se experimentó con BCI (Brain Computer Interface), en particular con EMOTIV [17], integrando la respuesta de biopotenciales eléctricos de individuos para el control de robots a través de Electromiograma (EMG), EEG y Electrooculograma (EOG), las cuales son bioseñales eléctricas generadas por los patrones de actividad de los músculos, el cerebro y los ojos del usuario respectivamente.

Líneas de Investigación, Desarrollo e Innovación

La hipótesis que guía este proyecto es: La

captura de estados emocionales de personas interactuando en distintos contextos (reales, simulados, virtuales o interactivos con robots) a través de un Framework MM, independiente del dominio, con la capacidad de registrar datos tanto fisiológicos como biométricos, contribuye a reflejar un estado emocional de mayor fidelidad que el obtenido cuando se utilizan pocos medios de recolección de datos. Para demostrar dicha hipótesis, se encuentran abiertas y en desarrollo las siguientes líneas de trabajo:

a) evolución del Framework MM: registro de datos biométricos y fisiológicos; integración distintas interfaces BCI (Neurosky y Emotiv), distintos sensores fisiológicos (presión arterial, pulsaciones, conductancia de la piel), y distintas herramientas que interactúen con estas interfaces.

b) inferencia emocional a través de la captura del rostro del usuario, correlacionándolo con su estado emocional empleando diversas APIS, por ejemplo, entre otras, Affective de MS, a fin de comparar con desarrollos reciente de autor llevados a cabo en el ISIER aplicando Machine Learning (ML) sobre las bases de técnicas de regresión logística.

c) Desarrollo de escenarios de aplicación (virtual, simuladores, gastronómico, interacción con robots). Para la validación emocional se aplica la prueba de Russel (modelo excitación-valencia), encuestas SAM (por sus siglas en inglés Self-Assessment-Manikin) propuesta por Bradley y Lang [18], banco de imágenes IAPS [19] (genéricas para distintos cuadrantes del circunflejo de Russel), además de aplicar las imágenes de OLAF [20] (alimentos y comidas) para el dominio gastronómico. En la actuación del usuario en el contexto virtual, se aplican casos de experimentación empleando el conjunto de datos recopilado por el Center for the Study of Emotion & Attention de la University of Florida en Gaines Ville, FL,

Estados Unidos. El conjunto contiene una base de datos de imágenes (IAPS) [19], que es utilizada en diversos estudios científicos que analizan emociones y es considerada un estándar de facto.

d) visualización y explotación. Se desarrollan consultas y filtros con integración de herramientas de explotación y descubrimiento sobre la base de datos que articula el Framework.

Objetivos y Resultados Esperados

Los objetivos planteados en el marco de este PICTO son: a) Integración de distintas herramientas BCI al Framework MM, b) Integración de distintos sensores fisiológicos (pulsera, conductancia de la piel), c) Captura y registración de ritmo cardiaco, d) Integración de Rostro a través de regresión logística, e) Integración de Voz, f) Integración de Test de Rusell y Encuestas SAM, g) Capturas ambientes (simulados o reales), h) Integración de herramientas para descubrir patrones, i) implantación de un espacio físico de inmersión emocional.

Resultados obtenidos

Hasta el momento los logros obtenidos se reflejan en el Framework MM, donde para su desarrollo se consideraron como entradas principales la captura de las ondas cerebrales, con el empleo del BCI (Neurosky, Emotiv), sus algoritmos, y la toma de “capturas de pantalla”, tanto del rostro del usuario interactuando con el contexto como las imágenes que observa el usuario. Se integra con Encuestas SAM. Se realizaron pruebas con ambientes estáticos de imágenes y ambientes dinámicos, juegos de computadora. Los resultados de las pruebas mencionadas fueron demostrados y publicados CACIC 2018 [21], CACIC 2019 [22] y CACIC 2020 [23], CACIC 2020 [24], y resultados en ambientes dinámicos en Revista Risti [25].

Recientes trabajos no publicados a la fecha obtuvieron resultados satisfactorios en el desarrollo de ML aplicado a la detección de emociones por voz; en particular este trabajo se orienta al diseño y desarrollo de redes neuronales para la clasificación de emociones en el discurso hablado, con la experimentación de diferentes métodos para convertir de un enfoque categórico de clasificación de emociones a uno dimensional.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por tres investigadores formados, tres investigadores en formación, y dos alumnos investigadores. Se defendieron en esta línea de investigación una tesis de doctorado en colaboración con FIUBA-UNLP, dos tesis de grado. Actualmente se desarrollan cuatro tesis de grado en el área de computación afectiva transfiriendo los resultados de investigación a la cátedra de trabajo de diploma de la carrera informática.

Reconocimiento

Los docentes investigadores integrantes del ISIER quieren expresar su reconocimiento al **Ing. Hugo Rene Padovani** quien en vida contribuyó a la creación del ISIER [26], como así también a la formación de RRHH en investigación en el área de robótica autónoma, y en el campo de la computación afectiva; sirva este espacio para expresar nuestro profundo agradecimiento, a nuestro querido Hugo, quien siempre estará presente en nuestros corazones.

Referencias

1. Baldassarri, S. (15 de 9 de 2016). Computación afectiva: tecnología y emociones para mejorar la experiencia del usuario. Obtenido de <http://hdl.handle.net/10915/53441>
2. A. Mehrabian, "Communication without Words", *Psychology Today*, vol. 2, no 4, (68)
3. Posner, J., Russell, J., & Peterson, B. (2005). "The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology.", *Development and psychopathology*, vol. 17, n° 3, págs. 715-34, 2005, ISSN: 0954-5794. *Development and psychopathology*, 715-34 Vol 17 n° 3.
4. Ekman, P. (2005 cap 3). "Basic Emotions": En e. T.Dalgleish y M.J.Power, *Handbook of Cognition and Emotion*, (págs. 45-60). Chichester, UK: John Wiley & Sons, Ltd.
5. van der Haar, D. T. (2019). Student Emotion Recognition. *International Conference on Human-Computer interaction* (págs. 301-311). Springer Chan.
6. Zatarain Cabada, R., Barron Estrada, M. L., Halor-Hernandez, G., & Reyes-García, C. A. (2014). Emotion Recognition in Intelligent Tutoring Systems. *Mexican International Conference on Artificial Intelligence* (págs. 494-504). Mexico: Springer.
7. Wei-Long, Z., & Bao-Liang, L. (2015). Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks. En I. X. Library, *IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT*, VOL. 7, NO. 3, SEPTEMBER 2015 (págs. 162-175). IEEE.
8. Xu, R., Chen, J., Han, J., Tan, L., & Xu, L. (2019). Towards emotion-sensitive learning cognitive state analysis of big data in education deep learning-based. En *Computing* (págs. 1-16). Austria: Springer Viena.
9. Alepis, E., & Virvou, M. (2006). User Modelling: An Empirical Study for Affect Perception Through Keyboard and Speech in a Bi-modal User Interface*. *International Conference on Adaptive Hypermedia and adaptive Web-Based Systems*. Berlin, Heidelberg: Springer.

10. Calot E., Ierache J., y Hasperu W. Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations. En *Intelligent Systems and Applications*, vol. 1037 de *Advances in Intelligent Systems and Computing*. Springer, 2019. ISBN 978-3-030-29515-8. doi: 10.1007/978-3-030-29516-5 15
11. Ierache, J., Dittler M., Pereira G., García Martínez R, "Robot Control on the basis of Bio-electrical signals" XV CACIC 2009, UNJu, ISBN 978-897-24068-3-9
12. Ierache J, Pereira G, Iribarren J, Sattolo I, "Robot Control on the Basis of Bio-electrical Signals": "International Conference on Robot Intelligence Technology and Applications" (RiTA 2012). Series *Advances in Intelligent and Soft Computing* of Springer.
13. Ierache., J, Pereira., G, Sattolo., I, Guerrero., A, D'Altto J, Iribarren, J. Control vía Internet de un Robot ubicado en un sitio remoto aplicando una Interfase Cerebro-Máquina". XVII CACIC 2011, UNLP, ISBN 978-950-34-0756-1, p 1373-1382.
14. Ierache J, Pereira G, Iribarren J "Demostración de los resultados en la integración de Interfases Lectoras de Bioseñales aplicadas al Control de un Robot" VII TEYET 2012 UNNOBA, 2012, demos educativas. ISBN 978-987-28186-3-0.
15. Ierache., J, Pereira., G, J, Iribarren Framework for Brain Computer Interface implemented to control devices in the context of home automation XIX CACIC 2013, CAECE Mar del Plata, ISBN 978-897-23963-1
16. Ierache J, Nervo F, Pereira G, Iribarren J Estado emocional centrado en estímulos, aplicando Interfase cerebro-maquina, XX CACIC (Buenos Aires, 2014) ISBN 978-987-3806-05-6
17. Emotiv Epoc Vigente 2018 <http://www.emotiv.com/>.
18. Bradley, M. M. & Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59 (1994).
19. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. NIMH Cent. Study Emot. Atten. 39–58, <https://doi.org/10.1027/0269-8803/a000147> (1997).
20. Miccoli, L., Delgado, R., Rodríguez-Ruiz, S., Guerra, P., García, E., & Fernández-Santaella, M. C. OLAF, the Open Library of Affective Foods in ADULTS:
- 21 J. Ierache, R. Nicolosi, G. Ponce, C. Cervino, E. Eszter. Registro emocional de personas interactuando en contextos de entornos virtuales. CACIC 2018, La Plata, pp 877-886, ISBN: 978-950-658-472-6
22. Ierache, J., Ponce, G., Nicolosi, R., Sattolo, I., & Chapperón, G. (2019). Valoración del grado de atención en contextos áulicos con el empleo de interfase cerebro-computadora. CACIC 2019. Rio Cuarto, pp: 417-426, ISBN: 978-987-688-377-1
23. Ierache J, Nervo F, Sattolo I, Ierache R, Chapperón G. Propuesta de un Modelo Multimodal de valoración emocional en el marco de la computación afectiva aplicado en ambientes gastronómicos. CACIC 2020 La Matanza. En prensa
- 24 Barrionuevo Carlos, Ierache. J. Sattolo I "Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística" CACIC 2020 La Matanza. En prensa
25. Ierache. J. Sattolo I, Chapperón G. Framework multimodal emocional en el contexto de ambientes dinámicos. DOI. 10.17013/risti 40. 45-59 ISSN: 1646-9895
26. www.isierum.cl.biz vigente febrero de 2021

Diseño y desarrollo de Serious Games para la rehabilitación de problemas cognitivos en pacientes neurológicos implementando VRPN para las comunicaciones entre las interfaces y los dispositivos.

Javier J. Rosenstein, Rodrigo Gonzalez, Juan Salvador Portugal, Julian Argañaráz

Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat, Belgrano 721, 5501 Godoy Cruz, Mendoza, Mendoza, Argentina.

rosensteinjavier@uch.edu.ar, gonzalezrodrigo@uch.edu.ar, thejuasz@gmail.com, hiory01@gmail.com

RESUMEN

En el desarrollo de sistemas de realidad virtual uno de los inconvenientes que se encuentran es la comunicación entre las aplicaciones y los dispositivos de adquisición de datos. Ya sea por no disponer de un método de acceso en forma directa o por necesitar independencia entre ambos, es decir que las aplicaciones corran en una plataforma y los dispositivos en otras. Para lograr esta independencia y a su vez permitir la integración de todo el sistema de realidad virtual, es necesario la implementación de algún protocolo de comunicaciones que permita esta vinculación heterogénea en tiempo real. Los dispositivos generalmente están asociados a funciones o características de los individuos que los utilizan y se necesita integrar los movimientos que estos representan hacia una interfaz gráfica que permita la realimentación neuronal del paciente y de este modo lograr la mejora cognitiva (efecto de neurofeedback deseado). El presente trabajo trata del análisis e implementación del protocolo de comunicaciones VRPN (Virtual Reality Peripheral Network) entre las partes de un entorno multimedia donde interactúan la adquisición de movimientos del usuario y la representación visual en un escenario virtual que permita la retroalimentación al usuario en tiempo real logrando una experiencia interactiva e inmersiva. Logrando de este modo la

producción de un sistema de capacitación / rehabilitación o mejor llamado Serious Game.

Palabras clave: VRPN, EOG, EEG, BCI, Serious Games, Mirror Neurons, Neurofeedback

CONTEXTO

El presente proyecto se desarrolla en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat (Godoy Cruz, Mendoza), en el marco de la Licenciatura en Sistemas de Información; en cooperación con el Laboratorio de I+D+i en Neurotecnologías de la empresa Neuromed Argentina S.A. (Godoy Cruz, Mendoza).

Este trabajo es parte del proyecto de investigación que dio inicio en Diciembre de 2017 denominado "Diseño y desarrollo de un prototipo de Serious Game destinado a la rehabilitación de problemas neurológicos implementando VRPN para la comunicación de la BCI".

1. INTRODUCCIÓN

El objetivo principal de la presente línea de investigación consiste en el diseño y desarrollo de una BCI (Brain Computer Interface) [1] [2], que permita interactuar

entre las señales generadas por un paciente neurológico [3] y una interfaz de realidad virtual. Este sistema debe permitir lograr el principio de neurofeedback [4], o retroalimentación hacia el paciente. De este modo se podrán mejorar sus capacidades cognitivas correspondientes. Esto tiene aplicación directa en los tratamientos de rehabilitación en pacientes de patologías neurológicas y cognitivas, principalmente en niños, este enfoque es importante ya que está demostrado que es la mayor causa de discapacidad en niños, según estadísticas como las publicadas por el Indec de su último informe del 2018 (ver figura1), en argentina dentro de las principales discapacidades, la mental-cognitiva es la mayor de las discapacidades entre las edades de 6 a 14 años.

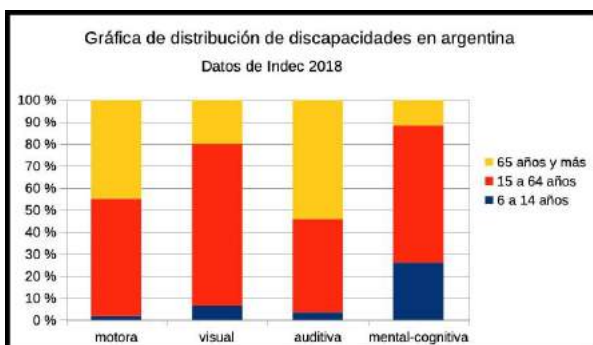


Figura 1: Discapacidades según Indec en argentina

Incluso, estudios indican la posibilidad de tratar patologías psiquiátricas como la depresión [5] [6].

La implementación del trabajo se organiza de acuerdo a las siguientes etapas:

- a) Adquisición de señales mediante técnicas de Electroencefalografía (EEG) [7] y adquisición de movimientos oculares mediante las técnicas de electrooculografía (EOG) [8] [9] [10] [11] [12].
- b) Análisis de estas señales en tiempo real para poder identificar la voluntad de movimiento del individuo, así como la dirección del movimiento.

- c) Transformar la voluntad de movimiento en comandos del protocolo VRPN que permitan transmitir la información al componente software / hardware que la requiera.
- d) Como continuación al resultado obtenido en el punto anterior, Diseñar, programar y poder comandar una interfaz gráfica de aprendizaje o interfaz Cerebro / Computadora (BCI – Brain Computer Interface).

Finalmente se presenta la BCI como un sistema de adquisición de datos, procesamiento del protocolo serie a VRPN y luego la representación en nuestro modelo de prototipo de Serious Game [14] [15]. Una vez adquiridas las señales EEG/EOG, estas se analizan y codifican con las bibliotecas desarrolladas como parte de este proyecto. Esto permite su comunicación mediante VRPN con las interfaces virtuales que interpretan este protocolo.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo está compuesto por tres etapas o fases de trabajo. La primera de ellas corresponde a la implementación de VRPN entre un sistema simulado de captura de datos provenientes de un paciente, y una interfaz virtual básica. La segunda etapa del proyecto pretende avanzar sobre la captura de datos reales para que, luego de procesados, se transfieran a través de VRPN hacia la interfaz virtual de neurofeedback. Una breve descripción de cada una de las etapas se describe a continuación:

1. La primer parte consiste en la implementación de un simulador de señales EEG y EOG necesarias para el análisis e interpretación de la voluntad del usuario de la BCI. Estas señales una vez procesadas se deben codificar en comandos de VRPN para poder ser transmitidas hacia una interfaz virtual. Esta debe poder interpretar

las señales transmitidas y representar la voluntad inicial del usuario correspondiente. De este modo se cumplen los objetivos de captura, análisis, procesamiento, transmisión, recepción y representación, lo cual produce el efecto de neurofeedback deseado sobre una interfaz virtual de capacitación a nivel prototipo.

2. La segunda etapa consiste en el desarrollo y mejora del proyecto mediante la adquisición real de señales EEG y EOG por electrodos ubicados superficialmente sobre la cabeza del paciente. Esta modificación al sistema requiere de un diseño e implementación electrónica así como del desarrollo del firmware que permita adquirir, analizar y procesar estas señales, que luego serán transmitidas a través del protocolo VRPN hacia una interfaz virtual. Así, el sistema diseñado se incorporará en forma transparente al proyecto implementado en la primer etapa.
3. La etapa final de este proyecto consiste en el diseño y desarrollo del escenario virtual de rehabilitación cognitiva para pacientes neurológicos según las indicaciones concretas por parte del especialista en neurología. Se partirá de un relevamiento de las técnicas de aprendizaje que se requieren implementar y los resultados que se pretenden obtener, indicados por el neurólogo o experto afin.

3. RESULTADOS ESPERADOS

Los resultados esperados se pueden dividir en dos grandes grupos:

1. BCI e interfaz virtual gráfica que produzca en el paciente el efecto de neurofeedback: se espera obtener como producto final un sistema de retroalimentación a un usuario, en principio simulando la generación de señales que serán interpretadas en el escenario de realidad virtual, se envía mediante comandos codificados en el protocolo VRPN que se transmiten a través de una red Ethernet.
2. Adquisición de datos reales de EEG y EOG: análisis y preprocesamiento; se identifican los comandos necesarios para su transmisión hacia el equipo generador de tramas VRPN para que finalmente se tenga un prototipo funcional que implemente la técnica de la terapia y se logre el efecto de neurofeedback propuesto.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de la tesis de maestría en teleinformática de la Universidad de Mendoza, de Javier J. Rosenstein, defendida durante el 2020. Dicha tesis se centró en la implementación del protocolo VRPN demostrando su uso en un sistema BCI.

Esta línea de I+D Además cuenta con la dirección del Mg. Lic. Javier J. Rosenstein (UCH) y como investigador principal el Dr. Rodrigo Gonzalez (UCH). En lo que respecta a la formación de estudiantes de la licenciatura, esta línea de investigación cuenta con dos tesis de grado en curso, pertenecientes a los estudiantes Juan Salvador Portugal y Julian Argañaráz, cuyos planes de tesis se encuentran específicamente dentro del marco de este proyecto. Ambos cursan la licenciatura en Sistemas de Información en la Universidad Champagnat.

Todos los avances logrados y las implementaciones realizadas relacionadas con el desarrollo del presente proyecto desde su primer etapa, se están utilizando como recurso para el dictado de talleres de comunicaciones, redes, programación de microcontroladores, programación en C/C++ y diseño y programación de interfaces virtuales de capacitación/rehabilitación en general, tanto para estudiantes de la universidad, como así también para alumnos externos a la UCH.

5. BIBLIOGRAFÍA

1. Y. Wang, X. Gao, B. Hong, and S. Gao, "Practical designs of brain-computer interfaces based on the modulation of EEG rhythms," in *Brain-Computer Interfaces*. Springer, 2009, pp. 137–154.
2. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
3. J. A. Pineda, "The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing"," *Brain Research Reviews*, vol. 50, no. 1, pp. 57–68, 2005.
4. S. Enriquez-Geppert, R. J. Huster, and C. S. Herrmann, "Boosting brain functions: Improving executive functions with behavioral training, neurostimulation, and neurofeedback," *International Journal of Psychophysiology*, vol. 88, no. 1, pp. 1–16, 2013.
5. R. Ramirez, M. Palencia-Lefler, S. Giraldo, and Z. Vamvakousis, "Musical neurofeedback for treating depression in elderly people." *Frontiers in neuroscience*, vol. 9, pp. 354–354, 2014.
6. W. Rief, "Getting started with neurofeedback," 2006.
7. J. D. Kropotov, *Quantitative EEG, event-related potentials and neurotherapy*. Academic Press, 2010.
8. A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 33, no. 4, pp. 741–753, 2011.
9. H. Singh and J. Singh, "A review on electrooculography," *International Journal of Advanced Engineering Technology*, vol. 3, no. 4, pp. 115–122, 2012.
10. D. P. Bautista, I. A. Badillo, D. De la Rosa Mejía, and A. H. H. Jiménez, "Interfaz humano-computadora basada en señales de electrooculografía para personas con discapacidad motriz," *ReCIBE*, vol. 3, no. 2, 2016.
11. S. Yathunathan, L. Chandrasena, A. Umakanthan, V. Vasuki, and S. Munasinghe, "Controlling a wheelchair by use of EOG signal," in *2008 4th International Conference on Information and Automation for Sustainability*, IEEE, 2008, pp. 283–288.
12. V. C. C. Roza, "Interface para tecnologia assistiva baseada em eletrooculografia," 2014.
13. A. C. Gaviria, I. C. Miller, S. O. Medina, and D. R. Gonzales, "Implementación de una interfaz hombre-computador basada en registros EOG mediante circuitos de señal mixta PSoC," in *V Latin American Congress on Biomedical Engineering CLAIB 2011 May 16-21, 2011, Habana, Cuba*. Springer, 2013, pp. 1194–1197.
14. P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: A survey and a classification towards a taxonomy," in *5th Iberian Conference on Information Systems and Technologies*. IEEE, 2010, pp. 1–6.
15. J. S. Breuer and G. Bente, "Why so serious? On the relation of serious games and learning," *Eludamos. Journal for Computer Game Culture*, vol. 4, no. 1, pp. 7–24, 2010.
16. Instituto Nacional de Estadísticas y Censos, INDEC, Argentina, 2018, ISBN:978-950-896-519-6, https://www.indec.gob.ar/ftp/cuadros/publicacion/estudio_discapacidad_07_18.pdf

DISEÑO Y DESARROLLO DE UNA APLICACIÓN MOVIL DE MONITOREO Y SEGUIMIENTO PARA ESCOLARES

Pablo M. Vera, Rocío A. Rodríguez, Hernán A. Viavattene, Cesar Daniel Delgado

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{pablomartin.vera, rocioandrea.rodriguez } @uai.edu.ar
{hernanalberto.viavattene, cesar.delgado } @alumnos.uai.edu.ar

RESUMEN

Hoy en día un amplio porcentaje de niños y jóvenes cuentan con teléfonos celulares con conectividad a internet que utilizan a diario y llevan consigo la mayor parte del tiempo. Estos equipos son en su mayoría de gama media y cuentan al menos con la posibilidad de ser geolocalizados por GPS y/o triangulación de señal celular. Sacando provecho de esas características se diseña y desarrolla una aplicación móvil denominada Vigia Escolar. Esta aplicación permite a los padres monitorizar de una forma no invasiva el recorrido de sus hijos para estar tranquilos de que llegan a destino y no se desvían de la ruta establecida. Utilizando la técnica de geofencing, Vigia Escolar permite definir corredores y zonas seguras emitiendo alertas para prevenir inconvenientes.

Palabras clave: Geolocalización, Geofencing Aplicaciones Móviles, APPs,

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado y una duración de 2 años. El proyecto se encuentra actualmente en segundo año.

1. INTRODUCCIÓN

La inseguridad es una cuestión que preocupa a toda la población y aún mas cuando se trata de los hijos que salen de la casa ya sea para ir a estudiar, hacer algún deporte o visitar algún amigo. No es posible estar todo el tiempo vigilando donde esta una determinada persona, pero si es posible tomar ciertos recaudos para prevenir inconvenientes. Por ejemplo, si tarda más de lo habitual en volver o se fue de la zona en la que debería estar puede indicar que se encuentra en problemas o frente a algún peligro. Muchos padres realizan llamadas o mensajes constantes a sus hijos lo que hace que se sientan invadidos o perseguidos. Y también genera un desgaste para los padres ante la preocupación constante. Para subsanar este inconveniente y evitar que ambos deban estar pendientes es posible recurrir a la tecnología.

La geolocalización permite saber dónde se encuentra un dispositivo en cada momento, pero esta es extendida mediante el concepto de geofencing al detectar cuando se entra o sale de determinada área. “Geofencing es una pequeña área geográfica que se describe para producir un evento de ubicación tan pronto como un usuario ingrese o salga de esta geovalla” [1]

Usando la geolocalización y el geofencing es posible desarrollar una aplicación de monitoreo no invasivo, donde la persona que está siendo monitoreada ni se entera ya que no es necesario que tenga una aplicación

abierta, sino que un servicio cada un tiempo determinado enviará su ubicación. Un algoritmo previamente configurado analizará si se va de una zona determinada antes de tiempo o ingresa a una zona no habitual lo que puede indicar algún problema. Solo en esos casos el sistema entonces enviará una alerta sobre un posible problema.

De esta forma los padres pueden estar tranquilos sabiendo que si ocurre algo el sistema les avisará sin tener que estar preguntando donde esta o fijándose mediante una aplicación si está donde debería.

Existen numerosos trabajos relacionados en el área como [2] que es similar a la propuesta, pero utilizando un dispositivo de rastreo propio en lugar de aprovechar el teléfono celular lo que requiere de un gasto adicional. También el trabajo [3],[4] se basa en el seguimiento de niños, pero en este caso mediante el uso de un smartwatch. Mientras que otros trabajos como [5] complementan la geolocalización con el uso de RFID para interiores y [6] incorpora la grabación de audio ante una situación de emergencia para evitar o detectar abusos.

Específicamente el trabajo [7] plantea el envío de notificaciones al realizar el seguimiento en distintas áreas de monitoreo, considerando restricciones de tiempo y movimiento entre las áreas.

Ya en el 2015 algunos trabajos como por ejemplo [8] veían la necesidad de ahorrar recursos en el monitoreo haciendo aplicaciones con clientes livianos, lo que ahorra también batería y permite realizar el monitoreo por más tiempo y en mayor cantidad de equipos con menor poder de procesamiento.

De hecho la solución denominada Vigia Escolar tendrá soporte en teléfonos celulares, que no necesitan ser de alta gama, haciendo uso de geofencing y analizando la reducción del consumo de batería en el proceso de monitoreo de las zonas delimitadas.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Este proyecto de investigación busca:

- Hacer uso de la geolocalización y de las técnicas de geofencing para aplicarlos a aplicaciones móviles.
- Diseñar una aplicación de monitoreo no invasiva que consuma pocos recursos.
- Investigar las distintas formas de envío de alertas disponibles para dar mayor flexibilidad al sistema (SMS, WhatsApp, etc)
- Complementar las funcionalidades de las aplicaciones mediante el aprovechamiento de otros sensores disponibles en los dispositivos móviles.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el primer año del proyecto se ha realizado la especificación funcional de la primera versión de la aplicación y se ha comenzado con el desarrollo.

Esta primera versión permite:

- Definir zonas y horarios de permanencia, por ejemplo, si el niño debe estar de 8 a 12 hs en el colegio y se detecta que sale de zona antes de horario. Las zonas pueden configurarse de forma repetitiva y agregar acepciones y lugar conocidos, por ejemplo, si luego del colegio va a la casa de un amigo o tiene una actividad en el gimnasio. Todas estas zonas pueden establecerse y configurarse desde una aplicación web que luego se sincroniza con la app móvil. La figura 1 muestra la pantalla de configuración de zonas donde se puede establecer el horario de cada una.
- Definir corredores seguros para llegar de una zona a otra.

- Enviar alertas en caso de que se salga de la zona establecida.
- Consultar la ubicación actual tal como puede verse en la figura 2

Tipo de Locación	Asignado A	Horario	Activo
Escuela	Juan	10-12hs	Si
Tenis	Juan	14-16hs	Si
Musica	Emilce	14-16hs	Si

Figura 1. Pantalla de configuración de zonas de monitoreo

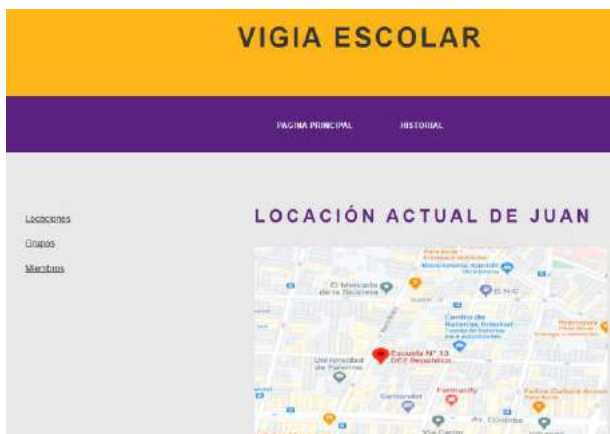


Figura 2. Visualización de la ubicación actual del dispositivo monitorizado

El desarrollo web para el backend se está realizando en C# utilizando la arquitectura MVC para su diseño utilizando el framework 4.8 de Microsoft con base de datos SQL Server version 13.0

Mientras que la parte móvil se está desarrollando con Android Studio aprovechando las APIs nativas de Google de geofencing. También se utilizan otras APIs como por ejemplo el estado de la batería para dar aviso sino se puede seguir monitoreando la ubicación o bajar la frecuencia de actualización para ahorrar batería.

Una vez finalizada y probada esta versión se planea incorporar nuevas características a la aplicación como por ejemplo determinar si se sube a un vehículo por la velocidad de traslado. También si el niño vuelve en colectivo a su casa al entrar en determinada

zona una alerta pueda avisarles a sus padres para que lo esperen en la parada antes de que baje.

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 1 tesis de maestría y 1 tesina de grado en la UAI (Universidad Abierta Interamericana).

5. BIBLIOGRAFIA

[1] Zuva, K., & Zuva, T. Tracking of Customers using Geofencing Technology. 2019

[2] Gilmore, J. N. (2020). Securing the kids: Geofencing and child wearables. *Convergence*, 26(5–6), 1333–1346. <https://doi.org/10.1177/1354856519882317>

[3] Indrayana, I. N. E., Sutawinaya, P., Pratiwi, N. M. W. D., Prihatini, P. M., & Asri, S. A. (2021, February). Android-Based Child Monitoring Application Using A Smartwatch and Geofence Service. In *Journal of Physics: Conference Series* (Vol. 1803, No. 1, p. 012024). IOP Publishing.

[4] Gupta, A., & Harit, V. (2016, February). Child safety & tracking management system by using GPS, GEO-fencing & android application: An analysis. In *2016 Second International Conference on Computational Intelligence & Communication Technology (CICT)* (pp. 683-686). IEEE.

[5] Chaudhari, T. R., Pathil, A. J., & Karhe, R. R. (2016). RFID and GPS Based Child Tracking System With Voice Recognition for Security. *International Journal of Scientific Development and Research (IJSDR)*, 1(12).

[6] Raflesia, S. P., & Lestarini, D. (2018, October). An integrated child safety using geo-fencing information on mobile devices. In 2018 International Conference on Electrical Engineering and Computer Science (ICECOS) (pp. 379-384). IEEE.

[7] Rodriguez Garzon, S., & Deva, B. (2014, September). Geofencing 2.0: taking location-based notifications to the next level. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 921-932).

[8] Garzon, S. R., Deva, B., Pilz, G., & Medack, S. (2015, March). Infrastructure-assisted geofencing: proactive location-based services with thin mobile clients and smart servers. In 2015 3rd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (pp. 61-70). IEEE.

Estrategias de análisis automático de datos urbanos: movilidad, turismo y desigualdad

Mauricio Savarro

Dpto. de Informática - Facultad de Ingeniería
UNPSJB
Puerto Madryn, Argentina
savarromauricio@gmail.com

Leo Ordinez

Laboratorio de Investigación en Informática
(LINVI) - Facultad de Ingeniería - UNPSJB
Puerto Madryn, Argentina
leo.ordinez@gmail.com

RESUMEN

Los territorios deben enfrentar nuevos desafíos para el diseño de estrategias de desarrollo dentro de un contexto de mayor complejidad, incertidumbre y velocidad de cambios, así como adquirir mayores competencias, adaptarse a las exigencias del mercado y avanzar hacia el desarrollo del mismo. El trabajo planteado aquí se asienta en la ciudad de Puerto Madryn, como caso de estudio de una ciudad intermedia. El objetivo general del proyecto es la investigación en técnicas, métodos, estrategias y teorías que asistan en la obtención y procesamiento de datos e información para la construcción de conocimiento de los procesos involucrados en el planeamiento urbano. Para el mencionado fin, se propone el estudio de la literatura relacionada, la construcción de aplicaciones de software que permitan la captura de datos por parte de los ciudadanos, el análisis de dichos datos, la validación de los resultados obtenidos y la sintetización de los mismos.

Palabras Clave: Scrapping - Análisis de Datos - Aprendizaje Automático - Ciudades inteligentes - Informática urbana.

CONTEXTO

El presente forma parte del trabajo en curso dentro del marco de la beca EVC CIN 2019 otorgada a Mauricio Savarro. La misma forma parte de una serie de líneas de investigación recientes desarrolladas en Laboratorio de Investigación en Informática y

en articulación con otros centros y facultades de la UNPSJB. En particular, la beca se enmarca en el proyecto de investigación UNPSJB-PI 1494 “Datos espacio-temporales en entornos urbanos” (2019/2020, prorrogado), cuyo Director es el Dr. Ing. Leonardo Ordinez. Este proyecto se propone investigar y aplicar técnicas, métodos y teorías que asistan en la obtención de información del ámbito urbano, caracterizada por ser heterogénea, proveniente de diferentes orígenes y especialmente ligada a datos espacio-temporales. Vale resaltar que este proyecto se complementa con el UNPSJB-PI 1541, titulado “Desigualdad, experiencia y movilidad urbana en una ciudad turística, Puerto Madryn, Chubut (2008-2018)”, el cual es dirigido por el Dr. Sergio Andrés Kaminker y co-dirigido por el Dr. Ordinez. En este caso, el abordaje es desde las ciencias sociales

1. INTRODUCCIÓN

Durante gran parte del siglo XX, la idea de que una ciudad podía ser inteligente era una ciencia ficción. Manuel Castells (1995) define a la sociedad actual en un imbricamiento interdependiente e inseparable de los procesos sociales y aquellos que se dan en el mundo digital, constituyendo una nueva forma de relacionamiento en lo que él denomina Era Informacional. La convergencia de las tecnologías de la información y la comunicación está produciendo entornos urbanos que son bastante diferentes de todo lo que hemos experimentado hasta ahora. Las ciudades se están volviendo inteligentes no

sólo en términos de la forma en que podemos automatizar funciones rutinarias que atienden a personas individuales, edificios, sistemas de tráfico, sino de forma que nos permitan monitorear, comprender, analizar y planificar la ciudad para mejorar la eficiencia, equidad y calidad de vida para sus ciudadanos en tiempo real (M. Batty et al., 2012). Las ciudades inteligentes a menudo se representan como constelaciones de instrumentos a través de muchas escalas, que están conectados a través de múltiples redes que proporcionan datos continuos sobre los movimientos de personas y materiales, en términos del flujo de decisiones sobre la forma física y social de la ciudad. Sin embargo, las ciudades solo pueden ser inteligentes si pueden integrar y sintetizar estos datos para algún propósito; así mejorar la eficiencia, la equidad, la sostenibilidad y la calidad de vida en las ciudades (Michael Batty, 2013).

La propuesta exploratoria de diferentes enfoques de Brigitte Lamy (2006), resulta básica para comprender las distintas aproximaciones al objeto de estudio de “lo urbano”. Lo trascendental, se encuentra en la mirada compleja sobre un fenómeno multicausal, dialéctico, imbricado y altamente dinámico, que dificulta la comparación o universalización dogmática de teorías. La creciente urbanización está haciendo que las autoridades de las ciudades, los gobiernos, las empresas e incluso los ciudadanos comiencen a pensar en formas alternativas de administrar los recursos dentro de una ciudad. Este objetivo se logra no solo buscando más eficiencia, es decir, hacer más con menos, sino también buscando formas de lograr un mayor nivel de satisfacción entre los ciudadanos y los agentes económicos dentro de una ciudad (Aguilera, Lopez-de-Ipina, & Perez, 2016).

Como explica Urry (2003), la mayoría de los estudios de transporte se han realizado desde la economía o la ingeniería, prestando poca atención a la dimensión subjetiva o la experiencia de la movilidad. Sin embargo, en las últimas décadas existió un cambio de paradigma, a partir del giro que las ciencias

sociales dieron hacia el estudio de los flujos, desplazamientos y movibilidades en general como crítica a una visión más estática de distintos fenómenos, como las distintas expresiones de la desigualdad urbana (Segura, 2012).

Hannerz (1986) introduce cinco dominios en el estudio de la movilidad urbana: doméstico, aprovisionamiento, recreación, vecindad y tránsito. Los estudios de la movilidad de las últimas décadas van desde los micromovimientos corporales, la movilidad cotidiana, residencial y profesional de las personas, hasta los flujos globales de bienes y servicios (Cresswell, 2006; Urry, 2007; Jensen, 2009; Gutiérrez, 2012). Allí se identifican tiempos, rutas, ritmos, experiencias, velocidades, etc. (Cresswell, 2010). Esto implica una diversidad de análisis que importan a esta investigación sobre movilidad urbana, por un lado, y turismo, por el otro.

Como explican Jirón e Imilán Ojeda (2018), en el caso de la movilidad urbana, analizar, tanto el movimiento como las experiencias y los significados que emergen, ha sido una forma de ampliar e incluso cuestionar las nociones clásicas sobre el transporte, centradas en la eficiencia del desplazamiento desde un punto A a un punto B.

Por su parte, los estudios de turismo relativos a transporte suelen referirse al vínculo entre los sistemas de transporte y las vías de circulación de larga o media distancia. En general, la gran mayoría de éstos analizan el vínculo entre transporte aéreo y turismo (Knobel, 2009; Wallingre, 2010, 2013), transporte terrestre y turismo (Ballent y Gorelik, 2001; Wallingre, 2015), transporte marítimo y turismo (Losano et al. 2008; Martínez, 2012). Por otro lado, recientemente se le ha comenzado a prestar atención a la movilidad urbana de las ciudades relativa al turismo y la capacidad de la población residente de los destinos de acceder a determinados bienes y servicios.

En cuanto a estudios de la desigualdad urbana, en América Latina, se ha prestado poca atención a la movilidad. Asimismo, se debe recordar que las movilidades son muy diversas y se experimentan dependientes de una serie de desigualdades (Jirón, 2007), de género, socioeconómica, de edad, origen, etc.

De lo anterior surge la necesidad de destacar la importancia de los datos en este proceso. En este escenario, las personas participan como sensores sociales que proporcionan voluntariamente datos que capturan sus experiencias de la vida cotidiana (Silva et al., 2016). Ello supone una posibilidad de análisis de información para la toma de decisiones de política pública inédita (De Dios Ortúzar y Willumsen, 2008). El entendimiento de los datos generados por la población en el medio urbano, permite corregir una gran cantidad de problemas que afectan la vida diaria y perjudican la eficiencia de las ciudades (BID, 2016), en tanto degradan la calidad de vida y promueven la inequidad. En este sentido, desde la Informática, se han planteado diferentes estrategias, técnicas y herramientas para la obtención de ese conocimiento sobre lo urbano de forma automática (Batty, 2018; Laurini, 2017; Bavsar, 2017; Abduljabbar, 2019; Zhang, 2011).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La masiva disponibilidad de incontables cantidades de datos y el gran poder de cómputo deslocalizado han sido factores determinantes para el surgimiento y auge de la denominada ciencia de datos. Este paradigma se basa en técnicas probadas, que se apoyan en modelos estadísticos de regresión. En virtud de ello, este proyecto de iniciación a la investigación busca profundizar en técnicas y tecnologías, como la extracción automática de datos, el aprendizaje de máquina y los análisis de regresión, con el objetivo de obtener conocimiento sobre el entorno urbano. En particular, se toma como caso de estudio la ciudad de Puerto Madryn. En este sentido, se

busca comprender cómo se expresa y experimenta la *desigualdad* en la *movilidad urbana* en la ciudad de Puerto Madryn, integrando a la diversidad socioeconómica y residencial propia, la mirada de los y las *turistas* y comprender cómo se entrelazan estas dimensiones con las políticas públicas relativas a movilidad y turismo en una ciudad costera patagónica en expansión.

3. RESULTADOS OBTENIDOS

Desde lo metodológico, se busca adquirir las habilidades básicas para el desarrollo de la actividad investigativa en informática. En términos generales, se busca investigar diferentes estrategias, técnicas y modelos, basados en el procesamiento de datos de fuentes heterogéneas, para el análisis de la movilidad urbana en ciudades intermedias y su relación con el turismo y la desigualdad. Más específicamente, se pretende: 1) relevar técnicas y métodos para la sistematización de datos, mediante diferentes tipos de esquemas relacionales y no relacionales; 2) estudiar y desarrollar técnicas para el análisis automático de datos de disponibles en servicios externos de datos abiertos y fuentes primarias (encuestas, relevamientos, datos propios generados en el proceso de investigación); 3) profundizar el conocimiento en el análisis y modelado de datos espacio-temporales.

En esta primera etapa del proyecto se decidió reorientar el trabajo, postergando aspectos de movilidad y comenzando por el análisis de precios de inmuebles (casas y departamentos, en particular) en la ciudad de Puerto Madryn, Chubut. Esto se debió a que el período de trabajo del becario inició en julio de 2020, en el marco de la pandemia por COVID-19. Con este objetivo, el primer punto de este trabajo consistió en obtener información actualizada de valores inmobiliarios. Mediante el uso de scraping se obtuvo información de diferentes webs de diferentes sitios inmobiliarios. La herramienta utilizada fue Web Scraping, un plugin para navegador gratuito con componentes de pago.

Aquí surgió la primera complejidad del trabajo. Ya que la mayoría de las webs poseen información incompleta e inconsistente entre ellas. Es decir, la falta de algunos parámetros útiles para la valuación de un Inmueble como número de ambientes, si posee cochera, etc. Con lo cual el set de datos inicial debió ser trabajado arduamente hasta conseguir un punto de partida para la implementación de los diferentes algoritmos de Aprendizaje Automático. Para esto se revisaron los datos para corregir cualquier fallo en el Scrapping o faltante de información en las webs y se seleccionaron los parámetros finales que formarán parte del conjunto de entrenamiento.

El siguiente punto fue el análisis de los datos donde destaca la generación de una Matriz de Correlación, para ver rápidamente correlaciones entre los diferentes parámetros del modelo, aspecto que sirvió para depurar aún más los datos, junto con la eliminación de *Outliers*, Valores Atípicos que no representan la realidad.

A partir de aquí, se decidió implementar algoritmos de Aprendizaje Automático Supervisados, para generar un modelo de predictor. Se dividieron los datos en dos conjuntos, de entrenamiento para y de prueba. Los algoritmos elegidos fueron Regresión Lineal Múltiple, Regresión Keras (Redes Neuronales), K-Vecinos más cercanos y *Random Forest*.

Como conclusión general de esta primera etapa se destaca la experiencia adquirida en el análisis y depuración de la datos, para una posterior implementación de algoritmos de Machine Learning. Eran campos completamente nuevos que fueron explorados a través de una inmersión directa, que sin duda sirvió para motivar y marcar los aspectos más débiles de cara a continuar en futuras etapas con otros ámbitos de estudio dentro del marco del proyecto. Los siguientes puntos de análisis serán sobre datos recopilados a través encuestas realizadas en la ciudad sobre el turismo y un conjunto de datos provistos por el gobierno nacional sobre el uso de la tarjeta Sube en la ciudad, con el

objetivo de obtener información sobre los recorridos más comunes realizados por los ciudadanos.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo primario sobre esta temática está conformado por los autores del presente. No obstante, debido a los proyectos de investigación en los que se enmarca el trabajo de beca EVC CIN de Mauricio Savarro, las y los integrantes de dichos proyectos hacen las veces de soporte y expertos a disposición para consultas. En particular, entre ambos proyectos se contabilizan siete investigadores/as, cuatro becarios/as y siete estudiantes, que se encuentran haciendo pasantías o tesinas de grado.

5. BIBLIOGRAFÍA

1. Abduljabbar, Rusul; Dia, Hussein; Liyanage, Sohani; Bagloee, Saeed A. (2019). "Applications of Artificial Intelligence in Transport: An Overview." *Sustainability* 11, no. 1: 189.
2. Aguilera, U., Lopez-de-Ipina, D., & Perez, J. (2016). Collaboration-Centred Cities through Urban Apps Based on Open and User-Generated Data. *Sensors*, 16(7), 1–25. <https://doi.org/10.3390/s16071022>
3. Ballent A. y A. Gorelik (2001): "País urbano o país rural: La modernización territorial y su crisis", en Alejandro Cattaruzza, Crisis económica, avance del Estado e incertidumbre política, Sudamericana, Buenos Aires.
4. Batty, M. (2013). Urban Informatics and Big Data, 1–36.
5. Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Portugali, Y. (2012). Smart cities of the future. *European Physical Journal: Special Topics*, 214(1), 481–518. <https://doi.org/10.1140/epjst/e2012-01703-3>
6. Batty M. (2018) "Artificial intelligence and smart cities". *Environment and Planning B: Urban Analytics and City Science*, SAGE Journals, (2018), Vol. 45(1) pp. 3-6.
7. Castells, Manuel (1995). "La ciudad informacional. Tecnologías de la

- comunicación, reestructuración económica y el proceso urbano-regional”
8. Cresswell, T. (2006) *On the Move: Mobility in the Modern Western World*, Routledge, New York.
 9. Cresswell, Tim (2010) “Towards a politics of mobility”, *Environment and Planning D*, N° 28, pp. 17-31.
 10. De Dios Ortúzar, J. y L. Willumsen (2008): *Modelos de transporte (Vol. 1)*. Ed. Universidad de Cantabria.
 11. Gutiérrez, A. (2012): “¿Qué es la movilidad? Elementos para (re) construir las definiciones básicas del campo del transporte”, *Bitácora*, N° 21, pp. 61-74.
 12. Hannerz, U. (1986). *La exploración de la ciudad. Hacia una antropología urbana*. Buenos Aires, Fondo de Cultura Económica.
 13. Jensen, O. (2009): “Urban Mobility as Meaningful Everyday Life Practice”, *Mobilities*, N°4, pp. 139-158.
 14. Jirón, P. (2007): “Prácticas de Movilidad Cotidiana Urbana: Un Análisis para Revelar Desigualdades en la Ciudad”, en Tironi, M. y Pérez, F. *Espacios, Prácticas y Cultura Urbana*, ARQ Ediciones, Escuela de Arquitectura, pp. 176-189.
 15. Jirón, P. y W. Imilán Ojeda (2018), “Moviendo los estudios urbanos. La movilidad como objeto de estudio o como enfoque para comprender la ciudad contemporánea” *Quid 16*, Revista del Área de Estudios Urbanos del Instituto Gino Germani, N°10 – Dic. 2018-Mayo 2019, Facultad de Ciencias Sociales, UBA, pp.17-36.
 16. Knobel, Horacio (2009). “El transporte aéreo de pasajeros y sus equipajes” Editorial Ladevi, Buenos Aires.
 17. Lamy, Brigitte (2006). “Sociología urbana o sociología de lo urbano”, en “Estudios Demográficos y Urbanos”, Vol. 21, No 1.
 18. Laurini, Robert. (2017). *Geographic Knowledge Infrastructure: Applications to Territorial Intelligence and Smart Cities*.
 19. Losano, P., Trifaro, N., Arévalo, A. y A. Tagliorette, (2008): “Implicancias económicas, sociales y ambientales derivadas de la actividad de los cruceros turísticos en la ciudad de Puerto Madryn, Chubut, Patagonia Argentina”, Trelew, Facultad de Ciencias Económicas, Universidad Nacional de la Patagonia San Juan Bosco.
 20. Martínez, C. (2012). “Perspectivas del Turismo de Cruceros en Argentina en el Marco de las Tendencias Mundiales”, *Notas en Turismo y Economía*, Año III. Nro. IV, Universidad del Caribe, México, pp. 44-71.
 21. Segura, Ramiro (2012) “Elementos para una crítica de la noción de segregación residencial socio-económica: desigualdades, desplazamientos e interacciones en el periferia de La Plata”, *Quid 16*, N° 2, pp. 106-132.
 22. Silva, Thiago & Celes, Clayson & Borges, João & Mota, Vinícius & Felipe, Cunha & Ferreira, A.P.G. & Ribeiro, A.I.J.T. & Vaz de Melo, Pedro & Almeida, J.M. & Loureiro, A.A.F.. (2016). *Users in the urban sensing process: Challenges and Research Opportunities*. 10.1016/B978-0-12-803663-1.00003-6.
 23. Urry, J. (2003): *Global Complexity, Polity*, Cambridge.
 24. Urry, J. (2007): *Mobilities*, Polity Press, Cambridge.
 25. Wallingre, N. (2010): “Limitaciones del transporte aerocomercial al desarrollo del turismo. Análisis del caso Transporte interno en Argentina”, *Anuario Turismo y Sociedad*, Vol. X.
 26. Wallingre, N. (2013): “Transporte aéreo, su relación con el turismo y los servicios”, Ediciones Turísticas Fundación ProTurismo, Buenos Aires.
 27. Wallingre, N. (2015): *Transporte carretero argentino*, 4ª edición. Ediciones Turísticas LADEVI, Buenos Aires.
 28. Zhang, J. Wang, F. Wang, K. Lin, W. Xu, X. Chen, C. (2011)“Data-Driven Intelligent Transportation Systems: A Survey,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624-1639.

Extracción de datos, análisis de información y predicción del comportamiento en el rubro económico chubutense en contexto de COVID-19

Matías Ducid

Dpto. de Informática
Facultad de Ingeniería - UNPSJB
Trelew, Argentina
matias.ducid08@gmail.com

Leo Ordinez

Laboratorio de Investigación en Informática
(LINVI) - Facultad de Ingeniería - UNPSJB
Puerto Madryn, Argentina
leo.ordinez@gmail.com

Resumen

En virtud del Informe del segundo relevamiento del impacto social de las medidas de aislamiento en la provincia del Chubut, se indica que desde el sector empresarial-comercial y de servicios, se expresan dificultades de variada índole, que van desde preocupaciones individuales como miedo al contagio, hasta la situación económica general de la comunidad y específica de cada sector. Manifiestan preocupación por los cierres de comercios y microempresas, la disminución de la demanda, los cambios en los patrones de consumo y problemas financieros por la ruptura de la cadena de pagos. Además, se plantean dificultades por el cambio de la dinámica comercial: la innovación forzada en la modalidad de ventas –virtual, web-, el desplazamiento de comercios hacia la periferia de las localidades por no poder afrontar los montos de alquileres del centro, el cambio en la modalidad de oferta de los servicios y la re-planificación y/o menor requerimiento de la fuerza laboral, entre otras.

La escasez, dispersión, inconsistencia y desactualización de la información dificulta definitivamente la toma de decisiones. En términos generales, este proyecto busca sentar una base de construcción de información sólida sobre cuestiones específicas en las ciudades chubutenses, que permitan desarrollar estrategias y acciones tendientes a la comprensión y mejoramiento de la actividad económica comercial de las comunidades, en el contexto sanitario por COVID-19.

Contexto

El presente trabajo sintetiza el plan propuesto a la convocatoria 2020 de las becas EVC CIN. El mismo se enmarca en un conjunto de proyectos articulados interdisciplinariamente de los cuales forman parte diversas instituciones académicas y científicas, los cuales se sintetizan a continuación. El proyecto “Datos espacio-temporales en entornos urbanos” (UNPSJB-PI 1494), el cual se propone investigar y aplicar técnicas, métodos y teorías que asistan en la obtención de información del ámbito urbano, caracterizada por ser heterogénea, proveniente de diferentes orígenes y especialmente ligada a datos espacio-temporales. El proyecto “Desigualdad, experiencia y movilidad urbana en una ciudad turística, Puerto Madryn, Chubut” (UNPSJB-PI 1541), el cual busca comprender cómo se expresa y experimenta la desigualdad en la movilidad urbana en la ciudad de Puerto Madryn, integrando a la diversidad socioeconómica y residencial propia, la mirada de los y las turistas y comprender cómo se entrelazan estas dimensiones con las políticas públicas relativas a movilidad y turismo en una ciudad costera patagónica en expansión. El proyecto “Análisis prospectivo inteligente del impacto social, económico y productivo del COVID-19 en la provincia de Chubut” (RESOL-2020-170-APN-MCT), cuyo objetivo es la construcción de conocimiento acerca del impacto social y económico del COVID-19 en la provincia del Chubut desde un enfoque multidimensional en pos de evaluar prospectivamente escenarios de acción. El proyecto “Plataforma de Datos Abiertos Enlazados para la Gestión y

Visualización de Datos Primarios de Ciencias del Mar” (UNPSJB-PI 1562), se propone el desarrollo de una plataforma para la especificación, modelado, generación y publicación de los conjuntos de datos primarios como datos abiertos enlazados, con visualizaciones que faciliten su interpretación y comparación, como así también su explotación basada en semántica. En la misma línea, otros dos proyectos conforman catálogo de trabajos articulados en los que se enmarca la presente propuesta: el proyecto “Soporte inteligente para la planificación en entornos urbanos”, en evaluación en la convocatoria 2020 de la Secretaría de Ciencia y Técnica de la UNPSJB y el proyecto “Sistema de soporte inteligente para desarrollo urbano y ordenamiento territorial de Puerto Madryn”, el cual fue presentado y se encuentra en evaluación en la convocatoria PDTs UNPSJB 2019.

1. Introducción

La vida de las personas y el curso de las actividades y servicios no esenciales han sufrido cambios drásticos a partir del Decreto DECNU-2020-297-APN-PTE y sus normativas anexas de nivel provincial y/o municipal en Chubut. Las medidas de Aislamiento Social, Preventivo y Obligatorio (ASPO) han generado y continuarán generando desencadenantes de impacto social y económico que necesitan ser identificados y monitoreados para proveer información a los formuladores de políticas públicas. Si bien existen herramientas y emprendimientos a nivel nacional, la provincia del Chubut necesita fortalecer el conjunto de instrumentos que sean capaces de determinar el cúmulo de variables involucradas, medirlas y transformarlas en indicadores cuantitativos y cualitativos que aporten a un tablero de análisis de impacto provincial.

La motivación principal de la propuesta se sustenta en el contexto de las ciudades intermedias, que presentan características remarcables en términos de ordenamiento territorial y movilidad, en particular. Estas ciudades, a diferencia de las grandes

metrópolis sobre las que se centra gran parte de la investigación en transporte, movilidad y planificación urbana, se destacan por un sostenido crecimiento poblacional, bajo desarrollo de infraestructura y una expansión territorial alta (Bolay, 2004). A la vez, funcionalmente estas ciudades juegan un rol de intermediación fundamental en el crecimiento de sus zonas de influencia: “A medida que estas ciudades pequeñas y medianas crecen rápidamente, ofrecen oportunidades críticas para eludir las tecnologías antiguas e implementar prácticas eficientes y ecológicamente racionales que pueden contribuir a dar forma a un futuro más sostenible, con un impacto no solo a nivel local sino también en las zonas rurales del interior.” (Bolay, 2019). En este contexto, la dinámica e interpretación de sucesos ocurridos en espacios geográfico-temporales, hacen que tanto la recolección de los datos como su transformación, análisis, interpretación y explotación de los mismos sea un desafío. De este modo, se pretende hacer un uso racional y sustentable de esta información vital para la toma de decisiones.

En este marco, surge la necesidad de construir conocimiento a partir de diferentes estrategias y herramientas de relevamiento de información y datos, que posibiliten un análisis y ponderación, así como una predicción futura de la situación específica del rubro comercial, en virtud de una circunstancia sanitaria como la generada por el COVID-19. En particular se considerará la región delimitada por los límites geográficos de la provincia del Chubut, ajustando la escala territorial a nivel de ciudades, pueblos, comunas rurales y potencialmente barrios y áreas rurales dispersas. La construcción de conocimiento, así como su modelado y sistematización tienen el objetivo de servir de insumo para la toma de decisiones de los actores políticos y sociales con capacidades para accionarlas. La construcción de modelos y la utilización de técnicas estadísticas y computacionales proporcionará las bases para el tratamiento espacio-temporal de los datos.

Asimismo, el desarrollo de herramientas informáticas de captura de datos en campo para la implementación de relevamientos, encuestas y entrevistas, junto con herramientas informáticas de análisis y visualización de información espacio-temporal, permitirá generar reportes, mapas, figuras y gráficas explicativas para los tomadores de decisiones. A este ecosistema se integrarán diferentes tipos de aplicaciones que permitirán ensayar escenarios y realizar simulaciones a partir de los datos reales disponibles en la plataforma. Esto involucra la construcción de modelos matemáticos y utilización de técnicas estadísticas y computacionales para el análisis inferencial de información, así como la utilización de algoritmos y estrategias predictivas basadas en técnicas computacionales de aprendizaje de máquina, inteligencia artificial y procesamiento de lenguaje natural. En todos los casos los modelos y aplicaciones generadas serán trabajadas interdisciplinariamente considerando características sociales de la población y sus dinámicas, en un contexto sanitario extraordinario como es el generado por el COVID-19.

A modo de síntesis, en el presente plan se propone la construcción de conocimiento a partir de diferentes estrategias y herramientas de relevamiento de información y datos, que posibiliten un análisis situacional de un rubro económico particular como es el comercial en el contexto de la pandemia por COVID-19. El territorio a abordar se limita a la provincia del Chubut, ajustando la escala territorial de acuerdo a la información disponible.

2. Líneas de I+D

Algunos interrogantes a desarrollar tienen que ver con el modelado de las “escalas territoriales”, la categorización de “los impactos”, la incorporación de condicionantes externos y la integración de datos en diferentes contextos.

Esta propuesta asume tres puntos básicos respecto al ámbito de aplicación de la misma:

- Baja calidad de información sobre el medio productivo y comercial de la ciudad: la escasez de información actualizada y pertinente perjudica los procesos de toma de decisiones de inversión. No se verifican fuentes de información de acceso abierto que contribuyan a los procesos de inversión y a un desarrollo territorial equilibrado.
- Aparente sobre oferta de ciertos rubros y escasez en otros: la información del mercado es imperiosa en la generación de una unidad de negocio. En el marco de un estudio de mercado, el análisis de los competidores es crítico y esta dimensión presenta tensiones al observarse una alta repetición de rubros comerciales.
- Bajo nivel de planificación comercial y productiva de la ciudad, con alto grado de dispersión geográfico de los sectores comerciales: la invisibilización del ordenamiento comercial dificulta la identificación de problemas y la planificación. La carencia de un diagnóstico y planes de acción en el mediano plazo impactan negativamente sobre la capacidad de generar políticas públicas de promoción económica.

En función de estos supuestos, se identifican las siguientes causas y efectos:

- *Causa*: bajo nivel de información del medio productivo local, en general. *Efecto*: imposibilidad de tomar decisiones estratégicas.
- *Causa*: sobre oferta comercial en ciertos rubros. *Efectos*: alta tasa de mortandad de comercios que replican casos “exitosos” y degradación de la oferta por aplicación de propuestas de valor de tipo “menor precio”
- *Causa*: escasez de oferta productiva y comercial local de ciertos rubros. *Efecto*: movimiento de la demanda a otros mercados cercanos.

Nuestra base de trabajo es que mediante la obtención de datos del impacto en el medio comercial del COVID-19 en diferentes escalas territoriales se pueden extraer patrones generales de tratamiento del tema, que asistan al diseño de políticas y estrategias de abordaje del mismo en el contexto de la Provincia del Chubut. Como objetivo general, se busca profundizar el conocimiento en el análisis y modelado de los factores que afectan al ámbito comercial local en el contexto sanitario mencionado, a fin de promover estrategias superadoras, mediante la utilización de técnicas de análisis inteligente de datos.

3. Resultados esperados

A lo largo del período de beca (1 año) se esperan alcanzar los siguientes objetivos:

- Generar conocimiento en el análisis y modelado de los factores que afectan el ámbito comercial local en el contexto sanitario mencionado, a fin de promover estrategias superadoras, mediante la utilización de técnicas de análisis inteligente de datos.
- Realizar una investigación de diferentes estrategias, técnicas y modelos, basados en el procesamiento de datos de fuentes heterogéneas, para el análisis de la situación del rubro comercial en los entornos locales enmarcada en el contexto de la pandemia por COVID-19.
- Releva técnicas y métodos para la sistematización de la obtención de datos, mediante diferentes tipos de esquemas relacionales y no relacionales.
- Generar productos de software visuales, de usuario final, que permitan representar la información y ensayar escenarios de acción.

4. Formación de recursos humanos

El diseño metodológico planteado para el caso de un becario en iniciación busca, por un lado, introducirlo en los aspectos generales de la investigación (relevamiento de fuentes, contrastación, ensayo, experimentación, verificación, etc.). Mientras que por otro lado,

se pretende realizar un proceso de inmersión disciplinar a través de la experiencia concreta, en este caso, del desarrollo de software. Se propone un abordaje de ambos procesos en forma iterativa e interrelacionada, disponiendo puntos de control en los que se establecen instancias de validación en conjunto con actores externos como autoridades municipales, expertos en planificación urbana, investigadores de la Facultad de Humanidades y Ciencias Sociales de la UNPSJB y del Instituto Patagónico de Ciencias Sociales y Humanas del CCT-CENPAT CONICET.

El equipo de trabajo primario sobre esta temática está conformado por los autores del presente. No obstante, debido a los proyectos de investigación en los que se enmarcará el trabajo de beca EVC CIN de Matías Ducid, las y los integrantes de dichos proyectos harán las veces de soporte y expertos a disposición para consultas. En particular, entre los proyectos satélites se contabilizan siete investigadores/as, cuatro becarios/as y siete estudiantes, que se encuentran haciendo pasantías o tesinas de grado.

5. Bibliografía

- [1] Informe del segundo relevamiento del impacto social de las medidas de aislamiento en la provincia del Chubut (a julio 2020). (2020) Sergio Kaminker, Renata Hiller. <http://covid19.mdn.unp.edu.ar/?p=30>
- [2] Covid y trabajo doméstico en argentina (abril 2020). (n.d.). <https://cieg.unam.mx/covid-genero/pdf/reflexiones/gobierno/covid-y-trabajo-domestico-argentina.pdf>
- [3] Bolay, Jean-Claude and Adriana Rabinovich. 2004. "Intermediate Cities in Latin America: Risks and Opportunities of Coherent Urban Development." *Cities: The International Journal of Urban Policy and Planning*, 21(5): 407–421. DOI: 10.1016/j.cities.2004.07.007.
- [4] Bolay, J. and Kern, A. (2019). Intermediate Cities. In *The Wiley Blackwell Encyclopedia of Urban and Regional Studies*, A. M. Orum (Ed.). doi:10.1002/9781118568446.eurs0163
- [5] Impacto Social y Económico de la COVID-19 y Opciones de Políticas en Argentina (abril 2020). (n.d.). Retrieved July 13, 2020, from

- <https://www.undp.org/content/dam/rblac/Policy%20Papers%20COVID%2019/finaldocuments/UNDP-RBLAC-CD19-PDS-Number6-ES-Arg.pdf>
- [6] “Goal 11: Make cities inclusive, safe, resilient and sustainable,” *Metas del Desarrollo Sostenible*. Naciones Unidas. [Online]. Available: <https://www.un.org/sustainabledevelopment/cities/>. [Accessed: 12-Oct-2019].
- [7] “Programa Ciudades Emergentes y Sostenibles,” BID. [Online]. Available: <https://www.iadb.org/es/desarrollo-urbano-y-vivienda/programa-ciudades-emergentes-y-sostenibles>.
- [8] “Programa de Ciudades,” CIPPEC. [Online]. Available: <https://www.cippec.org/programas/ciudades/>.
- [9] “World Cities 2015 - 2018,” European Union. [Online]. Available: <http://world-cities.eu/>.
- [10] “Smart Growth in Small Towns and Rural Communities,” United States Environmental Protection Agency. [Online]. Available: <https://www.epa.gov/smartgrowth/smart-growth-small-towns-and-rural-communities>.
- [11] “Smart Communities & Smart Cities,” ESRI. [Online]. Available: <https://www.esri.com/en-us/smart-communities/overview>.
- [12] Ballent A. y A. Gorelik (2001): "País urbano o país rural: La modernización territorial y su crisis", en Alejandro Cattaruzza, *Crisis económica, avance del Estado e incertidumbre política*, Sudamericana, Buenos Aires.
- [13] Batty, M. (2013). *Urban Informatics and Big Data*, 1–36.
- [14] Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Portugali, Y. (2012). Smart cities of the future. *European Physical Journal: Special Topics*, 214(1), 481–518. <https://doi.org/10.1140/epjst/e2012-01703-3>
- [15] Batty M. (2018) “Artificial intelligence and smart cities”. *Environment and Planning B: Urban Analytics and City Science*, SAGE Journals, (2018), Vol. 45(1) pp. 3-6.
- [16] M. Castells, *La Cuestión urbana*. 2005.
- [17] D. Cravacuore, “Gobiernos Locales en Argentina,” *Manual de Gobiernos Locales*. CLAD Centro Latinoamericano de Administración para el Desarrollo, Universidad Autónoma de Chile, Santiago, Chile, pp. 15–40, 2016.
- [18] B. Warf, “Geographies of E-Government in Latin America and the Caribbean,” *Journal of Latin American Geography*, vol. 13, no. 1. pp. 169–185, 2014.

Framework para la generación de Procedimientos Semánticos aplicados a Navegadores de Realidad Aumentada

Martin Becerra ¹, Jorge Ierache ¹, Nahuel Mangiarua ¹, María José Abasolo ^{2,3},

¹Universidad Nacional de La Matanza, DIIT, Grupo de Realidad Aumentada Aplicada
Florencio Varela 1903, La Matanza, Buenos Aires, Argentina
{mabecerra, jierache, nmangiarua}@unlam.edu.ar

²Universidad Nacional de la Plata, Facultad de Informática, III-LIDI

³Comisión de Investigaciones Científicas de la Pcia. de Bs. As. CICPBA
mjabasolo@lidi.info.unlp.edu.ar

Resumen

En este trabajo se enuncian la línea de investigación aplicada al desarrollo de un framework que permitirá a usuarios sin conocimientos específicos de programación, tener la capacidad de crear procedimientos o series de pasos a realizar en entornos físicos mediante el uso de navegadores de Realidad Aumentada. Se utilizarán tecnologías de la web semántica para enriquecer los procedimientos con información de la web y lograr una interoperabilidad semántica con otras aplicaciones mediante el uso de ontologías.

Palabras clave: Realidad Aumentada ubicua, Catálogo Virtual Aumentado, Ontologías, Web Semántica

Contexto

La investigación presentada es desarrollada por el grupo de investigación de Realidad Aumentada Aplicada del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, como trabajo de tesis doctoral que expande el marco del proyecto PROINCE C-231 2019-2020 Comandos de Voz y Reconocimiento Facial para Aplicaciones de Realidad Aumentada.

Introducción

La Realidad Aumentada (RA) permite la fusión de datos virtuales sobre el mundo físico, enriqueciendo con información virtual la

percepción de la realidad [1]. En los últimos años, la RA se ha expandido a diferentes campos de aplicación tales como educación, salud, industria, turismo, marketing y entretenimiento. Nuestro equipo de investigación desarrolló diferentes aplicaciones, como juegos de tablero [2],[3], herramientas para la generación de materiales didácticos para el área educativa [4] y juegos didácticos [5], o sistemas de aumentación de información de salud mediante una tarjeta aumentada basadas en conocimiento para la asistencia médica en emergencias [6],[7]. El proyecto de investigación aplicada se encuadra en el contexto de la aplicación de tecnologías de RA en la vida cotidiana de las personas que contribuye a la participación en el ámbito tecnológico generando así un impacto significativo en la sociedad.

En la actualidad existen diferentes navegadores de Realidad Aumentada (En inglés AR Browsers) populares en el mercado como LayAR[8], wiktitude[9] para proveer experiencias de Realidad Aumentada. Estas son limitadas ya que le permiten a un usuario consumir pasivamente un conjunto delimitado de funciones. Existen diferentes alternativas como ARCAMA3D [10], T. Matuszka et. al. [11] y SmartReality[12] que ofrecen una experiencia ubicua mediante la integración de tecnologías de web semántica para integrar información de la nube de datos interconectados (En inglés *Linked data Cloud*) para enriquecer las descripciones de puntos de interés cercanos a la posición de un usuario.

Estas aplicaciones permiten crear contenidos, pero son consumidos estáticamente. En otras palabras, solamente pueden ver descripciones sin poder realizar ninguna acción sobre las mismas. Resulta de utilidad que el usuario pueda definir un procedimiento como conjunto de acciones a realizar en un entorno enriquecido por la Realidad Aumentada.

En la próxima sección se describen las líneas de investigación de “Navegadores de Realidad Aumentada Semántico” y “Sistema de catálogos virtuales aumentados semánticos y templates” que tienen como finalidad crear procedimientos y marcadores de Realidad Aumentada que puedan ser reutilizadas por otras aplicaciones de RA. El objetivo es permitir la creación de fuentes de datos accionables e interoperables mediante el uso de tecnologías de web semántica para dirigirnos hacia una Realidad Ubicua en donde los datos son consumidos independientemente de la aplicación que los genere.

Líneas de Investigación, Desarrollo e Innovación

La línea de investigación y desarrollo presentadas tienen por objetivo desarrollar aplicaciones de Realidad Aumentada en dirección a:

- Navegador de Realidad Aumentada Semántico.
- Sistema de Catálogos aumentados semánticos y templates.

Navegador de Realidad Aumentada Semántico

El objetivo de esta línea de investigación aplicada se centra en el desarrollo de herramientas que asistan a las personas en la realización de tareas en sus entornos físicos mediante el empleo de tecnologías de Realidad Aumentada como interfaz.

Nuestra contribución consiste en que los

contenidos de Realidad Aumentada sean accionables mediante la asistencia paso a paso a la persona mediante procedimientos enriquecidos con información de la web. Esta línea de investigación se puede aplicar en varios contextos que aplican tecnologías de Realidad Aumentada en especial en la industria 4.0 en la asistencia de tareas de trabajadores inteligentes.

El objetivo de esta línea de investigación aplicada se centra en el desarrollo de un framework que permitirá a usuarios sin conocimientos específicos de programación, tener la capacidad de crear procedimientos o series de pasos a realizar en entornos físicos para su explotación con navegadores de Realidad Aumentada. Se utilizarán tecnologías de la web semántica para enriquecer dichos procedimientos con información de la web y lograr que sean interoperables semánticamente. Esto permitirá la integración de datos de nuestro sistema a otras aplicaciones de Realidad Aumentada mediante el uso de la ontología resultado de nuestra investigación. Se espera que el framework disponga de un editor que le permita armar procedimientos a un usuario creador de contenidos para articular tareas en un entorno real como por ejemplo la creación de tareas que debe realizar un operador inteligente en su puesto de trabajo en el contexto de la industria 4.0.

Para la explotación de contenidos el prototipo dispondrá de un navegador de Realidad Aumentada semántico que permita al usuario buscar y utilizar dichos procedimientos. En una instancia preliminar el prototipo se implementará para teléfonos móviles. Se contemplará el desarrollo de una expansión del framework para que pueda ejecutarse con gafas (Epson Moveiro) y comandos de voz para agilizar su interacción.

Sistema de catálogos aumentados semánticos y templates

Se desarrolló el Sistema de Catálogos Virtuales Aumentados [13], [14] el cual permite la generación, distribución y explotación de contenidos de Realidad Aumentada. Dichos catálogos están compuestos por un conjunto de marcadores que son aumentados con información provista por los usuarios al momento de su creación, la cual es visualizada utilizando una aplicación para teléfono inteligente conectada a internet. El sistema de catálogos virtuales permite predefinir la cantidad y tipos de contenidos asociados a cada marcador junto con sus transformaciones geométricas (posición, rotación, escala) y su orden de aparición en el editor.

En [15] se simplificó el flujo trabajo del usuario para construir y generar contenido aumentado sin la necesidad de tener que contar con conocimientos específicos del dominio de la RA mediante el uso de templates. En la línea presentada se busca extender el sistema de catálogo para estructurar el acceso a datos mediante el uso de ontologías que permita que el sistema sea interoperable semánticamente funcionando como un repositorio de datos universal para aquellas aplicaciones de Realidad Aumentada que puedan interpretar el modelo ontológico proporcionado.

Resultados y Objetivos

En relación con la línea de Navegador de Realidad Aumentada Semántico se está trabajando en el editor y en navegador de Realidad Aumentada semántico. Se espera que el usuario utilice diferentes métodos de entrada en el navegador que dispara la búsqueda de procedimientos a ser visualizados y completados en los diferentes visores semánticos aumentados:

- Realizar una búsqueda de texto para seleccionar manualmente procedimientos que se necesite utilizar.

- Consumir un catálogo virtual aumentado para descubrir procedimientos que se pueden realizar a partir de las entidades virtuales de ese catálogo.

- Buscar procedimientos por voz en particular en resulta útil con el empleo de gafas y cascos de Realidad Aumentada (ya que suele ser engorroso la operación con estas interfaces a partir de mecanismos de interacción convencionales).

- Detección el objeto a aumentar para poder obtener procedimientos e información (por ejemplo: el rostro o la foto de la tarjeta médica para asistir a una persona en situaciones de emergencia).

Para realizar la búsqueda de procedimientos el usuario utilizará alguno de los diferentes métodos de entrada ya mencionados. En la figura 1 se observa que una vez que se dispara alguno de ellos se procederá a buscar procedimientos comunicándose con un middleware semántico. Este middleware tiene las responsabilidades de almacenar los procedimientos creados, exponerlos públicamente mediante un endpoint SPARQL [16] para ser consumidos por otras aplicaciones.

Cada procedimiento almacenado está compuesto por pasos que puede involucrar entidades existentes de DBpedia [17] para enriquecer su descripción. Una vez que el middleware encuentra resultados en su base de datos, se procede a enviarlos hacia el visor de Realidad Aumentada correspondiente para su explotación por parte de los usuarios.

Además, el middleware a desarrollar proveerá una ontología que permitirá que nuestro sistema de catálogos virtuales aumentados [14], [15] pueda ser interoperable semánticamente con otros sistemas de Realidad Aumentada. El objetivo de esta ontología es brindar un modelo semántico para el acceso a datos unificado para aplicaciones de Realidad Aumentada. Ambas ontologías se conectarán

para asociar procedimientos a catálogos virtuales aumentados.

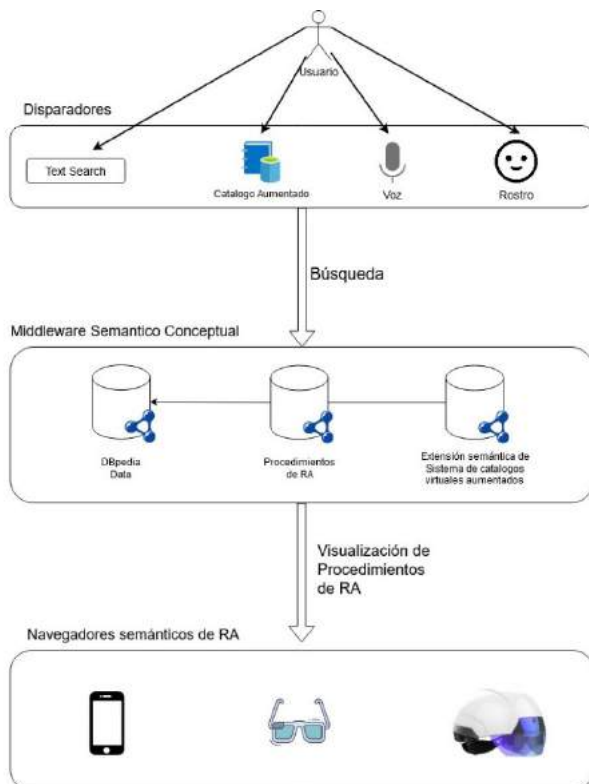


Figura 1 Esquema conceptual de la búsqueda de un procedimiento

En cuanto a la línea de Sistema de Catálogos aumentados semánticos y templates se trabajará en la ontología y el servicio web correspondiente para que el sistema sea una fuente de datos interoperable y accesible para otras aplicaciones de Realidad Aumentada.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por tres investigadores formados y dos investigadores en formación, trabajando en el área de RA. Uno de los investigadores en formación se encuentra realizando el Doctorado en Ciencias Informáticas en la UNLP, particularmente en el área específica del

presente trabajo. Durante el año 2020 se defendió la tesis de doctorado titulada “Integración escalable de Realidad Aumentada basada en imágenes y rostros” (Mangiarua Nahuel, 2020), [18] [19], que conforma una de las líneas de investigación y desarrollo del Grupo de Realidad Aumentada.

Referencias

- [1] Yee C., Abásolo M. J., Más Sansó R. y Vénere M. (2011). “Realidad Virtual y Realidad Aumentada. Interfaces avanzadas.” ISBN 978-950-34-0765-3.
- [2] Ierache J., Mangiarua N., Verdicchio N., Sanz D., Montalvo C., Petrolo F., Igarza S. (2015). “Sistema de Catálogo Virtual Aumentado. Integración de Framework Especializado orientado a juegos didácticos” TEYET, pp 350-356, ISBN 978-950-656-154-3.
- [3] Verdicchio N., Sanz D., Igarza S., Mangiarua N., Montalvo C., Ierache J. (2016) “Sistema de Catálogo Virtual Aumentado Integración de Framework Especializado Orientado a Juegos Didácticos”. TE&ET, pp 597-604, ISBN 978-987-3977-30-5.
- [4] Ierache J., Mangiarua N., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. (2014). “Herramienta de Realidad Aumentada para facilitar la enseñanza en contextos educativos mediante el uso de las TICs”. Revista Latinoamericana de Ing de Software,1(1): -3, ISSN 2314-2642.
- [5] Ierache J., Mangiarua N., Becerra M., Igarza S. Sposito O. Framework for the Development of Augmented Reality Applications Applied to Education Games. In: De Paolis L., Bourdot P. (eds) Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2018. Lecture Notes in Computer Science, vol 10850. Springer, Cham.

p. 340-350.

[6] Ierache N., Mangiarua N., Verdicchio D., Sanz D., Montalvo C., Petrolo F. and Igarza S., "Augmented. Card System Based on Knowledge for Medical Emergency Assistance". IEEE Xplore ISBN 978-1-5090-2938-9 2016.

[7] Ierache J., Verdicchio N., Duarte N., Montalvo C., Petrolo F., Sanz D., Mangiarua N., Igarza S., "Augmented Reality Card System for Emergency Medical Services", IWBBIO 2016 Proceedings Extended abstracts 20 - 22 abril 2016 Granada (SPAIN), pp 487-494, ISBN 978-84-16478-75-0.

[8] LayAR. Disponible en <https://www.layar.com/>. Accedido febrero 2021.

[9] Wikitude. Disponible en <https://www.wikitude.com/>. Accedido febrero 2021.

[10] Aydin B., Gensel J., Genoud P. Extending Augmented Reality Mobile Application with Structured Knowledge from the LOD Cloud. Disponible en https://www.researchgate.net/publication/278241470_Extending_Augmented_Reality_Mobile_Application_with_Structured_Knowledge_from_the_LOD_Cloud. Accedido febrero 2021.

[11] T. Matuszka et. al. The Design and Implementation of Semantic Web-Based Architecture for Augmented Reality Browser. Disponible en https://www.researchgate.net/publication/280068074_The_Design_and_Implementation_of_Semantic_Web-Based_Architecture_for_Augmented_Reality_Browser. Accedido febrero 2021

[12] Nixon L., Grubert J. Reitmayr G. SmartReality: Integrating the Web into Augmented Reality. Disponible en <https://www.researchgate.net/publication/25720>

7430_SmartReality_Integrating_the_Web_into_Augmented_Reality. Accedido en febrero 2021.

[13] Ierache J., Mangiarua N., Verdicchio N., Becerra M., Duarte N., Igarza S. (2014). "Sistema de Catálogo para la Asistencia a la Creación, Publicación, Gestión y Explotación de Contenidos Multimedia y Aplicaciones de Realidad Aumentada". CACIC 2014 Red UNCI ISBN 978-987-3806-05.

[14] Ierache J., Mangiarua N., Bevacqua S., Verdicchio N., Becerra M., Sanz D., Sena M., Ortiz F., Duarte N., Igarza S. (2015). "Development of a Catalogs System for Augmented Reality Applications". World Academy of Science, Engineering and Technology, International Science Index 97, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9(1), 1 - 7. ISSN 1307:6892.

[15] Mangiarua, N., Ierache, J.S., Becerra, M.E., Maurice, H., Igarza, S., & Sposito, O. (2018). Templates Framework for the Augmented Catalog System. CACIC 2018 Red UNCI ISBN 978-3-030-20786-1.

[16] SPARQL. Disponible en <https://www.w3.org/TR/sparql11-overview/>. Accedido en febrero 2021

[17] DBpedia. Disponible en <https://www.dbpedia.org/>. Accedido en febrero 2021

[18] Mangiarua, Nahuel, Jorge Ierache, and María José Abasolo. "Scalable Integration of Image and Face Based Augmented Reality." In Augmented Reality, Virtual Reality, and Computer Graphics, edited by Lucio Tommaso De Paolis and Patrick Bourdot. Cham: Springer International Publishing, https://doi.org/10.1007/978-3-030-58465-8_18

[19] Mangiarua, Nahuel, Jorge Ierache, and María José Abasolo. "Implementation of an Open Source Based Augmented Reality Engine

for Cloud Authoring Frameworks.” *Journal of Computer Science and Technology* 19, no. 2 (October 2019):
<https://doi.org/10.24215/16666038.19.e16>.

Investigación, desarrollo y publicación de un prototipo de segmento terreno satelital

Pablo Soligo , German Merkel, Jorge S. Ierache 

Grupo de Investigación y Desarrollo de Software Aeroespacial (GIDSA)
Departamento Ingeniería e Investigaciones Tecnológicas (DIIT)
Universidad Nacional de la Matanza (UNLaM)

psoligo@unlam.edu.ar; gmerkel@alumno.unlam.edu.ar; jierache@unlam.edu.ar

RESUMEN

Las líneas de investigación y desarrollo presentadas tienen por objetivo demostrar la factibilidad de desarrollar sistemas de segmento terreno satelital costo-efectivos utilizando exclusivamente componentes dentro de los denominados, de estantería, en cualquiera de sus variantes OTS (Del Inglés Off-the-Shelf), COTS (Del Inglés Commercial-Off-The-Shelf) y OSS (Del Inglés Open Source Software) prescindiendo de soluciones propias o de herramientas de escasa penetración en la industria de software de propósito general.

Palabras clave: *Segmento Terreno, Software, Costo-Efectivo.*

CONTEXTO

Las experiencias realizadas en la Maestría en Desarrollos Informáticos de Aplicación Espacial (MDIAE) (Comisión Nacional de Actividades Espaciales-Universidad Nacional de La Matanza), tanto de manera directa, operando unidades de software de segmento terreno 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 de la Universidad Nacional de La Matanza (GIDSA), en el marco de proyecto “Proince C-211: Sistemas de segmento terreno satelital de próxima generación” y actualmente se radica en el proyecto “Proince C-230,

Aprendizaje automático para el control del estado de salud en sistemas Aeroespaciales”. El GIDSA [1] está dedicado a investigar e implementar prototipos de software alternativos de bajo costo basados en las soluciones ampliamente aceptadas, de probada madurez y con penetración en la industria de software de propósito general.

1. INTRODUCCIÓN

El alto costo asociado a las misiones espaciales y la baja propensión a tomar riesgos determina el enfoque en las soluciones implementadas en la industria espacial ([2], [3]). El desarrollo de software está fuertemente orientado al cumplimiento de los requerimientos dificultando una estrategia más amplia y de visión de largo plazo. El grupo de investigación GIDSA desarrolla un prototipo de segmento terreno alternativo genérico, costo efectivo y basado completamente en COTS, OSS o OTS, minimizando no solo el desarrollo sino también el costo de mantenimiento.

La solución que el grupo GIDSA propone una alternativa que prescinde de software propietario o de propósito específico, haciendo a esta misma portable y poderosa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación en desarrollo proponen la exploración de alternativas basadas exclusivamente en técnicas y herramientas de alta penetración en la industria del software, en

particular, aquellas que puedan ser aplicadas en la industria espacial. El objetivo principal es demostrar que las herramientas usadas en la industria de software de propósito general pueden ser aplicadas en la industria espacial, no solamente en el segmento terreno sino también en el segmento de vuelo. La aplicación de estas técnicas puede ayudar a desarrollar soluciones de bajo costo con un alto nivel de mantenibilidad.

Capa de visualización y operación

La capa de visualización del Unlam Ground Segment (UGS) está basada en el NASA OpenMCT (Open Mission Control Technologies). El NASA OpenMCT es el framework de visualización de control de misión de código abierto de la NASA.

Se evaluaron múltiples alternativas, partiendo de la necesidad de tener acceso desde múltiples plataformas. Las soluciones nativas fueron descartadas al requerir un desarrollo específico por plataforma y no mostrarse como una alternativa costo efectiva. Se han explorado soluciones basadas en paneles de comandos, como Grafana (<https://grafana.com/>). Estas alternativas resuelven de manera satisfactoria la visualización, pero son pretendidamente genéricas y carecen de necesidades propias de la industria aeroespacial. La actualización en tiempo real, paneles de comandos, si bien son técnicamente posibles, su implementación no resulta transparente, siendo estas necesidades en un software de control de misión.

Por otro lado, las soluciones SPA se presentan como la alternativa tecnológica que permite un único desarrollo compatible con prácticamente cualquier plataforma, ofreciendo un nivel de interactividad comparable, en determinados escenarios, con una aplicación nativa. Se exploró un desarrollo SPA propio invirtiéndose más de 120 horas alcanzando pobres resultados [4].

Finalmente, la implementación de la capa de visualización sobre OpenMCT fue considerada la mejor opción, ofreciendo una SPA con funcionalidad específica del área espacial y un costo de adaptación inferior al costo de desarrollo de una SPA propia.

Las interfaces del Unlam Ground Segment (UGS) están completamente basadas en la comunicación HTTP/HTTPS, siendo este un estándar en la industria del software. Toda la telemetría es incorporada y puede ser extraída del sistema mediante un servicio REST lo que permitió una integración transparente con el OpenMCT

El prototipo actual (Disponible en <https://ugs.unlam.edu.ar>), [5] permite visualizar telemetría de distintas formas, tales como gráficos y tablas, crear múltiples tableros y paneles, guardar y compartir. La visualización se conecta vía HTTP y JSON modularmente, permitiendo añadir nuevos satélites o nuevas variables de telemetría sin necesidad de modificar la aplicación web en sí. OpenMCT provee al desarrollador un framework con tableros predeterminados, almacenamiento local en el navegador y un módulo de complementos.

Recuperación y persistencia

La definición de datos como alarmas, variables de telemetría, comandos, tipos de datos y formatos son almacenados sobre un RDBMS. El UGS utiliza desde su primera versión un RDBMS con el objetivo de estandarizar el almacenamiento, recuperación, seguridad e integridad de los datos en cualquier nivel de procesamiento. Los datos son accedidos por medio de un ORM, esta capa intermedia provee productividad en el desarrollo e independencia del proveedor del motor de base de datos, haciendo sencillo, dentro de ciertos límites, el cambio de RDBMS.

Por otro lado, el uso de ORMs y RDBMS, con modelos estrictamente normalizados puede establecer un límite en el rendimiento cuando se trabaja con grandes volúmenes de datos. Las tecnologías híbridas, montadas sobre motores relacionales, han sido testeadas en busca de mejores rendimientos tanto en el almacenamiento como en la recuperación para grandes volúmenes de datos.

Detección de fallas

Existen tres métodos comunes de análisis de telemetría y control de salud: control de límites, sistemas expertos y sistemas basados en modelos. El primer método es el más simple y

el más común del grupo, consiste en establecer un valor aceptable máximo y un valor aceptable mínimo para un sensor con la ayuda de un experto, y verificar que los valores de telemetría de dicho sensor se mantengan entre los valores definidos. Este método es completamente insensible al contexto y puede ser tedioso establecer valores límites a cada sensor. El segundo método obtiene mejores resultados, aunque no puede encontrar tipos de fallas no consideradas o indefinidas, y requiere una laboriosa configuración. La idea detrás del tercer método es detectar anomalías y sus razones comparando simulaciones computacionales con el verdadero comportamiento del sistema.

Los tres métodos requieren que un experto continuamente complete ciertas tareas, tales como actualizar límites, crear reglas o parametrizar situaciones. Por otro lado, el aprendizaje automático provee un largo rango de posibilidades para la predicción de comportamientos, y, por tanto, detección de fallas. En lugar de un experto infiriendo reglas y desarrollando modelos, el aprendizaje automático puede ofrecer una manera más eficiente de capturar conocimiento y aplicarlo.

Desde el comienzo, el UGS verifica la salud de un satélite utilizando control de límites tanto como para variables directas como para variables derivadas. Trabajos anteriores [7] muestran la creación de una entidad asociada al tipo de telemetría donde un modelo de predicción es automáticamente creado, por tanto, presentando las habilidades del aprendizaje automático. Estableciendo dinámicamente los valores máximos y mínimos admisibles utilizando la predicción obtenida, se puede lograr que el sistema obtenga controles actualizados sensibles al contexto para el satélite.

El caso de estudio cubre el análisis de un sensor de tensión de baterías durante un eclipse, teniendo en cuenta el tiempo que el satélite se encuentra eclipsado y la tensión de la batería. La tensión de la batería normalmente decrece durante un eclipse dado que los paneles solares del satélite se encuentran parcial o totalmente

tapados por la Tierra. Una vez que los paneles solares enfrentan el Sol, la tensión de la batería se recupera rápidamente.

Usando el método de control de límites, los valores mínimos y máximos deben tener en cuenta el eclipse, y, por tanto, la amplitud de estos debe ser lo suficientemente grande. Sin embargo, existe la posibilidad de que el satélite no se encuentre eclipsado por la Tierra y que su batería se encuentre con baja tensión. Dado que el método de control de límites no conoce cuándo el satélite está o no eclipsado, no detectará ninguna anomalía (aunque en realidad puede haber un serio problema con la carga de la batería). Un grupo de reglas proveídas por un experto puede mejorar esta situación, aunque en este caso un experto es necesario y los valores mínimos y máximos finales son tan solo conocidos con el sistema en vuelo.

Con el aprendizaje automático, un proceso distribuido prueba los tipos de telemetría que tienen modelos de predicción expirados. Por cada tipo de telemetría, el sistema crea una nueva predicción de datos acorde a su configuración, crea un nuevo modelo de predicción y, en caso de ser lo suficientemente preciso, persiste el modelo, por el contrario, se señala una alarma.

Se representa en la Figura 1 - Arquitectura Conceptual muestra los principales módulos presentes en el prototipo.

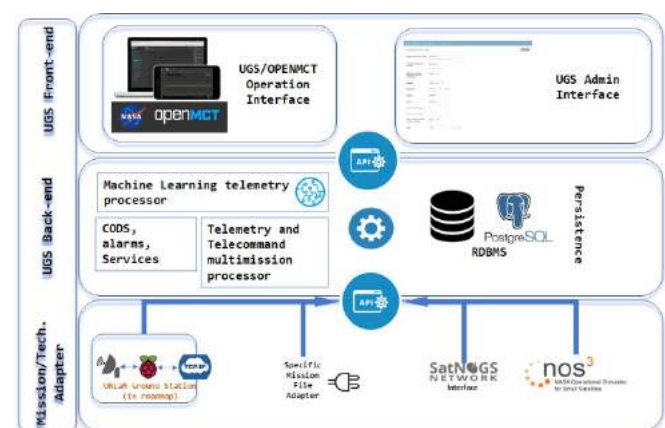


Figura 1 - Arquitectura Conceptual

La Figura 2 muestra un dashboard ejemplo con datos del satélite de satellogic butsat1 (Tita).

Accesible en <https://ugs.unlam.edu.ar/#/browse/mine/demodash>



Figura 2 – UGS/NASA OPENMCT Front-end

3. RESULTADOS OBTENIDOS/ESPERADOS

El UGS ha sido publicado [8] en el año 2020, donde la telemetría de los satélites BugSat1/TITA y CSIM-FD puede ser explorada y analizada. La telemetría de dichos satélites ha sido provista mayormente por la red SatNOGS. La actual capa de visualización provee una manera interactiva de explorar la telemetría de ambos satélites de manera histórica y parcialmente en tiempo real. Los paquetes son decodificados usando el procesador de telemetría descripto. Características específicas son procesadas utilizando scripts programados en lenguajes de propósito general, cargados en tiempo de ejecución [3]. Se espera añadir a la capa de visualización del UGS una interfaz completa de telemetría en tiempo real y de envío de comandos. Los scripts de comandos serán desarrollados, como se ha dicho, un lenguaje de programación de propósito general. Aunque el OpenMCT sea un framework completo, se han requerido no menos de 100 horas de desarrollo para lograr una integración parcial. Se requiere entrenamiento en el lenguaje, herramientas y estructuras necesarias para la adaptación. El framework está bien documentado, pero carece de tutoriales útiles y ejemplos, y su desarrollo puede ser complejo. Con todo, OpenMCT fue considerado la mejor opción dado el alto costo del desarrollo de SPA: problemas que el framework soluciona parcialmente.

En términos de persistencia el uso de soluciones híbridas no han mostrado una clara ventaja que justifique su implementación [6].

Las primeras implementaciones de aprendizaje automático aplicado a la telemetría satelital han dado resultados prometedores [3] aunque no está exenta de problemas o de puntos de investigación por explotar. La búsqueda de relaciones o las variables de contexto todavía requieren del asesoramiento de un experto para su correcta identificación. Una de las futuras líneas de investigación es la implementación de una solución al problema de encontrar correlaciones y la corrección automática de estas, esperando utilizar completamente el poder computacional, no para reemplazar al experto, sino para encontrar en él un aliado en la búsqueda e identificación de patrones escondidos.

4. FORMACIÓN DE RECURSOS HUMANOS

Los prototipos desarrollados presentan una plataforma realista de experimentación. Le permite a investigadores y estudiantes probar soluciones de software, obtener límites, comparar alternativas y establecer criterios de decisión. La posibilidad de trabajar con datos de varias misiones espaciales desde pequeñas misiones universitarias hasta grandes misiones científicas permite responder a la premisa de desarrollar un sistema terreno transparente al satélite en órbita. Actualmente el grupo de investigación este compuesto por un investigador formado, un investigador en formación y un alumno investigador becario BIC (Beca de investigación científica UNLaM).

5. BIBLIOGRAFÍA

- [1] <https://gidsa.unlam.edu.ar/>, *Grupo de Investigación y Desarrollo de Software Aeroespacial de la Universidad Nacional de La Matanza*, 2020.
- [2] P. Soligo y J. S. Ierache, «Segmento Terreno Para Misiones Espaciales de Próxima Generación,» *WICC 2019*.
- [3] P. Soligo y J. S. Ierache, «Software de segmento terreno de próxima generación,»

- de *XXIV Congreso Argentino de Ciencias de la Computación (La Plata, 2018)*, 2018.
- [4] P. Soligo y J. S. Ierache, «Experiences and lessons learned developing a next-generation ground segment prototype,» *2nd IAA Latin American Symposium on Small Satellites*.
- [5] *openmct nasa open mission control technologies*.
- [6] P. Soligo, G. Merkel y J. S. Ierache, «Telemetría de altas prestaciones sobre base de datos de serie de tiempos,» *REDDI 2021*, 2021.
- [7] P. Soligo y J. S. Ierache, «Arquitectura de segmento terreno satelital adaptada para el control de límites de telemetría dinámicos,» de *XXV Congreso Argentino de Ciencias de la Computación (La Plata, 2019)*, 2019.
- [8] <https://ugs.unlam.edu.ar/>, *Unlam Ground Segment*, 2020.
- [9] J. J. Ramos Pérez, «A design for an advanced architecture of satellite ground segments,» 2014.

ACRÓNIMOS

COTS
Commercial-Off-The-Shelf.....

GIDSA
Grupo de Investigación y Desarrollo de Software Aeroespacial.....

HTTP
Protocolo de transferencia de hipertexto, del Inglés Hypertext Transfer Protocol.....

HTTPS
Protocolo de transferencia de hipertexto seguro, del inglés Hypertext Transfer Protocol Secure.....

NASA
Administración Nacional de Aeronáutica y el Espacio, del Inglés, National Aeronautics and Space Administration.....

OpenMCT
Software de control de misión abierto, del Inglés Open Source Mission Control Software.....

ORM
Mapeo objeto-relacional, del inglés Object-Relational Mapping.....

OSS
Open Source Software.....

OTS
Off-the-Shelf.....

RDBMS
Sistema de administración de base de datos relacional, del Inglés Relational Database Management System.....

REST
Transferencia de Estado Representacional, del inglés representational state transfer.....

SatNOGS
Satellite Networked Open Ground Station.....

SPA
Aplicación de página única, del Inglés Single Page Application.....

UGS
UNLaM Ground Segment.....

JUEGOS SERIOS DE REALIDAD AUMENTADA ORIENTADOS A ENTRENAR Y RECUPERAR MOVIMIENTO EN NIÑOS CON DISCAPACIDAD

Cruz Alejandro¹, Acosta Nelson²

¹Escuela de Ingeniería, Universidad Nacional de Chilecito (UNdeC), Chilecito, La Rioja

²Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN), Tandil, Buenos Aires

acruz@undec.edu.ar, el.nelson.acosta@gmail.com

RESUMEN

El sector de videojuegos viene transitando un crecimiento sostenido desde hace varios años. El estilo de vida sedentario es peligroso para la salud, y jugar videojuegos en muchos casos contribuye a este problema. La evolución de los teléfonos inteligentes y el avance de la tecnología de realidad aumentada, permiten el desarrollo de nuevos juegos combinando la diversión y el entretenimiento con el ejercicio físico. Caminar es una actividad recreativa muy recomendable no solo para mantenerse en forma, sino también para combatir problemas de circulación, enfermedades del corazón, sobrepeso, entre otras. En este artículo presentamos el proyecto Ragamese, el cual tiene como objetivo el desarrollo de juegos de realidad aumentada para incentivar la actividad física al aire libre y el posterior análisis de los datos obtenidos de los entrenamientos.

Palabras Clave: Juegos de Realidad Aumentada, Juegos Serios, Actividad Física. Rehabilitación Mediante Juegos.

CONTEXTO

Los videojuegos se han convertido en una de las principales industrias del entretenimiento y cada vez cobran una mayor importancia en el ámbito del ocio.

El sector de los videojuegos es uno de los sectores tecnológicos con mayor proyección de crecimiento tanto a nivel nacional como a escala mundial. Según el informe Global Game Market Report 2020 [1], el 49 % de los ingresos mundiales de la industria de los videojuegos corresponde al segmento de los dispositivos móviles (smartphone y tablets), lo que supone 77.300 millones de dólares (el 82% proviene de juegos para smartphones); el 28 % corresponde a los juegos de consola, generando ganancias por 45.200 millones de dólares y el 23% restante pertenece a los juegos para PC, obteniendo ingresos por 36.900 millones de dólares.

En Argentina, la industria de videojuegos conforma un sector relativamente nuevo comenzando a principios de este siglo. En menos de 20 años, este sector ha crecido de manera sostenida, situándose en una posición competitiva a nivel regional con empresas radicadas en distintos puntos del país [2].

1. INTRODUCCIÓN

El estilo de vida sedentario se ha vuelto bastante común hoy en día y jugar videojuegos, se puede considerar como uno de los comportamientos que contribuyen a este problema de salud.

Según la Organización Mundial de la Salud (OMS), al menos un 60% de la población mundial no realiza actividad física para obtener beneficios para la salud [3].

La mayoría de los videojuegos que encontramos en el mercado, no están orientados al movimiento del participante sino más bien a un juego pasivo, de movimiento reducido que obliga al jugador a estar sentado para ejecutarlo, imposibilitándolo de explorar su entorno y objetos que lo rodean.

Los niños se han vuelto menos activos físicamente en las últimas décadas. La expansión de la tecnología de entretenimiento en el hogar ha provocado que dediquen una mayor parte de su tiempo libre a actividades de tipo sedentarias. Estos factores están afectando las experiencias de la infancia al disminuir el tiempo para jugar al aire libre y pueden contribuir a tendencias indeseables en la salud de los niños.

La inactividad física durante los primeros años de vida es un factor que colabora en el incremento de los niveles de obesidad y de otros trastornos médicos que se observan en niños, niñas y adolescentes [4][5].

Con este proyecto surge Ragamese [6] una herramienta para crear videojuegos de Realidad Aumentada (RA), innovadores y entretenidos para dispositivos móviles compatibles con el sistema operativo Android.

El avance tecnológico que han experimentado los smartphones y la tecnología de la RA, nos ha permitido desarrollar videojuegos con el objetivo de incentivar la actividad física al aire libre.

Investigaciones indican que los niños son más activos físicamente cuando juegan al aire libre, particularmente en espacios naturales, estimulando la actividad física lo que potencialmente reduce enfermedades como la obesidad infantil [7].

El entorno exterior brinda oportunidades de juego que difícilmente se puedan reproducir en el interior y donde es posible experimentar libertad y estar en contacto con elementos naturales [8].

El juego es una actividad exploradora, de aventura y experiencia, indispensable para el desarrollo físico, intelectual y social del niño [9].

Los niños, en promedio, son aproximadamente dos veces más activos cuando están fuera de sus hogares. Caminar y jugar fuera de casa pueden contribuir significativamente al volumen de actividad física de los niños [10].

El uso de la tecnología de la RA, permite ampliar el universo del juego añadiendo modificaciones virtuales al mundo real, que solo ocurren a través del dispositivo con el que se está observando el escenario.

El niño deberá caminar, recorrer y explorar determinadas áreas, permitiendo la interacción con el mundo real y los objetos que lo rodean, ayudando de esta manera a fomentar la actividad física y eliminar las limitaciones de los videojuegos basados en ubicaciones fijas. Caminar es una forma común, accesible y económica de actividad física, que proporciona numerosos beneficios para la salud [11].

Además de entretener, incentivando el movimiento y la actividad física, los videojuegos desarrollados fomentan el aprendizaje, persiguiendo la educación del jugador a través de diferentes preguntas que aparecen a lo largo del juego. Un juego serio es aquel en el que la educación, en sus diversas formas, es el principal objetivo [12]. Son juegos que se usan para educar, entrenar e informar [13].

Los videojuegos están basados en un motor de RA de diseño propio que nos permite acceder a los diferentes servicios del smartphone, tales

como GPS, cámara, giróscopo, acelerómetro, entre otros. Además, se aplican un conjunto de filtros y funciones para el tratamiento de los datos obtenidos, facilitando la construcción del juego, manteniendo objetos geo-localizados, cálculos de ángulos, posiciones y tamaños de los distintos objetos de acuerdo a los grados de libertad con los que cuenta el dispositivo móvil.

También se destaca el control parental, que es una característica especialmente útil para padres. Permite configurar los temas sobre los cuales se realizarán las diferentes preguntas en el juego (geometría, animales, sumas, restas, colores, frutas, verduras, entre otros) y delimitar el tiempo de uso del juego por parte del niño para que el mismo sea consumido en forma óptima y regulada.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La recuperación de movimiento en niños con problemas motrices, es una rama en la cual no existen o son muy pocos los juegos serios desarrollados que utilicen la tecnología de RA y los dispositivos móviles para incentivar el entrenamiento.

Por lo general, los juegos que encontramos no son tan atractivos como la mayoría de los juegos que existen en el mercado. Esto se debe principalmente al acotado presupuesto con los que generalmente se cuenta.

De cualquier manera, con el trabajo y la vinculación que se tiene día a día con este sector, hemos encontrado necesidades que van transformando nuestro proyecto a diario.

En la actualidad estamos trabajando en tres líneas de investigación:

- Motor de videojuegos (serios) de RA con algunos desarrollos para incentivar la actividad física en niños. Se realizó una primera versión del

juego, para establecer las características que debía tener un motor, luego se llevó a cabo el desarrollo del motor, y por último un juego utilizando el motor propio. Esto nos va a permitir crear juegos a medida que pueden ser adaptados para diferentes patologías con distintos tipos de entrenamientos.

- Analizar los datos obtenidos de los entrenamientos. Este juego genera muchos datos que aportan información valiosa para considerar posteriormente. Nos permite definir ciertos parámetros relacionados con la dispersión en la distribución de descargas y lugares donde se juega; distancia recorrida en cada uno de los entrenamientos realizados por cada niño, los cuáles nos brindan una pauta sobre cómo fue la evolución con respecto a su estado físico, entre otros.
- Generar nuevos juegos enfocándonos en diferentes patologías. La inactividad física es un problema de salud mundial que se acentúa aún más en niños con diagnósticos dentro del espectro autista, diabetes, asperger, sobrepeso, obesidad infantil, entre otras, donde el sedentarismo tiene graves consecuencias.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se ha logrado desarrollar dos videojuegos de realidad aumentada para smartphones compatibles con Android, denominados "Cofre del Tesoro Perdido" y "Buscando a Bobby". Están basados en el tradicional juego de la búsqueda del tesoro. Para moverte en el mundo virtual obligan al jugador a desplazarse por el mundo real. El número de descargas de ambos juegos es de aproximadamente 14160, en donde el 95% corresponde al juego "Buscando a Bobby".

Con respecto a las zonas de influencia, podemos mencionar países de Latinoamérica (Colombia, México, Brasil, Chile, Perú, Argentina) y algunos países del sur asiático como por ejemplo la India y Pakistán, entre otros.

La RA es una tecnología que nos permitió crear videojuegos innovadores, con el objetivo de fomentar la actividad física en los niños. Nuestros juegos permiten a los niños adquirir nuevas experiencias a nivel motor abandonando el sedentarismo que hasta ahora fueron propuestos por otros videojuegos. Además, generan un ambiente de aprendizaje que les permite adquirir nuevos conocimientos a través del uso del juego.

Actualmente las acciones se encuentran focalizadas en un análisis detallado de los datos obtenidos del uso de los videojuegos. Esto nos permitirá detectar patrones de entrenamiento de nuestros jugadores (tiempo de uso, lugares de juego, mapas creados, distancia recorrida, entre otros). De esta forma, en base a los resultados obtenidos, se podrán diseñar nuevas estrategias que nos permitan incentivar la actividad física en los niños.

Al contar con nuestro propio motor de RA, tenemos la posibilidad en futuros proyectos de crear videojuegos diseñados especialmente para niños con diferentes patologías (sobrepeso, obesidad, diabetes, etc.) donde se recomiende el ejercicio de caminata controlado.

En este momento estamos trabajando de forma conjunta con un instituto de rehabilitación definiendo líneas de trabajo para utilizar nuestros juegos en la rehabilitación de niños con distintas patologías, principalmente motoras y neuronales. De manera de poder adaptar los juegos a los distintos niveles de entrenamientos requeridos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de este proyecto trabaja desde hace tres años en el desarrollo de software orientado al entretenimiento. Uno de los integrantes del grupo obtuvo una Beca Doctoral de Agencia.

El proyecto cuenta con la participación de dos alumnos de grado de la carrera de Ingeniería en Sistemas de la UNICEN.

Actualmente se esperan avances del Ing. Cruz quien está haciendo el Doctorado en Matemática Computacional e Industrial de la Facultad de Ciencias Exactas de la UNICEN dirigido por el Dr. Acosta.

5. BIBLIOGRAFÍA

- [1] Newzoo, "Newzoo Global Games Market Report 2020 | Light Version". Recuperado de: <https://newzoo.com/insights/trend-reports/newzoo-global-games-market-report-2020-light-version/>, (Marzo de 2021).
- [2] Gala, R. (2019). Mapa de la Industria Cultural de Videojuegos en Argentina: un estado de situación. SADIO EJS, 18 (2), pp. 103-118.
- [3] OMS, "Inactividad física: un problema de salud pública mundial". Recuperado de: https://www.who.int/dietphysicalactivity/factsheet_inactivity/es/, (Marzo de 2021).
- [4] Speiser PW, Rudolf MC, Anhalt H, Camacho-Hubner C, Chiarelli F, Eliakim A, Freemark M, Gruters A, HersHKovitz E, Iughetti L et al. Childhood obesity. J Clin Endocrinol Metab 2005, 90:1871-1887.
- [5] Livingstone MB. Childhood obesity in Europe: a growing concern. Public Health Nutr 2001, 4:109-116.
- [6] Ragamese. <http://www.ragamese.com/es/inicio/>

[7] Herrington, Susan & Brussoni, Mariana. (2015). Beyond Physical Activity: The Importance of Play and Nature-Based Play Spaces for Children's Health and Development. *Current obesity reports*. 4. 10.1007/s13679-015-0179-2.

[8] Maynard, Trisha & Waters, Jane. (2007). Learning in the outdoor environment: A missed opportunity?. *Early Years*. 27. 255-265. 10.1080/09575140701594400.

[9] Roncancio Melgarejo C. R, Sichacá Ávila E. G, La actividad Física Como Juego en la Educación Inicial de los Niños Preescolares, Trabajo de Grado: Especialistas en Epidemiología, Universidad de Antioquia - Facultad Nacional De Salud Pública Hector Abad Gómez, 2009.

[10] Mackett, Roger & Paskins, James. (2008). Children's Physical Activity: The Contribution of Playing and Walking. *Children & Society*. 22. 345 - 357. 10.1111/j.1099-0860.2007.00113.x.

[11] Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U. Wells, J. C.: Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*, 380: 247–257 (2012).

[12] Deterding S, Dixon D, Khaled R, Nacke L.: From game design elements to gamefulness: defining “gamification”. *Proceedings of the 15th International Academic MindTrek Conference, Tampere, Finland (2011)*.

[13] MICHAEL, D. Y CHEN, S. (2006). *Serious Games. Games that educate, train and infoms*. Canadá: Thonsom.

Mejorando la performance en Aplicaciones Web Progresivas mediante estrategias de utilización de la Cache

Rocío A. Rodríguez, Pablo M. Vera, M. Roxana Martínez, Franco Cifuentes, Claudia G. Alderete, Mariano G. Dogliotti

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{rocioandrea.rodriguez, pabломartin.vera, roxana.martinez, claudia.alderete, mariano.dogliotti}@uai.edu.ar

francoignacio.cifuentes@alumnos.uai.edu.ar

RESUMEN

Las aplicaciones web progresivas (PWA) han permitido trasladar las ventajas de las aplicaciones nativas (funcionamiento desconectado, acceso a hardware, uso de almacenamiento local, notificaciones, icono de acceso, etc...) a las aplicaciones web. Se cuenta entonces con aplicaciones que como toda solución web destacan por su portabilidad, pero con los beneficios y forma de uso de las aplicaciones nativas. En esta línea de investigación, el foco está puesto en el uso de almacenamiento interno en el dispositivo (cache) para generar estrategias que permitan optimizar esa cache disminuyendo la necesidad de consumir datos de la red. Para lo cual se diseñó una estrategia que une varios enfoques de cache existentes para generar una solución integral. Lo que permitirá comparar la performance de las soluciones existentes que utilizan una solución tradicional de manejo de cache, con la propuesta optimizada del equipo de investigación.

Palabras clave: Web Móvil, Dispositivos Móviles, PWA, Cache

CONTEXTO

Esta línea de 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 tanto de sede Centro como de la Castelar (ambas en la provincia de Buenos Aires). El proyecto cuenta con financiamiento asignado y una duración de 2 años.

1. INTRODUCCIÓN

“La rápida expansión y adopción de los teléfonos móviles han generado cambios sociales y culturales en la sociedad, han modificado las formas de comunicación, de acceder a la información y las maneras en que los individuos se relacionan entre sí. La telefonía móvil ha producido una transformación en los ritos sociales de interacción. El uso de las TIC ha modificado la forma de trabajar, aprender, colaborar, jugar, pasar el tiempo y socializar de los individuos” [1]. Es por esto que resulta necesario que la web también de respuesta a nuevas necesidades, primeramente, a través del diseño adaptativo [2] en donde la misma solución podía ser visualizada sin dificultades en diversos dispositivos y actualmente derivando en las aplicaciones web progresivas (PWA).

Las PWA presentan diversas ventajas, por supuesto la portabilidad por ser aplicaciones web pero además incorporan las características de las aplicaciones nativas, entre lo que se destaca poder trabajar en forma desconectada y accediendo a la cache del dispositivo.

“Las aplicaciones web progresivas son una evolución natural de las aplicaciones web que difuminan la barrera entre la web y las aplicaciones, pudiendo realizar tareas que generalmente solo las aplicaciones nativas podían llevar a cabo. Algunos ejemplos son las notificaciones, el funcionamiento sin conexión a Internet o la posibilidad de probar una versión más ligera antes de bajarte una aplicación nativa de verdad” [3].

Para lograr esto debe modificarse la cache cuando los datos se modifican y no se cuenta con conexión al servidor para actualizar, también es necesario poder detectar sino existieron cambios y en ese caso usar la cache sin recurrir al servidor... es decir en cada situación habrá una estrategia para el acceso y recuperación de datos. Actualmente existen enfoques para el acceso a los datos: (1) Solo cache (2) Solo red (3) Cache y si falla red (4) Red y si falla cache (5) Carrera entre cache y red (6) Cache y después red (7) Contenido de reserva

Para una solución que es absolutamente estática se podría usar el esquema “sólo cache”. Para una solución que siempre es dinámica se podría usar “solo red”. Para una solución integral, empieza a ser importante establecer estrategias de uso de la cache en las distintas situaciones que pueden ocurrir. Por otra parte utilizar la cache del dispositivo en vez de descargar todo desde la red, permitirá disminuir el tráfico y consumo de datos. Nicolas Gallagher un desarrollador de la red social Twitter en su artículo “How we built Twitter Lite” [4], indica que gracias a la versión PWA consiguieron reducir el consumo de datos hasta fracciones irrisorias en comparación con sus aplicaciones nativas.

Para que una aplicación pueda tener las características previamente mencionadas es preciso tomar en cuenta diversas pautas al momento de su desarrollo. Resulta interesante un checklist provisto por google [5], en donde se destacan 8 pautas:

- Hosteado en https

- Usa diseño adaptativo y se visualiza correctamente en mobile y tablets
- Todas las páginas deben funcionar cuando se no se tiene conexión
- Debe tener metadatos para dar la opción “Agregar a la pantalla de Inicio” lo que permite “instalar” la aplicación.
- Debe tener un inicio rápido aún en redes lentas (<10 seg en redes 3g)
- Crossbrowsing (visualizarse correctamente independientemente del navegador).
- Los cambios de página deben ser rápidos
- Cada página debe tener su url y si es una app de una solo página con distintas vistas se debe poder reconstruir para permitir acceder directamente a cada vista.

Los componentes principales de una PWA son:

1. Archivo de Manifiesto
2. Service Worker
3. Almacenamiento Local
4. Notificaciones

En la figura 1, se presenta una captura de pantalla en donde puede observarse que, al ingresar por primera vez a la web, en la parte inferior la pantalla aparece un mensaje con el ícono de la aplicación en donde se ofrece agregarla a la pantalla principal.

Al instalar la aplicación se puede acceder desde un ícono a la misma y ya se no se muestra la barra del navegador (se ha elegido ocultarla) y en el caso de la pantalla de la derecha de la figura 2 se ha integrado la barra de título con la de estado con los íconos clásicos del sistema operativo. La forma final que se visualizará será configurada por los desarrolladores según un archivo de manifiesto, donde entre otras cosas se puede seleccionar: modo de visualización, orientación de la pantalla, íconos, color de fondo, color de tema, idioma, nombre corto para la aplicación...



Figura 1. Ejemplo Instalación de una PWA

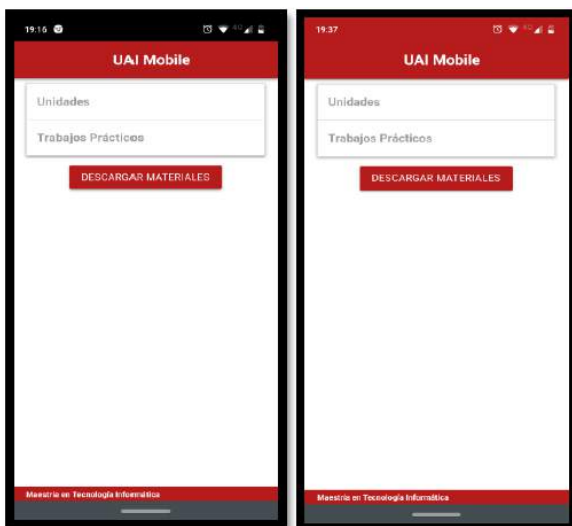


Figura 2. Ejemplo de Visualización de una PWA

Existen diversos artículos académicos que implementan PWA en una amplia variedad de temáticas, en menor medida hay trabajos en el área que se enfocan en el manejo y optimización de cache [6], [7], [8], [9] y [10].

El manejo de cache de forma optimizada permite mejorar la performance de las PWA, disminuyendo el consumo de datos de la red. En base a esta premisa se establecen las líneas de investigación y desarrollo para este año del proyecto de investigación.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Analizar los esquemas de cache utilizados por las PWA actuales existentes en el mercado.
- Evidenciar mejora de performance optimizando el uso de la cache en las PWA existentes.
- Disponibilizar APIs que permitan implementar la estrategia de manejo de cache propuesta por el equipo de investigación

3. RESULTADOS OBTENIDOS/ESPERADOS

En el primer año del proyecto de investigación se realizó una propuesta de estrategia de manejo de cache que combina los esquemas existentes, dicha propuesta es desarrollada e implementada para la cual se definieron pruebas de funcionamiento.

Las pruebas planteadas fueron: (1) Activación de service worker y descarga de archivos, (2) Aumentar versión de caché en 1 y agregar nuevo archivo modificado (3) Aumentar versión de caché en 1 y no modificar arreglo de paginasModificadas (4) Aumentar versión de caché en 2 para simular salteo de versiones y agregar nuevo archivo modificado (5) Aumentar versión de caché en 2 para simular salteo de versiones y no agregar archivo al arreglo de paginasModificadas (6) Traer listado de clientes por primera vez (7) Traer listado de clientes con fecha de caché menor a fecha de actualizacion de la base de datos (8) Traer listado de clientes con fecha de caché

mayor a fecha de actualización de la base de datos (9) Traer listado de clientes sin servicio de internet.

Todas estas pruebas junto a sus resultados esperados fue lo que se validó en cuanto al funcionamiento de la estrategia de cache propuesta. En esta segunda etapa se comparará soluciones de terceros y su impacto al aplicar esta metodología integral de cache para evaluar la performance.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está formado por 5 docentes: 2 de ellos doctores en Ciencias Informática graduados en la Universidad Nacional de La Plata (UNLP), 1 con maestría finalizada (UAI) con la cursada del doctorado finalizada y que se encuentra actualmente haciendo su tesis doctoral (UNLP) y 1 realizando actualmente su maestría en UAI. Esto implica que 4 de los 5 docentes que componen el grupo tienen estudios de posgrados finalizados o en progreso.

Este proyecto también cuenta con la participación de alumnos de grado y posgrado de la UAI (actualmente en el proyecto se encuentran vinculados 4 alumnos).

En el área de dispositivos móviles se encuentran en realización 3 tesis de maestría (2 en la UAI y 1 en UNLaM – Universidad Nacional de La Matanza) y 1 tesina de grado (UAI), siendo directores de dichas tesis miembros del equipo de investigación.

5. BIBLIOGRAFIA

- [1] Castro Rojas, S. "Ubicuidad y comunicación: los Smartphones." Chasqui: Revista Latinoamericana de Comunicación 118 (2012): 91-95.
- [2] Rodríguez, R. A., Vera, P. M., Ramirez, M. A., Alderete, C. G., Conca, A. G., Dogliotti, M. G., & Zain, G. A. (2020). Análisis del Diseño Web Adaptativo Caso de estudio: Universidad Argentinas. Revista Abierta de Informática Aplicada (RAIA), (4), 51-62.
- [3] Ramírez Ivan (2018), ¿Qué es una aplicación web progresiva o PWA?. <https://www.xataka.com/basics/que-es-una-aplicacion-web-progresiva-o-pwa>
- [4] N. Gallagher, "How We Built Twitter Lite," Apr. 2017. https://blog.twitter.com/engineering/en_us/topics/open-source/2017/how-we-built-twitter-lite.html
- [5] Google Developers. "Progressive Web Apps" <https://developers.google.com/web/progressive-web-apps/>
- [6] Rojas, C. (2020). Caching Strategies. In Building Progressive Web Applications with Vue.js (pp. 67-81). Apress, Berkeley, CA.
- [7] Hajian, M. (2019). Advanced Angular Service Worker and Runtime Caching. In Progressive Web Apps with Angular (pp. 107-139). Apress, Berkeley, CA.
- [8] Kim, H. G., Park, J. T., Choi, M. H., & Moon, I. Y. (2019). Design of Web Content Update Algorithm to Reduce Communication Data Consumption using Service Worker and Hash. The Journal of Advanced Navigation Technology, 23(2), 158-165.
- [9] Gambhir, A., & Raj, G. (2018). Analysis of Cache in Service Worker and Performance Scoring of Progressive Web Application. 2018 International Conference on Communications (COMM), Communications (COMM), 2018 International Conference On, 01–06. <https://doi.org/10.1109/ICComm.2018.8430158>

- [10] Malavolta, I., Chinnappan, K., Jasmontas, L., Gupta, S., & Soltany, K. A. K. (2020, July). Evaluating the impact of caching on the energy consumption and performance of progressive web apps. In *Proceedings of the IEEE/ACM 7th International Conference on Mobile Software Engineering and Systems* (pp. 109-119).

PROCESAMIENTO DE DATOS METEOROLÓGICOS PARA DETERMINAR LA OCURRENCIA DE HELADAS EN LA AGRICULTURA.

María Masanet*, Flavio Capraro **, Raúl Klenzi*, Martín Muñoz*,

*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: {mimasanet, rauloscarklenzi,}@gmail.com

** Instituto de Automática (INAUT), UNSJ - CONICET
Av. Lib. Gral. San Martín 1112 (o), San Juan, ARGENTINA,
Teléfonos: 264-4213303, sitio web: www.inaut.unsj.edu.ar
e-mail: fcapraro@inaut.unsj.edu.ar

RESUMEN

Este trabajo aborda el procesamiento de datos climáticos y la investigación, desarrollo y análisis de algoritmos que contribuyan a determinar el pronóstico del fenómeno meteorológico de la ocurrencia de helada. Se realizará en base al análisis de distintas variables asociadas al clima, registradas en estaciones meteorológicas automáticas (con fines agronómicos) ubicadas estratégicamente en una zona agroproductiva de la provincia de San Juan. El objetivo de este estudio es desarrollar una herramienta moderna para que el productor agropecuario pueda tomar decisiones acertadas ante la ocurrencia de este tipo de fenómeno para evitar o mitigar el daño que ocasiona en los cultivos.

Palabras clave: Agricultura Inteligente, Ciencia de Datos, Meteorología, Heladas.

CONTEXTO

La propuesta de investigación se enmarca dentro de los objetivos planteados en el proyecto PIO N°84 "TELEMETRÍA AGRÍCOLA, una herramienta tecnológica para la gestión eficiente del riego, supervisión de cultivos, y generación de alertas" que se

lleva delante en dependencias del Instituto de Automática (INAUT), de la Facultad de Ingeniería de la Universidad Nacional de San Juan (UNSJ); planteándose como una propuesta de trabajo final de alumnos pertenecientes a la carrera Licenciatura en Ciencias de la Computación del Departamento de Informática FCFN UNSJ, ámbito en el cual se llevó adelante también, el proyecto "Visualización y Deep Learning en Ciencia de Datos" bienio 18-19 que dio origen a la propuesta y que se extiende en el actual proyecto bienio 20-21 "Evaluación de visualizaciones eficientes en ciencia de datos" ambos enmarcados en el Laboratorio de Sistemas Inteligentes para Extracción de Conocimiento en Datos Masivos del Instituto de Informática de la Facultad de Exactas Físicas y Naturales de la UNSJ –FCFN– UNSJ– desde donde surgirán propuestas de preprocesamiento, análisis y procesamiento de datos así como alternativas de visualización que permitan presentar adecuadamente la predicción de ciertas variables climáticas asociadas a los fenómenos meteorológicos bajo estudio. Es de destacar que las actividades relacionadas con ambos proyectos se demoraron y postergaron producto de la pandemia y protocolos de

aislamiento y distanciamiento dispuestos por los gobiernos nacionales, provinciales y la propia autoridad universitaria que contiene a los integrantes de los citados proyectos.

1. INTRODUCCIÓN

La actividad agrícola debe hacer frente a la problemática que ocasionan los distintos fenómenos meteorológicos adversos que suceden, como es el viento, el granizo, las heladas, entre otros. En algunos casos el productor agrícola se vale de la experiencia propia para pronosticar un factor meteorológico adverso, lo cual implica gran esfuerzo, dedicación y conocimiento de campañas anteriores. Una alternativa moderna es usar una herramienta tecnológica que monitoree de forma automatizada los registros de distintas variables meteorológicas de la zona donde se encuentra el cultivo, por tener algunos de estos fenómenos características locales, y así pronosticar la ocurrencia y magnitud de la adversidad climatológica para activar mecanismos de resguardo del cultivo; con ello se optimiza el uso de recursos, el trabajo de los productores y mejora la certeza del pronóstico.

Proteger las plantas contra los efectos letales de las bajas temperaturas es importante en la agricultura [1]. Las heladas generan daños significativos en esta actividad, causando pérdidas de cosechas de todo un año y comprometiendo los ingresos del año siguiente.

Los valores de los datos meteorológicos son particulares de una zona, presentando variaciones entre zonas a pesar de que se encuentren próximas, por ejemplo, los balances de radiación en una zona de la superficie terrestre dependen de la ubicación sobre la Tierra, porque la inclinación de los rayos solares que llegan a la zona influye en la cantidad de energía que ésta recibe [2].

Resulta necesario registrar y analizar los datos meteorológicos específicos del lugar donde se encuentra el cultivo para estimar la ocurrencia de un factor meteorológico. Estos datos son medidos por estaciones meteorológicas automáticas montadas para fines agronómicos. Para este estudio se consideran los datos de dos estaciones ubicadas en la provincia de San Juan. Una de ellas situada en el predio de la Estación Experimental Agropecuaria San Juan del Instituto de Tecnología Agropecuaria en el departamento de Pocito y la otra situada en el Establecimiento San Francisco S.A. (explotación privada), ubicado en la localidad de Cañada Honda, departamento Sarmiento, ambas se encuentran separadas por una distancia de 37km aproximadamente. Estas estaciones registran distintas variables meteorológicas cada 10 minutos, proveyendo datos que son la fuente de información para el análisis de fenómenos meteorológicos en la zona.



Fig 1: Ubicación Geográfica del centro de recepción de datos INAUT y las estaciones meteorológicas INTA EEA San Juan y Est. San Francisco S.A.

El objetivo del presente trabajo es el de llevar a cabo la investigación, desarrollo y análisis de algoritmos de predicción ocurrencia de heladas a partir de los datos (previamente procesados y filtrados) obtenidos en las estaciones meteorológicas de la INTA EEA San Juan y del Establecimiento San Francisco, los cuales se envían

telemétricamente al Instituto de Automática, de la Facultad de Ingeniería de la UNSJ.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La realización de este trabajo ha sido posible debido al financiamiento del proyecto PIO N°84, “TELEMETRÍA AGRÍCOLA, una herramienta tecnológica para la gestión eficiente del riego, supervisión de cultivos, y generación de alertas” cofinanciado por CONICET y la Secretaria de Ciencia, Tecnología e Innovación (SECITI) del gobierno de San Juan y con el apoyo brindado por INTA EEA San Juan.

El objetivo del proyecto es desarrollar un sistema de telemetría y generar una herramienta tecnológica que permita al productor un acercamiento a las nuevas tecnologías disponibles para gestionar el riego con precisión (sensores, aplicaciones informáticas, gestión en línea del riego, reportes sobre aplicación del riego, consumo hídrico de los cultivos, etc.) [3][4][5].

En lo referente al área de Ciencia de Datos y en el ámbito del Departamento e Instituto de Informática se lleva adelante el proyecto CICITCA_UNSJ “Evaluación de visualizaciones eficientes en ciencia de datos” donde se procesan, analizan y visualizan diferente tipología de datos entre otras, series temporales que es el formato de representación que se utiliza en la presente propuesta y desde donde se intentará realizar tareas de pronóstico mediante la utilización de algoritmos de predicción aplicados a las diferentes variables medidas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Las tareas llevadas adelante se han materializado en diferentes entornos de software y lenguajes de programación (KNIME Analytics, JS, Python, PHP entre

otros). Se está analizando el algoritmo de regresión lineal, a través de la librería la librería PHP-ML en el lenguaje PHP. Por otra parte, se está investigando la implementación de redes neuronales en Python usando librerías como sklearn y keras.

Inicialmente se han llevado adelante instancias de preprocesamiento sobre los datos relevados por las estaciones meteorológicas durante los períodos comprendidos entre el mes de enero del año 2016 al mes julio del año 2020 para una de ellas; y desde el mes de abril del 2013 al mes de julio del año 2020, para la otra. Lo que involucra, entre otros aspectos, la unificación del formato horario, dado que si bien ambas estaciones registran datos cada diez minutos una lo hace con el formato de 24hs y otra en formato de 12 hs (am y pm) respectivamente. Además, el reporte generado por la estación provee la fecha y hora como datos separados (dos atributos o columnas), el procesamiento de la serie temporal, requiere disponer en un único dato la fecha y hora, esto también implica unificar el formato en cuanto a cantidad de dígitos y orden de los datos (dd-mmm-aaaa hh:mm:ss).

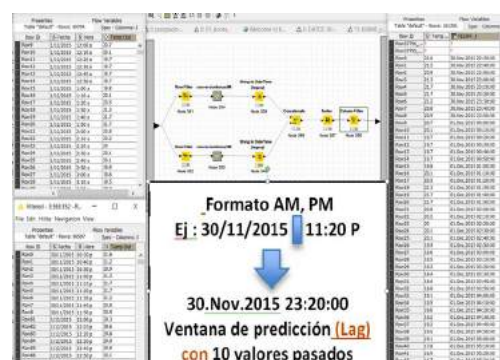


Figura 2: Diagrama en bloques de conversión de horarios AM, PM a formato de 24hs y con la fecha incorporada en el mismo string de salida. Entorno KNIME Analytics 4.3.1

Otra tarea muy importante durante el preprocesamiento es identificar los periodos de tiempo en los cuales la estación no registró

los valores de las variables. Estas situaciones se producen eventualmente ocasionando la interrupción de la serie temporal, identificar estos casos es fundamental para un adecuado procesamiento.

Además, se renombraron variables, se reconocieron zonas de datos de potenciales registros de heladas, conforme reuniones mantenidas con el especialista del INTA y conocedor de los datos generados en las estaciones manteniendo aquellos atributos (variables medidas) estrictamente necesarios para las tareas de predicción en las que se aplica, inicialmente, un algoritmo de regresión lineal.

El objetivo es lograr un pronóstico del fenómeno de ocurrencia de helada a partir de los valores previos de la temperatura medidos por las estaciones. Específicamente, es necesario determinar la cantidad de valores de entrada, es decir, la cantidad de lecturas previas, con la cual el algoritmo genera el mejor resultado.

Se está analizando la calidad de las respuestas obtenidas por el algoritmo de regresión lineal considerando como entrada los valores de la temperatura durante las últimas tres horas. Como salida se obtiene el pronóstico de la ocurrencia o no del fenómeno. Esto opera para el rango horario comprendido entre la 0hs hasta las 8hs, período donde usualmente ocurre la helada.

También se ha llevado a cabo la revisión de la literatura [6][7][8][9][10] respecto del uso de las redes neuronales para el pronóstico de temperatura. Existiendo antecedentes de la utilización distintos tipos de redes, multicapa perceptrón (MLP), convolucionales (CNN) y de memoria a largo-corto plazo (LSTM) con buenos resultados de pronóstico. La principal variable de entrada es la temperatura, pero en algunos casos también se considera la

humedad, las características del viento y la presión atmosférica.

Resultados esperados:

Se espera, realizar el preprocesamiento y análisis de los valores de las variables meteorológicas relevados por las estaciones mencionadas, para un posterior procesamiento a través de distintos algoritmos que permiten el pronóstico, como son la regresión lineal y las redes neuronales. Con el fin de determinar cuál de ellos otorga el mejor resultado respecto al pronóstico de las heladas teniendo en cuenta las necesidades de los productores que es conocer de antemano la hora de ocurrencia de una helada, la duración y su magnitud. Se llevará a cabo un análisis comparativo de los resultados obtenidos por los algoritmos a través de distintas métricas y criterios de información.

Finalmente, se analizará la forma de visualización más adecuada, considerando las características de los valores temporales y las necesidades de información del usuario [11], para facilitarle la interpretación y evaluación de los resultados obtenidos. Esto permitirá que, a través de una herramienta tecnológica se le informe al productor agropecuario, y le ayude en la toma de decisiones en cuanto a la protección de los cultivos.

De este modo, en un trabajo de cooperación interinstitucional se generarán nuevos conocimientos y tecnologías de gran utilidad para el sector agrícola.

4. FORMACIÓN DE RECURSOS HUMANOS

La temática planteada en esta presentación es llevada adelante como trabajos finales de grado en Licenciatura en Ciencias de la Computación y hasta el momento, los alumnos que la elaboran trabajan en dependencias del Instituto de Automática (Unidad de doble dependencia UNSJ-

CONICET) en donde los mismos encuentran un asesoramiento de calidad sobre los aspectos físicos y biológicos de las aplicaciones a realizar. Los aspectos referidos a los conceptos de presentación y desarrollo web como así también lo referido a predicciones de variables en el contexto del Data Science, la realizan con el apoyo de docentes investigadores del Departamento de Informática DI_FCEFN y el Laboratorio de Sistemas Inteligentes para la Búsqueda de Conocimiento en Datos Masivos del Instituto de Informática, conformando con ello un marco de trabajo de mucha sinergia y enriquecimiento mutuo.

Así mismo, en el marco de la Maestría de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNSJ, se desarrolla la tesis “Análisis de fenómenos en estaciones agrometeorológicas mediante Ciencia de Datos” cuya autora encabeza la presente propuesta y como parte de su formación y a través de un curso de posgrado brindado por la UNLP, se accedió a un conocimiento más exhaustivo y profundo sobre Series Temporales, formato de los datos de entrada, utilizados en su tesis.

5. BIBLIOGRAFÍA

- [1] P. Möller-Acuña, R. Ahumada-García, and J. Reyes-Suárez, “Predicción de Episodios de Heladas Basado en Información Agrometeorológica y Técnicas de Aprendizaje Automático,” in *2016 IEEE International Conference on Automatica, ICA-ACCA 2016*, 2016.
- [2] M. Ángel, O. Chong, S. De Gobernación, L. Felipe, and P. Espinosa, “SERIE Fascículos - Heladas,” Mexico, 2014.
- [3] F. Capraro, S. Tosetti, F. Rossomando, V. Mut, and F. V. Serman, “Web-based system for the remote monitoring and management of precision irrigation: A case study in an arid region of Argentina,” *Sensors (Switzerland)*, vol. 18, no. 11, Nov. 2018.
- [4] F. Capraro, S. Tosetti, and V. Mut, “Telemetría Agrícola. Un acercamiento hacia las nuevas tecnologías disponibles en riego de precisión,” *Congr. Argentino Agroinformática, 47º JAIIO.*, 2018.
- [5] M. Masanet, F. Capraro, R. Klenzi, M. Muñoz, and C. Suarez, “Entorno web de visualización de información meteorológica para el uso agrícola y de generación de alertas ante eventos climáticos,” Argentina, 2019.
- [6] R. M. A. Latif, S. B. Brahim, S. Saeed, L. B. Imran, M. Sadiq, and M. Farhan, “Integration of Google Play Content and Frost Prediction Using CNN: Scalable IoT Framework for Big Data,” *IEEE Access*, vol. 8, pp. 6890–6900, 2020.
- [7] S. Lee, Y.-S. Lee, and Y. Son, “Forecasting Daily Temperatures with Different Time Interval Data Using Deep Neural Networks,” *Appl. Sci.*, vol. 10, no. 5, p. 1609, Feb. 2020.
- [8] A. Castañeda-Miranda and V. M. Castaño, “Smart frost control in greenhouses by neural networks models,” *Comput. Electron. Agric.*, vol. 137, pp. 102–114, May 2017.
- [9] M. Fuentes, C. Campos, and S. García-Loyola, “Application of artificial neural networks to frost detection in central chile using the next day minimum air temperature forecast,” *Chil. J. Agric. Res.*, vol. 78, no. 3, pp. 327–338, Sep. 2018.
- [10] E. B. Abrahamsen, O. M. Brastein, and B. Lie, “Machine Learning in Python for Weather Forecast based on Freely Available Weather Data.”
- [11] E. Ben Mohamed, H. Ltifi, and M. Ben Ayed, “Integration of temporal data visualization techniques in a KDD-based DSS Application in the medical field,” 2014.

Prototipo de Sistema para la Gestión de Controles de Tránsito Vehicular

Darío Propato, Marisa Daniela Panizzi

Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales. Universidad de Morón.
Cabildo 134. Morón, Argentina.

dariopropato@gmail.com, marisapanizzi@outlook.com

Resumen

La línea de investigación de este trabajo se enmarca en la aplicación de las Tecnologías de la Comunicación e Información (TICs) para optimizar los controles vehiculares, agilizando la operatoria de los agentes de tránsito y por consiguiente generando una reducción en los tiempos de verificación.

La construcción de rutas, caminos y el tránsito vehicular se incrementa año tras año en Argentina, esto conlleva a un aumento considerable en las infracciones y accidentes de tránsito. Esto genera una complicación y una demora en los controles de verificación vehicular, retrasando tanto los conductores como a las autoridades de tránsito. Este trabajo tiene como objetivo desarrollar un prototipo de sistema para la gestión de tránsito vehicular. Este prototipo presenta como funcionalidades principales, el escaneo y reconocimiento de patentes vehiculares.

Palabras clave: Sistema de Gestión, Control Vehicular, Escaneo de patentes.

Contexto

La línea de investigación que se reporta en este artículo es desarrollada en el marco de una tesis de grado de la carrera Licenciatura en Sistemas de la Universidad de Morón

Este trabajo de desarrollo tiene por objetivo contribuir a la mejora de los servicios públicos que se proveen a los ciudadanos en Argentina mediante la aplicación de las TICs que permitan gestionar de manera eficiente los controles vehiculares.

Este trabajo se propone atender a una problemática que requiere una solución centralizada a nivel nacional.

Introducción

En el período 2007 y 2012 la asociación sin fines de lucro “Luchemos por la Vida”, comprobó una drástica reducción en la cantidad de actas labradas en los controles de tránsito en la Ciudad Autónoma de Buenos Aires. La disminución de los controles se opone a las actuales estrategias de los países modelo para el logro de la seguridad vial y que Naciones Unidas recomienda para reducir sustancialmente los muertos y heridos en el tránsito (Aciti et al., 2012).

Como se menciona en la Estadística de incidentes viales con fallecidos y lesionados en 2017 de la Provincia de Buenos Aires (ver Figura 1), se considera importante disponer de datos confiables sobre los incidentes de tránsito para evaluar el impacto y establecer estrategias que permitan reducirlos (Ministerio de Salud, 2019, Gobierno de la Provincia de Buenos Aires).

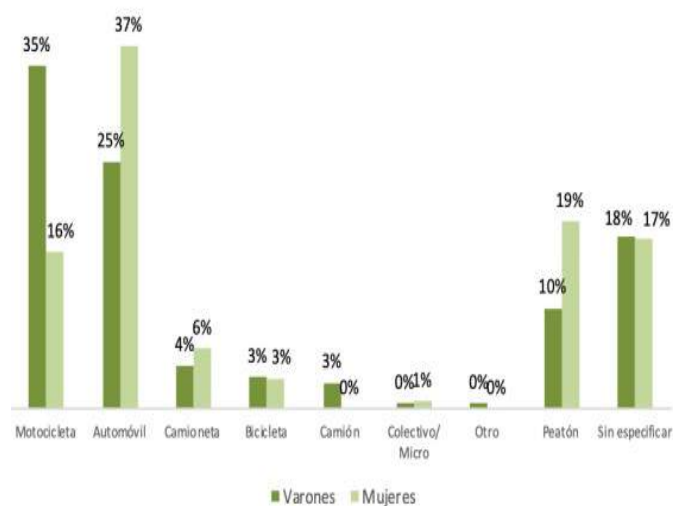


Figura 1. Distribución por vehículo de la víctima y sexo de los fallecidos en incidentes de tránsito. Provincia de Buenos Aires. Año 2017.

De acuerdo al informe de Eficacia de los controles de tránsito del año 2002, las actas labradas representan el 16% del total de infracciones graves que se producen durante el mes analizado. En la Provincia de Buenos Aires se cometen unos 1067 millones de infracciones graves por mes, de las cuales se labran 36893 actas. (Luchemos por la Vida, 2002). Desde octubre de 2017, bajo la modalidad Agentes 2.0, los agentes de tránsito de la Ciudad Autónoma de Buenos Aires (Gobierno de la Ciudad de Buenos Aires, 2008), cuentan con celulares exclusivamente laborales, adaptados especialmente para optimizar sus funciones:

- Digitalizar y mejorar el proceso de infracciones,
- Coordinar cambios en sus recorridos y
- Optimizar el trabajo del agente mediante el sistema de geoposicionamiento que facilita la ubicación de los Agentes en tiempo real.

En la Argentina existe un vehículo por cada 3,1 habitantes. Las estadísticas cuentan solo a los autos y utilitarios. El parque automotor creció un 6,4 con respecto a 2016 y subió un 29,8% con respecto a la primera medición, realizada en 2011 (10,24 millones) (AFAC, 2017)

Esto incrementa considerablemente la cantidad de vehículos a controlar, la congestión vehicular y la consecuente demora en el control y labrado de actas en caso de cometerse infracciones.

De acuerdo al último informe de la Super Intendencia de Seguros de la Nación (SSN) surge que en los doce meses del año 2018 se denunciaron 60832 casos de robo total de vehículos, lo que promedia 167 episodios por día o uno cada 9 minutos (SSN, 2018).

La identificación de los vehículos robados en los controles vehiculares en base a la estadística referenciada requiere de una base de datos que contenga la información actual.

Hoy en día las actas labradas por los agentes de tránsito se labran de forma manual. El agente de tránsito debe consultar al 911 la situación actual de los vehículos que se están

controlando. Las imágenes tomadas de los vehículos en infracción se realizan con un dispositivo móvil que luego son adjuntadas al acta digitalizada. Los informes actuales se envían al Ministerio de Seguridad.

Argentina carece de sistemas de gestión que permitan agilizar los controles de tránsito vehiculares de forma centralizada, rápida y efectiva. Tanto municipios, como provincias, recurren a métodos de control limitados por el conocimiento y el presupuesto que disponen. Se cuenta actualmente con la Licencia Nacional de Conducir, en la cual el conductor registra sus datos y asocia su licencia de conducir física a la aplicación Mi Argentina.

Líneas de Investigación, Desarrollo e Innovación

La línea de investigación aborda un problema que enfrentan las grandes ciudades que incrementan su población día a día, por consiguiente, su tránsito vehicular y los controles a realizar sobre los mismos. Se toma como punto de partida los antecedentes sobre escaneo de matrículas o patentes vehiculares, sus posibles usos y en especial los destinados al control vehicular.

Se pretende lograr una mejora en la realización de los controles vehiculares realizados por las autoridades de tránsito, evitando los controles repetitivos, documentación en papel y el acceso de forma descentralizada a la información.

Por las razones mencionadas, se pretende que las TICs contribuyen a agilizar los controles vehiculares beneficiando tanto a las autoridades de tránsito como a los usuarios conductores.

Para dar respuesta a nuestra pregunta de investigación planteada en la sección Introducción, se proponen los siguientes objetivos:

Objetivo general: obtención de un prototipo de un sistema de gestión de controles de tránsito vehicular.

Objetivos específicos:

1. Elaborar el estado del arte sobre la gestión de controles de tránsito vehicular.

2. Identificar y especificar los requisitos funcionales y no funcionales para el prototipo de gestión del tránsito vehicular.
3. Conceptualizar el sistema de gestión del tránsito vehicular. Mediante la construcción de modelos.
4. Implementar el prototipo.
5. Probar la solución.
6. Evaluar la viabilidad del prototipo en un caso real.

Para el desarrollo del trabajo, se seguirá un enfoque de investigación clásico (Riveros H. et al., 1985) con énfasis en la producción de tecnologías (Sábato J. et al., 1982); identificando los métodos y materiales necesarios.

- Métodos: estudio de mapeo sistemático de la literatura (en inglés, Systematic Literature Mapping o SMS) método DESMET (Kitchenham B. et al., 1996), estudios de casos (Runeson P. et al., 2012), encuestas (Genero Bocco M. et al., 2014), prototipado evolutivo experimental (Basili V., 1993). Para el desarrollo de los formalismos y procesos de Ingeniería de software, se utilizarán los modelos conceptuales propuestos en el Lenguaje de Modelado Unificado (UML) (Booch G. et al., 2006) y los procesos técnicos, de soporte y de gestión definidos en el estándar ISO/IEC/IEEE 12207 (ISO/IEEE 12207, 2017).
- Materiales. Para llevar a cabo la búsqueda de la literatura se emplearán las librerías digitales que se encuentran disponibles en la Biblioteca de la Universidad de Morón: IEEE Xplore, Scopus y repositorios y/o bibliotecas abiertas como SEDICI¹ y Semantic Scholar. Para la documentación de los modelos conceptuales del prototipo se utilizará una herramienta CASE (Software Engineering Asisted by Computer), la cual será evaluada en ese hito del proyecto.

A continuación, se presenta el abordaje metodológico para el desarrollo del trabajo:

- Para el Objetivo Específico 1, se realizará un estudio de mapeo sistemático (SMS).
- Para los Objetivos Específicos 2, 3, 4 y 5, se seguirán los procesos técnicos, de soporte y de gestión propuestos en la ISO/IEC/ IEEE 12207 (IEEE, 2017), se emplearán los formalismos de UML se utilizará una herramienta CASE. Para el Objetivo Específico 6, se utilizarán estudios de casos.

Resultados y Objetivos

Dado que es una línea de investigación que se inicia en la UM, no se presentan **resultados concretos** a describir.

Los **resultados estimados** a lograr son:

1. Académicos: lograr un trabajo de fin de carrera de grado.
2. Producción Científica: presentar los avances de la investigación en eventos científicos de alcance nacional (CACIC 2021²), y en el ámbito internacional, CIACA 2021³, y, etc.
3. Formación en I+D+I: es un desafío para el tesista afrontar la resolución de un problema de la sociedad mediante la aplicación de las TICS. Se adquirirán competencias en el uso de métodos de investigación experimentales como, revisiones sistemáticas, estudios de casos.
4. Transferencia al medio socioeconómico: un prototipo que permita la gestión de los controles de tránsito vehicular.
5. Otro de los ejes a cumplir, consiste en la concientización a la comunidad UM y extramuros sobre la utilización de TICs que permitan involucrar e interrelacionar a las autoridades de tránsito y a los conductores en la colaboración de un problema que afecta a ambas partes.

Formación de Recursos Humanos

¹ SEDICI: Repositorio Digital de la Universidad Nacional de La Plata.

² Congreso Argentino de Ciencias de la Computación.

³ Conferencia de Computación Aplicada.

El grupo se encuentra conformado por un Director, y un tesista de grado.

Se estima la formación de un Licenciado en Sistemas de la UM.

Referencias

Aciti, C., Marone, J. A., Capra, J., & Capra, B. (2012, October). Una aplicación móvil para el reconocimiento automático de matrículas de automóviles argentinos. <http://sedici.unlp.edu.ar/handle/10915/23802>

Patricia. (n.d.). Drástica reducción de los controles de tránsito en la Ciudad de Buenos Aires - Estudios e Investigaciones - Luchemos por la Vida. Obtenido de <http://www.luchemos.org.ar/es/investigaciones/drastica-reduccion-de-los-controles-de-transito-en-la-ciudad-de-buenos-aires>

Ministerio de Salud, M. de S. (2019). Estadísticas de Incidentes Viales con Fallecidos y Lesionados 2017. Gobierno de La Provincia de Buenos Aires. Obtenido de <http://www.gob.gba.gov.ar/UF/Informe.pdf>

Luchemos por la Vida (2002). Eficacia de los controles de Tránsito. Obtenido de <http://www.luchemos.org.ar/revistas/articulos/rev22/pag08.pdf>

Cuerpo de Agentes de Tránsito. (n.d.). Obtenido de <https://www.buenosaires.gob.ar/movilidad/plan-de-seguridad-vial/cuerpo-de-agentes-de-transito>

Asociacion de Fábricas Argentinas de Componentes (05, 2018). Flota Circulante en Argentina 2017. Obtenido de www.afac.org.ar

Nicolas Jasper. (2018). Axeo I, Vehiculos Expuestos a Riesgo. Recuperado de https://www.argentina.gob.ar/sites/default/files/ssn_20172018_vehiculo_expuesto_riesgo_a_nexo.pdf

[Clarín.com. \(n.d.\). Licencia de conducir digital: para qué sirve y por qué no reemplaza al carnet tradicional. Clarín. Recuperado del 20/02/2020, https://www.clarin.com/ciudades/licencia-conducir-digital-sirve-reemplaza-carnet-tradicional_0_KT2gpeaK.html](https://www.clarin.com/clarin-com/n.d./licencia-de-conducir-digital-para-que-sirve-y-por-que-no-reemplaza-al-carnet-tradicional-clarin-recuperado-del-20-02-2020-https-www-clarin-com-ciudades-licencia-conducir-digital-sirve-reemplaza-carnet-tradicional_0_KT2gpeaK.html)

Genero, M., Piattini, M., & Cruz Lemus, J. A. (2014). Métodos de investigación en

Ingeniería del Software. Madrid: Ra-Ma S.A. Editorial y Publicaciones (2014).

IEEE ISO/IEC/IEEE 12207:2017(E). Systems and software engineering — Software life cycle processes (2017).

Kitchenham, B. (1996). "DESMET: A Method for Evaluating Software Engineering Methods and Tools " Department of Computer Science, University of Keele, Technical Report TR96-09, August 1996. ISSN: 1353-7776

Kitchenham, B. y Chartes, S. (2007). Guidelines for performing systematic literature reviews in Software engineering, Keele University, EBSE-2007-01.

Realidad aumentada y realidad virtual aplicadas a proyectos con fines sociales

Claudia Banchoff, Laura Fava, Sofía Martín

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
{cbanchoff, lfava}@info.unlp.edu.ar, smartin@linti.unlp.edu.ar

RESUMEN

Continuando una línea de investigación y desarrollo que se viene trabajando en el LINTI, Laboratorio de Investigación y Nuevas Tecnologías Informáticas de la UNLP (Argentina), en este artículo se presenta la concreción de dos proyectos que utilizan Realidad Aumentada (RA) con fines sociales y el estado de avance de nuevos proyectos que utilizan tanto RA como Realidad Virtual (RV) para potenciar aplicaciones móviles. Dentro de esta línea se abordan aspectos relacionados al desarrollo de aplicaciones educativas interactivas y materiales didácticos. Las dos proyectos que se describirán corresponden a dos aplicaciones, una destinada a la recreación de niños durante tratamientos terapéuticos y/o de rehabilitación prolongados y otra para usar en el Museo de Ciencias naturales de La Plata, orientada a la experimentación y comprensión de los cambios que sufrió la Tierra milenios atrás.

En este artículo, se continúa con las líneas de I+D presentadas en WICC 2020 [1], donde se analizaron y evaluaron entornos de desarrollo para aplicaciones de gamificación usando RA y RV y se presentan los avances en los desarrollos de aplicaciones usando RA y RV.

Palabras clave: juegos serios, proyectos sociales, realidad aumentada, realidad virtual, gamificación.

CONTEXTO

Esta línea de investigación incluye el desarrollo de aplicaciones lúdicas interactivas vinculadas a distintas problemáticas sociales. Se utilizan tecnologías de RA y/o RV como elementos motivadores para complementar las actividades que se trabajan en distintos contextos.

En el LINTI se viene trabajando, desde hace más de 10 años en proyectos relacionados a demandas de la sociedad. En los últimos años, el enfoque estuvo dado por desarrollos pensados para distintos ámbitos educativos, como ser el videojuego RAÍCES [2] y aplicaciones más sencillas desarrolladas en el marco de la asignatura Seminario de Lenguajes, opción Python [3]. Esto sentó las bases para nuevos desarrollos y experiencias que incluyen tecnologías de RA y RV.

La línea de investigación que se presenta en este trabajo viene desarrollándose desde hace dos años y se encuentra enmarcada en el proyecto: "De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región" del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI.

1. INTRODUCCIÓN

La realidad aumentada (RA) y la realidad virtual (RV) se han utilizado en múltiples áreas, aunque en nuestra región, no ha sido sencilla su incorporación. En el ámbito educativo, desde hace unos años, estas tecnologías han sido consideradas como una inclusión innovativa en los procesos de enseñanza-aprendizaje, aunque no han podido ser aplicadas en nuestra región. Sólo algunas experiencias aisladas y proyectos específicos.

En el caso de RV, la adopción y aplicación es más compleja debido al requerimiento adicional de los cascos y otros recursos de hardware.

Si bien existen experiencias con cascos de bajo costo como los Google Cardboard, las mismas suelen estar asociadas a entornos de educación no formal [4]. El ámbito de la escuela sigue siendo un terreno complejo para su adopción. Las experiencias conocidas han sido llevadas adelante por docentes “curiosos” que se animan a probar y experimentar con aplicaciones existentes. Estas aplicaciones, en muchos casos, se encuentran en idiomas extranjeros o con licencias privativas que limitan su uso.

No hay duda, que estas tecnologías pueden resultar altamente atractivas a docentes y estudiantes, pero no se cuenta con estudios y resultados registrados en nuestra región que sirvan de modelo que retroalimentan nuevas experiencias.

Otros dos contextos donde se ha trabajado con RA es en el ámbito de visitas a museos y en tratamientos médicos. Respecto al primer campo, hay numerosas experiencias que demuestran que el uso de RA permite integrar objetos del mundo real con animaciones e información adicional, favoreciendo su difusión, y preservando el patrimonio cultural tangible e intangible [5].

Respecto al uso de RA en tratamientos médicos, se pueden mencionar estudios con pacientes amputados o con pacientes infantiles con quemaduras agudas [6].

Los ámbitos de implementación de aplicaciones en RV se han ido ampliando a lo largo de los años. Se pueden encontrar experiencias de aplicación en medicina, telepresencia, sociedad virtuales, empatía para tomar conciencia sobre situaciones críticas suceden en comunidades lejanas [7][8]. También se puede encontrar su uso en simuladores de vuelo, cuidado de personas con problemas médicos, además de los ámbitos más conocidos como educación y videojuegos. El uso de las características de la RV permite posicionarte en situaciones diferentes pero es importante tener en cuenta varios aspectos que hacen a la experiencia RV. Si bien el hardware y software es importante, no deja de ser menos relevante el estudio de la percepción humana para generar entornos realmente inmersivos[9].

Respecto a las herramientas para el desarrollo, si bien la más utilizada es Unity 3D, se está trabajando con otras tales como Unreal engine y Godot.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes en los que se está investigando están relacionados con el desarrollo de videojuegos y aplicaciones interactivas usando RA en ámbitos educativos de nivel inicial y primario, ámbitos educativos no formales y recreación en general. Asimismo se está trabajando con RV en un proyecto pensado para crear entornos que faciliten el uso de modelos arquitectónicos para su visualización y manipulación antes de la construcción. Esta aplicación se trabaja con el aporte de

un profesional del área de Arquitectura y Urbanismo y, en un principio, está pensada para ser utilizada en la etapa de formación de los futuros arquitectos.

A continuación se detallan los ejes de investigación:

- Estudio y análisis de nuevas tecnologías para RA y RV, tanto relacionadas a entornos para el desarrollo de software como las vinculadas con hardware específico.
- Análisis de nuevos contextos donde potencialmente se podría aplicar RA y RV.
- Realización de pruebas y testeos de las aplicaciones finalizadas.

3. RESULTADOS Y OBJETIVOS

El objetivo general de esa línea de trabajo es aplicar técnicas de RA y RV en proyectos sociales, en particular orientadas educación y videojuegos serios.

Para cumplir con el objetivo general, se proponen los siguientes objetivos específicos:

- Analizar los distintos contextos de uso de las aplicaciones, los cuales permiten direccionar las tecnologías a utilizar. En este sentido, se están explorando distintas herramientas alternativas a las actualmente en uso, como Godot, entre otras.
- Desarrollar aplicaciones interactivas que puedan utilizarse en el ámbito de la escuela y entornos de educación no formales.
- Construir casos de prueba de las aplicaciones desarrolladas con el fin de aplicarlos en los entornos para los cuales fueron desarrolladas y comprobar su adecuación y usabilidad.

- Continuar promoviendo esta temática en el marco del desarrollo de las tesis de grado.

En este apartado se describen las dos aplicaciones móviles con RA finalizadas en los últimos meses y puestas a disposición en la plataforma de distribución digital de aplicaciones móviles para los dispositivos con sistema operativo Android operada por Google.

Una de las aplicaciones móviles finalizada es “**Sinfonía Terrestre**”, cuyo objetivo es acompañar a los visitantes del Museo de Ciencias naturales de La Plata en su recorrido de las salas, brindando información suplementaria acerca de las exhibiciones de manera atractiva. Para la primera sala, “la Tierra, una historia de cambios”, una de las actividades desarrolladas consistió en modelar las placas tectónicas y animarlas para simular los choques entre ellas. La Fig. 1 muestra el uso de la app apuntando sobre los gráficos de las placas para lograr la animación de las mismas por texto, sonido y vibraciones de los dispositivos móviles.



Fig. 1. Actividad en Sala 1 del Museo.

Para la sala 2 creó un personaje diplodocus con animaciones para ejemplificar sus movimientos y un ambiente con el cual interactuar. La mayor dificultad fue respetar las exactitudes científicas solicitadas por el museo al mismo tiempo que mantener la

interactividad con el personaje. La imagen de la Fig. 2, muestra niños interactuando con el personaje.



Fig. 2. Actividad en Sala 2 del Museo.

Finalmente, en la sala 3 se trabajó sobre una simulación del experimento de Miller. Éste es un experimento complejo del que solo existe una imagen en el museo por lo que animarlo y aumentarlo favoreció a su comprensión. La mayor dificultad fue lograr transmitir las instrucciones para lograr la animación, ya que no solo implica una RA sino que los componentes pueden manipularse. La Fig. 3 muestra una imagen de la animación del experimento.



Fig. 3. Actividad en Sala 3 del Museo.

La segunda experiencia desarrollada en el marco de esta línea de investigación se enmarca en la tesina de grado denominada “**ERA: entretenidos con realidad aumentada**”. Este trabajo enmarcado presentado en diciembre de 2020, está destinado a niños y niñas que deben permanecer en tratamientos

prolongados. En el marco de esta tesina, se desarrolló la aplicación móvil “Entretenidos con Kota” que acompaña al libro “Kota Corta” de Katia Maria Soto Kiewit, con arte de Tatiana Zanelli, como un paratexto del mismo¹. Incluye no sólo la lectura aumentada del libro sino también un desafío para acceder a minijuegos relacionados al mismo. La Fig. 4 muestra una captura de algunas de las escenas aumentadas del libro.

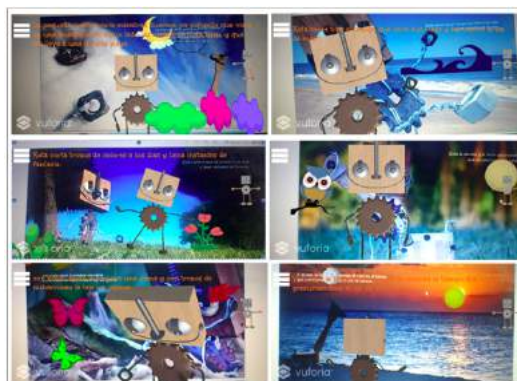


Fig 4. Escenas de aumentaciones del libro.

“Entretenidos con Kota” fue evaluada con distintos grupos de usuarios en distintos contextos. La Fig. 5 muestra alguna de estas experiencias.



Fig. 5. Experiencias de Entretenidos con Kota.

Se entregaron varias copias impresas del libro en el Hospital Garrahan y en el Hospital de Niños de la ciudad de La Plata. Si bien la situación planteada por la pandemia de COVID no permitió supervisar las pruebas, se recibieron muy buenos comentarios por parte de los médicos que la analizaron y utilizaron.

¹ En este desarrollo se trabajó en conjunto con la cátedra Lenguaje Visual 3 de la Facultad de Artes de la UNLP, donde se elaboró el libro Kota Corta.

Los trabajos presentados en este artículo se llevan a cabo de manera interdisciplinaria entre docentes y estudiantes de la Facultad de Informática y Artes de la UNLP y, en el caso de Sinfonía terrestre, la colaboración de los directivos del Museo de Ciencias Naturales.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I+D+i presentada en este artículo está formado por docentes investigadores categorizados del LINTI y estudiantes de la Facultad de Informática. Se destaca la formación en equipos interdisciplinarios en los cuales participan docentes y profesionales de la Facultad de Artes, de la Facultad de Arquitectura y Urbanismo y del Museo de Ciencias Naturales de la UNLP.

A través de la generación permanente de conocimiento por medio de líneas de investigación y desarrollo de aplicaciones vinculadas al sector productivo y su aplicación en el ámbito social, el LINTI promueve el uso innovador de las tecnologías informáticas en la región.

5. BIBLIOGRAFÍA

- [1] Banchoff C, Fava, L., Schiavoni, A., Martin S. (2020). *Realidad Aumentada y Realidad Virtual: experiencias en diferentes ámbitos de aplicación*. XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020). Abril, 2020, Calafate, Argentina.
- [2] Nomdedeu, L. RAÍCES: un juego serio social para revalorizar las culturas originarias. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/48087>. Último acceso: marzo de 2021.
- [3] Banchoff C., Vazquez M. R., Martin S. y Harari V. (2016). *Va de vuelta: un dispositivo lúdico-educativo para abordar la alfabetización digital de niños y niñas*. IV Congreso de Videojuegos y Educación (CIVE 2016). Universidad de Vigo.
- [4] Chirinos Delfino, Y. (2020). La realidad virtual como mediadora de aprendizajes: desarrollo de una aplicación móvil de realidad virtual orientada a niños. Revista TE&ET no. 27. ISSN: 1850-9959. Páginas: 98-99.
- [5] Shengli Xu (2018). *Intangible Cultural Heritage Development Based on Augmented Reality Technology*. 2018 International Conference on Robots & Intelligent System (ICRIS). DOI: 10.1109/ICRIS.2018.00094.
- [6] Jara J. (2020) Tesina de Grado, *ERA: Entretenidos con Realidad Aumentada*.
- [7] Craig, A, Sherman, B., William R., Jeffrey D. (2009). *Introduction to Virtual Reality, Developing Virtual Reality Applications, Morgan Kaufmann, Boston, Pages 1-32, ISBN 978-0-12-374943-7, DOI: 0.1016/B978-0-12-374943-7.00001-X*.
- [8] Das, P.; Zhu, M.; McLaughlin, L.; Bilgrami, Z.; Milanaik, R.L. (2017). *Augmented reality video games: new possibilities and implications for children and adolescents*. Multimodal Technologies and Interaction, 2017; 1:8.
- [9] M. Lavalley, M., *Virtual Reality, 1st ed. Cambridge: Cambridge University Press, 2019. Disponible en: <http://vr.cs.uiuc.edu/vrbook.pdf>. Último acceso marzo de 2021.*

Reconocimiento de emociones en la voz empleando redes neuronales y su integración en frameworks multimodales de educación emocional

Jorge Ierache , Fernando Elkfury 

Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica (ISIER)

Escuela superior de Ingeniería, Informática y Ciencias Agroalimentaria

Secretaría de Ciencia y Tecnología Universidad de Morón Cabildo 134, Buenos Aires, Argentina

jierache@unimoron.edu.ar

Resumen

El habla es una de las formas naturales para que los humanos expresen sus emociones. Es fácil de obtener y procesar en escenarios en tiempo real, pero, sin embargo, el reconocimiento automático del habla emocional implica muchos problemas que necesitan ser estudiados cuidadosamente, tales como: qué emociones podemos identificar realmente, definir concretamente qué se entiende por cada emoción descripta, cuáles son las mejores características para la identificación y qué clasificadores dan el mejor rendimiento. En este trabajo se describe el diseño y desarrollo de redes neuronales para la clasificación de emociones en el discurso hablado (voz), se proponen diferentes métodos para convertir un enfoque categórico de clasificación de emociones a uno dimensional y la integración del clasificador con frameworks multimodales de captura de emociones.

Palabras clave: Aprendizaje automático, redes neuronales, reconocimiento de emociones en la voz

Contexto

Esta investigación aplicada se desarrolla en el contexto del Proyecto de Investigación Científica Tecnológica Orientado (PICTO) aprobado por la Agencia Nacional de promoción de la investigación, el desarrollo tecnológico y la innovación (ANPCyT), denominado

“Influencias del estado biométrico emocional de personas interactuando en contextos de entornos simulados, reales e interactivos con robots”. El mismo se desarrolla dentro del Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica de la Universidad de Morón (ISIER-UM) y está auspiciado por la Secretaría de Ciencia y Tecnología con una duración de tres años. Este trabajo se sustenta sobre las bases iniciales de las investigaciones realizadas en el marco de los proyectos: “Influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales” Ping/17-03-JI-004,(2017-2019), el proyecto denominado “Explotación de datos EEG y parámetros fisiológicos de usuarios interactuando en contextos virtuales” (DC diálogo con las ciencias 2018-2020) - UM-2019 código 80020190100007 UM y el proyecto presentado “Valoración Emocional Multimodal aplicada en contextos gastronómicos” convocatoria PIO 2019 UM.

Introducción

El análisis de las emociones en la voz humana es una tarea poco trivial, incluso para el propio ser humano. Si bien el habla es la forma tradicional de comunicación, no es una característica sensible a los cambios emocionales y por lo tanto la educación emocional a partir de la misma, cuando no se posee contexto semántico ni de otra clase, resulta parcial. Según Albert Mehrabian, el tono de la voz expresa solo un 38% de las emociones que pueden transmitir las personas

en un momento dado [1].

Deep Learning ha sido considerado como un campo de investigación emergente en el aprendizaje automático y ha ganado más atención en los últimos años. Las técnicas de aprendizaje profundo para los sistemas de reconocimiento de emociones tienen varias ventajas sobre los métodos tradicionales, dada su capacidad para detectar la estructura compleja y sus características asociadas, sin la necesidad de extracción y ajuste manual de estas. Lo cual es un aspecto clave en el desarrollo, dado que la precisión de los clasificadores suele estar ligada a la selección de las características de la voz que se usaran para el entrenamiento.

En este trabajo se optó por el uso de espectrogramas, en los cuales las frecuencias fueron convertidas a escala de Mel [2] [Figura 1], y se evalúa el desempeño de redes neuronales convolucionales (CNN) [3] [4] y redes neuronales recurrentes (RNN) [5] para la construcción de un clasificador de emociones.

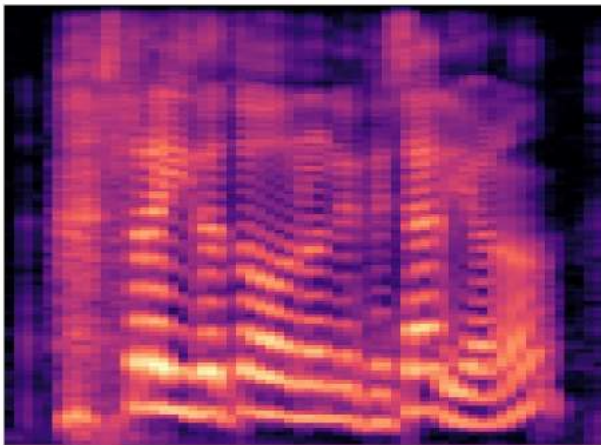


Figura 1. Ejemplo de espectrograma con las frecuencias en escala de Mel.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se desarrolla en base al aporte potencial que puede tener la educación emocional a partir del discurso hablado (voz) en un framework multimodal de análisis emocional, como el visto en [6] [7] [8] [9] [10] [11] [12] que está asociado al contexto

de esta investigación.

La comunicación entre las personas y las máquinas o sistemas es cada vez más frecuente por los avances tecnológicos, sin embargo, los desarrollos son en su mayoría carentes de la componente afectiva. Por tanto, uno de los objetivos recientes de la comunicación persona-máquina es la mejora de la experiencia de usuario, intentando conseguir que esta comunicación sea lo más parecida a la interacción entre personas. [13].

Dicho esto, se plantean los siguientes problemas para los cuales se propuso y desarrollaron soluciones. En primer lugar, la falta de una arquitectura de reconocimiento de emociones en el discurso que se adapte a la variante de español latino rioplatense. Seguido de la ausencia de un API moderna de uso libre que pueda ser utilizada para la investigación y desarrollo de software. Y, por último, la falta de un método de conversión de enfoques que nos permita trabajar la voz en un framework multimodal de análisis emocional.

Objetivos y Resultados Obtenidos

Con intención de dar respuesta a los problemas planteados se propuso:

1. Desarrollar un clasificador de emociones que pueda reconocer emociones en el discurso hablado en español rioplatense.
2. Establecer una comparativa de performance entre clasificadores basados en RRN y CNN.
3. Integrar el clasificador con un framework multimodal de captura de emociones.
 - a. Desarrollar una API que facilite la explotación del modelo diseñado.
 - b. Proponer e implementar un método para pasar del enfoque categórico a un enfoque dimensional.

4. Sentar las bases para la mejora continua a partir de la adquisición de nuevos datos para el entrenamiento del modelo elegido.

Los modelos de redes neuronales construidos y entrenados con muestras de audios provenientes de dos sets de datos, INTERSIP (ELRA) [14] y el EMOFILM [15], obtuvieron resultados a la altura del estado del arte [16] [17] [18], en relación al valor de precisión en el set de pruebas.

A continuación, se presenta una tabla con los valores de precisión obtenidos para un modelo puramente convolucional y otro que combina capas convolucionales con capas LSTM [5]:

Modelo	Prec. En el set de pruebas
1 CNN	92.53
2 CNN+LSTM	82.62

Tabla 1. Precisión de ambos modelos de clasificadores construidos

Uno de los desafíos de integración con frameworks multimodales es pasar de un enfoque categórico a uno dimensional que nos permita contrastar equitativamente los datos de todos los métodos de educación emocional.

En este aspecto se trabajó con 2 enfoques. El primero se basa en obtener valores de valencia a partir de la resta de la probabilidad obtenida para la etiqueta “alegría” y la emoción negativa más probable, como lo plantean Leanne Loijens et al. [19] Para los valores de excitación se plantea utilizar la diferencia entre los valores medios de dB de las subsecuentes muestras tomadas durante una sesión de evaluación de un sujeto y los valores promedio de dB de la muestra particular evaluada

Otra forma presentada, llamada vectorial, es usar el circunflejo de Russel [20] para tomar una coordenada de origen asociada a la emoción más probable según el clasificador y, a partir de ahí, desviarse en dirección y magnitud proporcional (a la probabilidad

correspondiente) hacia los puntos asociados a las 2 emociones siguientes más probables predichas.

Para esto se optó por elegir una expansión del modelo de Russel presentada por Klaus Sherer [21], de forma de tener las coordenadas necesarias para asociar las 8 emociones con las cuales trabaja el clasificador en base al modelo categórico de Ekman.

Solo se extrajeron las emociones de las cual carece el circunflejo de Russel por lo que el modelo utilizado para trabajar es el que se muestra en la figura 2 debajo.



Figura 2 Circunflejo basado en publicaciones Russel y Sherer

Dado que este trabajo se desarrolla bajo las líneas de investigación del ISIER UM [22], para las pruebas se implementó el API del clasificador con el framework visto en [6] [7] [8] [9] [10] [11] [12]. Haciendo uso de la última versión en desarrollo del framework, se hicieron pruebas con una metodología muy similar a la planteada en los artículos mencionados. Se pretende capturar la voz del sujeto de prueba y extraer su estado emocional a partir de la misma mientras se lo estimula con imágenes obtenidas del IAPS [23].

De esta forma se pueden obtener datos sincronizados de excitación valencia desde diferentes sensores para su contrastación y consecuente validación.

Para evaluar la capacidad de generalización del clasificador incorporado se recolectaron 22 audios de novelas argentinas y se las clasificó con un grupo de 8 personas por medio de encuestas SAM [24]. Se obtuvo un 72% de precisión, que, si bien es alentador, está sujeto a la necesidad de una evaluación más extensiva dado el volumen reducido de muestras disponibles para las pruebas.

Este proyecto sienta las bases para la continuación de la línea de investigación que vincula la educación de emociones en el discurso hablado (voz) con un framework multimodal de análisis emocional. Se abre el camino para la mejora continua de clasificadores basados en redes neuronales a partir de la recolección de muestra de audio en castellano rioplatense, como también para validar los avances hasta ahora conseguidos.

Formación de Recursos Humanos

El grupo de investigación se compone de un investigador formado y tres investigadores en formación. En el marco de la investigación se finalizó una tesis de grado en ingeniería en informática y se encuentra en etapa inicial el proyecto de una tesis de doctorado.

Referencias

- [1] A. Mehrabian, «Communication Without Words,» de *communication theory*, Routledge, 2017, p. 193–200.
- [2] J. Volkman, S. S. Stevens y E. B. Newman, «A Scale for the Measurement of the Psychological Magnitude Pitch,» *The Journal of the Acoustical Society of America*, vol. 8, p. 208–208, 1 1937.
- [3] K. Fukushima, «Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,» *Biological Cybernetics*, vol. 36, p. 193–202, 4 1980.
- [4] I. Shafkat, «Intuitively Understanding Convolutions for Deep Learning,» Junio 2018. [En línea]. Available: <https://towardsdatascience.com/intuitively-understanding-convolutions-for-deep-learning-1f6f42faee1>. [Último acceso: 10 2020].
- [5] A. G. Josh Patterson, «Deep Learning,» [En línea]. [Último acceso: 10 2020].
- [6] J. Ierache, G. Ponce, R. Nicolosi, C. Cervino y E. Eszter, «Registro emocional de las personas interactuando en contextos de entornos virtuales,» *CACIC 2018*, La Plata, Plata, pp 877-886, ISBN: 978-950-658-472-6
- [7] J. Ierache, G. Ponce, R. Nicolosi, I. Sattolo y G. Chapperon, «Valoración del grado de atención en contextos áulicos con el empleo de interface cerebro-computadora en el marco de la computación afectiva,» *CACIC 2019*, Rio Cuarto, pp: 417-426, ISBN:978-987-688-377-1
- [8] J. Ierache, I. Sattolo, G. Chapperon, R. Ierache, F. Nervo, F. Elkfury, G. Ponce y R. Nicolosi, «Computación afectiva aplicada a la valoración emocional en contextos gastronómicos,» *WICC 2020*, El Calafate, pp: 664-668. ISBN: 978-987-3714-82-5
- [9] J. Ierache, F. Nervo, I. Sattolo y G. Chapperon, «Propuesta de un Modelo Multimodal de valoración emocional en el marco de la computación afectiva aplicado en ambientes gastronómicos,» *CACIC 2020*, La Matanza. En Prensa.
- [10] C. Barrionuevo, J. Ierache y I. Sattolo, «Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística,» *CACIC 2020*, La

- Matanza. En prensa.
- [11] J. Ierache, I. Sattolo, G. Chapperon, R. Ierache, F. Elkfury, C. Barrionuevo y F. Nervo, «Captura multimodal de estados emocionales aplicado a contextos de computación Afectiva,» *WICC 2021*, Chilecito, 2021. Comunicado 24/2/2021.
- [12] J. Ierache, I. Sattolo y G. Chapperon, «Framework multimodal emocional en el contexto de ambientes dinámicos». DOI. 10.17013/risti 40. 45-59 ISSN: 1646-9895.
- [13] S. P. Garcia, *reconocimiento afectivo automático mediante el análisis de parámetros acústicos y lingüísticos del habla espontánea*.
- [14] ELRA, «Emotional speech synthesis database,» [En línea]. Available: <http://catalog.elra.info/en-us/repository/browse/ELRA-S0329/>.
- [15] E. Parada-Cabaleiro, G. Costantini, A. Batliner, A. Baird y B. Schuller , «EmoFilm - A multilingual emotional speech corpus,» [En línea]. Available: <https://zenodo.org/record/1326428#.XoyMIlgzbc>.
- [16] L. Kerkeni, Y. Serrestou, M. Mbarki, K. Raoof, M. A. Mahjoub y C. Cleder, «Automatic Speech Emotion Recognition Using Machine Learning,» 2018.
- [17] Mustaqeem y S. Kwon, «A CNN-Assisted Enhanced Audio Signal Processing for Speech Emotion Recognition,» *Sensors*, vol. 20, p. 183, 12 2019.
- [18] M. Sánchez-Gutiérrez, E. M. Albornoz, F. Martínez, L. Rufiner y J. Goddard, «Deep Learning for Emotional Speech,» 2014.
- [19] L. Loijens y O. Krips, «FaceReader Methodology Note,» [En línea]. Available: <https://www.noldus.com/facereader/resources>. [Último acceso: 10 noviembre 2020].
- [20] J. A. Russell, «A circumplex model of affect.,» *Journal of Personality and Social Psychology*, vol. 39, p. 1161–1178, 1980.
- [21] K. R. Scherer, «What are emotions? And how can they be measured?,» *Social Science Information*, vol. 44, p. 695–729, 12 2005.
- [22] «Instituto De Sistemas Inteligentes y Enseñanza De La Robótica (ISIER),» [En línea]. Available: <http://isierum.c1.biz/>.
- [23] P. J. Lang, M. M. Bradley y B. N. Cuthbert, *International Affective Picture System*, American Psychological Association (APA), 2005.
- [24] P. J. Lang, «The Cognitive Psychophysiology of Emotion,» de *Anxiety and the Anxiety Disorders*, Routledge, 2019, p. 131–170.

REDES DE CONOCIMIENTO PARA LA AYUDA DE CERCANÍA - INSTRUMENTO PARA LA REACTIVACIÓN MICROECONÓMICA

Gustavo Tripodi¹ <g.d.tripodi@gmail.com>; Jorge Sagula² <jorgesagula@gmail.com>;

Walter Pannesi³, <wpanessi@unlu.edu.ar>; Josefina Tripodi⁴ <tripodijosefina@gmail.com>.

¹Instituto 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).

^{2,3}Departamento de Ciencias Básicas (DCB), División Computación, Universidad Nacional de Luján (UNLu)

⁴Facultad de Ciencias Económicas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA)

RESUMEN

Nos hemos enfocado en trabajar en un Proyecto que mitiga las consecuencias devastadoras que la Pandemia ha infligido en la economía de familias, emprendedores y microempresas. Para su consecución conformamos un equipo interdisciplinario con la participación de diferentes profesiones, especialidades y unidades académicas. El objetivo es la construcción de una Red de Ayuda de Cercanía para conectar beneficiarios con benefactores y emprendedores y microempresas con clientes. La Metodología propuesta contiene los conocimientos producidos en la secuencia de Investigaciones que hemos realizado en el últimos años. Además en este Proyecto abordaremos la modelización y el desarrollo de una Plataforma para la implementación dinámica y ad-hoc de Procesos. En ellos incluiremos la configuración de objetos específicos para cada solución con sus estados, sus cambios, su transformación y su evolución en nuevos Objetos. Para lograr el objetivo Tecnológico se sincronizarán tres Sistemas: La Red de Conocimiento, la búsqueda de Productos y Servicios de Cercanía más Transferencias y Pagos.

Para dar un respaldo consensuado y sólido trabajamos en el Plan Estratégico del Proyecto que nos orientó en los Objetivos, Misión y Visión que nos regirá en la Investigación y Desarrollo.

Palabras clave: *Redes de Conocimiento, Ayuda de Cercanía, Gestión del Conocimiento, Dinámica de Procesos.*

CONTEXTO

En la “infraestructura global de la información” existen instrumentos para la Gestión del Conocimiento sustentados en las TICs. En este contexto el Análisis de Redes de Conocimiento ha dejado de ser una metáfora sugerente para constituirse en un enfoque analítico y un paradigma, con sus principios teóricos, métodos de software para análisis de Redes de Conocimiento y líneas de Investigación propias.

Una de las líneas de investigación de las Redes de Conocimiento es examinar cómo los Usuarios de una Organización se relacionan y cómo las Organizaciones interactúan unas con otras, caracterizando las múltiples conexiones formales e informales que vinculan a los Usuarios, así como sus asociaciones y conexiones.

El fundamento de esta propuesta de solución se sustenta en **Redes de Conocimiento para la Gestión**, pues brindan la arquitectura fundamental y apropiada para abordar soluciones complejas, como lo son las evoluciones socioculturales; los cambios operativos, tácticos y estratégicos; y las transformaciones digitales, como aspectos más salientes, pero basadas en sistemas de planificación que permiten atender situaciones extraordinarias y catástrofes, como la situación crítica actual donde somos más que observadores pues estamos inmersos.

1. INTRODUCCIÓN

Con espíritu de Responsabilidad Social Universitaria y considerando las necesidades de los grupos de riesgo y microempresas, es nuestro Objetivo poner a disposición nuestro compromiso social como facilitadores tecnológicos. La percepción que la sociedad necesita I+D+i para obtener soluciones en esta nueva realidad nos estimula a transferir

Conocimiento, Tecnología y Formación de Talentos. Este enfoque nos permitirá elaborar un Modelo para prestar servicios a los más necesitados y visualizar las microempresas mediante la colaboración ciudadana, empresas, organizaciones intermedias y autoridades. Tomando en consideración las consecuencias de la pandemia que nos aqueja y mostrando que un fuerte pilar de una ciudadanía comprometida reside en **hacer**.

Para conseguir el Objetivo modelaremos una solución general de cercanía, poniendo a disposición instrumentos tecnológicos para acompañar tanto a individuos y familias con Necesidades Básicas Primarias Insatisfechas como así también facilitar la visualización y el vínculo con Comercios y microempresas a aquellas personas con Necesidades de Productos o Servicios Específicos.

La propuesta incorpora soluciones evolutivas, escalables, a través de: **Ayuda de Cercanía, Coordinación de la Ayuda y Control y Auditoría** en el seguimiento de las actividades de la Red mediante **Informes e Indicadores Automatizados**, de los distintos niveles de actores en la red. Las etapas de escalabilidad corresponden a los diferentes estadios de la Red, a fin de brindar acceso a cada usuario a datos e información, necesarios para cumplir efectivamente con las funciones posibles. El Sistema nos ubica en las puertas de “un círculo virtuoso” integrando el Proceso de Mejora Continua, las Buenas y Mejores Prácticas y la Gestión de Conocimiento encontrando un ambiente de convergencia natural.

El epicentro del Proyecto reside en **traccionar la demanda a partir de herramientas tecnológicas**, en pos de impactar positivamente en la productividad, y orientándose a generar fuentes de trabajo. El Proyecto busca “generar un ambiente de reactivación” mediante redes de desarrollo a nivel local, con alto nivel de interconexión.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

PROYECTO: Gestión Informática del Conocimiento como soporte para la toma de decisiones Organizacionales

Línea 1: Analítica de Datos para la gestión del conocimiento orientado a la toma de decisiones en espacios de Educación Superior.

Línea 2: Redes de Conocimiento para la Ayuda de Cercanía a Ciudadanos, Comercios y Microempresas - Instrumentos para la Reactivación Microeconómica

Línea 3: Capital Social y Redes Sociales Solidarias: Análisis y Gestión del Conocimiento mediante TICs.

La Metodología transita por caminos que parten desde diferentes unidades académicas y confluyen en un núcleo consolidado desde las diferentes perspectivas. Esto es posible por la experiencia, interdisciplina y transdisciplina del Equipo.

El Paradigma en el que se sustenta la Metodología son las Redes de Conocimiento. Estas se apoyan sobre la Infraestructura global de la Información, cuyas herramientas informáticas ponen de manifiesto el potencial y la efectividad de las Redes de Conocimiento (‘software social’), operando en tres ámbitos, “las 3Cs”, de forma relacionada:

- Comunicación (nos ayudan a poner en común conocimientos).
- Comunidad (nos ayudan a formar e integrar comunidades).
- Cooperación (nos ayudan a hacer cosas juntos).

Estos conceptos pueden ser extendidos a las relaciones entre Organizaciones para la consecución de Proyectos en forma conjunta, donde la interdisciplinariedad es moneda corriente. Para generar estos escenarios es necesario utilizar herramientas que faciliten la Gestión del conocimiento, como los tópicos que se presentan a continuación.

Planeamiento Estratégico

Para sustentar el Proyecto en sus procesos de Innovación, Transformación Digital y Soluciones Integradoras nos apoyamos en un Planeamiento Estratégico Sistémico e Interdisciplinar. Recorreremos desde el Diseño, hasta el Mapa Estratégico para establecer las Hipótesis de Trabajo, los Talentos y Planificación. El Plan nos mostrará el camino a seguir a través de la Misión y nos

enfocará según la Visión consensuada y compartida.

Ingeniería del Conocimiento

Nos valdremos de la ingeniería del conocimiento que parte de la Teoría actor-red, ya que pone de manifiesto redes y genera otras nuevas, y además hace uso de la Teoría de la Traducción-Traslación, que mediante el acercamiento y puesta en relación de actores, se produce un modelado para llevar los enunciados a nuevos estadios de evolución.

El desarrollo de estos modelos nos permitirá plasmar mediante iconos y gráficos representaciones formales de la Red que se quiere Implementar. Se trata de transformar una realidad dada y los Modelos Naturales que de ella existan, en distintos pero conexos y relacionados Modelos Formales Evolutivos.

TICs

Las Tecnologías de la Información y las Comunicaciones no constituyen un fin en sí mismo, son un medio, sin duda un medio particular que afecta nuestra manera de pensar y constituye uno de los caminos para mejorar la calidad de la Investigación. Dentro de este contexto, nos corresponde como docentes investigadores la generación de las ideas, el diseño de experiencias, la aplicación y la reflexión evaluativa que aporte conocimiento para el mejoramiento de la acción. En este sentido, el trabajo que proponemos se basa en el modelado de una solución integrada e integradora para la implementación de Redes de Gestión complejas.

En este punto cabe destacar que la solución tiene la posibilidad de interactuar con diferentes sistemas e incorporar desarrolladores para brindar soluciones específicas y apropiadas en cada contexto. Un Proyecto de esta envergadura permite sentar bases adecuadas en escenarios de borde, con el objeto de minimizar diferencias, enfocándonos en el tratamiento metodológico y sistémico en aras de conseguir y luego, coordinar trabajos interdisciplinarios en equipo.

Además, facilita la creación de formularios operativos personalizables, y la recolección de datos distribuidos a través de ellos, con el

fin de analizar los resultados y obtener la información deseada.

Buscamos modelar soluciones que integren sistemas, asegurando la fiabilidad en la información, reduciendo tareas duplicadas y optimizando el tiempo. Es decir, brindar la centralización de datos, sin necesidad de distribuir información y procesos en múltiples sistemas imposibles de gestionar. Autonomía, agilidad, usabilidad, generación de informes, alertas y avisos, todo en un mismo entorno facilitando la premisa del acceso a la información a un click de distancia.

Tecnologías de Gestión

Las Redes de Conocimiento para la Gestión propician el Trabajo en Equipo, dando el marco adecuado tanto para la Gestión de Operaciones, de Proyectos distribuidos, de Equipos de Trabajo, de Áreas dentro de una Organización, de una Organización y Proyectos Interorganizacionales. Porque comunicarse es clave, destacamos la comunicación que permite generarse interna y externamente a través de la Solución, estableciendo canales de información corporativa, formales e informales, integrados y al alcance de todos los involucrados. La Solución posibilita la definición de Ámbitos de Trabajo para que los Usuarios cuenten con los recursos necesarios y suficientes, de esta forma podrán agregar valor a los procesos, obtener la información requerida y brindar la información adecuada a quien corresponda. Así publicar y enviar avisos, noticias, vídeos y cualquier otro contenido de interés para la organización es realmente sencillo, ágil y fomenta la participación e involucramiento. En este punto es importante destacar que la Solución posee una fuerte configuración de ámbitos de acceso y modificación a la información del tipo pública, privada y restringida.

Gestión del Conocimiento

El análisis para la especificación de indicadores y agentes inteligentes (softbots) permitirán ayudar a los decisores a enfocar el Proyecto rumbo a la Optimización en un Ambiente de Reactivación no sólo de recursos sino de interacción entre los distintos actores.

Como el epicentro del Proyecto consiste en “traccionar la demanda a partir de herramientas tecnológicas” en pos de impactar positivamente en la productividad, se hace imprescindible la Gestión basada en Indicadores para generar Conocimiento.

De la misma forma se necesitará monitorear la generación de fuentes de trabajo, en una interacción compleja de redes de desarrollo a nivel local, con alto nivel de interconexión zonal y regional que se constituirá en una muestra elocuente de la Escalabilidad del Proyecto.

3. RESULTADOS OBTENIDOS/ESPERADOS

Brindar modelos para consolidar la gestión y la coordinación de protocolos tanto con el estado como con organizaciones intermedias para proveer un control de satisfacción y funcionamiento. Además, en el umbral de un Modelo consolidado, se propugna trabajar específicamente en sentar las bases estructurales y analíticas para brindar los Indicadores del Proceso de Mejora Continua, en términos de optimización y evolución del Proyecto Global.

En la etapa de Gestión del Conocimiento nos enfocaremos en continuar y agudizar la transferencia de conocimiento a los destinatarios finales, beneficiarios e interesados a partir de casos de uso, capacitaciones, cursos, charlas y workshops. Estas intervenciones estarán a cargo de los especialistas que trabajaron en la consecución del Proyecto. Como complemento a esta actividad el material necesario y producido quedará almacenado en la misma Red de Conocimiento y accesible a quien corresponda. La Transformación Digital basada en el paradigma de Redes de Conocimiento para la Gestión nos permite transferir conocimiento en diferentes formas: dentro de un ámbito específico al embeber el Sistema con las mejores prácticas. Luego, diferentes ámbitos podrán comunicarse dentro de la plataforma para cooperar en una transferencia cercana.

Formulación del Plan Estratégico

- Enfoque: Entrevista orientada a las Redes de Conocimiento; Análisis Interno y Externo;

Misión y Visión; One Page; Brochure; Diagnóstico.

- Estrategias: Construcción; Categorización; Ponderación; Planificación.
- Dirección Estratégica: Proyectos Estructuradores; Líneas Estratégicas; Mapa Estratégico. Bloques Estratégicos.

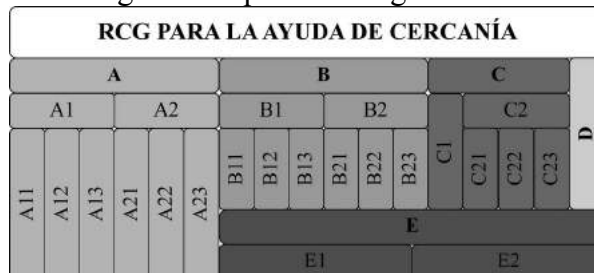


Imagen 1 - Bloques estratégicos

Referencias de la imagen 1:

- A. Relaciones Interinstitucionales
 - A1. Triple Hélice
 - A11. Universidad
 - A12. Gobierno
 - A13. Empresas
 - A2. Solidarias / Intermedias
 - A21. ONG
 - A22. Fundaciones
 - A23. Ayuda Social
- B. Sistemas
 - B1. Integración
 - B11. Red de Conocimiento
 - B12. Comercios y Pymes
 - B13. Medios de Pago
 - B2. Customización
 - B21. Análisis
 - B22. UX
 - B23. Desarrollo
- C. Capacitación
 - C1. Interna
 - C2. Externa
 - C21. Comercios y Micropymes
 - C22. Benefactores
 - C23. Beneficiarios
- D. Marketing
- E. Financiamiento
 - E1. Público
 - E2. Privado

Construcción de los Modelos de Redes de Conocimiento

- MR - Modelos de Relevamiento: Natural; Flujo de Trabajo; Básico; Estímulo Comunicacional.
- MA - Modelos de Análisis: Escenario; Contenido.
- MD - Modelos de Diseño: Estratégico - Negocio; Táctico - Proceso
- DCR - Deployment de la Red Comunicacional.

Modelos de Interacción entre Sistemas y Procesos

Las soluciones propuestas en el marco de este Proyecto se asientan sobre tres pilares tecnológicos fundamentales:

1. La plataforma que reúne e interrelaciona en Red a los tres grupos de interesados principales (benefactores o donantes, beneficiarios o receptores y comercios adheridos)
2. La App de transferencia de dinero desde el benefactor o donante al comercio (o microempresa) adherido
3. La búsqueda geo-referenciada de comercios y microempresas de cercanía.

Diagramas de Actividades, Procesos y Estados; Formularios de Comunicación entre Aplicaciones; Objetos; Estados de los Objetos; Diagramas de Transición de Estados de los Objetos.

Definición de los instrumentos para la Toma de decisiones Operacionales y Tácticas

- *Ámbito de los Usuarios*
- *Informes: Display; Reportes Básicos; Reportes Avanzados; Reportes Favoritos*
- *Trazabilidad*
- *Notificaciones*

Indicadores para una Gestión Estratégica

Se brindarán herramientas para que los involucrados puedan calificar y hacer recomendaciones sobre el funcionamiento general de la **Ayuda de Cercanía**.

Para la etapa donde la Red se encuentre consolidada se analizarán y recomendarán Indicadores para Procesos de Mejora Continua, Optimización y Evolución del Proyecto.

Para efectuar la Gestión de Conocimiento se proveerán herramientas basadas en: Reportes Generales; Informes Estadísticos; Informes Geo-referenciados; Gestión de Indicadores automatizados.

4. FORMACIÓN DE RRHH

- Tesis Doctoral de Doctorado en Administración: Mgr. Gustavo Tripodi, doctorando del Programa Doctoral en Administración de la Facultad de Ciencias Económicas-UNICEN. Temática: “Componentes holísticas de las redes de

conocimiento para el diagnóstico organizacional”

- Tesis de Grado: Josefina Tripodi en la Carrera de Licenciada en Administración de Empresas de la Facultad de Ciencias Económicas de la UNICEN, dirigida por Fernando Mascetti y Asesorada por Jorge Sagula.

- Ing. de Sistemas Santiago Carliski, Ayudante de la Cátedra de Fundamentos de Economía y Proyectos de Inversión de quinto año de la Carrera de Ing. de Sistemas de la Facultad de Ciencias Exactas de la UNICEN.

- Lic. Victoria Navarro, alumna del curso de Posgrado “Metodología para la gestión de proyectos de base tecnológica” de la UNLu del Departamento de Ciencias Básicas

- Dana Alvarez, Horacio Casado, Palmira Benavente, Valentina Besonart, Juan Viviant, son Alumnos de la cátedra “Fundamentos de Economía y Proyectos de Inversión” de quinto año de la Carrera de Ing. de Sistemas de la Facultad de Ciencias Exactas de la UNICEN.

5. BIBLIOGRAFIA

- The Strength of Weak Ties (1973). Granovetter, M. American Journal of Sociology, Vol. 78, No.6..
- Economic Action and Social Structure: The Problem of Embeddedness (1985). Granovetter, M. American Journal of Sociology,
- Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness (1997). Uzzi, B. Administrative Science Quarterly,
- Fundamentals Business Process Management (Second Edition 2018). Marlon Dumas M., La Rosa M., Mendling J., Reijers H. Springer
- Knowledge Management (Second Edition 2018). North K., Kumta G.
- Digital Business Models (2019). Wirtz B. Springer.

Seguimiento de Pacientes en Rehabilitación Robótica

ACOSTA, Nelson (FCs Exactas – UNICEN) Email: el.nelson.acosta@gmail.com
CHIODI, Andrea (CIANE) Email: andrebchiodi@gmail.com
DÍAZ, Alejandro (FCs Salud – UNICEN/CONICET) Email: alejandrounicen@gmail.com
CRUZ Alejandro (Dto Informática – UN Chilecito) Email: alejandrocruz1987@gmail.com

Resumen

El entrenamiento en cinta con apoyo de peso corporal asistido manualmente es un enfoque contemporáneo de rehabilitación de la marcha, donde un individuo camina en una cinta con peso corporal parcialmente soportado por un arnés superior, mientras uno a tres terapeutas facilitan el control de las extremidades. La robótica aplicada al entrenamiento intensivo o recuperación del movimiento permite que el patrón de pasos rítmicos y repetitivos proporcionado, combinado con la carga activa de las extremidades, promueve la plasticidad neuronal cerebral y de la médula espinal. Con esta tecnología podemos aumentar el número de repeticiones, cambiar la velocidad, o variar el soporte del peso corporal para contribuir al aprendizaje motor estableciendo nuevas conexiones con el cerebro. Este desarrollo impone una cuidadosa interfaz entre el usuario, quién maneja el equipo, y el personal médico que dirige y controla los ejercicios. Se generan todos los informes necesarios para que personal médico pueda saber y evaluar exactamente lo que sucede con su paciente a la distancia. El objetivo del proyecto es evaluar los parámetros directos e indirectos que puede aportar información completa al equipo médico, que permita un seguimiento cercano de los pacientes a distancia.

Contexto

La terapia convencional está limitada por el tiempo, el número de repeticiones, la falta de

calidad de movimiento reproducible y que es agotadora tanto para los terapeutas como para los pacientes. Existe un desequilibrio entre la terapia que se debe ofrecer según los principios del aprendizaje motor, y todos los factores que impiden alcanzar ese objetivo. Estudios demuestran que el ejercicio intenso es la clave para mejorar la función motora y produce resultados sostenibles [1].

Según la *Organización Mundial de la Salud* el 7% de la población mundial tiene discapacidad en miembros inferiores [2]. De 15 millones de personas por año que sufren un ACV, 30% quedarán con una discapacidad permanente [3], y 25% de los supervivientes acaba siendo dependiente por su movilidad [4].

El entrenamiento en cinta con apoyo de peso corporal asistido manualmente es un enfoque actual de rehabilitación de la marcha [5], donde un individuo camina con peso corporal parcialmente soportado por un arnés, mientras 1-3 personas facilitan el control de las extremidades [6, 7]; hecho muy difícil y que presenta riesgo de lesión para el personal, o en el mejor de los casos agotamiento y frustración por lo lento del avance de los pacientes.

Por esto se desarrolla la robótica aplicada al entrenamiento intensivo o recuperación del movimiento. El patrón de pasos rítmicos y repetitivos proporcionado por la asistencia robótica, combinado con la carga activa de las extremidades, promueve la plasticidad neuronal cerebral y de la médula espinal. Con esta tecnología podemos aumentar el número

de repeticiones, cambiar la velocidad, o variar el soporte del peso corporal para contribuir al aprendizaje motor estableciendo nuevas conexiones con el cerebro [8]. Dispositivos usados para: ACV, parálisis cerebral, lesiones medulares, Parkinson, distrofia muscular, varios tipo de parálisis, y toda condición que mantenga en silla de rueda o cama al paciente.

1.-Introducción

En este proyecto usamos *Realwalker* (www.realwalker.com.ar), diseño disruptivo por la tecnología patentada que baja su precio permitiendo que el paciente entrene intensivamente, agilizando la recuperación y bajando los costos, siempre bajo coordinación de personal médico que debe saber y evaluar exactamente lo que sucede con su paciente a la distancia [9, 10].

Está demostrado que los dispositivos robóticos como medio para automatizar e intensificar el entrenamiento locomotor en neurorehabilitación mejora la recuperación del movimiento y acorta el tiempo de los tratamientos [11].

Hay tres motivaciones principales para desarrollar nuevas tecnologías para asistir en rehabilitación [12]. Primero, la tecnología tiene el potencial de permitir más terapia con menos supervisión, mejorando los perfiles de costo-beneficio de la rehabilitación. Este objetivo puede expresarse como el desarrollo de una tecnología que promueva de forma óptima la plasticidad dependiente del uso, a la vez que reduce el coste de la terapia [7].

Segundo, la tecnología tiene el potencial de cuantificar la terapia con mayor precisión, incluyendo las características del paciente que predicen el éxito de la terapia, la dosis y el contenido de la terapia y los resultados clínicos [13, 14]. Esta propiedad de cuantificación de la tecnología es importante

para mejorar la comprensión mecanicista de la ciencia de la rehabilitación, la toma de decisiones clínicas y la retroalimentación y motivación del paciente [15, 16].

Tercero, la tecnología tiene el potencial de permitir tipos de terapia completamente nuevos. Un ejemplo es el concepto de proporcionar terapia continua con dispositivos robóticos o que se pueden llevar puestos [17].

Los terapeutas de rehabilitación no pueden ser omnipresentes, pero la tecnología robótica e inteligente si puede, proporcionando terapia durante mayor cantidad de veces, con mayores repeticiones, con mayor posibilidad de repetibilidad, y con mayor cantidad de variaciones. Y todo esto en un ambiente donde el personal médico conoce exactamente lo que entrena cada paciente [10, 18, 19].

Estas tres motivaciones son los grandes motores de este proyecto, y generan estos grandes interrogantes que pretendemos resolver con este proyecto son:

1. ¿Es posible detectar antes que suceda alguna complicación derivada del ejercicio?
2. ¿Es posible (1) sin necesidad de cargar gran cantidad de planillas?
3. ¿Cuáles parámetros es posible inferir con los que genera el equipo?

2.-Líneas de Investigación y Desarrollo

Actualmente se está trabajando en 4 líneas de investigación y desarrollo:

- a) Selección de parámetros (Proyecto de incentivos de la UNICEN). Este proyecto se comenta en los resultados obtenidos y esperados.
- b) Medición física de efectos del entrenamiento robótico pasivo en pacientes con discapacidad motor

(Proyecto de investigación CIANE - UNICEN). Se realizan mediciones: de flujo y caudal sanguíneo, fuerza y movilidad (escala de Daniels), tono muscular (escala de Ashworth, y escala de Tardieu), calidad de materia fecal (escala de Bristol), parámetros en metabolismo fosfocalcico (escala FIM), independencia funcional (escala de Barthel). Todas estas mediciones se realizan antes y al finalizar el periodo de entrenamiento planificado con el sistema robótico pasivo, de esa forma se puede analizar el efecto del entrenamiento intensivo en el paciente.

- c) Actualización de sistema de generación de movimiento (Patente en trámite).
- d) Ortesis dinámica de rodilla (Proyecto de investigación en conjunto con UN San Martín, evaluado y financiado por Procodas). Desarrollo de exoesqueleto de rodilla que soporta el peso del paciente permitiendo un desplazamiento en caminata.

3.-Resultados Obtenidos y Esperados

Este artículo plantea la situación actual, a 3 años de comenzar a trabajar en conjunto. Donde se ha realizado el análisis de los parámetros del sistema robótico (Línea 2.a), y los resultados han sido: se ha definido un conjunto de parámetros a tener en cuenta, y como respuesta ya está planteándose un plan para su adopción escalonada en los dispositivos. Se espera obtener una evolución de los dispositivos que permita evaluar varios parámetros adicionales a los planteados originalmente.

En el caso de la medición física de efectos del entrenamiento robótico pasivo en pacientes con discapacidad motor (Línea 2.b), se ha comenzado a trabajar en enero de 2021. Hasta

el momento hemos realizado algunas tareas: 1) Analizar los grupos de pacientes, 2) Seleccionar quienes pueden participar del proyecto, y 3) Avanzar con el documento de consentimiento informado para la práctica de los entrenamientos en el marco del proyecto de investigación. Este proyecto se espera sea completado en el transcurso de 2021.

Por otra parte, los otros proyectos (las Líneas 2.c y 2.d) avanzan por carriles diferentes con gran independencia y con una problemática totalmente diferente. Una basada en trámites de patentes, mientras que el otro estamos buscando financiar motores que permitan un movimiento más suave y preciso.

En este proyecto se desarrolló un conjunto de parámetros que permitan a personal médico analizar y evaluar de forma certera la evolución (o involución) de un paciente entrenando caminata en un equipo robótico. Un conjunto de parámetros son generados, de forma automática (*paciente, diagnóstico, personal médico, velocidad, aceleración, tiempo, fecha y hora, entre otros*) o manual (*peso descargado*), por el equipo; mientras que el resto de los parámetros son calculados en base a las evaluaciones médicas, la búsqueda de información y el análisis de tendencias en los datos generados.

Estas tecnologías permitirán asistir en rehabilitación logrando más terapia con menos supervisión, cuantificando la terapia con mayor precisión, incluyendo las características del paciente que predicen el éxito de la terapia (la dosis, el contenido de los ejercicios, y los resultados clínicos). Todas estas características tienen dos efectos principales; para las clínicas y obras sociales mejora la relación costo-beneficio de la rehabilitación, mientras que, para el paciente reduce los tiempos de recuperación del movimiento. Los terapeutas no pueden ser omnipresentes, pero con esta tecnología se

permite que sin su presencia física permanente conozca todos los datos del entrenamiento de cada paciente.

La universidad por parte de la Facultad de Ciencias Exactas tanto como la Facultad de Ciencias de la Salud tiene un equipo multidisciplinario que trabaja en dispositivos médicos basados en tecnología informática. El impacto en nuestras cátedras es directo, ya que el proyecto tiene directa relación con las materias de todos los involucrados en el proyecto, provocando un enriquecimiento y su directo impacto a los alumnos de grado y postgrado. También los integrantes del proyecto por parte de las instituciones contarán con el uso del equipo para sus pacientes, y desarrollarán sus propias técnicas de tratamiento o evaluarán las publicadas por otros colegas.

Se espera tener capacidad de atención de al menos 20 personas diarias, que se beneficiarán tanto del entrenamiento como de toda la información y estudios realizados (donde la principal limitación está dada por las horas del personal médico y de asistencia en la institución). La vinculación del CIANE tiene un impacto directo, ya que el proyecto nace vinculando a dicha institución, y dentro de los objetivos está el poder disponer del uso del dispositivo por parte de sus pacientes.

4.-Formación de Recursos

Humanos

El equipo formato tiene 4 partícipes muy comprometidos: la Facultad de Ciencias Exactas-UNICEN con su núcleo NICE, la Facultad de Ciencias de la Salud-UNICEN, el centro de atención a niños CIANE, y la empresa Tecnología LINDA SRL que fabrica los entrenadores.

El equipo del NICE tiene experiencia en desarrollo de sistemas, tanto el software como el hardware, elementos básicos de los

sistemas robóticos. La patente del desarrollo del sistema robótico de entrenamiento pertenece a un miembro activo del núcleo.

El personal involucrado de la Facultad de Ciencias de la Salud, tiene amplia experiencia profesional y científica en salud humana. El equipo cuenta con las especialidades vinculadas a la rehabilitación humana, médico fisiatra, neurólogo, y se destacan las contribuciones del Dr. Alejandro Díaz en cardiología, quien ha sido promovido a la categoría de Investigador Independiente de CONICET luego de obtener importante premio otorgado por la Sociedad Argentina de Hipertensión Arterial.

El CIANE es una asociación civil sin fines de lucro creada en el año 2006 por un grupo de padres de niños con discapacidad para la creación de un Centro Integral de Atención a Niños Excepcionales. Este 2021 inaugura su nueva sede.

Hace ya 2 años que estamos trabajando juntos (gran parte del equipo). Hoy en día casi todos los equipos de salud tienen en su interior hardware y software, y en conjunto somos un equipo humano, científico y técnico, con la capacidad de realizar especificaciones y diseño de equipos de alto desempeño aplicables a la salud humana.

Se esperan avances en cuanto al doctorado de:

- Ing. Cruz quien está haciendo el Doctorado en Matemática Computacional e Industrial de la Facultad de Ciencias Exactas de la UNICEN dirigido por Acosta.
- Med. Chiodi se inscribirá próximamente al doctorado en Ciencias Aplicadas, mención Ambiente y Salud, de la UNICEN dirigido por Acosta.

La médica Andrea Chiodi ha terminado como postgrado la especialidad de Medicina Física y Rehabilitación otorgada por el Colegio de Médicos Distrito I (La Plata). También se

cuenta con la participación de alumnos de grado en el proyecto.

5.-Bibliografía

- [1] Andrews AW, Li D, & Freiberger JK, **Association of Rehabilitation Intensity for Stroke and Risk of Hospital Readmission**. Phys Ther 2015 [4].
- [2] World Health Organization (WHO). **Neurological Disorders: public health challenges**. ISBN 92 4 156336 2. ISBN 978 92 4 156336 9. Switzerland (2006) [99].
- [3] World Health Organization. **Neurological Disorders: Public Health Challenges**. 2006, Geneva, Switzerland: World Health Organization [2].
- [4] Hendricks, H. T., J. van Limbeek, A. C. Geurts and M. J. Zwarts (2002). **Motor recovery after stroke: a systematic review of the literature**. Arch Phys Med Rehabil 83(11):1629-1637 [3].
- [5] Kelly P Westlake1 and Carolyn Patten. **Pilot study of Lokomat versus manual-assisted treadmill training for locomotor recovery post-stroke**. Journal of NeuroEngineering and Rehabilitation Research, (2009) 6:18 [98].
- [6] R Salvatore, A Naro, M Russo, P Bramanti, L Carioti, T Balletta, A Buda, A Manuli, S Filoni & A Bramanti. **Shaping neuroplasticity by using powered exoskeletons in patients with stroke: a randomized clinical trial**. Journal of NeuroEngineering & Rehabil. (2018) 15:35 [70].
- [7] H van Hedel, G Severini, A Scarton, A O'Brien, T Reed, D Gaebler-Spira, T Egan, A Meyer-Heim, J Graser, K Chua, D Zutter, R Schweinfurter, JC Möller, L Paredes, A Esquenazi, S Berweck, S Schroeder, B Warken, A Chan, A Devers, J Petioky, NJ Paik, WS Kim, & P Bonato. **Advanced Robotic Therapy Integrated Centers (ARTIC): an international collaboration facilitating the application of rehabilitation technologies**. J. NeuroEngineering & Rehabil. (2018) 15:30 [80].
- [8] JC Moreno, F Barroso, D Farina, L Gizzi, C Santos, M Molinari & JL Pons. **Effects of robotic guidance on the coordination of locomotion**. J. NeuroEngineering & Rehabil. (2013), 10:79 [83].
- [9] KY Nam, HJ Kim, BS Kwon, JW Park, HJ Lee & A Yoo. **Robot-assisted gait training (Lokomat) improves walking function and activity in people with spinal cord injury: a systematic review**. Journal of NeuroEngineering and Rehabilitation (2017) 14:24 [96].
- [10] K Kammen, AM Boonstra, L van der Woude, H Reinders & R Otter. **Differences in muscle activity and temporal step parameters between Lokomat guided walking and treadmill walking in post-stroke hemiparetic patients and healthy walkers**. Journal of NeuroEngineering and Rehabilitation (2017) 14:32 [97].
- [11] K Bustamante Valles, S Montes, M Madrigal, A Burciaga, ME Martínez & MJ Johnson. **Technology-assisted stroke rehabilitation in Mexico: a pilot randomized trial comparing traditional therapy to circuit training in a Robot/technology-assisted therapy gym**. Journal of NeuroEngineering and Rehabilitation (2016) 13:83 [69].
- [12] DJ Reinkensmeyer & ML Boninger. **Technologies and combination therapies for enhancing movement training for people with a disability**. Journal of NeuroEngineering and Rehabilitation (2012), 9:17 [78].
- [13] Elena Kokkoni, E Mavroudi, A Zehfroosh, JC Galloway, R Vidal, J Heinz & HG. Tanner. **GEARING smart environments for pediatric motor rehabilitation**. Jour. NeuroEngineering & Rehabilitation (2020) [105].
- [14] S. Fricke, C Bayón, H van der Kooij & E. van Asseldonk. **Automatic versus manual tuning of robot-assisted gait training in people with neurological disorders**. J. NeuroEngineering and Rehabilitation (2020) 17:9 [100].
- [15] N Fitter, M Mohan, KJ. Kuchenbecker & MJ. Johnson. **Exercising with Baxter: preliminary support for assistive social-physical human-robot interaction**. J NeuroEngineering and Rehabilitation (2020) 17:19. [103].
- [16] S Guillén-Climent, A Garzo, MN Muñoz-Alcaraz, P Casado, J Arcas, M Mejías & F Mayordomo-Riera. **A usability study in patients with stroke using MERLIN, a robotic system based on serious games for upper limb rehabilitation in the home setting**. J NeuroEngineering Rehabil (2021). [113].
- [17] Rachel L. Hawe, Andrea M. Kuczynski, Adam Kirton & Sean P. Dukelow. **Robotic assessment of rapid motor decision making in children with perinatal stroke**. Journal of NeuroEngineering and Rehabilitation (2020). [108]
- [18] I Bortone, M Barsotti, D Leonardis, A Crecchi, A Tozzini, L Bonfiglio & A Frisoli. **Immersive Virtual Environments and Wearable Haptic Devices in rehabilitation of children with neuromotor impairments: a single-blind randomized controlled crossover pilot study**. J NeuroEngineering Rehabil (2020). [111].
- [19] R. Rodríguez, M. Torres, M. Zayas, A. Montoya, D. Milánés & R. Zamora. **Robotic therapy for the hemiplegic shoulder pain: a pilot study**. J NeuroEngineering & Rehabilitation (2020) [106]

Sistema multiplataforma para el registro y asistencia de pacientes con síndrome de intestino irritable

Nicolas Domínguez, Jorge Ierache ¹, Juan Zaragoza¹, Hernan Merlino¹, Milagros Celleri², Florencia Koutsovitiss², Maria Piskorz³, Jorge Olmos³, Cristian Garay²

¹ Laboratorio de Sistemas de Información Avanzados, Departamento Computación, Facultad de Ingeniería, Universidad de Buenos Aires, Argentina

{jierache,jzaragoza, dominguez,hmerlino}@fi.uba.ar

<http://lsia.fi.uba.ar/>¹

² Facultad de Psicología, Universidad de Buenos Aires, Argentina

{mcelleri, fkoutsovitiss, cristiangularay}@psi.uba.ar

³ Facultad de Medicina, Universidad de Buenos Aires, Argentina

{maipiskorz}@me.com

Resumen

En este trabajo de investigación aplicada se centra en el desarrollo de un prototipo de sistema multiplataforma en particular se presentan las características del sistema que se puede usar desde aplicaciones móviles o desde un ambiente web. El desafío del mismo es contribuir al tratamiento de lo que se denomina trastornos gastrointestinales funcionales, particularmente el más prevalente, el "síndrome de intestino irritable", el que en la actualidad tiene un elevado impacto en la calidad de vida de las personas que sufren este síndrome.

Palabras clave: sistema de registro; factores psicológicos; intestino irritable; trastornos gastrointestinales funcionales.

Contexto

Esta investigación se encuentra dentro del contexto del Proyecto de Desarrollo Estratégico (PDE) 43-2019: "Desarrollo de una herramienta para el registro por parte de los pacientes de los factores psicológicos que inciden en los trastornos gástricos funcionales", Facultad de

Psicología, Universidad de Buenos Aires (UBA). En el proyecto también participa el Sector de Neurogastroenterología, Hospital de Clínicas San Martín, Facultad de Medicina, UBA, con la participación conjunta del laboratorio de sistemas de información avanzados y la cátedra de Aplicaciones informáticas del Departamento de Computación, Facultad de Ingeniería, UBA.

Introducción

Los trastornos gastrointestinales funcionales constituyen cuadros altamente prevalentes en la población y que generan altos costos económicos y de recursos en la salud pública. Dentro de estos cuadros, los más prevalentes son el Síndrome de Intestino Irritable (SII) y la Dispepsia Funcional (DF) [1]. Hasta el momento, no se conoce una fisiopatología clara que explique el cuadro [2]. El síndrome de intestino irritable [3] se caracteriza por la presencia de dolor abdominal, constipación, diarrea o ambas (alternancia entre constipación y diarrea), al menos una vez por semana los últimos tres meses. Es un trastorno sumamente frecuente que lo padece cerca del 12% de la población general, siendo más prevalente en mujeres (15%). Se diagnostica en base a

criterios clínicos llamados criterios de Roma IV [4]. En pacientes menores de 50 años, que no presentan síntomas de alarma (sangre en la materia fecal, pérdida de peso significativa, fiebre), tienen un examen físico y análisis de sangre normales, no hace falta hacer ningún estudio diagnóstico específico y el diagnóstico se realiza en base a los criterios clínicos. En pacientes mayores a 50 años o con síntomas de alarma se debe realizar una serie de test diagnósticos que incluyen las endoscopias para descartar una serie de confundidores como el cáncer de colon, las enfermedades inflamatorias intestinales y la enfermedad celíaca entre otras. En el caso que los estudios sean normales, nuevamente se hace diagnóstico de Síndrome de intestino irritable en base a los criterios clínicos. El problema de este síndrome es su impacto negativo en la calidad de vida del paciente. Tanto es así, que este trastorno es la segunda causa de ausentismo laboral después del resfriado común y es uno de los principales motivos de consulta en la práctica clínica y gastroenterología. Luego del diagnóstico, se requiere una evaluación médica para identificar los posibles desencadenantes de los síntomas. Entre los que se encuentran principalmente el estrés en personas predispuestas y otros factores emocionales, también los hábitos alimenticios, entre otros. Los síntomas se logran controlar haciendo algunas modificaciones del estilo de vida, la alimentación, el manejo del estrés y síntomas psicológicos. En algunos casos es necesario utilizar medicación, que será indicada por el médico gastroenterólogo, para controlar los síntomas. Entre los factores involucrados, los factores psicológicos ocupan un lugar relevante siendo los más frecuentes, la ansiedad, la depresión, el neuroticismo y una elevada somatización [5]. Los aspectos emocionales acompañan estos síntomas con mucha frecuencia presentando altas comorbilidades con los mismos. El estrés cotidiano, emociones normales como la ansiedad, la angustia, el enojo y la tristeza

pueden incrementarse por diversas causas, incluyendo haber vivido momentos difíciles a lo largo de la vida, e influir en la intensidad de los síntomas de los trastornos funcionales. En estos pacientes en los que la sintomatología ansiosa y depresiva es elevada el tratamiento habitual suele no ser suficiente y presentarse de manera recurrente, además de no responder al tratamiento de forma adecuada, por lo que el abordaje psicológico es esencial [6]. El tratamiento psicológico de primera elección para el síndrome de intestino irritable es la terapia cognitivo-conductual (TCC) [7] y, en la dispepsia funcional también este tipo de terapia ha demostrado buenos resultados [8]. La TCC supone que los pensamientos, las emociones y las conductas de las personas se encuentran relacionados y se influyen mutuamente. El automonitoreo de las mismas y su registro para su posterior análisis junto al terapeuta y la utilización de estrategias y técnicas que permitan a los pacientes regular sus emociones de forma funcional son centrales en este tipo de tratamiento psicológico [9].

Diferentes aplicaciones en la actualidad contribuyen con la problemática planteada en otras podemos considerar las siguientes aplicaciones: La app B.Energy, de ALVUM [10] incorpora la escala de Bristol, tiene como objetivo ayudar a monitorizar la condición de una persona, compartir su evolución con los profesionales de la salud y los pacientes. FODMAP para la asistencia en dieta baja caloría, ayudan a disminuir los síntomas SII, con un control de los alimentos que se consumen [11]. Otra app es mySymptoms Food Diary & Symptom Tracker permite [12], permite hacer seguimiento de los síntomas y de la alimentación, La aplicación Bowelle realiza el seguimiento de síntomas y los relaciona también con las comidas. [13], incorpora una evaluación a nivel de valencia del estado emocional. Digestivamente en comparación con las aplicaciones anteriormente indicadas integra

información del paciente, historia clínica, permite la explotación de la información y brinda respuesta de la evolución del paciente, motivando al mismo, particularmente integra la información del estado emocional del paciente considerando además de la valencia, la intensidad del mismo, sumando esta información con la escala de Bristol, y los síntomas presentados como hinchazón, dolor abdominal etc, además de resúmenes de estados semanales y estresores del paciente, futuras líneas de trabajo contemplan incorporar información de la alimentación que consume el paciente rutinariamente.

Líneas de Investigación, Desarrollo e Innovación

La hipótesis que guía este proyecto es: Los registros de pacientes en tiempo real con el empleo de aplicación “Digestivamente”, interactuando con el contexto real del paciente en términos de estresores, manifestaciones fisiológicas y sus estados emocionales contribuye a reflejar a una mejora en su calidad de vida.

Para demostrar dicha hipótesis, se encuentran abiertas y en desarrollo las siguientes líneas de trabajo:

a) evolución del desarrollo de la herramienta Digestivamente.

b) visualización y explotación. Se desarrollan consultas y filtros con integración en futuras líneas con herramientas de explotación y descubrimiento sobre la base de datos que articula el sistema.

Objetivos y Resultados Esperados

El objetivo ulterior del presente trabajo es facilitar el registro de los síntomas influyentes en la sintomatología en pacientes con diagnóstico de intestino irritable que acuden a

un servicio de Neurogastroenterología en un hospital público

El desarrollo tiene como objetivo el diseño y creación de una aplicación PWA [14]. para el registro de síntomas psicológicos en pacientes con Síndrome de Intestino Irritable asociado a los factores psicológicos, principalmente el estrés, la ansiedad y la depresión. Se trata de un trabajo descriptivo y transversal. Se espera que el uso de la aplicación web ayude en la reducción de la sintomatología en los pacientes evaluados a través del registro de emociones, pensamientos, conductas y la psicoeducación.

La Infraestructura tecnológica del prototipo consiste en: a) base de datos a AWS RDS para un despliegue sencillo de la base de datos en MySQL. El Frontend / Backend, b) Node.js Node.js [15], como entorno en tiempo de ejecución multiplataforma, de código abierto, para la capa del servidor, c) React React.js [16] biblioteca Javascript de código abierto empleada en el desarrollo de las interfaces de usuario. Ambos alojados en la nube usando el servicio Heroku [17].

Resultados obtenidos

Hasta el momento los logros obtenidos se reflejan en el desarrollo del prototipo de sistema Digestivamente, se realizaron pruebas con usuarios finales (pacientes, médicos, psicólogos). El sistema desarrollado denominado “Digestivamente” permite entre sus principales funcionalidades: a) Registrar los valores diarios de Escala de Bristol de los pacientes, sus sensaciones físicas y emocionales, como así también registros de pensamientos y actividades, b) Brindar una interfaz que le permite a los pacientes determinar los valores de defecación de acuerdo con la escala de Bristol, c) Identificar las emociones que enfrentan los pacientes al finalizar el día, d) Reconocer los estresores semanales que los pacientes han vivido durante el transcurso de la semana., e) Evaluar, por

parte de los profesionales, los resultados de cada paciente a través de gráficos que muestran las relaciones entre cada una de las variables mencionadas anteriormente y facilitar la explotación de datos del paciente, f) Realizar un seguimiento de los valores de la escala Bristol, las emociones, y estresores de los pacientes, g) Cargar y administrar los diferentes usuarios: profesionales y pacientes, se presenta en la figura 1, h) Visualizar y analizar la información recolectada para cada uno de los pacientes, asistiendo así a la identificación de los factores involucrados en la presentación o exacerbación de los síntomas encontrados, i) Visualizar y realizar un seguimiento de los diferentes diagnósticos de cada paciente, brindando información “no apreciable” a simple vista por el médico a la hora de observar los datos.

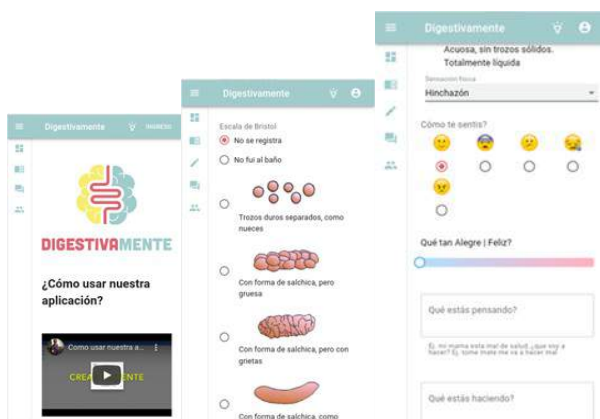


Fig. 1a. Dashboard paciente en mobile Fig. 1: b) Carga de registro: referencia de escala Bristol y 1c) Carga de registro, selección de sensaciones y pensamientos.

Se presenta en la figura 2 la arquitectura del prototipo junto con los componentes que lo identifican.

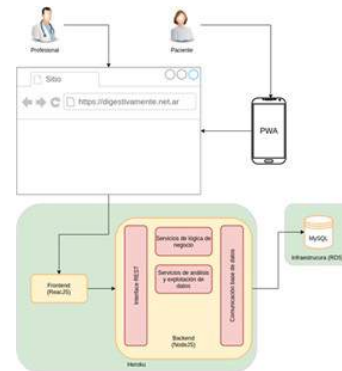


Fig. 2 Diagrama la arquitectura del prototipo junto con los componentes que lo identifican.

Recientes trabajos no publicados a la fecha obtuvieron resultados satisfactorios en el desarrollo Digestivamente aplicado a la explotación y representación de información a médicos, psicólogos como así también a los pacientes a fin de que estos puedan ver su evolución, sus registros frente a distintas situaciones que actúan como estresores ocasionando síntomas de intestino irritable que afectan su calidad de vida.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por dos investigadores formados, del LISA FIUBA y dos alumnos que se encuentran finalizando su trabajo profesional en el marco de la carrera de ingeniería informática de la FIUBA. Por dos estudiantes de posgrado y un investigador formado de la Facultad de Psicología, de la Universidad de Buenos Aires, y dos médicos gastroenterólogos pertenecientes a la división de Neurogastroenterología del Hospital de Clínicas.

Referencias

1. Juan Zaragoza , Jorge Ierache , Nicolas Domínguez , Hernan Merlino , Milagros Celleri , Florencia Koutsovitis , Cristian

- Garay (2020) Sistema para el registro de factores psicológicos que inciden en los trastornos: gastrointestinales funcionales. "Digestivamente". XX Congreso Argentino de Ciencias de la Computación (CACIC) Octubre 2020, en prensa.
2. Mahadeva S, Goh KL. Epidemiology of functional dyspepsia: A global perspective. *World J Gastroenterol.* 2006;12(17):2661–6.
 3. Brizuela Quintanilla RA. Síndrome de intestino irritable. *Rev Cuba Med Mil.* 1997;26(1):63–8.
 4. Sebastián Domingo JJ. The new Rome criteria (IV) of functional digestive disorders in clinical practice. *Med Clin (Barc)* [Internet]. 2017;148(10):464–8. Disponible en: <http://dx.doi.org/10.1016/j.medcli.2016.12.020>
 5. Whitehead WE, Palsson O, Jones KR. Systematic review of the comorbidity of irritable bowel syndrome with other disorders: What are the causes and implications *Gastroenterology.* 2002;122(4):1140–56.
 6. Jia L, Jiang SM, Liu J. Behavioral gastroenterology: An emerging system and new frontier of action. *World J Gastroenterol.* 2017;23(33):6059–64.
 7. American Psychological Association. Diagnosis: irritable bowel syndrome, psychological treatments [Internet]. [Consultado feb 2021]. Disponible en: <https://www.div12.org/diagnosis/irritable-bowel-syndrome/>
 8. Haug TT, Mykletun A, Dahl AA. Are anxiety and depression related to gastrointestinal symptoms in the general population? *Scand J Gastroenterol.* 2002;37(3):294–8.
 9. Beck J. Terapia cognitiva: conceptos básicos y profundización. Editorial Gedisa; 1995.
 10. B.Energy, disponible en : <https://www.alvumhealth.com/2019/02/21/alvum-lanza-b-energy-aplicacion-movil-sindrome-intestino-irritable-sii-colon-irritable/>, [Consultado feb 2021]
 11. Figueroa, Carolina. (2015). Dieta baja en fodmap en el síndrome de intestino irritable. *Revista Médica Clínica Las Condes.* 26. 628-633. 10.1016/j.rmcl.2015.08.002
 12. mySymptoms Food Diary & Symptom Tracker permite , disponible en: https://colon-irritable.es/5-apps-android-iphone-intestino-colon-irritable/#mySymptoms_Food_Diary_Symptom_Tracker [Consultado feb 2021].
 13. Bowelle disponible en : <https://colon-irritable.es/5-apps-android-iphone-intestino-colon-irritable/#Bowelle> [Consultado feb 2021]
 14. PWA. Plataformas web en dispositivos móviles. [Internet]. [Consultado 28 may 2020]. Disponible en: https://en.wikipedia.org/wiki/Progressive_web_application
 15. NodeJS. Entorno de ejecución para JavaScript construido con el motor de JavaScript V8 de Chrome. [Internet]. [Consultado feb 2021]. Disponible en: <https://nodejs.org/>
 16. React. Biblioteca de JavaScript para desarrollo de clientes web. [Internet]. [Consultado feb 2021]. Disponible en: <https://reactjs.org/>
 17. Heroku. Plataforma como servicio de computación en la nube. Disponible en: <https://www.heroku.com/>

SISTEMAS DE CONTROL DE ROBOTS SOBRE PLATAFORMA DE INTERNET DE LAS COSAS

Azurra, Diego Santos, Damián Fernández, Gustavo Amaro, Marcos
dazcurra@yahoo.com damiansantos@gmail.com gusferna@yahoo.com.ar marcos.n.amaro@gmail.com

Dpto. Desarrollo Productivo y Tecnológico
Universidad Nacional de Lanús
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús
Buenos Aires, Argentina.

Resumen

Este proyecto se inscribe en una línea de investigación desarrollada en el marco de la Carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús, con foco en sistemas de control de robots sobre plataformas de IoT.

En este trabajo se presentan los supuestos que guían al proyecto, así como los objetivos, actividades y resultados esperados.

Palabras clave: robótica – sistemas de control - Internet de las Cosas - telecomunicaciones

Contexto

Este proyecto de investigación integra la línea de trabajo en aplicaciones de sistemas industriales, robótica y telecomunicaciones desarrollada en el marco de la carrera de la Carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús, en particular continuando al proyecto de investigación bianual (2018-2019) ROBÓTICA BASADA EN INTERNET DE LAS COSAS, presentado en [6] y [12].

Introducción

Según [1], en los últimos años la robótica ha tenido aplicaciones en campos tan diversos y críticos como la medicina, la exploración planetaria y submarina, automatización de procesos industriales, seguridad,

entretenimiento, entre otros. Sin embargo, es en la educación donde ha dado uno de los aportes de mayor impacto, donde los robots al integrarse al grupo de estudiantes y tutores, propician el aprendizaje y el fortalecimiento de habilidades cognitivas.

Por su parte, según la definición de la Internet of Things Community del Instituto de Ingeniería Eléctrica y Electrónica (IEEE), Internet de las Cosas (IoT, por sus siglas en inglés) es un concepto de computación en el que todas las cosas, incluyendo todo objeto físico, pueden ser conectadas, lo que los hace inteligentes, programables y capaces de interactuar con los seres humanos.[2]

Avanzando con el concepto, en [3] se destaca que la IoT involucra sensores, circuitos, sistemas embebidos, comunicaciones, interfaces inteligentes, gestión de energía, gestión de datos, fusión de datos, gestión de conocimiento, sistemas en tiempo real, procesamiento distribuido, diseño de sistemas y técnicas sofisticadas de software que se relacionan con la llamada Big Data. También prevé que su desarrollo podría cambiar profundamente desde los procesos productivos hasta la salud-electrónica, ya que ofrece nuevas formas para el cuidado de humanos y el tratamiento de dolencias.

Un análisis desde el punto de vista económico es realizado en [4], definiéndolo como un mercado en expansión, donde cada vez son más las empresas que valoran la conectividad entre dispositivos y con la Red.

En [5] se definen a las plataformas de IoT como la base para que los dispositivos estén interconectados y se genere un ecosistema propio. Detalla que constituyen el software al que se conectan los dispositivos de hardware, brindando comunicación y puntos de acceso para el desarrollo de aplicaciones.

Líneas de investigación y desarrollo

Este proyecto se inscribe en una línea de investigación relacionada con sistemas industriales, robótica y telecomunicaciones, enfocada en tecnologías con potencial de transferencia a la Industria, particularmente al sector PyME. [6] [12]

Entre los supuestos que guían al proyecto se encuentran:

- I. Existen plataformas de IoT robustas, fiables y de acceso libre o bajo costo que podrían ser utilizadas en modo local en la administración y control de robots.
- II. Es posible adaptar los mecanismos de administración y control de robots para que interactúen en este entorno.
- III. Es factible desarrollar un modelo de transferencia a la Industria, particularmente al sector PyME.

Resultados y Objetivos

El objetivo general de este proyecto es desarrollar y sistematizar el cuerpo de conocimiento vinculado a la administración y control de robots utilizando plataformas de Internet de las Cosas, buscando: [a] Implementar un entorno local de pruebas, y [b] Desarrollar y probar prototipos de software y hardware de mecanismos de administración y control de robots sobre el entorno local.

Se plantean entonces los siguientes objetivos específicos:

- I. Relevar y seleccionar una plataforma de IoT de acceso libre apta para ser instalada en modo local y para implementar mecanismos de administración y control de robots autónomos.
- II. Implementar un entorno local de pruebas. Instalar la plataforma de IoT local.
- III. Desarrollar y probar prototipos de software y hardware de mecanismos de administración y control de robots sobre el entorno local. Evaluar los resultados
- IV. Desarrollar un modelo transferencia a la industria de esta tecnología. Evaluarlo en un caso identificado en la región de influencia de la UNLa (interno de la universidad o PyME).

Formación de Recursos Humanos

El grupo de trabajo está formado por cuatro docentes-investigadores y tres alumnos avanzados de la carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús.

Los tres alumnos están desarrollando su trabajo de fin de carrera (TFI –*Trabajo Final Integrador*) en el marco del proyecto.

Referencias

- [1] J. González E.; B. Jovani A. Jiménez ; (2009); “LA ROBÓTICA COMO HERRAMIENTA PARA LA EDUCACIÓN EN CIENCIAS E INGENIERÍA”; Revista Iberoamericana de Informática Educativa, 10: 31-36; ISSN: 1699-4574
- [2] IEEE Internet of Things Community ; <https://www.ieee.org/>
- [3] Vincenzo Piuri, Roberto Minerva, Construyendo la Internet de las Cosas, Julio 2015;

<https://www.computer.org/web/computingnow/archive/building-the-internet-of-things-july-2015-spanish-version>

[4] El Internet de las Cosas de código abierto: plataformas y aplicaciones para desarrolladores, Agosto 2015; <https://bbvaopen4u.com/es/actualidad/el-internet-de-las-cosas-de-codigo-abierto>

[5] ¿Qué es una plataforma IoT?, 2016 ; <https://secmotic.com/blog/plataforma-iot/>

[6] Azcurra, D., Santos, D., Fernández, G., Fernández, S., (2018), "ROBÓTICA BASADA EN INTERNET DE LAS COSAS". Proceedings del XX Workshop de Investigadores en Ciencias de la Computación. Pág. 835 – 837. ISBN 978-987-3619-27-4.

[7] Kranz, M. (2017). "Building the Internet of Things". ISBN 978-1-119-28566-3. Ed. Wiley.

[8] Sarangi, S., R., Sethi, P. (2016). "Internet of Things: Architectures, Protocols, and Applications"

[9] Al-Fuqaha, Ala., Guizani, Mohsen., Mohammadi, Mehdi., Aledhari, Mohammed., Ayyash, Moussa. (2015). "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications".

[10] McEwen, Adrian., Cassimally, Hakim. (2014), "Designing the Internet of Things". ISBN 978-1-118-43063-7. Ed. Wiley.

[11] Azcurra, D., Rojo, S., Rodríguez, D., (2013), "AVANCES EN EL PROYECTO ARQUITECTURAS DE CONTROL PARA ROBOTS AUTÓNOMOS MÓVILES DIDÁCTICOS BASADAS EN SISTEMAS EMBEBIDOS", Proceedings del XV Workshop de Investigadores en Ciencias de la Computación. Pág. 835 – 837. ISBN 978-987-3619-27-4. Pág. 987-989. ISBN: 9789872817961

[12] Azcurra, D., Santos, D., Fernández, G., Fernández, S., Higa, E., Amaro, M., Wieilly, A., (2019), "AVANCES EN EL PROYECTO

ROBÓTICA BASADA EN INTERNET DE LAS COSAS". Proceedings del XXI Workshop de Investigadores en Ciencias de la Computación. Pág. 673-675. ISBN 978-987-3984-85-3

Sitios web:

IBM. "IaaS PaaS SaaS – Modelos de servicio Cloud": <https://www.ibm.com/cloud-computing/es-es/learn-more/iaas-paas-saas/>

http://tinyos.stanford.edu/tinyos-wiki/index.php/FAQ#What_is_TinyOS.3F

<https://riot-os.org/#features>

<http://www.contiki-os.org/#why>

<https://thinger.io/>

<https://www.freertos.org/about-RTOS.html>

<https://thingsboard.io/docs/getting-started-guides/what-is-thingsboard/>

<https://devicehive.com/>

MyDevices Cayenne: [https://mydevices.com/AWS IoT Core "What is AWS IoT?":](https://mydevices.com/AWS_IoT_Core_What_is_AWS_IoT?) <https://docs.aws.amazon.com/iot/latest/developerguide/what-is-aws-iot.html>

Microsoft Azure. (17/05/2018) "Introduction to Azure and the Internet of Things". Microsoft Docs: <https://docs.microsoft.com/en-us/azure/iot-fundamentals/iot-introduction>

Microsoft Azure. (17/05/2018) "Azure IoT technologies and solutions: PaaS and SaaS". Microsoft Docs: <https://docs.microsoft.com/en-us/azure/iot-fundamentals/iot-services-and-technologies>

Google Cloud "Google Cloud IoT": <https://cloud.google.com/solutions/iot/?hl=es>

IBM Cloud. (08/05/2018). "Watson IoT Platform Feature Overview": https://console.bluemix.net/docs/services/IoT/feature_overview.html#feature_overview

SUPERVISIÓN Y CONTROL DE PROCESOS

Enrique E. Tarifa^{1,4}, Sergio L. Martínez¹, Samuel Franco Domínguez¹,
Susana A. Chalabe¹, Luis E. Ituarte¹, Álvaro F. Núñez¹,
Jorgelina F. Argañaraz¹, Adolfo N. Riveros Zapata², Julieta Martínez²,
Juan P. Gutierrez^{2,4}, Lara V. Lescano Farias³, Ubaldo J. M. Aramayo¹

⁽¹⁾ Facultad de Ingeniería / Universidad Nacional de Jujuy / CONICET
Ítalo Palanca N° 10 / S. S. de Jujuy / Provincia de Jujuy

⁽²⁾ Facultad de Ingeniería / Universidad Nacional de Salta
Av. Bolivia 5150 / Salta Capital / Provincia de Salta

⁽³⁾ Facultad de Agronomía y Agroindustrias / Universidad Nacional de Santiago del Estero
Avda. Belgrano (Sud) N° 1912 / Capital / Provincia de Santiago del Estero

⁽⁴⁾ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

eetarifa@fi.unju.edu.ar; smartinez@fi.unju.edu.ar; sfdominguez@yahoo.com.ar;
susana.chalabe@gmail.com; luisituarte@hotmail.com; afnunez@fi.unju.edu.ar;
jfarganaraz@hotmail.com; ariveros@unsa.edu.ar; julemartinez@unsa.edu.ar;
gutierrezjp@unsa.edu.ar; lescanolaravaleria@gmail.com; uba.aramayo@gmail.com

Área temática: Agentes y Sistemas Inteligentes (ASI)

RESUMEN

El actual contexto productivo exige la optimización tanto del diseño como de la operación de las plantas industriales. Basados en ese enfoque, en el proyecto “Supervisión y Control de Procesos”, se desarrollarán, adaptarán y aplicarán herramientas propias de la Operabilidad de Procesos (flexibilidad, controlabilidad, confiabilidad, robustez). Los procesos que tendrán prioridad son los que están implementados en la planta piloto de la Facultad de Ingeniería de la UNSa (extracción líquida-líquida, absorción gas-líquida, producción de vapor, pasteurización, reacción, entre otros), y los procesos vinculados al gas, al petróleo, al litio y a las energías no convencionales —especialmente a la energía solar—. Para alcanzar los objetivos del proyecto, se emplearán técnicas del campo de la Ingeniería de Procesos (simulación, optimización, control, diseño) y de la Inteligencia Artificial (sistemas expertos fuzzy, redes neuronales, minería de datos).

Palabras clave: Operabilidad, Flexibilidad, Controlabilidad, Confiabilidad, Robustez.

CONTEXTO

El grupo responsable de la realización del proyecto “Supervisión y Control de Procesos”, es el grupo de investigación *IngProAr* (Ingeniería de Procesos Argentina), con lugar de trabajo en la Facultad de Ingeniería de la UNJu (Universidad Nacional de Jujuy). El grupo *IngProAr* fue creado en 1995 por el Dr. Enrique Tarifa (investigador de CONICET y de la UNJu), y está constituido por docentes investigadores, profesionales y estudiantes de diversas áreas disciplinares.

El proyecto al que hace referencia este trabajo, es continuación del proyecto anterior denominado “Desarrollo de Herramientas para la Operación de Procesos”. En el presente proyecto, se propone continuar con la línea de investigación emprendida en el proyecto anterior, concentrándose en la supervisión y el control de procesos. Para ello, se desarrollarán, adaptarán y aplicarán herramientas propias de la Operabilidad de Procesos (flexibilidad, controlabilidad, con-

fiabilidad, robustez). Estas herramientas serán aplicadas en los procesos implementados en la planta piloto de la Facultad de Ingeniería de la UNSa. Por otra parte, debido a la importancia que tienen los sistemas de energía no convencionales para Jujuy y para la UNJu, en este proyecto se considerarán también los sistemas de aprovechamiento de energía solar.

Para cumplir con las actividades educativas propuestas en el proyecto, se incorporó al grupo a la Dra. Jorgelina Argañaraz, Profesora en Ciencias de la Educación y Doctora en Ciencias Sociales. En cuanto a la parte experimental del proyecto, estará a cargo de los siguientes investigadores de universidades de otras provincias: Adolfo Néstor Riveros Zapata (UNSa), Julieta Martínez (UNSa), Juan Pablo Gutierrez (UNSa) y Lara Valeria Lescano Farias (UNSE).

El proyecto de referencia es financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales (SeCTER) y la Facultad de Ingeniería de la Universidad Nacional de Jujuy, y fue reconocido por la SeCTER-UNJu con el código D/0164 y por el Programa de Incentivos de la Secretaría de Políticas Universitarias del Ministerio de Educación de la Nación.

1. INTRODUCCIÓN

El proyecto de referencia, propone el desarrollo, adaptación y aplicación de herramientas propias de la Operabilidad de Procesos. La operabilidad de un proceso productivo comprende las siguientes propiedades:

Flexibilidad: Implica la capacidad de un sistema de operar en estado estacionario estable para un rango de condiciones inciertas que pueden encontrarse durante la operación de la planta [1].

Controlabilidad: Puede ser definida como la capacidad de la planta para alcanzar y mantener un estado deseado [2].

Confiabilidad: Hace referencia a la capacidad de todos los elementos que conforman la planta de funcionar continuamente satisfaciendo un conjunto de especificaciones o condiciones [3].

Robustez: Es la capacidad de la planta de hacer mínima la variación de la medida de calidad de los productos ante variaciones de las condiciones de operación [4].

Los actuales estudios de operabilidad involucran cada vez más el uso de técnicas de Inteligencia Artificial (redes neuronales, sistemas expertos con lógica fuzzy, entre otras), siendo los sistemas de control inteligentes los mejores exponentes de esta situación [5].

Para poder desarrollar herramientas destinadas a la operación de un proceso, es necesario contar con un conocimiento acabado del comportamiento estacionario y dinámico que el mismo presenta. Por ello, se desarrollará un modelo estacionario y otro dinámico para cada proceso a estudiar. Para lograr esto, se modelará el proceso escogido utilizando HYSYS —un simulador dinámico de propósito general— o, cuando sea necesario, se desarrollarán modelos orientados a ecuaciones por medio de utilitarios matemáticos apropiados como MATLAB, Simulink, Mathcad, Berkeley Madonna. En los casos que se justifique, se programará utilizando el lenguaje de programación más adecuado: Delphi, Visual Basic, C++.

Las licencias de las aplicaciones citadas tienen un costo considerable; por ese motivo, se hará un esfuerzo por sustituirlas con las opciones libres y de código abierto que brinden funcionalidades equivalentes: Lazarus, Octave, Julia, wxMaxima. La evaluación de aplicaciones libres con código abierto es un tema que está cobrando importancia por sí mismo [6] y será un tema a analizar en el proyecto propuesto.

Una vez que se cuente con el simulador del proceso seleccionado, se procederá a realizar un conjunto de estudios con el fin de evaluar la operabilidad del proceso (flexibilidad, controlabilidad, confiabilidad y robustez). También, se llevarán a cabo estudios de estabilidad y de optimización. Los modelos de optimización serán implementados en LINGO. La información obtenida de los estudios realizados se empleará para determinar las modificaciones que deberán efectuarse en el diseño, en la operación y en la gestión del proceso con el fin de optimizar su desempeño.

Para la supervisión del proceso, se diseñará el sistema de instrumentación y de procesamiento de datos que permita proveer al operador toda la información necesaria para la correcta operación del proceso. El método que se empleará para diseñar el sistema de instrumenta-

ción fue desarrollado por el grupo *IngProAr*. Otro método que se empleará también será la reconciliación de datos.

En cuanto a los sistemas de control, se desarrollarán sistemas de control avanzado empleando técnicas de Inteligencia Artificial (e.g., sistemas expertos fuzzy, redes neuronales, redes neuronales wavelets). Un punto importante a desarrollar serán los controladores MIMO (Multiple-input, Multiple-output) y los controladores predictivos. Para la implementación de técnicas de Inteligencia Artificial, se empleará preferentemente MATLAB con licencia académica y, eventualmente, sus alternativas libres y abiertas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto de referencia se enmarca en las siguientes líneas prioritarias de la UNJu (Res. CS N°168/93):

Desarrollo Económico–Social Regional Sustentable: Las herramientas a desarrollar en el proyecto servirán para aumentar la operabilidad de los procesos productivos de la región. En consecuencia, se favorecerá el desarrollo económico de la región.

También, el proyecto se enmarca principalmente en la siguiente línea prioritaria de la Facultad de Ingeniería (Res. CAFI N°393/13):

Línea 3 - El estudio de procesos específicos o integrados que contribuyan a la cadena de valor de los productos obtenidos: Esta línea comprende los siguientes temas: Análisis, diseño y síntesis óptima de procesos; Simulación y optimización de procesos; Operación y mantenimiento de procesos; Supervisión y control de procesos; Sistemas de apoyo a la toma de decisiones en procesos; Integración de masa y energía; Análisis de la Performance de procesos. Los resultados a obtener con el proyecto propuesto están vinculados a todos los temas citados.

El proyecto en consideración además está vinculado con las siguientes líneas prioritarias de la Facultad de Ingeniería:

Línea 1 - La exploración, manejo, valorización, obtención, transformación y aprovechamiento de recursos naturales renovables, no renovables y energéticos. La Gestión ambiental relacionada:

La vinculación con esta línea se presenta debido a que las herramientas a desarrollar en el presente proyecto pueden ser aplicadas a cualquier tipo de proceso productivo. Por ejemplo, en un proyecto anterior, las herramientas desarrolladas se aplicaron en la gestión óptima de fincas en la cuenca de la Quebrada de Humahuaca. Además, entre los procesos a estudiar, están los que involucran petróleo, gas, energía solar y litio; los cuales son los recursos energéticos más ampliamente utilizados en la actualidad.

Línea 5 - El desarrollo de la Informática, los sistemas y la tecnología derivados de ésta: Algunos de los productos del proceso son aplicaciones informáticas —simuladores, optimizadores, sistemas de control—, para las cuales se deben desarrollar nuevos algoritmos numéricos o técnicas de Inteligencia Artificial.

Línea 6 - El desarrollo de la calidad educativa y el estudio de los procesos de enseñanza y aprendizaje: Algunos productos del proyecto —simuladores, optimizadores, sistemas de control, publicaciones, apuntes— tienen impacto directo sobre la calidad educativa.

3. OBJETIVOS Y RESULTADOS

El proyecto en cuestión, tiene como objetivo general desarrollar herramientas para la operación de procesos. Para ello se diseñarán, desarrollarán, adaptarán y aplicarán herramientas propias de la Operabilidad de Procesos.

Los objetivos particulares del proyecto propuesto son los siguientes: 1) Desarrollar métodos para determinar y aumentar la flexibilidad de procesos; 2) Desarrollar sistemas de supervisión; 3) Desarrollar sistemas de control avanzados; 4) Desarrollar material para la enseñanza de Ingeniería Química (simuladores, optimizadores, sistemas de control, publicaciones y apuntes).

A continuación, se comentan los principales resultados obtenidos en este último año. En el caso de los sistemas de control avanzado, se han realizado modelado y pruebas de sistemas de control inteligentes, a partir de arquitecturas fuzzy, como así también con arquitecturas neuronales. Los modelos bajo prueba se aplicaron a un sistema industrial de mezcla de caudales, donde las simulaciones realizadas mostraron resultados satisfactorios en ambas arquitecturas. La arquitectura

neuronal, cuyo desempeño es estrictamente numérico, fue trabajada con una nueva estrategia de generación de datos de aprendizaje —diferente a la configuración tradicional—, consistente en la implementación de un entorno reducido y dinámico de datos, que simplifica sustancialmente el proceso de entrenamiento. A publicarse en el libro del *VI Simposio Internacional de Investigación (en prensa)*.

Dentro del mismo contexto, también se han desarrollado simulaciones y pruebas con modelos de control más avanzados, en este caso controladores inteligentes híbridos. Dada la amplia variedad de modelos que pueden ser implementados, las primeras pruebas estuvieron orientadas a la hibridización de un modelo tradicional y un modelo inteligente, aplicado a un proceso industrial. Los resultados preliminares fueron prometedores, por lo que se sigue trabajando en esta línea. Resumen publicado en los anales de las *XIV Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA*.

En la publicación mencionada también están contenidos los siguientes trabajos directamente vinculados con los objetivos del proyecto: *Análisis comparativo de entornos numéricos gratuitos para simulación*; *Optimización de un tren contracorriente de unidades mezcladoras-decantadoras*, y *Propagación de errores en simulación estacionaria*.

En 2020, las especiales condiciones sanitarias han orientado a los investigadores de *IngProAr* a la investigación y desarrollo de un equipo especial; se trata de un equipo esterilizador eficiente de aire, diseñado particularmente para la esterilización de aire de un cuarto de aislamiento de un paciente con COVID-19. El equipo consta de un soplador, un economizador y una celda de calentamiento. El aire que ingresa al economizador es precalentado por el aire que egresa. La celda de calentamiento es la responsable de alcanzar la temperatura de esterilización. Para realizar el diseño, se desarrolló un modelo estacionario de simulación y se obtuvo la solución analítica. Los resultados de las simulaciones realizadas muestran que el equipo cumple con su objetivo y mantiene un bajo consumo de electricidad en el punto de diseño. El análisis de flexibilidad muestra que el desempeño del equipo sigue siendo satisfactorio cuando es sometido a condiciones diferentes a

las del punto de diseño. Este trabajo será publicado en el N°6 la revista *CODINOA (en prensa)*.

Los aspectos académico-docentes tampoco han sido descuidados durante el tiempo que lleva de ejecución este proyecto, como se ha contemplado en uno de sus objetivos secundarios.

Considerando que la emergencia sanitaria obligó a las universidades a adaptarse rápidamente al dictado remoto de emergencia. En un trabajo realizado sobre este tema, se presentó la experiencia de la cátedra *Simulación y Optimización*, del último año de la carrera Ingeniería Química que se dicta en la Facultad de Ingeniería de la UNJu. En ese trabajo, se analizaron todas las decisiones tomadas para implementar el dictado virtual de la materia, como así también el impacto que tuvieron esas decisiones en los resultados del dictado virtual. Este trabajo será publicado en el N°6 la revista *CODINOA*.

En otro trabajo, se analizó la incorporación del aula virtual al modelo pedagógico de la citada materia. Esta incorporación demandó un gran esfuerzo de la cátedra, pero los resultados de la experiencia fueron destacables. El análisis y descripción completa del proceso se puede conocer en libro *La Educación en Prospectiva. Prácticas Disruptivas mediadas por Tecnologías*, editado por RUEDA (Red Universitaria de Educación a Distancia Argentina) y la UNJu en 2020.

Finalmente, se desarrolló un sistema de asignación óptima de aulas para la toma de parciales para la asignatura *Introducción a la Informática* (primer año de las carreras de informática de la Facultad de Ingeniería de la UNJu). Esta materia cuenta con una matrícula promedio superior a 900 estudiantes, por lo que la organización de las evaluaciones es compleja. A fin de hacer más eficiente la organización de los exámenes parciales, se planteó un modelo matemático de optimización para minimizar la cantidad de aulas requeridas durante la evaluación, con lo que se logró reducir el tiempo de preparación y control preliminar de una hora a un poco más de cinco minutos. El planteo del modelo y la simulación del proceso se puede conocer en libro *La Educación en Prospectiva. Prácticas Disruptivas mediadas por Tecnologías*, editado por RUEDA y la UNJu en 2020.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo que forma el grupo de investigación *IngProAr* cuenta con profesores, auxiliares, graduados y pasantes de la Facultad de Ingeniería, quienes desarrollan diversas actividades en la formación de recursos humanos, durante la ejecución del proyecto. Particularmente, los docentes investigadores del grupo *IngProAr* están a cargo de diversas cátedras en distintas carreras de la Facultad de Ingeniería de la UNJu: *Simulación y Optimización de Procesos, Ingeniería de Procesos, Electricidad y Electrónica e Introducción a la Informática* —en Ingeniería Química—; *Métodos de Simulación, Inteligencia Artificial, Técnicas y Estructuras Digitales, Metodología de la Programación* —en Ingeniería Informática—; *Electrotecnia* en Ingeniería Industrial; *Geología Ambiental y Cartografía especial* en Ingeniería de Minas. Es de destacar que los contenidos de dichas materias están directamente vinculados con los temas considerados en este proyecto. Como actividades de formación de recursos humanos, se pueden destacar:

Tesis de posgrado

- Tarifa E. E., codirección de la tesis doctoral *Secado en lecho de chorro bidimensional para la deshidratación de proteínas del plasma y porción globina de sangre bovina*, Ing. Lara Valeria Lescano Farías, Directora: Dra. Eve Liz Coronel. Res. F.A.A. N°641/2013, UNSE.
- Tarifa E. E., dirección de la tesis doctoral *Control Inteligente con Algoritmos Híbridos Optimizados aplicados a Modelos de Procesos Productivos*, Ing. Sergio L. Martínez, Codirector: Dr. Juan P. Gruer (UNT), Facultad de Ciencias Exactas y Tecnología, UNT, Res. FACET 620/2017, desde may/2017.
- Tarifa E. E., dirección de la tesis doctoral *Diseño óptimo de un lecho fluidizado cónico para la inactivación y deshidratación del grano entero de soja*, Ing. Gustavo Salcedo, UNSE, beca doctoral de CONICET, desde abr/2020.

- Martínez S. L., dirección del trabajo final de especialización *Modelo de espacio de estados*, Dr. Enrique E. Tarifa, Facultad de Humanidades y Ciencias Sociales, UNJu, desde 2020.

Tesis de grado

- Martínez S. L., Dirección de la tesis de grado *Desarrollo de una Herramienta con Interfaz Gráfica para Implementación de Redes Neuronales Feedforward*. Tesista: David Llusco, codirector Ing. Jorge J. Gutiérrez, Carrera de Ingeniería Informática, UCSE-DASS, San Salvador de Jujuy. En ejecución.

5. BIBLIOGRAFÍA

1. Reay D., Ramshaw C.; Harvey A., *Process Intensification: Engineering for efficiency, sustainability and flexibility*. Butterworth-Heinemann, 2013.
2. Ramos W. B., Figueirêdo M. F., Brito K. D., Cianella S., Vasconcelos L. G. S., Brito R. P., Effect of Solvent Content and Heat Integration on the Controllability of Extractive Distillation Process for Anhydrous Ethanol Production, *Ind. Eng. Chem. Res.*, 55 (43), pp 11315–11328, 2016.
3. Roy A., Srivastava P., Sinha S., Risk and reliability assessment in chemical process industries using Bayesian methods, *Reviews in Chemical Engineering*, 40(5), 479-499, 2014.
4. Kharrazi A., Kumar P., Saraswat C., Avtar R., Mishra B. K., Adapting Water Resources Planning to a Changing Climate: Towards a Shift from Option Robustness to Process Robustness for Stakeholder Involvement and Social Learnings, *Journal of Climate Change*, vol. 3, no. 2, pp. 81-94, 2017.
5. Santos M., Aplicaciones Exitosas de Control Inteligente a Casos Reales, *Revista Iberoamericana de Automática e Informática Industrial*, V0-0, pp. 1–8. 2011.
6. Sharma N., Gobbert M. K., A comparative evaluation of Matlab, Octave, FreeMat, and Scilab for research and teaching. Technical Report HPCF–2010–7, *UMBC High Perform-*

mance Computing Facility, University of Maryland, Baltimore County, 2010.

Tecnologías 4.0 para la Industria Argentina - Análisis y herramientas de I&D+i para la formulación de políticas y el desarrollo de la Industria 4.0 – Proyecto PICT

**María Belén Odena¹, Ulises Girolimo¹, Patricio Feldman¹,
Susana Finquelievich¹, Pablo Fillottrani², Elsa Estévez², Sebastián Cortez Oviedo**

¹**Instituto de Investigación Gino Germani (IIGG-UBA)
Facultad de Ciencias Sociales – Universidad de Buenos Aires
Pres. José Evaristo Urriburu 950 - Ciudad Autónoma de Buenos Aires
(belen.odena, ugirolimo, patafeldman, sfinquel, pcortezoviedo) @gmail.com**

²**Laboratorio de Ingeniería de Software y Sistemas de Información (LISSI)
Departamento de Ciencias e Ingeniería de la Computación – Universidad Nacional del Sur
Av. San Andrés 800 – Campus de Palihue - Bahía Blanca, Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 - La Plata, Buenos Aires
(ece, prf.) @cs.uns.edu.ar**

RESUMEN

El proyecto “Tecnologías 4.0 para la Industria Argentina - Análisis y herramientas de I&D+i para la formulación de políticas y el desarrollo de la Industria 4.0” tiene como objetivo construir el estado del arte de la investigación de origen público y privado en C&T+i vinculada con la Industria 4.0 en Argentina, en el periodo 2013-2020, y su vinculación con las Pequeñas y Medianas Empresas (PyMEs) industriales. Asimismo se propone desarrollar instrumentos que permitan la formulación y optimización de las políticas públicas dirigidas tanto a la investigación científico-tecnológica en biotecnología, robótica, Inteligencia Artificial (IA), Internet de las Cosas (IOT) y nanotecnología, como a la aplicación en el sector productivo del conocimiento generado.

Palabras Clave

Industria 4.0 - Tecnología 4.0 - Políticas públicas - Impactos sociales - Desarrollo sociotecnológico

CONTEXTO

La línea de investigación y desarrollo presentada forma parte de las tareas planificadas en el proyecto PICT “Tecnologías 4.0 para la Industria Argentina

- Análisis y herramientas de I&D+i para la formulación de políticas y el desarrollo de la Industria 4.0”. El Proyecto, financiado por la Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación, a través del Fondo para la Investigación Científica y Tecnológica (FONCyT), cuenta con la dirección de la Dra. Elsa Estévez y la Dra. Susana Finquelievich.

Este proyecto formaliza las relaciones científicas entre dos equipos: el Laboratorio de Investigación & Desarrollo en Ingeniería de Software y Sistemas de Información (LISSI) junto con el Instituto UNS-CONICET de Ciencias e Ingeniería de la Computación (ICIC) en la Universidad Nacional del Sur (UNS); y el Programa de Investigaciones sobre la Sociedad de la Información, del Instituto de Investigaciones Gino Germani, de la Facultad de Ciencias Sociales, Universidad de Buenos Aires, a través del equipo I-Polis.

La formulación y el desarrollo del proyecto se basan en la adopción de una perspectiva transdisciplinaria (uniendo las Ciencias e Ingeniería de la Computación con las Ciencias Sociales), así como en un enfoque interinstitucional de los equipos de LISSI-ICIC - UNS e IIGG - UBA.

El Proyecto se identifica con el número 2018-03627

1. INTRODUCCIÓN

La Industria 4.0 constituye una nueva oportunidad para Argentina de avanzar en un camino de desarrollo socio-tecnológico-económico, a partir de la optimización de los procesos de producción y distribución, la generación de nuevos mercados y el desarrollo de capacidades para producir bienes diferenciados, intensivos en conocimiento.

El Banco Interamericano de Desarrollo - BID (2018) analiza los posibles impactos sociales del desarrollo de la Industria 4.0, especialmente en el ámbito de la salud, la educación, el empleo y las brechas de género. El BID afirma que la Cuarta Revolución Industrial y la transformación digital “no sólo contribuyen a la productividad y a la eficiencia, sino también al desarrollo socioeconómico más amplio. Pueden dar lugar a una sociedad más inclusiva y diversa así como a mejores acuerdos de gobernanza; mejorar el acceso a servicios claves tales como la salud, la educación y los servicios bancarios; mejorar la calidad y la cobertura de los servicios públicos y la participación política y ampliar la manera en la que las personas colaboran y crean contenidos” (Pombo *et al*, 2018: 71). La OCDE (2017) plantea que políticas sólidas de ciencia e I+D son especialmente relevantes en esta área para este período.

En el contexto actual, la mayoría de las empresas y países son usuarios de tecnologías 4.0, pero algunos de ellos van más allá: son desarrolladores. Argentina posee el potencial para jugar un rol importante en el grupo de los desarrolladores, si bien existen diversos desafíos a resolver. La mayoría de los desafíos en investigación, críticos para la Industria 4.0, son multidisciplinares. Las políticas de ciencia, tecnología e innovación (C&T+i) necesitan intensificar las investigaciones multidisciplinares, la ampliación e incremento de sus investigaciones aplicadas, y los vínculos entre los diversos sectores:

estatal, empresario, científico-tecnológico y ciudadano. La Industria 4.0 precisa de políticas públicas 4.0, en las que se alienten las innovaciones y se regulen las nuevas realidades sin inhibir procesos de innovación productiva. Esta necesaria transformación incluiría a las instancias estatales, al sector empresarial, a la academia y la sociedad civil.

El concepto de Industria 4.0 ha sido acuñado en Alemania, desde la feria Hannover Messe en el 2011 (Pfeifer, 2017). En el año 2014 se impulsó la iniciativa “Science Year 2014 - The Digital Society” con la finalidad de visualizar los cambios impulsados por la incorporación de tecnologías 4.0, que pueden enfrentar la sociedad y la industria para el año 2020. Los estudiosos sobre este nuevo paradigma industrial plantean que se basa en producción individualizada, integración horizontal en redes colaborativas e integración digital en la cadena de producción (Brettel *et al.*, 2014; Kagermann *et al.*, 2013). De todas maneras, los avances tecnológicos pueden no ser el elemento fundamental de cambio disruptivo por sí mismos (Drath y Horch, 2014), sino por su impacto en la concepción, producción y distribución del producto, y especialmente en la manera en que las empresas crean, distribuyen y se apropian valor. También son relevantes las transformaciones en las relaciones interinstitucionales, en la organización del trabajo, y finalmente, en la sociedad (Klingenberg, 2017).

La Industria 4.0 usa la electrónica, la robótica, y la IA para automatizar la producción. Esta etapa se caracteriza por la convergencia y fusión de diversas áreas de CyT. En ella se desvanecen las fronteras entre la informática, la biología, la genética y la nanotecnología. Desde un enfoque económico, el concepto corresponde a una nueva forma de organizar los medios de producción, a través de la generación y desarrollo de un alto número de “fábricas inteligentes” más adaptables a las necesidades y a los procesos de producción, así como a una asignación y gestión más eficientes de los recursos. El término enfatiza

la idea de una creciente y adecuada digitalización y coordinación cooperativa en todas las unidades productivas de la economía, conduciendo a nuevos modelos económicos (Challenges, 2013; Rinn y Kube, 2014). Como señala Schroeder (2016), trata de establecer una red de interconexiones entre todos los elementos del proceso de creación de valor: desde materias primas y productos de escaso valor agregado, pasando por el proceso de producción, las redes de clientes y los procesos logísticos y de prestación de servicios; todo anclado en la conversión digital de los datos.

Algunas de las tecnologías que integran la Industria 4.0 son: 1) Internet de las Cosas (Internet of things, IoT); 2) Sistemas Ciber-Físicos (SCF); 3) Computación en la Nube – capacidad de procesamiento y acceso a recursos digitales compartidos a través de Internet y de forma ubicua, 4) Fabricación Aditiva e Impresión 3D; 5) Robótica Colaborativa - robots industriales colaborando con trabajadores humanos; 6) Ciberseguridad – sistemas y herramientas para detectar, prevenir y neutralizar amenazas contra los sistemas de información de las industrias, el Estado y los ciudadanos; 7) Nuevos Materiales - materiales emergentes fuertes y livianos que recuerdan su forma, aleaciones exóticas, se auto-reparan o se integran en componentes y materiales que responden a la luz y al sonido; 8) Biotecnología, Bioinformática y Medicina Personalizada – La biotecnología incluye investigación y desarrollo de sustancias bioactivas y alimentos funcionales para bienestar de organismos acuáticos, diagnóstico celular y molecular, y manejo de enfermedades asociadas a la acuicultura, toxicología y genómica ambiental, manejo ambiental y bioseguridad, biocombustibles, y gestión y control de calidad en laboratorios. La bioinformática es la aplicación de tecnologías computacionales y la estadística a la gestión y análisis de datos biológicos. La industria de la salud vislumbra que la digitalización de sus procesos posibilitará la medicina personalizada. Surgen herramientas que facilitan que el tratamiento adecuado (diferente para cada quien) llegue a la

persona indicada en el momento en que lo necesita (Ramírez Coronel, 2018); y 9) Nanotecnología – manipulación de la materia para producir productos a micro-escala. Es un campo muy amplio que incluye disciplinas como química orgánica, biología molecular, microfabricación, etc.

Feldman y Girolimo (2021) analizan las políticas públicas de promoción de la Industria 4.0 en América Latina, tomando los casos de México, Argentina y Brasil. Si bien reconocen que los avances hasta el momento son modestos, los gobiernos han dado los primeros pasos en el diagnóstico y formulación de planes para su promoción. En términos generales, se reconoce una situación de déficit en la incorporación de tecnologías 4.0 en los sistemas productivos, escasez de capital humano especializado en algunos campos como el Big Data y la IoT, y un fuerte desequilibrio territorial en cuanto a capacidades tecnológicas. El caso de México, muestra una orientación hacia la conformación de redes multiactorales mediante el estímulo a clústeres tecnológicos en distintas regiones del país, y la formación de capital humano en áreas prioritarias. En Brasil, en 2019 se conformó una alianza público-privada federal para dinamizar la agenda de transformación digital de la industria, y se creó la Cámara Brasileña para la Industria 4.0, que agrupa a diferentes actores. En Argentina, se observa la existencia del Plan Industria Argentina 4.0, lanzado en 2019, que registra avances todavía preliminares; y se cuenta con la Ley de Economía del Conocimiento, que otorga exenciones impositivas a empresas productoras de tecnologías 4.0.

Una breve investigación sobre la producción de información y conocimiento en Argentina y América Latina sobre la Cuarta Revolución Industrial y/o la Industria 4.0, ha dado como resultado la existencia de varios estudios de organismos internacionales (BID, CEPAL, OECD, etc.), una profusión de artículos periodísticos, algunos estudios de consultoras generalmente extranjeras, y muy escasos trabajos científicos. Este proyecto de investigación plantea entre otros objetivos

contribuir a llenar ese vacío. El aporte fundamental que se propone es generar conocimientos sobre el estado actual de las investigaciones académicas y empresariales sobre la Industria 4.0 y desarrollar una caja de herramientas que permitan contribuir a la formulación de nuevas políticas públicas en este campo para favorecer su desarrollo.

El resto de este artículo se organiza de la siguiente manera: la sección 2 explica los ejes que guían el desarrollo del proyecto, la sección 3 presenta los resultados esperados/obtenidos y, por último, la sección 4 discute la formación de recursos humanos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los objetivos del proyecto son:

- Caracterizar las investigaciones realizadas durante los últimos 6 años en centros de investigación públicos y privados (universidades y centros de I+D en empresas privadas) sobre temas vinculados con la Industria 4.0 y dar cuenta del grado de desarrollo, perfil y orientación de las mismas.
- Caracterizar el perfil productivo de las PyMEs vinculadas al sector de la Industria 4.0 en base a los siguientes criterios: a) patrón de especialización; b) desarrollo de nuevos productos y mercados; c) capacidad exportadora; d) composición y tipo de participación del sector privado interviniente (empresas internacionales, nacionales, PyMEs, etc.); e) composición del mercado de trabajo, con énfasis en un enfoque de género; y f) inversión en I+D pública y privada, entre otros; y reconocer las oportunidades y sus limitaciones para vincularse al sistema científico-tecnológico.
- Identificar y evaluar las capacidades de los actores públicos y privados relacionados a la Industria 4.0 y las interacciones entre el sistema de Ciencia y Tecnología (C&T) y el sistema productivo, e identificar los impactos en materia económica y social que dichas capacidades generan.
- Identificar y verificar el estado de actualización de las políticas públicas que

rigen o estimulan la investigación científico-tecnológica en tecnologías 4.0 en el país, así como las dirigidas a incentivar al sector productivo; y desarrollar los instrumentos necesarios para que, en base a los conocimientos producidos y por medio de un estudio prospectivo, se pueda facilitar la formulación y optimización de las mismas.

3. RESULTADOS ESPERADOS/OBTENIDOS

Los resultados obtenidos durante el primer semestre son:

- Se ha desarrollado un relevamiento sobre el estado del arte de la Industria 4.0 tanto a nivel nacional como internacional y de las políticas públicas vinculadas con la temática. A partir de ello, se analizaron las trayectorias de los países líderes en el desarrollo e incorporación de tecnologías 4.0 (Alemania, Estados Unidos y China), y las trayectorias de los países latinoamericanos que han comenzado a avanzar en este campo (Argentina, México y Brasil).
- Se han publicado dos artículos en revistas científicas sobre el lugar que ocupan las tecnologías 4.0 en la agenda de las ciudades argentinas, en el contexto de la segunda oleada informacional, y sobre los principales desafíos para el desarrollo y la incorporación de tecnologías 4.0 en Argentina.
- Miembros del equipo han participado en eventos científicos vinculados al tema de investigación.
- Se delinearón las pautas para dar comienzo al trabajo de campo propuesto en el proyecto.

Como trabajo a futuro, los planes incluyen:

- Conceptualización de capacidades para la Industria 4.0: se creará un marco conceptual que permita la tipificación de los diferentes actores vinculados con la Industria 4.0 en función de las capacidades tecnológicas y organizacionales que poseen actualmente. A continuación, éstas serán contrastadas con las capacidades requeridas

para el desarrollo de las tecnologías que forman parte de este nuevo paradigma. Mediante la identificación de la brecha entre capacidades actuales y capacidades requeridas, se diseñarán propuestas de política pública que contribuyan a achicar dicha distancia.

- Construcción de indicadores para la identificación y medición de impactos de la Industria 4.0 en el campo social, económico, ambiental e industrial, entre otros, en el país.

- Diseño de instrumentos para la formulación de políticas públicas. Los instrumentos desarrollados se empaquetarán junto con los procedimientos correspondientes para su utilización de manera que constituyan herramientas de apoyo para el monitoreo continuo y permanente actualización de políticas públicas relacionadas a la Industria 4.0.

- Síntesis de recomendaciones para la formulación de políticas públicas que promuevan el desarrollo de capacidades de los diferentes actores vinculados a la Industria 4.0, que faciliten las interacciones entre ellos, el desarrollo tecnológico y la adopción de tecnologías 4.0 en distintos sectores productivos, con especial atención a la situación de las PyMEs.

- Diseminación de resultados a través de publicaciones científicas en revistas nacionales e internacionales, participación en eventos científico-tecnológicos y otras reuniones, para presentar resultados y difundir las actividades del proyecto. Asimismo, se buscará la difusión al público, a través del sitio web del proyecto, que se desarrollará para cumplir tal propósito.

4. FORMACION DE RECURSOS HUMANOS

- Capacitación de los miembros del proyecto en tareas de investigación asociadas al universo de las tecnologías 4.0.

- Desarrollo de una tesis doctoral y actividades de becarios posdoctorales en el área.

- Participación de los integrantes de esta línea de investigación en el dictado de

asignaturas y cursos de posgrado en la Facultad de Ciencias Sociales, Pontificia Universidad Católica Argentina; Facultad de Ciencias Económicas, Universidad del Salvador; en la Universidad Pontificia Bolivariana y en el Departamento de Ciencias e Ingeniería de la Computación de la UNS.

5. BIBLIOGRAFÍA

Basco, A.I., G. Béliz, D. Coatz, P. Garneró (2018): Industria 4.0. Fabricando el futuro, UIA - BID - INTAL

Brettel, M., Friederichsen, N., Keller, M., y Rosenberg, M. (2014). How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective. *International Journal of Mechanical, Industrial Science and Engineering*, 8(1), 37-44.

Drath, R., y Horch, A. (2014). Industrie 4.0: Hit or hype? [industry forum]. *IEEE industrial electronics magazine*, 8(2), 56-58.

Feldman, P. y Girolimo, U. (2021 - en prensa) La Industria 4.0 en perspectiva latinoamericana: limitaciones, oportunidades, y desafíos para su desarrollo.

Klingenberg, C. (2017): Industry 4.0: what makes it a revolution?, Conference paper.

OECD (2017): The next production revolution: a report for the G20, Paris.

OECD (2016): Enabling the next production revolution: the future of manufacturing and services - interim report, Meeting of the OECD Council at Ministerial Level Paris, 1-2 June 2016,

Pombo, C, Gupta, R, y Stankovic, M. (2018). Servicios sociales para ciudadanos digitales: oportunidades para América Latina y el Caribe. Publicación del BID.

Ramírez Coronel, M. (2018): La industria 4.0 y la medicina personalizada, en: *El Economista*, México DF.

Schroeder, W. (2016). La estrategia alemana Industria 4.0: el capitalismo renano en la era de la digitalización. Friedrich-Ebert-Stiftung, Madrid.

Schwab, K. (2016): La Cuarta Revolución Industrial, Ed. Debate, Buenos Aires.

Tecnologías de sistemas de QA aplicadas a la Web Semántica.

Matias Oyarzun and Sandra Roger

email: matias.oyarzun@est.fi.uncoma.edu.ar
roger@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

Los sistemas de búsqueda de respuestas (Question Answering - QA) afrontan el problema de localizar, extraer y presentar al usuario única y exclusivamente aquella información que desea conocer, evitando así la ardua tarea de recopilación de información a través de la lectura de documentos relevantes. En sus inicios, en un sistema de QA tradicional se realizaba una consulta en lenguaje natural y se trataba de localizar la respuesta en bases de conocimientos disponibles en textos de manera no estructurada.

A causa del rápido crecimiento de la información en la Web, cada vez se hace más difícil encontrar y gestionar las respuestas concisas a diversas preguntas planteadas. Con la evolución de la Web Semántica, una gran cantidad de datos estructurados están disponibles en forma de bases de conocimiento (KB) para ser utilizados, lo que permite minimizar la posibilidad de existencia de ambigüedades, facilitando así el trabajo necesario para el desarrollo de aplicaciones que hagan uso de los datos.

El propósito principal de un sistema de QA de calidad es recuperar la información deseada de una o varias KB, utilizando preguntas en lenguaje natural. En su mayoría, los enfoques actuales no abordan el desafío de la multilingüedad o la independencia de KB tanto para pre-

guntas completas como para preguntas de palabras clave [1]. Por ello, el objetivo principal de este plan es la investigación y desarrollo de soluciones basadas en tecnologías de sistemas de QA que permitan reducir la búsqueda de información para extraer las respuestas, sobre tecnologías de la Web Semántica a través de herramientas de Lenguaje Natural, lo que contribuye al desarrollo de agentes inteligentes inmersos en la Web.

Palabras Clave: Sistemas de Búsqueda de Respuestas, Question Answering, Generalización de Texto, Web Semántica, Procesamiento de Lenguaje Natural.

Contexto

Este trabajo está parcialmente financiado por la UNCo, en el marco del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014). Como así también, lo financia parcialmente el Consejo Interuniversitario Nacional (CIN) con una Beca de Estímulo a las Vocaciones Científicas 2021. El proyecto de investigación tiene una duración de cuatro años y ha comenzado en 2017, con una extensión para el año 2021. El proyecto se desarrolla en forma colaborativa con docentes-investigadores de la UNS.

1. Introducción

Actualmente, es esperable que las aplicaciones accedan a grandes cantidades de datos que se modifican en tiempo real, de manera constante y que se presentan en diferentes formatos, especialmente disponible a través de la Web. Muchas de estas aplicaciones ocurren bajo restricciones de tiempo críticas y en intensa interacción con el usuario. En este contexto, la representación conceptual de los dominios para la generación y extracción de información y conocimiento, es central en la toma de decisiones. Bajo este aspecto, el objetivo general que persigue el proyecto de investigación es el de generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica. Para ello, es necesario profundizar el estudio de técnicas de representación de conocimiento y razonamiento, tecnologías del lenguaje natural, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, tanto a nivel de procesos como de datos.

Asimismo, la Web es un ambiente ideal para estos agentes que, junto con la generación y representación de conocimiento, las ontologías y el razonamiento automático, son fundamentales en la evolución de dicha Web, denominada Web Semántica. La meta de la Web Semántica es crear una Web de conocimiento en la cual la semántica del contenido es explicitada, permitiendo novedosas aplicaciones que combinan datos de sitios heterogéneos para, entre otros objetivos, mejorar la experiencia de los usuarios de acuerdo a sus necesidades. Esta red de funcionalidades basadas en conocimiento hará a dicho conocimiento procesable por una computadora, para soportar conductas inteligentes por parte de dichas máquinas.

Hacer que este valioso conocimiento semántico sea accesible y utilizables para los usuarios finales es uno de los principales objetivos de los sistemas de QA sobre KB. La mayoría de los sistemas de QA actuales consultan una KB en un idioma. Los enfoques existentes no están diseñados para adaptarse

fácilmente a nuevas KB e idiomas.

El desarrollo del plan de trabajo se realizará en el marco del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014). En dicho proyecto de investigación se desarrolla una línea de investigación que explora sobre temas afines tanto al análisis y desarrollo de técnicas y herramientas útiles para la aplicación del Procesamiento en Lenguaje Natural con el fin de dar soporte al modelado conceptual basado en ontologías; como también al estudio y aplicación de técnicas de Generación del Lenguaje Humano aplicado a tecnologías de la Web Semántica. Particularmente, se ha escogido experimentar sobre herramientas de QA que den soporte a la búsqueda de información en el ámbito de la Web Semántica.

2. Línea de Investigación y Desarrollo

El proyecto de investigación *Agentes Inteligentes y Web Semántica* tiene como objetivo general, generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir la Web Semántica.

Para ello, en esta línea de investigación, se procederá a diseñar un sistema de QA semántico modular. Además, se prevé la implementación de algunos de estos módulos que conformarán un primer prototipo del sistema de QA desarrollado dentro del marco de nuestro proyecto de investigación anteriormente mencionado. Se pretende relevar metodologías de sistemas de QA basadas en la web semántica, y las condiciones de su aplicación propiciando un análisis y comparativa de tales herramientas.

Los sistemas de recuperación de información (*Information Retrieval* - IR), son aquellos encargados de la organización, almacenamiento, recuperación y evaluación de la información de documentos principalmente en formato de texto en lenguaje natural. Éstos buscan encontrar documentos, dados en una colección, que tratan acerca de un mismo tema o que satisfacen una

consulta generada por un usuario. Por lo general, suelen utilizar esta última para clasificar los documentos de la colección, retornando al usuario un subconjunto que satisfacen un criterio de clasificación. Se encuentran con mayor frecuencia en los motores de búsqueda, como Google, Yahoo, entre otros. El resultado con estos motores de búsquedas es una lista de enlaces, por lo que el usuario debería explorar cada enlace de la lista para obtener la respuesta deseada. Así, un sistema de búsqueda de respuestas no es una tarea simple de IR. De modo que, el objetivo de un sistema QA consiste en identificar la respuesta concreta a preguntas arbitrarias o necesidades precisas de información formuladas por los usuarios. Para éstos, es interesante encontrar información precisa, gracias a la explosión de la información en esta era de las tecnologías de la información y la comunicación, y debido a que suelen haber situaciones en las que el usuario final necesita conocer un dato muy específico y no dispone de tiempo o no puede leer toda la documentación referente al tema de la búsqueda para solucionar su problema. Por lo tanto, los sistemas QA son capaces de responder a cuestiones formuladas por los usuarios en lenguaje natural, realizando búsquedas en bases de datos no estructuras generalmente.

Gracias al desarrollo de la Web Semántica, una gran cantidad de datos estructurados se hicieron disponible en la web en forma de bases de conocimiento. Poner a disposición estos datos valiosos de manera accesible y utilizables para los usuarios finales es el principal objetivo de los nuevos sistemas de QA sobre KBs. Los enfoques existentes de QA no están diseñados para ser fácilmente adaptables a las nuevas KBs y nuevos idiomas.

Además de las técnicas de adquisición de datos, los datos (información) también son la parte más importante del proceso de búsqueda de información. A veces, el usuario usa palabras claves, en las cuales algunas de ellas ni siquiera están en el conjunto de datos disponibles, por lo que los resultados de la búsqueda no cumplen con los deseos del usuario. Por ello, se necesita tecnología de web semántica para superar esto. La Web Semántica proporciona la tecnología y los estándares necesarios para agregar un signi-

ficado que pueda ser entendido por la máquina en la web en general, para que la computadora pueda hacer un proceso de inferencia contra sus datos. Esto tiene algunas implicaciones, como ampliar el cuerpo de conocimientos de la informática, especialmente los campos de la web semántica y la ontología.

Los nuevos sistemas de búsquedas de respuestas semánticas (SQA) son definidos por los usuarios que preguntan en lenguaje natural usando sus propias terminologías en la que reciben una respuesta concisa generada al consultar una base de conocimiento RDF. De esta manera los usuarios quedan así liberados de dos requisitos principales de acceso a la Web Semántica: (1) el dominio de un lenguaje de consulta formal como SPARQL y (2) el conocimiento sobre los vocabularios específicos de la base o bases de conocimientos que desean consultar [1]. Dado que el lenguaje natural es complejo y ambiguo, los sistemas SQA confiables requieren de diferentes pasos y recursos, y aunque en algunos de ellos existen soluciones maduras, otras aun presentan grandes desafíos.

Linked Data creado por Berners-Lee puede ser categorizada como dato estructurado [2]. La Web Semántica a su vez, incluye más datos nuevos y también crea enlaces para establecer una conexión entre los datos de una fuente y los datos de otra fuente para que los usuarios finales y los sistemas de QA encuentren fácilmente los datos requeridos.

Según el origen de los datos, en términos generales, los datos se pueden clasificar en dos grupos [3]: datos centralizados y datos distribuidos. La ventaja de los datos centralizados es que acelera el proceso de recuperación de datos. Además, en los datos privados, los datos originales solo serán propiedad del propietario o proveedor de almacenamiento de datos. Mientras que, en los datos públicos, especialmente aquellos con fines promocionales, el propietario desea que los datos estén lo más difundidos posible. Por ello, para los datos públicos, es preferible un modelo de datos distribuido, de modo que, motores de búsqueda y sistemas de control de calidad, puedan utilizar los datos según su formato original. Esto conlleva una desventaja

sobre los datos distribuidos, ya que la recuperación de datos en aquí lleva más tiempo porque el motor realiza consultas remotas.

Han habido muchos trabajos relacionados que discuten las arquitecturas de los sistemas de QA. Arquitectura de QA para predecir la mejor respuesta [4], sistemas de QA centrado en el reconocimiento de imágenes [5, 6], sistemas de QA con *Deep Learning* [7], sistemas de QA [8] y YodaQA [9, 10], entre otros. Cada sistema de QA mencionado anteriormente tiene una arquitectura única por lo que es difícil compararlos entre sí. Pero los sistemas de QA de [11, 3] tienen las características más similares con nuestro objetivo. Las diferencias son los métodos utilizados y la fuente de datos. La fuente de datos de [3] provino de varias fuentes dispersas en Internet, mientras que [11] utilizó datos no estructurados, como párrafos en inglés.

El desarrollo del plan de trabajo se realizará en el marco del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014). En dicho proyecto de investigación se desarrolla una línea de investigación que explora sobre temas afines tanto al análisis y desarrollo de técnicas y herramientas útiles para la aplicación del Procesamiento en Lenguaje Natural con el fin de dar soporte al modelado conceptual basado en ontologías; como también al estudio y aplicación de técnicas de Generación del Lenguaje Humano aplicado a tecnologías de la Web Semántica. Particularmente, se ha escogido experimentar sobre herramientas de QA que den soporte a la búsqueda de información en el ámbito de la WS.

3. Resultados Obtenidos y Trabajos Futuros

Inicialmente, se está realizando un relevamiento de las diferentes estrategias de los sistemas de búsquedas de respuestas semánticos, herramientas de procesamiento de lenguaje natural y lenguajes de modelado de ontologías existentes, para poder crear un marco comparativo y poder evaluarlos.

Una vez planteadas las ventajas y desventa-

jas de cada estrategia, se pretende plantear la estrategia de diseño del modelo de búsqueda de respuestas a seguir, para luego dar comienzo a la implementación de un primer prototipo dentro del marco del proyecto de investigación, y su posterior validación correspondiente.

La finalidad de la proyecto es la implementación de un sistema de búsqueda de respuesta aplicado sobre la web semántica, que nos permita recuperar fragmentos de textos que contengan la respuesta precisa a una pregunta planteada en lenguaje natural, en lugar de una lista de enlaces a documentos como lo hacen los motores de búsquedas tradicionales. Además, busca ampliar el conocimiento en el área de Ciencias de la Computación, y por sobre todo en los campos de la Web Semántica y la Ontología.

4. Formación de Recursos Humanos

Durante la realización de esta investigación se espera lograr, como mínimo, la culminación de 2 tesis de grado dirigidas y/o codirigidas por los integrantes del proyecto. Uno de los autores de este trabajo posee una Beca de Estímulo a las Vocaciones Científicas (CIN) 2021.

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

Referencias

- [1] Dennis Diefenbach, Andreas Both, Kamal Singh, and Pierre Maret. Towards a question answering system over the semantic web. *Semantic Web*, (Preprint):1–19, 2020.
- [2] GM Rasiqul Islam Rasiq, Abdullah Al Sefat, Tanjila Hossain, Md Israt-E-Hasan Munna, Jubayeath Jahan Jisha, and Mohammad Moinul Hoque. Question answering system over linked data: A detailed survey. *ABC Research Alert*, 8(1):32–47, 2020.

- [3] M Syarief, W Agustiono, A Muntasa, and M Yusuf. A conceptual model of indonesian question answering system based on semantic web. In *Journal of Physics: Conference Series*, volume 1569, page 022089. IOP Publishing, 2020.
- [4] Dalia Elalfy, Walaa Gad, and Rasha Ismail. A hybrid model to predict best answers in question answering communities. *Egyptian informatics journal*, 19(1):21–31, 2018.
- [5] Caiming Xiong, Stephen Merity, and Richard Socher. Dynamic memory networks for visual and textual question answering. In *International conference on machine learning*, pages 2397–2406, 2016.
- [6] Zichao Yang, Xiaodong He, Jianfeng Gao, Li Deng, and Alex Smola. Stacked attention networks for image question answering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 21–29, 2016.
- [7] Yashvardhan Sharma and Sahil Gupta. Deep learning approaches for question answering system. *Procedia computer science*, 132:785–794, 2018.
- [8] Mourad Sarrouti and Said Ouatik El Alaoui. A passage retrieval method based on probabilistic information retrieval model and umls concepts in biomedical question answering. *Journal of biomedical informatics*, 68:96–103, 2017.
- [9] Petr Baudiš. Yodaqa: a modular question answering system pipeline. In *POSTER 2015-19th International Student Conference on Electrical Engineering*, pages 1156–1165, 2015.
- [10] Petr Baudiš and Jan Šedivý. Modeling of the question answering task in the yodaqa system. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 222–228. Springer, 2015.
- [11] Raju Barskar, Gulfishan Firdose Ahmed, and Nepal Barskar. An approach for extracting exact answers to question answering (qa) system for english sentences. *Procedia Engineering*, 30:1187–1194, 2012.

Tecnologías emergentes aplicadas a problemas agrícolas en regiones áridas

Alberto Eduardo Riba¹, Fernanda Beatriz Carmona¹, Jorge Damián Tejada¹, Emmanuel Alejandro Portugal Murcia¹, Matías Agustín Pérez¹, Cristian Emmanuel Ríos¹, Adriel Aaron Riperto¹

¹ Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina
{ariba, fbcarmona, jtejada, mperez, eportugal, crios}@undec.edu.ar
aaron.ariporto@gmail.com

Resumen

El estudio, análisis y proyección de datos agrometeorológicos de los valles de nuestra provincia permite ofrecer asesoramiento a los sectores público y privado para el desarrollo de la agricultura sustentable, de manera tal que los sistemas de producción tengan la capacidad de mantener su productividad, rentabilidad y desarrollo en sus comunidades a largo plazo y además, realizar investigaciones agronómicas, meteorológicas, biológicas y ecológicas de interés en temas tales como el cambio climático, la prevención de plagas, entre otros.

El sector agrícola regional se esfuerza por aplicar prácticas modernas que garanticen una provisión segura y sostenible de alimentos de calidad, fomentar la eficiencia del uso de recursos y desarrollar una economía rentable.

La agricultura de precisión cumple un rol muy importante en estas prácticas, los avances en herramientas de sensado, la generalización del uso de los sistemas de posicionamiento globales junto a la utilización de Sistemas de Información Geográfica (SIG) representa una potente herramienta para el procesamiento de los datos recolectados generando mapas que permitan visualizar e interpretar los resultados de manera gráfica, contribuyendo al proceso de toma de decisiones.

La captura automática de datos de la mano de Internet de las Cosas (IoT), los sistemas de bases de datos espaciales y NoSQL, el aumento en la velocidad de procesamiento High Performance Computing (HPC) y el nuevo paradigma de servicios de Cloud Computing (CC) han fomentado y potenciado la interrelación de los SIG con su entorno,

integrando cada vez más aplicaciones que gestionan y procesan este tipo de datos.

Esta línea de I+D+i de gran interés regional, se enfoca en el problema de la racionalización y el uso eficiente de los recursos agrícolas en regiones de climas áridos, con el objetivo de incrementar la productividad y combatir el cambio climático.

Palabras clave:

Agromática, Data-Driven Agriculture, Agricultura de Precisión, Sistemas GNSS, Sistemas de Información Geográficos, Series de Datos Agrometeorológicos.

Contexto

Esta línea de I+D+i corresponde al desarrollo e implementación de proyectos que fortalecen la inserción de la Universidad Nacional de Chilecito (UNdeC) en la comunidad y especialmente en el medio productivo de la región y refiere a proyectos aprobados por la Secretaría de Ciencia y Tecnología (SECYT), convocatoria para estímulo y desarrollo de la investigación científica y tecnológica (FICyT – UNdeC); “Programación y generación de pronósticos de riego presurizado para cultivo de olivo en regiones áridas, utilizando TIC” 2009-2011, “Integración y administración de índices de sensado en la programación y pronósticos de riego presurizado para cultivo de olivo en regiones áridas”, 2012-2014, “Utilización de métodos de diseño de software para desarrollar un sistema automatizado de riego”, 2011-2013, “Red de Sensores Inalámbricos basado en microcontroladores para la monitorización del riego presurizado en plantaciones de olivo”, 2013-2015,

“Incremento de la precisión posicional relativa utilizando receptores GPS de bajo costo” 2016-2018, “Sistemas de Información para el modelado y simulación de variable meteorológicas”, período de ejecución 2020-2022.

Esta línea se vincula con el proyecto “Software y aplicaciones en Computación de Altas Prestaciones” presentado en la Convocatoria Proyectos de I+D 2018 lanzada por la SECYT de la UNdeC con RR N° 883/18, para aquellos casos donde el cúmulo de datos a procesar necesite recurrir a técnicas de procesamiento paralelo para disminuir el tiempo de procesamiento de los datos.

Los proyectos enmarcados en esta línea convergen en Trabajos Finales de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC con la participación de docentes y alumnos avanzados de éstas.

Introducción

Los valles cordilleranos de la provincia de La Rioja integran una de las principales regiones frutícolas de la Argentina. Si bien estos frutales encuentran un ambiente propicio para desarrollarse vegetativa y productivamente, la condición climática de algunos sectores de los valles no garantiza que estos cultivos puedan ser actividades rentables y sustentables [1]. Por este motivo, los estudios de las condiciones climáticas como el comportamiento de las temperaturas en las distintas estaciones del año son de fundamental importancia para favorecer el desarrollo agropecuario en la región.

En la actualidad el proceso de recolección de datos climáticos se automatiza a través del uso sensores meteorológicos con capacidades de procesamiento y almacenamiento. La UNdeC dispone de una red de 80 sensores de temperatura distribuidos en tres valles (Valle La Costa, Valle Antinaco los Colorados y Valle Del Bermejo). Los datos generados son recolectados y guardados en una Base de Datos con capacidad de almacenar información georeferenciada. Los Sistemas de Información Geográfica (SIG) constituyen una herramienta fundamental para el análisis de este tipo información porque permiten

generar mapas que facilitan el análisis visual de los datos.

La actividad agrícola en la región sólo es posible con la ayuda del riego artificial. En un lote cultivado es posible encontrar sectores de alta productividad, muy próximos a sectores menos productivos. Esta variación espacial suele estar asociada a factores como la pendiente del suelo, la permeabilidad, el tipo del suelo y la fertilidad. Sin embargo, los agricultores manejan el cultivo de forma homogénea, aplicando dosis de fertilizante o irrigando de manera uniforme todo el lote.

La agricultura de precisión utiliza complejas fórmulas y modelos matemáticos para el análisis de los grandes volúmenes de datos geo-espaciales generados por las distintas tecnologías de sensado, convirtiéndose en un sistema de control en donde la retroalimentación de la información permite a los productores diseñar tratamientos específicos situados para incrementar la eficiencia en el uso de los recursos involucrados [2].

En regiones de climas áridos el costo final de explotación es afectado en un alto porcentaje por factores imputables al riego. Dentro de este costo se considera la inversión inicial del sistema y el costo energético para su extracción y distribución.

El sistema de irrigación más utilizado por los agricultores es el riego presurizado. Desde el punto de vista agronómico se denominan riegos localizados porque humedecen un sector de volumen de suelo suficiente para un buen desarrollo del cultivo.

Para incrementar la eficiencia en los sistemas de riego es necesario estudiar las distintas variables y como éstas se relacionan para determinar el uso adecuado del recurso. Estas variables son de naturaleza heterogénea y algunas de ellas pueden ser capturadas automáticamente mediante el uso de sensores para permitir su posterior tratamiento y análisis para una correcta programación del riego.

Como resultados de los proyectos desarrollados en esta línea de investigación se implementó un sistema de registro para el procesamiento de esta información

relacionada con el riego y la fertilización [3, 4, 5]. Se desarrollo una estación inalámbrica que automatiza la captura de la información de sensores ubicados en distintos sectores de la plantación. Cada nodo de la red está compuesto por un dispositivo inalámbrico autónomo y un conjunto de sensores para la recolección de datos de naturaleza agrometeorológica. Esta práctica ha sido implementada con éxito en diversos ámbitos como detección de incendios forestales [6] o la monitorización de viñedos [7]. Estas redes se caracterizan por su escalabilidad, ausencia de cableado y bajo consumo, lo que las vuelve muy interesantes para aplicaciones en agricultura, ya que serían más costosas y complejas de implementar con otras tecnologías.

La automatización de la recolección de información de sensores permitió advertir que las series agrometeorológicas resultantes presentan problemas de completitud, veracidad y exactitud.

La calidad de estas series depende de dispositivos electromecánicos, de redes de transmisiones inalámbricas, de personal de mantenimiento y de baterías. Cualquiera de estos factores resulta un punto de fallo ineludible que degrada la calidad [8]. Para mitigar estos se pueden establecer políticas y estándares de funcionamiento de hardware y software que tiendan a prevenirlos, minimizarlos y ante su aparición, controlarlos. A pesar de la implementación de tales políticas existen casos en los que simplemente los fallos ocurren, y se deben tomar medidas ante estas situaciones.

Los sistemas GNSS (Global Navigation Satellite System) son muy utilizados en la agricultura de precisión para georreferenciar sectores de parcelas, existen tecnologías implementadas como NAVSTAR-GPS y GLONASS, y otras en vías de implementación como GALILEO y COMPASS. [9, 10]

La agricultura de precisión necesita de sistemas que provean más precisión en sus localizaciones para optimizar el uso de los recursos, siendo normalmente estos dispositivos más costosos [11]. Por este

motivo es muy importante estudiar, diseñar y desarrollar algoritmos, técnicas y métodos que permitan disminuir el error en la posición entregada por receptores GNSS de bajo costo para mejorar la precisión del posicionamiento. Esta temática es abordada por uno de los integrantes del equipo de trabajo en el desarrollo de su tesis de maestría.

En la actualidad, surge la necesidad de integrar al conocimiento científico la información con base espacial y los Sistemas de Información Geográfica (SIG) son la herramienta idónea para este tipo de trabajo.

La principal característica de un SIG es que está diseñado para trabajar con datos referenciados con respecto a coordenadas espaciales o geográficas, así como trabajar con distintas bases de datos de manera integrada, permitiendo generar información gráfica útil para la toma de decisiones. Los mapas ayudan a condensar varios aspectos de la realidad de una zona cuyo objetivo es reconocer la existencia de patrones espaciales sobre algún fenómeno de interés. Estos son un poderoso instrumento para la organización de la información, y por lo tanto ayudan a la toma de decisiones [12] [13]. Como ya mencionamos en la actualidad el proceso de recolección de los datos se automatiza a través de una red de sensores agrometeorológicos conectados a Internet con mínimas capacidades de procesamiento y almacenamiento [14]. Los datos generados se transmiten y almacenan en una Infraestructura de Datos Espaciales, que debe tener una elevada capacidad de gestión y almacenamiento de información georeferenciada. Posteriormente, estos datos serán procesados por Sistemas de Información Geográfica [15] para alimentar modelos predictivos y realizar simulaciones sobre producción, mapas climáticos y analizar variaciones relacionadas con el cambio climático.

Líneas de Investigación, Desarrollo e Innovación

- Sistemas de Información Geográfica
- Posicionamiento de precisión

- Bases de Datos NoSQL
- Sistemas de Tiempo Real
- Comunicaciones inalámbricas
- Redes de sensores
- Ajuste de series de datos agrometeorológicos

Resultados y Objetivos

- Identificar y describir las principales magnitudes generadas por los sensores agrometeorológicos.
- Analizar y estudiar las diferentes tecnologías de microcontroladores, módulos de conexiones inalámbricas y tipos de sensores disponibles en el mercado.
- Monitorizar en forma centralizada, remota y en tiempo real las variables capturadas por los diferentes sensores.
- Analizar que tipo de Base de Datos se ajusta mejor para el almacenamiento y gestión de los datos bajo estudio.
- Generar mapas que representen los datos analizados.
- Mejorar la capacidad de administración y planificación de los recursos hídricos destinados al riego, a través del análisis de las variables obtenidas de los nodos instalados en sectores con diferentes características de suelo y clima.
- Configurar adecuadamente los algoritmos de control y aplicar estrategias de riego que optimicen la relación kg. producido por m³ de agua aplicada.
- Evaluar distintos esquemas de control, comparar los resultados e inferir en la elaboración de nuevas estrategias de riego.
- Desarrollar técnicas, métodos y algoritmos para mejorar a la precisión del posicionamiento utilizando receptores GNSS de bajo costo en un prototipo de GNSS diferencial.

Los resultados de esta línea de I+D+i permitirá consolidar un grupo de investigación, desarrollo y transferencia, explorar oportunidades de formación de recursos humanos, ofrecer servicios a la comunidad en el área de estudio y potenciar los vínculos de cooperación con otras

instituciones y grupos de docentes-investigadores de la UNdeC.

Como resultados se puede mencionar: los trabajos finales de grado de cuatro de los integrantes, denominados “Sistema de Gestión de Riego y Fertilización”, aprobado en el año 2014, “Estación inalámbrica basada en microcontroladores para la monitorización del riego en plantaciones de olivo” aprobado marzo de 2018, “Sistema de reacondicionamiento de series meteorológicas” aprobado en junio de 2018; “Sistema Colaborativo para el sector vitivinícola” febrero 2020; artículo presentado en el evento 45 JAIIO – Concurso de trabajos Estudiantiles 2016, titulado “Red de sensores inalámbricos basados en microcontroladores para la monitorización del riego presurizado en plantaciones de olivo”; artículo presentado en XXII CACIC en 2016 titulado “Procesamiento de sentencias NMEA-0183 para el análisis de la geometría satelital utilizando receptores GPS de bajo costo”; artículo presentado en XXIII CACIC en 2017 titulado “Estación de monitoreo en tiempo real de parámetros agrometeorológicos para determinar la necesidad de riego en plantaciones agrícolas”.

Formación de Recursos Humanos

El equipo de trabajo está formado por seis docentes investigadores y dos alumnos de grado avanzados. Dos de ellos son maestrands en informática con trabajo de tesis en desarrollo, seis están cursando la Especialización en inteligencia de datos orientada a Big Data.

Los alumnos de grado se encuentran desarrollando su trabajo final en esta línea de I+D+i.

Los integrantes son docentes de asignaturas que fomentan la participación en proyectos de investigación, por lo que pueden surgir nuevos trabajos en esta línea.

Referencias

[1] F. Vita Serman, “Estudio agroclimático en detalle para definir la aptitud frutícola (Olivo, Almendro, Pistacho, Pecán y Nogal) de los

valles cordilleranos de Antinaco-Los Colorados, Bermejo y Región Este, de la provincia de La Rioja”, Informe Final Provincia de La Rioja UNdeC, INTA, CFI, 2018

[2] E. W. Schuster, S. Kumar, S. E. Sarma, J. L. Willers and G. A. Milliken, “Infrastructure for data-driven agriculture”, IEEE 8th International Conference & Expo on Emerging Technologies for a Smarter World, New York, NY, 2011

[3] F. Carmona, A. Riba, A. Sfeir, and F. E. Frati, “Pronósticos de riego presurizado para cultivo de olivo en regiones áridas”, WICC 2008 eje: Ingeniería de Software y Base de Datos, 2008.

[4] F. Carmona, E. Riba, A. Sfeir, and F. Frati, “Pronósticos de riego presurizado para cultivo de olivo en regiones áridas,” WICC2010 eje: Innovación en Sistemas de Software, 2010.

[5] E. Riba, F. Carmona, F. Frati, J. Tejada, N. Acosta, and J. Toloza, “Integración y administración de índices de sensado en la programación y pronósticos de riego presurizado para cultivo de olivo en regiones áridas,” 2012, eje: Innovación en sistemas de software.

[6] J. Solebera, “Detecting forest fires using wireless sensor networks,” Sep. 2010.

[7] A. Bielsa, “Smart agriculture Project in galicia to monitor a vineyard with waspmote,” Jun. 2012.

[8] M. Khaliq, T. Ouarda, “Short communication on the critical values of the standard normal homogeneity test (SNHT)” International Journal of Climatology, Vol. 27 681-687p, 2007.

[9] J. Tolosa, “Algoritmos y técnicas de tiempo real para el incremento de la precisión posicional relativa usando receptores GPS estándar”, SEDICI, Universidad Nacional de La Plata, 2012.

[10] A. Riba, “Procesamiento de sentencias NMEA-0183 para el análisis de la geometría satelital utilizando receptores GPS de bajo

costo”. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), 2016.

[11] A. Riba, J. Tejada, F. Frati, N. Acosta, J. Toloza “Aumento de la precisión posicional empleando técnicas y algoritmos para el tratamiento del error en receptores GNSS de bajo costo”, XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017), 2017.

[12] C. García, S. Moral, R. Pisonero & J. Córdoba, “Creatividad e innovación aplicadas al estudio de la globalización: herramientas digitales como base del aprendizaje”. In La educación geográfica digital (pp. 599-614). Grupo de Didáctica de la Geografía (AGE), 2012.

[13] J. Membrado, “El lenguaje cartográfico en los mapas temáticos”, Estudios Geográficos, LXXV, I 278, 177-201, Doi: 10.3989/estgeogr.201506, 2015

[14] A. Botta, W. De Donato, V. Persico & A. Pescapé, “On the integration of cloud computing and internet of things”, In Future internet of things and cloud (FiCloud), 2014 international conference on (pp. 23–30), IEEE, 2014.

[15] P. Yue, H. Zhou, J. Gong & L. Hu, “Geoprocessing in cloud computing platforms—a comparative analysis”, International Journal of Digital Earth, 6(4), 404–425, 2013.

Transformación Digital en tiempos de la Industria 4.0

Sandra Oviedo, Daniel Díaz, Alejandra Otazú, Francisco Ibañez,
Pedro Daniel Zárate, Juan Aranda Romera

Laboratorio de Informática Aplicada a la Innovación Instituto de Informática /
Dpto. de Informática/ FCEFyN / UNSJ

CUIM – Av. Ignacio de la Roza 590 (O), Rivadavia – J5402DCS San Juan, 0264 4265101
{ soviedo, ddiaz, fibannez}@iinfo.unsj.edu.ar; pzarate.arg@gmail.com; otazuale@hotmail.com;
juanaranda@live.com

Resumen

Una nueva generación de Tecnologías de Información y Comunicación (TIC) esta dando comienzo a la cuarta revolución industrial o industria 4.0, es una nueva era que va a permitir a la organización dar un salto cuantitativo y cualitativo en gestión de cadenas de valor. La Transformación Digital (TD) es el camino para aprovechar las nuevas oportunidades que brindan las tecnologías de la industria 4.0 para la generación de nuevas estrategias de negocio. Este proyecto propone desarrollar un marco de trabajo que facilite a las organizaciones de nuestra región realizar su TD en el contexto de las oportunidades que brindan las tecnologías de la industria 4.0.

Palabras claves: Transformación Digital. Industria 4.0. Modelos de Negocios.

Contexto

Este trabajo describe los progresos en la investigación y estrategias que se están llevando a cabo encuadrados en el proyecto “Transformación Digital en tiempos de la Industria 4.0” que se desarrolla en el Laboratorio de Informática Aplicada a la Innovación del Instituto de informática de la Facultad de Ciencias

Exactas Físicas y Naturales de la Universidad Nacional de San Juan.

Introducción

Desde su surgimiento las TIC han influido en la vida de las empresas. En sus inicios el departamento encargado de las TIC solo prestaba servicios de soporte, tales como gestión de servidores, asistencia técnica a usuarios, entre otras. Con el tiempo las TIC transformaron las maneras de trabajar y gestionar los recursos de la empresa convirtiéndose en algunos caso en el core del negocio.

La incorporación de las TIC a las empresas ha sido un tema de estudio desde su surgimiento, para cada nueva tecnología que surge un nuevo proceso para la incorporación de dicha tecnología debe diseñarse. Este proceso en la literatura de hoy se lo conoce como transformación mediada por TIC.

En los últimos años, las nuevas oportunidades de estrategia de negocios que surgen gracias a la aparición de las nuevas tecnologías llevan a lo que se conoce bajo el nombre de Transformación Digital . De una manera más formal en [1] se define a la Transformación Digital como "un proceso que tiene como objetivo mejorar una entidad mediante la activación de cambios significativos en sus propiedades a través de combinaciones de tecnologías de información, informática, comunicación y conectividad".

Las nuevas tecnologías están dando lugar al desarrollo de nuevas estrategias de manufactura tales como industria 4.0 en Alemania, Internet Industrial en EEUU [2] y Made In China 2025 en China [3]. Según Hermann y otros [4] “la Industria 4.0 es un término colectivo para las tecnologías y conceptos de organización de la cadena de valor. Dentro de las fábricas inteligentes modulares y estructuradas de la Industria 4.0, los sistemas ciberfísicos (CPS) monitorean los procesos físicos, crean una copia virtual del mundo físico y toman decisiones descentralizadas. A través de Internet de las cosas (IoT), los CPS se comunican y cooperan entre sí y con los humanos en tiempo real.

A través de Internet de Servicios (IoS), los participantes de la cadena de valor ofrecen y utilizan servicios internos y de organización “cruzada”. Esta cuarta revolución industrial ya ha llegado a nuestro país, así lo demuestran las últimas conferencias de la Unión Industrial Argentina (UIA) [5], donde diversas temáticas relacionadas con la industria 4.0 han sido tratadas. Desde una perspectiva regional en [6] se presenta una visión general de la industria 4.0 en Argentina y en [7] se discute si la Industria 4.0 constituye un nuevo paradigma tecnoorganizacional, o si, por el contrario, representa una intensificación de los rasgos salientes del paradigma TIC.

En este contexto la TD y la Industria 4.0 presentan nuevos desafíos y oportunidades de innovación para las organizaciones de la región. Es por ello que esta propuesta plantea desarrollar un marco de trabajo que facilite a las organizaciones de nuestra región realizar su TD en el contexto de las oportunidades que brindan las tecnologías de la industria 4.0.

Líneas de Investigación

Industria 4.0. Conceptos

Según Hermann y otros [4] “la Industria 4.0 es un término colectivo para las tecnologías y conceptos de organización de la cadena de valor.

En [8] se analiza esta definición desde la perspectiva de los cuatro componentes claves de la Industria 4.0 que se describen a continuación:

- 1- **Sistemas ciberfísicos (CPS):** Los CPS pueden considerarse sistemas que unen el mundo físico y el virtual. Más precisamente, “los sistemas ciberfísicos son integraciones de procesos computacionales con procesos físicos. Dispositivos computarizados y redes integradas monitorean y controlan los procesos físicos, generalmente con bucles de retroalimentación donde los procesos físicos afectan los cálculos y viceversa. En el contexto de la manufactura, esto significa que la información relacionada con lo físico (el taller, la fábrica) y el espacio virtual de cómputo están altamente sincronizados. Esto permite un nuevo grado de control, vigilancia, transparencia y eficiencia en el proceso de producción. Con respecto a su estructura, CPS tiene dos redes paralelas para controlar, a saber, la red física de componentes interconectados de la infraestructura y la red cibernética compuesta por controladores inteligentes y los enlaces de comunicación entre ellos. CPS realiza la integración de estas redes mediante el uso de múltiples sensores, actuadores, unidades de procesamiento de control y dispositivos de comunicación [9].
- 2- **Internet de las cosas (IoT):** El término no tiene una sola definición, tal esta expresado en [10, 11], la definición del Estándar de IEEE dice “Una red de elementos, cada uno integrado con sensores, que están conectados a Internet” [10]. Porter [12] “son los productos inteligentes y conectados que

ofrecen posibilidades de expansión exponenciales para nuevas funcionalidades, mayor confiabilidad, mayor utilidad del producto y capacidades que atraviesan y trascienden los límites de los productos tradicionales “. En resumen podría decirse que la IoT puede hacer que todas las cosas físicas puedan convertirse en cosas inteligentes conectadas a internet.

- 3- Internet de servicios (IoS): De manera similar al IoT, está surgiendo una Internet de servicios (IoS), basada en la idea de que los servicios se ponen a disposición fácilmente a través de tecnologías web, lo que permite a las empresas y a los usuarios privados combinar, crear y ofrecer un nuevo tipo de servicios de valor agregado [13]. Los servicios de la nube, así como las aplicaciones basadas en arquitectura orientada servicios o en arquitectura de microservicios son parte del internet de servicios.
- 4- Fabricación inteligente: Los conceptos de CPS, IoT e IoS se introdujeron como componentes principales de Industria 4.0. Cabe señalar que estos "conceptos" están estrechamente relacionados entre sí, ya que CPS se comunica a través de IoT e IoS, lo que permite la llamada "fabricación inteligente", que se basa en la idea de un sistema de producción descentralizado, en el que “Los seres humanos, las máquinas y los recursos se comunican entre ellos de manera tan natural como en una red social.

Tecnologías Industria 4.0

En [14] se presenta un muy completo radar de tecnologías asociadas a la Industria 4.0. A continuación se describen someramente las más destacadas basadas en el trabajo de [15].

- Simulación: Permite la virtualización del diseño de productos, procesos y

diseño de fábrica. Las herramientas de simulación permiten probar modelos virtuales de productos o procesos antes de aplicarlo en soluciones reales, optimizando el desarrollo de nuevas tecnologías.[15].

- Big Data & Analytics : Conjunto de tecnologías y herramientas capaces de procesar y analizar grandes volúmenes de datos que son: generados continuamente; compuesto de textos, imágenes, etc.; y de múltiples departamentos. El análisis de dichos datos puede ayudar a identificar las fallas existentes en detalle, aumentando el conocimiento disponible sobre los hábitos y preferencias de los consumidores, entre otras [16] [15].
- Cloud Computing: Servicios que brindan acceso a máquinas, sistemas, software y herramientas a través de redes como Internet. Reemplaza la necesidad de adquirir productos, conocimientos y / o infraestructuras costosas [17], [18].
- Cyber Physical Systems : Integración de entidades informáticas físicas colaborativas (máquinas, robots, sensores, etc.) que interactúan a través de una red virtual. Incluye máquinas inteligentes, sistemas de almacenamiento e instalaciones de producción que pueden intercambiar información con autonomía e inteligencia, son capaces de decidir y desencadenar acciones, y pueden controlarse mutuamente de forma independiente [9].
- Cybersecurity : Servicios y tecnologías con el objetivo de proteger a usuarios, sistemas, equipos, redes y datos industriales de intrusión ilícita [19].
- Robótica colaborativa: Robots (móviles y / o fijos) que operan en procesos físicos automatizados e interactúan con operadores humanos u otros robots en

un comportamiento intuitivo de autoaprendizaje [20].

- Realidad Aumentada: Integración de información virtual con el mundo real a través de la combinación de elementos 3D con el contexto espacial de la fábrica. Permite una interactividad y un procesamiento en tiempo real de la proyección de imágenes, que se utiliza para mejorar el proceso de fabricación o para probar nuevos productos [21].
- Fabricación Aditiva: Tecnología que permite imprimir objetos mediante la composición de capas de plástico o metal, evitando el desperdicio de material en procesos como el corte. Utilizado inicialmente para producir prototipos o pequeñas series de piezas complejas, ahora se está utilizando para la producción a gran escala [22].
- Integración de Sistemas: Integración de datos en todos los niveles (desde la gerencia hasta el taller) de una compañía y entre (desde proveedores hasta clientes) compañías en la cadena de suministro de acuerdo con sus patrones de transferencia de datos. Generalmente conectado a través de aplicaciones de Internet de las cosas [23].

Transformación Digital

En [1] se presenta una revisión completa de definiciones de TD (DT), se destaca la que resulta de un análisis semántico que realiza el autor de las definiciones encontradas y define a la TD como "un proceso que tiene como objetivo mejorar una entidad mediante la activación de cambios significativos en sus propiedades a través de combinaciones de tecnologías de información, informática, comunicación y conectividad", donde con la palabra entidad se da un sentido amplio al alcance, pudiendo tratarse de una organización, una industria o una sociedad. Los autores destacan además

que se espera alcanzar mejoras a través de la DT aunque no se puede garantizar.

La TD se vale de las tecnologías de la industria 4.0 para hacer una evolución en su concepto atendiendo demandas de los tiempos que corren, como puede ser la generación de valor a partir de grandes volúmenes de datos o la necesidad de plantear nuevos modelos de negocios o productos resultantes de innovación disruptiva.

Existen tres factores externos principales que impulsan la necesidad de la TD [24].

- 1- desde la llegada de la WWW y su adopción mundial, un número creciente de tecnologías que lo acompañan (por ejemplo, Internet de banda ancha, teléfonos inteligentes, Web 2.0, SEO, computación en la nube, reconocimiento de voz, sistemas de pago en línea y criptomonedas) han aumentado y han fortalecido el desarrollo del comercio electrónico. Se prevé que la omnipresencia del big data y el advenimiento de las tecnologías digitales emergentes, como la inteligencia artificial (IA), blockchain, internet de las cosas (IoT) y la robótica, tengan efectos de gran alcance en los negocios.
- 2- Debido a estas nuevas tecnologías digitales, la competencia está cambiando dramáticamente. En el comercio minorista, las tecnologías han alterado el panorama de la competencia, desplazando las ventas a empresas digitales relativamente jóvenes. La competencia no solo se ha vuelto más global, sino que la intensidad también ha aumentado a medida que grandes empresas ricas en información de los EE. UU. (Por ejemplo, Amazon, Alphabet, Apple y Facebook) y China (por ejemplo, Alibaba y JD) comienzan a dominar numerosas industrias.

3- El comportamiento del consumidor está cambiando como respuesta a la revolución digital. Las cifras del mercado muestran que los consumidores están transfiriendo sus compras a tiendas en línea, y los puntos de contacto digitales tienen un papel importante en el recorrido del cliente que afecta tanto a las ventas en línea como fuera de línea. Con la ayuda de nuevas herramientas de búsqueda y redes sociales, los consumidores se han vuelto más conectados, informados, capacitados y activos. Las tecnologías digitales permiten a los consumidores co-crear valor diseñando y personalizando productos, realizando actividades de distribución de última milla y ayudando a otros clientes compartiendo reseñas de productos. Los dispositivos móviles se han vuelto importantes en el comportamiento del consumidor de hoy y facilitan el comportamiento de exhibición, la práctica de examinar la mercancía fuera de línea y luego comprarla en línea.

Resultados y Objetivos

El proyecto que se describe acá tiene por objetivo “desarrollar un marco de trabajo que facilite a las organizaciones de nuestra región realizar su TD en el contexto de las oportunidades que brinda las tecnologías de la industria 4.0.

Como un primer abordaje se está trabajando en la obtención de un marco de referencia que permita definir un modelo de negocios donde se introduzcan tecnologías de industria 4.0 de manera gradual. Se trata de plantear un modelo basado en Canvas Business Model [25] Para lo cual se propone el siguiente plan de labor:

1- Hacer una revisión acerca de las estrategias, framework, metodologías, procesos de implementación de TD

Considerando además, en especial, los que hagan referencia a modelo de negocio.

2. Proponer un framework y/o una estrategia propia que combine la TD con Canvas Business Model.
3. Aplicar el Framework en un caso de estudio.

Formación de Recursos Humanos

El equipo de trabajo que lleva adelante este proyecto se compone de:

- 6 docentes investigadores,
- 1 tesistas de grado en período iniciación.
- 2 tesistas de posgrado (maestría) iniciando sus trabajos.












Referencias

1. Vial, G., *Understanding digital transformation: A review and a research agenda*. Journal of Strategic Information Systems 2019. **28**(2): p. 118-144.
2. Internet, I., *The Industrial Internet Consortium: A Global Nonprofit Partnership Of Industry, Government And Academia*. 2021.
3. Li, L., "China's manufacturing locus in 2025: With a comparison of "Made-in-China 2025" and "Industry 4.0". Technological Forecasting Social Change, 2018. **135**: p. 66-74.
4. Herman, M., T. Pentek, and B. Otto, *Design principles for industrie 4.0 scenarios*. 49th Hawaii International Conference on System Sciences (HICSS), 2016.
5. UIA2018, *24 Conferencia Industrial* 2018.
6. Basco, A.I., et al., *Industria 4.0 Fabricando el Futuro*. UIA-BID-INTAL, Buenos Aires, Argentina, 2018.
7. Brixner, C., et al., *Industria 4.0: ¿Intensificación del paradigma TIC o nuevo paradigma tecnoorganizacional?* CENTRO INTERDISCIPLINARIO DE ESTUDIOS EN CIENCIA, TECNOLOGÍA E INNOVACIÓN- MinCyT, 2019.

8. Hofmann, E. and M. Rüsçh, *Industry 4.0 and the current status as well as future prospects on logistics*. Computer and Industry, 2017. **89**: p. 23-34.
9. Leitão, P., A.W. Colombo, and S. Karnouskose, *Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges*. Computers in Industry, 2016. **81**: p. 11-25.
10. Minerva, R., A. Biru, and D. Rotondi, *Towards a Definition of the Internet of Things (IoT)*. IEEE Internet Initiative, 2015.
11. Ibarra-Esquer, J.E., et al., *Tracking the Evolution of the Internet of Things Concept Across Different Application Domains*. MDPI Sensors (Basel), 2017.
12. Porter, M.E. and J.E. Heppelmann, *How smart, connected products are transforming competition*. Harvard business review, 2014. **92**,: p. 64-88.
13. Wahlster, W., et al., *Towards the internet of services*. The THESEUS research program: Springer, 2014.
14. Schallmo, D., C.A. Williams, and L. Boardman, *Digital transformation of bussines models- Best Practice, Eneblers and roadmaps*. International Journal of Innovation Management, 2017. **21**(8).
15. Dalmarco, G., et al., *Providing industry 4.0 technologies: The case of a production technology cluster*. The Journal of High Technology Management Research, 2019. **30**(2).
16. Gölzer, P. and A. Fritzsche, *Data-driven operations management: organisational implications of the digital transformation in industrial practice*. Production Planning & Control 2017. **28**(16).
17. Marston, S.a., et al., *Cloud computing— The business perspective*. Decision Support Systems 2011. **51**(1): p. 176-189.
18. Antonopoulos, N. and L. Gillam, *Cloud computing*. Springer, 2010.
19. Niekerk, J.v. and R. von Solms, *From information security to cyber security*. Computers & Security, 2013. **38**: p. 97-102.
20. Bonkenburg, T., *Robotics in logistics - A DPDHL perspective on implications and use cases for the logistics industry*. 2016.
21. Paelke, V., *Augmented reality in the smart factory: Supporting workers in an industry 4. 0. Environment*. 19th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2014, 2015.
22. Laureijs , R.E., et al., *Metal Additive Manufacturing: Cost Competitive Beyond Low Volumes*. Journal of Manufacturing Science and Engineering, 2017.
23. Thoben, K.-D., S.A. Wiesner, and T. Wuest, *Industrie 4.0 and smart manufacturing-a review of research issues and application examples*. International Journal of Automation Technology, 2019. **11**(1): p. 4-19.
24. Verhoefa, P.C., et al., *Digital transformation: A multidisciplinary reflection and research agenda*. Journal of Business Research 2021. **122**: p. 889-901.
25. Osterwalder, A. and Y. Pigneur, *Generación de Modelos de Negocios*. Ed. DEUSTO, 2011.

Procesamiento Distribuido y Paralelo

ALGORITMOS PARALELOS Y EVALUACION DE RENDIMIENTO EN PLATAFORMAS DE 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⁾ , María José Basgall⁽¹⁾⁽²⁾ , Mariano Sánchez⁽¹⁾ , Manuel Costanzo⁽¹⁾,
Silvana Gallo⁽¹⁾⁽²⁾ , Erica Montes de Oca⁽¹⁾ , Emmanuel Frati⁽¹⁾, Adriana Gaudiani⁽⁴⁾

¹Instituto 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

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
³CIC – Comisión de Investigación Científica de la Provincia de Buenos Aires

⁴Universidad Nacional de General Sarmiento
{mnaiouf, degiusti, ldgiusti, francoch, vsanz, apousa, erucci, mjbasgall,
msanchez, mcostanzo, sgallo, emontesdeoca, 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) del Dpto. de Arquitectura de Computadores y Sistemas Operativos de la Universidad Autónoma de Barcelona; con el Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid; y con el grupo Soft Computing and Intelligent Information Systems (SCI2S) de la Universidad de Granada, 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: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales 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 por AECID, CyTeD, OEI y CIC y becas de Telefónica de Argentina. 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]. En la actualidad se comercializan placas de bajo costo como Raspberry PI [4] u Odroid [5] que poseen múltiples núcleos de baja complejidad y en algunos casos son procesadores multicore asimétricos (AMPs) con el mismo repertorio de instrucciones. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones [6], así como las características de scheduling en los mismos [7]. 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.

La creación de algoritmos paralelos en arquitecturas multiprocesador no es un proceso directo [8]. 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 [9].

Muchos problemas algorítmicos se vieron impactados por los multicore y clusters de multicore. A partir de incorporar varios chips multicore dentro de un nodo y conectar múltiples nodos vía red, se puede crear una arquitectura NUMA, de modo que los cores en un chip compartan memoria principal, y puedan acceder remotamente a la memoria dedicada de otro chip, aunque ese acceso sea más costoso, surgiendo así varios niveles de comunicación. Esto impacta sobre el desarrollo de algoritmos que aprovechen adecuadamente las arquitecturas, y motiva el estudio de performance en sistemas híbridos. Además, 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 [10][11] o los más recientemente explorados UPC, Chapel y Titanium del modelo PGAS [12].

La combinación de arquitecturas de múltiples núcleos con aceleradores dio lugar a plataformas híbridas con diferentes características. 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 [13]. 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 aunque queda mucho por hacer en cuanto al diseño, lenguajes y programación

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* [14]. 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, placas de bajo costo y entornos cloud.
- 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 [15] 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 [16]. Los resultados revelaron que este algoritmo supera en rendimiento a trabajos previos, desarrollados para sistemas multicore y GPUs, para datos de tamaño considerable. En [34] 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.

➤ **Alineamiento de secuencias biológicas.** Esta operación consiste en comparar dos o más secuencias biológicas, como pueden ser las de ADN o las de proteínas, y resulta fundamental en investigaciones de la bioinformática y la biología molecular. El algoritmo de Smith-Waterman es considerado el método de alineamiento más preciso. Desafortunadamente, este algoritmo resulta costoso debido a su complejidad computacional cuadrática mientras que la situación se agrava aún más a causa del crecimiento exponencial de datos biológicos en los últimos años [17]. El reciente surgimiento de aceleradores en HPC (GPU, Xeon Phi, FPGA, entre otros) da la oportunidad de acelerar los alineamientos sobre hardware comúnmente disponible a un costo accesible, como se ha mostrado en [18][19][20][21]. A futuro, interesa explorar las fortalezas y debilidades del uso de nuevas tecnologías de software para arquitecturas paralelas como, por ejemplo, oneAPI [22].

➤ **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 $O(n^3)$. Se desarrollaron implementaciones optimizadas para dos arquitecturas HPC recientes -como son Intel Xeon Phi KNL y NVIDIA Pascal- y se analizó comparativamente su rendimiento y eficiencia energética (teórica) en diferentes escenarios. Como trabajo futuro, interesa incorporar otros modelos de las arquitecturas elegidas al estudio para robustecer el análisis [33].

➤ **Simulaciones utilizando modelos basados en agentes.** El objetivo de esta línea son las simulaciones basadas en agentes sobre infraestructuras de altas prestaciones con modelos más cercanos a la realidad que ayuden a la toma de decisiones a científicos de otras disciplinas, y no expertos en el área de la informática, (biología, ecología, física, etc.). Para ello, y por ser modelos complejos que necesitan mucha potencia de cómputo, resulta imprescindible el uso de soluciones de HPC con el fin de lograr tiempos de respuesta reducidos para entornos complejos. Se busca desarrollar algoritmos y simuladores como soluciones HPC que sean eficientes y escalables y por ello una cuestión importante, tratada en esta línea de trabajo, es lograr que además de estas premisas, la simulación sea realizada con el menor consumo posible de energía. Dado que son simulaciones con un gran número de ejecuciones por escenario, no solo es necesario que la solución tenga buenas prestaciones sino que sea posible predecir, mediante un modelo energético, la energía necesaria para llevar a cabo diferentes escenarios de simulación.

➤ **Simulación de enfermedades infecciosas transmitida por vectores utilizando modelos basados en agentes.** El Dengue, Zika y Chikungunya, son las enfermedades reemergentes de mayor preocupación a nivel mundial. La carencia de tratamientos médicos obliga a los agentes de salud a abordar la contención de los focos infecciosos desde la identificación y eliminación de los criaderos del vector transmisor. Una herramienta tecnológica que ayude en la toma de decisiones representa la solución a los costos asociados al tiempo de recolección de muestras y el análisis de datos, además de la consiguiente reducción de los gastos económicos. Se implementó un modelo basado en agentes en GPU para la evaluación de la reproducción del vector *Aedes aegypti*, orientado a la toma de decisiones. El modelo ha sido validado con datos de Santo Tomé, Corrientes, dando excelentes resultados [23]. Se realizará un estudio y análisis de performance en otras arquitecturas paralelas basadas en GPU. Además, se evaluará el consumo energético para determinar las relaciones entre escalabilidad, consumo y eficiencia.

➤ **Simulación de N cuerpos computacionales con atracción gravitacional.** Su propósito es aproximar en forma numérica la evolución de un sistema de cuerpos en el que cada uno interactúa con todos los restantes. El uso más conocido de esta simulación quizás sea en la astrofísica, donde cada cuerpo representa una galaxia o una estrella particular que se atraen entre sí debido a la fuerza gravitacional. Si bien existen diferentes métodos para procesar la simulación de los N cuerpos, en todos los casos se requiere alta demanda computacional. Se ha utilizado este problema para analizar diferentes aspectos y/o arquitecturas como: cluster de GPUs homogéneas [24]; análisis de diferentes distribuciones de trabajo en

cluster de GPUs heterogéneas [25]; desarrollo de modelos de estimación de tiempo de ejecución y de energía para GPU, cluster de GPU y multiGPU [26]. Además, se estudió la paralelización de la versión directa de esta simulación sobre diferentes arquitecturas Intel considerando procesadores de última generación Xeon y Xeon Phi. Se focalizó en cómo diferentes técnicas de optimización logran mejorar el rendimiento final y en la comparación de rendimiento y eficiencia energética entre ambas arquitecturas. A futuro, interesa extender el estudio incorporando arquitecturas de GPUs [27].

➤ **Problemas de optimización de simulación de sistemas dinámicos complejos mediante heurísticas.** La búsqueda de un conjunto de parámetros de entrada que optimicen el funcionamiento de un simulador de un sistema físico es un proceso de alto costo computacional que puede considerarse intratable y requiere de heurísticas que permitan disminuir el tiempo de ejecución. En los nuevos avances realizados en esta línea se aprovechó la continuidad en los valores de los parámetros físicos distribuidos sobre el dominio del sistema, de manera de realizar búsquedas de parámetros ajustados sobre espacios de búsquedas de tamaño mucho más reducidos que los utilizados en una primera etapa del trabajo. Esta metodología mucho más eficiente se puede extender a otros simuladores de fenómenos físicos y en particular con simuladores de inundaciones de ríos con las que se llevaron adelante las experiencias. Las experiencias se pueden correr de manera colaborativa beneficiándose ampliamente del uso de plataformas de clusters de procesadores [28] [29].

➤ **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 [30]. 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 [31]. Se ha ampliado para incorporar conceptos básicos de computación en la nube (Cloud Computing) [31]. 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.

➤ **Aplicaciones en Big Data utilizando HPC.** A partir de distintas fuentes de datos tales como sensores, simulaciones científicas, internet, dispositivos móviles, entre otras, se genera continuamente un enorme volumen de datos haciendo que este tipo de escenarios Big Data sea cada vez más común en la actualidad. Para poder trabajar con problemas de este tipo, se requiere de cómputos de altas prestaciones y de herramientas de software específicas capaces de procesar Big Data en tiempos razonables, y que puedan ser ejecutados de manera paralela y distribuida. Existen distintos tipos de herramientas para trabajar con Big Data utilizando HPC, y la mayoría de ellas se basan en el paradigma de programación llamado MapReduce, propuesto por Google en los inicios de esta gran explosión de datos. La más ampliamente utilizada en la actualidad, es conocida como Apache Spark y, entre sus más importantes características se encuentran los mecanismos eficientes de tolerancia a fallos y el uso intensivo de memoria RAM que emplea, siendo una de las más rápidas. Además, dispone de una amplia variedad de librerías que posibilitan utilizar algoritmos de Machine Learning, procesamiento en Streaming, etc. En esta línea de trabajo, se están desarrollando nuevas técnicas para el preprocesamiento de datos para problemas de clasificación en Big Data, utilizando el framework Apache Spark sobre Cómputo de Altas Prestaciones. Como parte del trabajo realizado se lleva a cabo el estudio de la calidad de los datos actualmente disponibles de manera pública en el ambiente científico relacionados a este tipo de problemas de clasificación [32]. Dicho estudio se basa en el análisis de las características intrínsecas de los conjuntos de datos disponibles donde se analizan aspectos tales como redundancia, outliers, áreas sobrepuestas del problema, porcentaje de desbalance entre clases, entre otros factores que, estando presentes, pueden generar una degradación en el rendimiento del clasificador a utilizar.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D se concluyó 1 tesis doctoral, 1 Trabajo Final de Especialización y 1 Tesina de Grado de Licenciatura. Se encuentran en curso en el marco del proyecto 3 tesis doctorales, 2 de maestría, 3 trabajos de Especialización y 3 Tesinas de grado. Se participa en el dictado de las carreras de Doctorado en Cs. Informáticas y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática UNLP, por lo que potencialmente pueden generarse más Tesis y Trabajos Finales. Hay cooperación con grupos de otras Universidades del país y del exterior, y tesis 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, James; Reinders, James. "Intel Xeon Phi Coprocessor High Performance Programming". Morgan Kaufmann. 2013.

[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]. Raspberry PI. <https://www.raspberrypi.org/> Accedido 21 de marzo de 2016.

[5]. Odroid <http://www.hardkernel.com> Accedido 21 de Marzo de 2016.

[6]. Annamalai A., Rodrigues R., Koren I., Kundu S., "Dynamic Thread Scheduling in Asymmetric Multicores to Maximize Performance-per-Watt," 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, pp. 964-971, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012.

[7]. Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias. "ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems". ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.

[8]. McCool, Michael. "Structured Parallel Programming: Patterns for Efficient Computation", Morgan Kaufmann, 2012

[9]. 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.

[10]. Chapman, B., Jost, G. & Van der Pas. "Using OpenMP – Portable Shared Memory Parallel Programming". (2008). UK: MIT Press.

[11]. Hager, G. & Wellein, G. "Introduction to HPC for Scientists and Engineers". (2011) EEUU: CRC Press.

[12]. De Wael, M; Marr, S; De Fraine, B; Van Cutsem, T; De Meuter, W. "Partitioned Global Address Space Languages". ACM Computing Surveys (CSUR) 47 (4), 2015.





















[13]. Amazon Elastic Compute Cloud (Amazon EC2). <http://aws.amazon.com/es/ec2/>. Febrero 2013.

[14]. 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

[15]. V. Sanz, A. Pousa, M. Naiouf, and A. De Giusti, "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.
- [16]. Sanz V., Pousa A., Naiouf M., De Giusti A. (2020) 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.
- [17]. Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sánchez, Guillermo Botella Juan, Manuel Prieto-Matías. “State-of-the-art in Smith-Waterman Protein Database Search”. Big Data Analytics in Genomics. Ka-Chun Wong (Editor). ISBN: 978-3-319-41278-8 (print) 978-3-319-41279-5 (online), Springer, págs. 197-223, 2016.
- [18]. E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. “Accelerating Smith-Waterman Alignment of Long DNA Sequences with OpenCL on FPGA”. En: Bioinformatics and Biomedical Engineering. IWBBIO 2017. Lecture Notes in Computer Science, vol 10209, ISBN: 978-3-319-56154-7, Springer, Cham, págs. 500-511, 2017.
- [19]. E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. “SWIFOLD: Smith-Waterman Implementation on FPGA with OpenCL for Long DNA Sequences”. BMC Systems Biology. ISSN: 1752-0509. In press. 2018.
- [20]. E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. “First Experiences Accelerating Smith-Waterman on Intel’s Knights Landing Processor”. Algorithms and Architectures for Parallel Processing. ICA3pp 2017. Lecture Notes in Computer Science, vol 10393, ISBN: 978-3-319-65482-9, Springer International Publishing, págs. 569-579, 2017.
- [21]. Enzo Rucci, Carlos García, Guillermo Botella, Armando De Giusti, Marcelo Naiouf and Manuel Prieto-Matías. “SWIMM 2.0: enhanced Smith-Waterman on Intel’s Multicore and Manycore architectures based on AVX-512 vector extensions”. International Journal of Parallel Programming, Springer US, 2018. DOI: <https://doi.org/10.1007/s10766-018-0585-7>
- [22]. one API Programming Model, Disponible en www.oneapi.org, Accedido el 15/02/2021.
- [23]. Erica Montes de Oca, Remo Suppi, Laura De Giusti, Marina Leporace, María Victoria Micieli, María Soledad Santini, Marcelo Naiouf. “Simulación de altas prestaciones (GPU) para la reproducción del mosquito *Aedes aegypti* en el cementerio de Santo Tomé, Corrientes”. XXV Congreso Argentino de Ciencias de la Computación 2019 (CACIC 2019). Universidad Nacional de Río Cuarto, Argentina. pp. 183-193. ISBN 978-987-688-377-1.
- [24]. E. Montes de Oca, L. De Giusti, F. Chichizola, A. De Giusti, M. Naiouf. "Utilización de Cluster de GPU en HPC. Un caso de estudio". Proceedings del XX Congreso Argentino de Ciencias de la Computación (CACIC 2014), pp. 1220-1227, 2014.
- [25]. E. Montes de Oca, L. C. De Giusti, F. Chichizola, A. E. De Giusti, M. Naiouf. “Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo”. Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), pp. 159-168, 2016.
- [26]. Montes de Oca E. S. “Análisis de consumo energético en Cluster de GPU y MultiGPU en un problema de alta demanda computacional”. Trabajo Final Carrera de Especialista en Cómputo de Altas Prestaciones y Tecnología GRID. Facultad de Informática, Universidad Nacional de La Plata. (2018).
- [27]. “Optimization of the N-Body Simulation on Intel’s Architectures Based on AVX-512 Instruction Set”. E. Rucci, E. Moreno, A. Pousa, and F. Chichizola, En: Computer Science – CACIC 2019. Communications in Computer and Information Science., ISBN: 978-3-030-48325-8, Springer International Publishing, págs. 37-52, 2020.
- [28]. Trigila M., Gaudiani A., Luque E. (2018) “Agile Tuning Method in Successive Steps for a River Flow Simulator”. In: Shi Y. et al. (eds.) Computational Science – ICCS 2018. ICCS 2018. Lecture Notes in Computer Science, vol 10862. Springer, Cham
- [29]. Trigila, Mariano, Gaudiani, Adriana, Luque, Emilio. “Adjustment of a simulator of a complex dynamic system with emphasis on the reduction of computational resources”. Actas del Workshop. p. 142-147. 2018.
- [30]. J. Castro, L. D. Giusti, G. Gorga, M. Sánchez, and M. Naiouf, “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.
- [31]. L. C. De Giusti, F. Chichizola, S. Rodríguez Eguren, M. Sanchez, J. M. Paniego, A. E. De Giusti. “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. Octubre 2016. Pp 1357-1365.
- [32]. Basgall, M. J., Naiouf, M., Herrera, F., & Fernández, A. (2020). Towards Smart Data Technologies for Big Data Analytics. In VIII Conference on Cloud Computing, Big Data & Emerging Topics (Modalidad virtual, 8 al 10 de septiembre de 2020).
- [33]. “Comparación de Arquitecturas HPC para Computar Caminos Mínimos en Grafos. Intel Xeon Phi KNL vs NVIDIA Pascal”. Costanzo, M., Rucci, E., Costi, U., Chichizola, F., Naiouf, M. In: Actas del XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020). p. In press (2020).
- [34]. Sanz V., Pousa A., Naiouf M., De Giusti A. (2020) 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

Arquitecturas Edge- Fog - Cloud en Procesamiento Distribuido

Armando De Giusti ⁽¹⁾⁽²⁾ , Marcelo Naiouf⁽¹⁾ , Santiago Medina⁽¹⁾ , Joaquín De Antueno ⁽¹⁾ , Laura De Giusti⁽¹⁾⁽³⁾ , Julieta Lanciotti⁽¹⁾ , Fernando G. Tinetti ⁽¹⁾⁽³⁾ , Franco Chichizola⁽¹⁾ , Enzo Rucci⁽¹⁾⁽³⁾ , Fernando Romero⁽¹⁾ , Adrián Pousa⁽¹⁾ , Victoria Sanz ⁽¹⁾⁽³⁾ , Diego Montezanti ⁽¹⁾ , Diego Encinas ⁽¹⁾ , Ismael Rodríguez⁽¹⁾ , Sebastián Rodríguez Eguren⁽¹⁾ , Juan Manuel Paniego⁽¹⁾ , Martín Pi Puig⁽¹⁾ , Leandro Libutti⁽¹⁾ , Manuel Costanzo ⁽¹⁾ 

¹Instituto 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

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴Universidad Nacional del Comahue

{degiusti,mnaiouf, smedina, jdeantueno, ldgiusti, jlanciotti, fernando, francoch, erucci, fromero, apousa, vsanz, dmontezanti, dencinas, ismael, seguren, jmpaniego, mpipuig, llibutti, mcostanzo}@lidi.info.unlp.edu.ar;

Resumen

El eje de esta línea de I/D lo constituye el estudio de la integración de arquitecturas distribuidas que van desde el nivel de sensores (Edge Computing) a la capa de procesamiento en la nube (Cloud Computing), con una posible capa intermedia que se conoce como Fog Computing.

Los temas centrales son:

- Distribución de la capacidad de procesamiento en cada nivel.
- Análisis de performance en las comunicaciones, según el grado de distribución del procesamiento.
- Administración de recursos en cada nivel.
- Migración de “inteligencia” al nivel “Edge” para reducir consumo y comunicaciones.
- Desarrollo y evaluación de aplicaciones que integran niveles de procesamiento.
- Análisis de eficiencia en tiempo, consumo y comunicaciones.

Palabras clave: *Sistemas Distribuidos. Cloud Computing. Fog Computing. Edge Computing. IoT, Algoritmos distribuidos. Eficiencia.*

Contexto

Se presenta una línea de Investigación que es parte del proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos,

Métricas de Rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real.” del III-LIDI y de proyectos específicos apoyados por organismos nacionales e internacionales. También del proyecto “Arquitecturas Multiprocesador Actuales: Tendencias, Planificación, Análisis de rendimiento y Consumo Energético” financiado por la Facultad de Informática de la UNLP y el proyecto “Unidad Inteligente para Control de Consumo Energético” financiado por la Secretaría de Políticas Universitarias y la UNLP.

En el tema hay cooperación 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 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 tiene financiamiento de diferentes empresas de Argentina, en particular en la formación de recursos humanos en la temática de Cloud/Fog y Edge Computing.

Se participa en iniciativas como el Consorcio en temas de Cloud Computing-Big Data y Temas Emergentes, con Universidades de Argentina y España.

Asimismo, 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.

Introducción

Una de las áreas de creciente interés son los múltiples sensores/módulos de sentido “inteligentes” que recogen información para realizar su procesamiento en la nube. A medida que el número de señales/datos escala, la respuesta en tiempo del procesamiento centralizado en el Cloud está condicionado por las comunicaciones y la integración de información heterogénea.

La integración de capas de procesamiento, en lo que se denomina Edge y/o Fog Computing ha dado lugar a un nuevo modelo de arquitectura denominado “Edge-Fog Cloud Computing” [1] que trata de generar una serie de ventajas:

- Reducir la carga de comunicaciones en la red, con un mayor procesamiento en los mismos sensores o en una capa intermedia, anterior al Cloud.
- Recursos orientados a la integración con dispositivos móviles que transmiten desde diferentes puntos, en tiempo real.
- Posibilidad de procesar en función del contexto, incluyendo la ubicación física del sensor/sensores que recogen los datos en tiempo real.
- Minimizar el impacto de las fallas, ya que tenemos una arquitectura totalmente descentralizada.

Lógicamente este modelo de arquitectura trae nuevos desafíos, tales como [2], [3], [4], [5] :

- Administración de “capas” de procesamiento heterogeneo, con señales y datos en diferentes formatos.
- Definición de protocolos de interoperabilidad entre la capa “Edge” o de sensores inteligentes, la capa “Fog” o de procesamiento intermedio y los servicios en el Cloud.

- Análisis y recuperación de fallos en un sistema distribuido, débilmente acoplado.
- Análisis de la distribución óptima de tareas en cada capa, en función del tiempo de respuesta requerido, overhead de comunicaciones y el consumo energético.
- Análisis de la integración de diferentes tipos de sensores/módulos de sentido.
- Integración de diferentes herramientas para el desarrollo de aplicaciones y el almacenamiento de datos.
- Seguridad en los datos.
- Estudios de escalabilidad (no lineal) con el crecimiento de los sensores en la capa “Edge”.

Cloud Computing

Cloud Computing, proporciona grandes conjuntos de recursos físicos y lógicos (como pueden ser infraestructura, plataformas de desarrollo, almacenamiento y/o aplicaciones), fácilmente accesibles y utilizables por medio de una interfaz de administración web, con un modelo de arquitectura “virtualizada” [6][7]. Estos recursos son proporcionados como servicios (“as a service”) y pueden ser dinámicamente reconfigurados para adaptarse a una carga de trabajo variable (escalabilidad), logrando una mejor utilización y evitando el sobre o sub dimensionamiento (elasticidad) [8].

En nuestro modelo de procesamiento distribuido que integra desde sensores a la nube, Cloud Computing se reserva para el procesamiento centralizado de algoritmos complejos, con gran volumen de datos. Parte de estos datos podrán haber sido “preprocesados” en las capas “Edge” o “Fog”. [9][10].

Fog Computing

El modelo de Fog Computing surge como respuesta al crecimiento de los desarrollos relacionados con el Internet de las Cosas. Estos desarrollos requieren procesamiento en la nube, pero presentan características que hacen complejo su armado haciendo uso exclusivo de tecnologías de Cloud Computing. [11][12]. Fog Computing consiste entonces en una plataforma intermedia que provee procesamiento, almacenamiento y servicios de comunicación en red entre los dispositivos “Edge” que adquieren los datos en el modelo IoT. Las características de Fog Computing son las de una capa intermedia orientada a [13][14][15]:

- Aplicaciones de tiempo real distribuidas.
- Gran número de nodos, que puede escalar dinámicamente.
- Heterogeneidad de los nodos.
- Movilidad.
- Red predominante de comunicaciones inalámbricas.
- Baja latencia y conocimiento de la ubicación.

Edge Computing

El modelo de “Edge” computing está impuesto por el crecimiento exponencial del número de dispositivos sensores con inteligencia local disponibles. Internet “de las cosas” (IoT) crece y hoy hablamos de 50.000 millones de dispositivos conectados a Internet, con un tráfico del orden de 800 Zbytes. Se requiere entonces capacidad de procesamiento cerca de los sensores, integración de datos locales y geográficamente ubicados y posibilidad de pre procesar los mismos y enviarlos con un menor overhead de comunicaciones a las capas superiores (Fog o Cloud) [16][17][18][19]. Las ventajas del modelo Edge son directas [20][21]:

- Respuestas automáticas al usuario, en tiempo real, utilizando su capacidad local.
- Disminución del tráfico de datos.
- Mayor seguridad por el procesamiento local.
- Disminución del consumo energético, por las características de los componentes.
- Mejora de la eficiencia global para un sistema distribuido, débilmente acoplado.

Aplicaciones de tiempo real

Este nuevo modelo Edge-Fog Cloud Computing es especialmente aplicable a problemas de tiempo real (Cloud robotics, Vehículos autónomos, Monitoreos de salud personalizados, etc.), al permitir que las capas cercanas a los sensores y al usuario resuelvan en menor tiempo y con menor overhead de comunicaciones, la respuesta “inmediata”, dejando para el Cloud el procesamiento de los datos masivos “off-line” (por ejemplo para ajustar/perfeccionar un modelo de comportamiento que luego se transforma en ajustes a los algoritmos en las capas Edge y Fog. [10][22].

Eficiencia energética

La mejora de la eficiencia energética es un tema central en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores. Muchos esfuerzos están orientados a tratar la eficiencia energética como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirla. En el caso del modelo Edge-Fog Cloud Computing, hay una mejora del consumo por dos efectos:

- Normalmente los procesadores de la capa “Edge” son de bajo consumo.
- Al bajar el tráfico de comunicaciones, también se baja el consumo energético asociado al mismo.

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas aplicables en Edge Computing. Modelos de referencia.
- Arquitecturas aplicables en Fog Computing. Modelos de referencia.
- Administración de recursos y datos en Edge y Fog computing. Integración de los mismos.
- Seguridad en los datos.
- Manejo de aplicaciones móviles y relación entre las capas Edge y Fog.
- Vinculación de las capas Edge y Fog con el Cloud. Servicios requeridos.
- Distribución óptima de tareas entre capas, para aplicaciones complejas.
- Escalabilidad en aplicaciones Edge-Fog-Cloud.
- Métricas de eficiencia considerando tiempo de respuesta / costo comunicaciones / consumo energético.
- Aplicaciones: robots / drones / vehículos autónomos.
- Algoritmos colaborativos en tiempo real integrando las tres capas.
- Aplicaciones: sistemas inteligentes distribuidos para reducir el consumo energético.
- Integración de aplicaciones móviles y relación entre las capas Edge y Fog.
- Aplicaciones: robots / drones colaborativos trabajando en comunicación con una capa Fog y con el Cloud.
- Aplicaciones: Sistema de tableros inteligentes para optimización de consumo energético en edificios distribuidos.

Resultados obtenidos

- Se han analizado plataformas orientadas a la gestión de recursos en Edge y Fog Computing.
- Se han estudiado protocolos y tráfico de comunicaciones en aplicaciones distribuidas en tiempo real que requieren interactuar con el Cloud.
- Se desarrolló un modelo de tablero inteligente para reducción del consumo y su integración con la capa Fog y comunicación con la nube.
- Se está estudiando el tema de consumo distribuido, en el caso de los tableros inteligentes, de los robots distribuidos y del trabajo con drones.

Resultados y Objetivos

Investigación experimental por realizar

- Análisis comparativo de tecnologías empleadas en Edge/Fog Computing.
- Análisis comparativo de los servicios para arquitecturas Edge y Fog.
- Estudio de protocolos de comunicación con el Cloud, desde el nivel de Edge y/o Fog.
- Análisis y despliegue de plataformas de sensores/módulos de sensado.
- Análisis de performance de comunicaciones, en función del escalado de la arquitectura de sensores.
- Estudios de consumo energético en aplicaciones de Edge/Fog/Cloud.
- Estudio comparativo de la distribución de carga de procesamiento en aplicaciones distribuidas (en particular de tiempo real).

Organización de Eventos

En el año 2020 se han organizado las VIII Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2020) en Argentina [24][24], con participación de especialistas académicos del país y del exterior y de empresas con experiencia en Cloud Computing. En junio de 2021 se organizarán las IX JCC-BD&ET.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D están en curso 3 Tesis de Maestría.

Asimismo, se desarrollaron 2 trabajos finales de alumnos.

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 CONEAU), por lo que potencialmente pueden generarse Tesis de Doctorado y Maestría y Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior con posibilidad de realizar Tesis en colaboración.




















Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación dos asignaturas directamente relacionadas con los temas de la misma: "Cloud Computing y Cloud Robotics" y "Conceptos y Aplicaciones en Big Data". Asimismo todos los años se desarrollan proyectos con alumnos, relacionados básicamente con aplicaciones de tiempo real con robots y drones.

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, <https://doi.org/10.1016/j.sysarc.2019.02.009>.
- [10] Pi Puig M. et al. (2019) Intelligent Distributed System for Energy Efficient Control. In: Naiouf M., Chichizola F., Rucci E. (eds) *Cloud Computing and Big Data. JCC&BD 2019. Communications in Computer and Information Science*, vol 1050. Springer, Cham
- An Overview on Edge Computing Research
- [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] J. de Antueno, S. Medina, L. De Giusti and A. De Giusti, "Analysis, Deployment and Integration of Platforms for Fog Computing", *Journal of Computer Science and Technology*, 20(2), e12, October 2020.
- [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 Thingier.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] 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.
- [25] A. E. De Giusti, M. Naiouf, F. Chichizola, E. Rucci, and L. C. De Giusti. Short Papers of the 8th Conference on Cloud Computing Conference, Big Data & Emerging Topics (JCC-BD&ET 2020). Facultad de Informática (UNLP), ISBN: 978-950-34-1927-4, 2020.

Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones

Armando De Giusti ⁽¹⁾⁽²⁾ , Marcelo Naiouf⁽¹⁾ , Fernando G. Tinetti ⁽¹⁾⁽³⁾ , Horacio Villagarcía ⁽¹⁾⁽³⁾ , 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⁾ , Erica Montes de Oca⁽¹⁾ , Juan Manuel Paniego⁽¹⁾ , Martín Pi Puig⁽¹⁾ , César Estrebow⁽¹⁾ , Leandro Libutti⁽¹⁾ , Manuel Costanzo⁽¹⁾, Joaquín De Antueno⁽¹⁾, Julieta Lanciotti⁽¹⁾, Javier Balladini⁽⁴⁾ 

¹Instituto 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

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴Universidad Nacional del Comahue

{degiusti,mnaiouf,fernando,hvw,francoch,ldgiusti,erucci,apousa,vsanz,dmontezanti,dencinas,ismael,seguren,emont
esdeoca,jmpaniego,mpipuig,cesarest,llibutti, mcostanzo, jdeantueno,jlanciotti}@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), y asimétricas.
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas y su evaluación de rendimiento computacional y energético.
- 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. Clusters. Arquitecturas asimétricas. GPU, MIC, FPGA, TPU. Eficiencia energética. Resiliencia. Código heredado. E/S paralela.*

Contexto

Se presenta una línea de Investigación que es parte del proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de Rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real.” del III-LIDI y de proyectos específicos

apoyados por organismos nacionales e internacionales. También del proyecto “Arquitecturas Multiprocesador Actuales: Tendencias, Planificación, Análisis de rendimiento y Consumo Energético” financiado por la Facultad de Informática de la UNLP.

En el tema hay cooperación con varias Universidades de Argentina y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por CyTED, AECID y la OEI (Organización de Estados Iberoamericanos).

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado y se ha tenido el apoyo de diferentes empresas (IBM, Microsoft, Telecom, Intel) en la temática de Cloud Computing.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Asimismo, 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.

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 estas.

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 [1].

Lógicamente, esto trae aparejado una revisión de los conceptos del diseño de algoritmos paralelos (incluyendo los lenguajes mismos de programación y el software de base), así como la evaluación de las soluciones que éstos implementan. También resulta necesario investigar las estrategias de distribución de datos y de procesos a fin de optimizar la performance.

Además, el estudio del consumo y la eficiencia energética de los nuevos sistemas paralelos se vuelve tan importante como el de métricas clásicas (speedup, eficiencia, escalabilidad) debido a los costos económicos y los problemas operativos asociados [2].

GPUs y Cluster de GPUs

Las GPUs son el acelerador dominante en la comunidad de HPC hoy en día por su alto rendimiento y bajo costo de adquisición. En la actualidad, tanto NVIDIA como AMD trabajan especialmente en mejorar la eficiencia energética de sus placas y disminuir el alto costo de programación.

La combinación de GPUs con otras plataformas paralelas como clusters y multicores, brindan un vasto conjunto de posibilidades de investigación en arquitecturas híbridas, a partir de diferentes combinaciones como son:

- Máquinas multicore con más de una GPU, que combinan herramientas de programación paralela como OpenMP/CUDA o Pthread/CUDA.
- Cluster de máquinas multicore cada una con una o más placas de GPU, lo que permite combinar OpenMP/MPI/CUDA o Pthread/MPI/CUDA.

Los desafíos que se plantean 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.

MIC

En forma reciente Intel brinda una alternativa a partir de la arquitectura MIC (*Many Integrated Core Architecture*). Esta arquitectura permite utilizar métodos y herramientas estándar de programación con altas prestaciones (lo que los distingue especialmente de las GPUs). De esta forma, se remueven barreras de entrenamiento y se permite focalizar en el problema más que en la ingeniería del software. Xeon Phi es el nombre elegido por Intel para su serie de procesadores many-core. Recientemente, Intel ha lanzado Knights Landing (KNL), la segunda generación de Xeon Phi. A diferencia de sus predecesores que operaban como co-procesador a través del puerto PCI, los procesadores KNL pueden operar en forma autónoma. Además, integran las nuevas extensiones vectoriales AVX-512 y tecnología de memoria 3D, entre otras características avanzadas [3].

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. Tradicionalmente fueron utilizadas para el procesamiento digital de señales. Sin embargo, en los últimos años, existen dos tendencias claras para extender su uso a otros dominios. En primer lugar, el establecimiento de alianzas estratégicas entre fabricantes de procesadores y de FPGAs para integrar estos dispositivos en arquitecturas híbridas (Intel con Altera; IBM con Xilinx) [4][5]. En segundo lugar, el desarrollo de nuevas herramientas de programación para FPGAs empleando estándares familiares para HPC, con las cuales se espera reducir los tradicionales tiempos y costos de programación [6][7]. Por último, la

incorporación de FPGAs a los servicios de Cloud abre nuevas oportunidades para la explotación de esta clase de aceleradores.

TPUs

Las unidades de procesamiento tensorial (TPU) son una clase de circuitos integrados personalizados específicos de aplicaciones (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 [8]. 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.

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 como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirla.

Entre los puntos de interés pueden mencionarse:

- Análisis de metodologías y herramientas para medir y optimizar el consumo energético.
- Estudio de técnicas para reducir el consumo energético en aplicaciones de HPC de acuerdo con las arquitecturas utilizadas.
- Evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas.
- Optimización de la eficiencia energética. A partir de los valores de energía que brindan los contadores hardware es posible definir estrategias de programación que lleven a reducir el consumo, manteniendo a su vez el rendimiento en valores aceptables [9].

Código heredado

La mayoría de los programas de simulación numérica empleados hoy en día fueron

desarrolladas cuando las arquitecturas paralelas no existían. Es por lo que este conjunto de aplicaciones presenta la oportunidad de desarrollar técnicas y herramientas que permitan optimizar el código, tanto desde el punto de vista computacional como desde la ingeniería de software.

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 interesante desarrollar estrategias de detección y recuperación de fallos, especialmente a través de librerías de software.

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 dependerá de la carga de trabajo (patrones de E/S de las aplicaciones) y de su configuración (hardware y software). Contar con herramientas que permitan modelar y predecir el comportamiento de este tipo de aplicaciones en HPC resulta fundamental para mejorar su rendimiento [10].

Dispositivos de bajo costo con capacidades para cómputo paralelo

En la actualidad se comercializan placas de bajo costo como Raspberry PI [11] u Odroid [12] que poseen múltiples núcleos simples. Asimismo, existen diversos dispositivos móviles con capacidades similares. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones.

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas many-core (procesadores MIC, 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 y modelos de programación para HPC.
- Consumo energético en las diferentes arquitecturas paralelas, en particular en relación con los algoritmos paralelos y la configuración de la arquitectura. Análisis de metodologías y herramientas de medición. Modelado y estimación del consumo de potencia de arquitecturas HPC.
- 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 y costo de programación.
- Analizar las capacidades de lenguajes no convencionales para procesamiento paralelo, considerando rendimiento y costo de programación.
- 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 [13][9].
- Desarrollar técnicas de tolerancia a fallas que permitan aumentar la resiliencia de sistemas paralelos y distribuidos.
- Desarrollar técnicas de modelado y simulación de E/S en HPC que permita predecir cómo los cambios realizados en los diferentes componentes de éste afectan a la funcionalidad y al rendimiento del sistema.

Resultados obtenidos

- Se analizaron técnicas de optimización para código heredado relacionado a cómputo numérico [14][15].
- Se optimizaron soluciones paralelas para el procesamiento de *pattern matching* en arquitectura Xeon Phi KNL [16].
- Se optimizó una solución al problema de los N cuerpos computacionales con atracción gravitacional (versión directa) sobre diferentes arquitecturas Intel basadas en instrucciones AVX-512 [17].
- Se diseñó la herramienta SEDAR y evaluó su desempeño para detección y recuperación de fallos transitorios [18].
- Se desarrolló y validó un modelo estadístico para consumo de potencia en placas RPi de diferentes generaciones [19].
- Se realizaron y analizaron modificaciones al framework TensorFlow para permitir la maleabilidad de hilos [20].
- Se realizó una comparación entre las arquitecturas Intel Xeon Phi KNL y NVIDIA Pascal usando como caso de estudio el problema de caminos mínimos en grafos [21].
- Se analizó el rendimiento y consumo energético requerido para simulaciones basadas en agentes sobre arquitecturas de GPUs [22].
- Se avanzó en el modelado y simulación inicial del sistema de archivos paralelos PVFS2 y se implementó por medio del paradigma de agentes en un simulador [23].
- Se desarrolló una plataforma distribuida, escalable, híbrida y redundante que permite aprovechar recursos de cómputo basados en arquitecturas ARM a través de un servicio de cloud [24].
- Se desarrolló un modelo que permite predecir el consumo energético de un sistema ante diferentes estrategias aplicadas para reducirlo cuando ocurre una falla [25].
- Se compararon los enfoques de desarrollo HLS y HDL en FPGA en el ámbito del procesamiento de imágenes [26] [27][28].

Organización de Eventos

En el año 2020 se han organizado las VIII Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2020) en

Argentina, con participación de especialistas académicos del país y del exterior y de empresas con experiencia en Cloud Computing [29][30]. En junio de 2021 se organizarán las IX JCC-BD&ET.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyó 1 Tesis Doctoral. Al mismo tiempo se encuentran en curso 4 tesis de Doctorado en Ciencias Informáticas y 3 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 y Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior, y hay tesis 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 esta: “Taller de Programación sobre GPUs”, “Cloud Computing y Cloud Robotics” y “Conceptos y Aplicaciones en Big Data”.

Referencias

[1] Rucci, Enzo: “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.

[2] Ballardini, 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. Págs. 53-65. 2013.

[3] Reinders, J., Jeffers, J., Sodani, A. “Intel Xeon Phi Processor High Performance Programming Knights Landing Edition”. Morgan Kaufmann Publishers Inc., Boston, MA, USA, 2016

[4] 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>

[5] Intel. “Intel Acquisition of Altera”. Disponible en <http://intelacquiresaltera.transactionannouncement.com>

[6] Sean Settle: “High-performance Dynamic Programming on FPGAs with OpenCL”. In: IEEE High Performance Extreme Computing Conference. 2013.

[7] Xilinx Inc. “SDAccel Development Environment”. [Online]. Disponible en <http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html>

[8] Google Inc. “Cloud Tensor Processing Unit (TPU)”. Disponible en <https://cloud.google.com/tpu/docs/tpus?hl=es-419>

[9] Saez, J.C., Pousa, A., Rodríguez-Rodríguez, R., Castro, F., Prieto-Matias, M. “PMCTrack: Delivering performance monitoring counter support to the OS scheduler”. The computer journal Volume 60, Issue 1 January 2017.

[10] D. Encinas, M. Naiouf, A. De Giusti, S. Mendez, D. Rexachs, and E. Luque. “On the Calibration, Verification and Validation of an Agent-Based Model of the HPC Input/Output System”. Proceedings from The Eleventh International Conference on Advances in System Simulation (SIMUL 2019), November 24 - 28, 2019.

[11] Raspberry PI. <https://www.raspberrypi.org/>

[12] Odroid <http://www.hardkernel.com> Accedido 21 de marzo de 2016.

[13] 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.

[14] 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.

[15] F. J. Díaz and F. G. Tinetti, Short papers of the “Performance Analysis and Optimizations Techniques for Legacy Code Numerical Simulations”. 8th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2020), ISBN: 978-950-34-1927-4, págs. 18-21, 2020.

[16] V. Sanz, A. Pousa, M. Naiouf, and A. De Giusti, “Accelerating Pattern Matching on Intel Xeon Phi Processors”, Algorithms and Architectures for Parallel Processing. ICA3PP 2020., ISBN: 978-3-030-60245-1, págs. 262-274, doi. 10.1007/978-3-030-60245-1_18, 2020.

[17] “Optimization of the N-Body Simulation on Intel’s Architectures Based on AVX-512 Instruction Set”. E. Rucci, E. Moreno, A. Pousa, and F. Chichizola, En: Computer Science – CACIC 2019. communications in Computer and Information Science., ISBN: 978-3-030-

- 48325-8, Springer International Publishing, págs. 37-52, 2020.
- [18] 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.
- [19] J. M. Paniego, L. Libutti, M. P. Puig, F. Chichizola, L. De Giusti, M. Naiouf, and A. De Giusti, "Unified Power Modeling Design for Various Raspberry Pi Generations Analyzing Different Statistical Methods". En: *Computer Science – CACIC 2019. communications in Computer and Information Science.*, ISBN: 978-3-030-48325-8, Springer International Publishing, págs. 53-65, 2020.
- [20] L. A. Libutti, F. D. Igual, L. Piñuel, L. D. Giusti, and M. Naiouf, "Towards a Malleable Tensorflow Implementation", *Cloud Computing, Big Data & Emerging Topics. 8th Conference, JCC-BD&ET 2020, La Plata, Argentina, September 8-10, 2020, Proceedings*, págs. 30-40, doi. [10.1007/978-3-030-61218-4_3](https://doi.org/10.1007/978-3-030-61218-4_3), 2020.
- [21] Costanzo, M., Rucci, E., Costi, U., Chichizola, F., Naiouf, M.. "Comparación de Arquitecturas HPC para Computar Caminos Mínimos en Grafos. Intel Xeon Phi KNL vs NVIDIA Pascal" In: *Actas del XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020)*. p. In press (2020).
- [22] E. S. Montes de Oca, R. Suppi, L. C. De Giusti, and M. Naiouf, "Green High-Performance Simulation for AMB models of *Aedes aegypti*". *Journal of computer science and technology* (ISSN 1666-6038), vol. 20, num. 1, págs. 15-22, doi. [10.24215/16666038.20.e02](https://doi.org/10.24215/16666038.20.e02), 2020.
- [23] 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
- [24] D. Petrocelli, A. E. De Giusti, and M. Naiouf, "Collaborative, distributed and scalable platform based on mobile, cloud, micro services and containers for intensive computing tasks". *Short papers of the 8th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2020)*, ISBN: 978-950-34-1927-4, págs. 10-13, 2020.
- [25] M. Morán, J. Balladini, D. Rexachs, E. Rucci. "Towards Management of Energy Consumption in HPC Systems with Fault Tolerance" *Proceedings of the IV IEEE ARGENCON 2020 CONGRESS*, En prensa, 2020.
- [26] R. Millón, E. Frati, and E. Rucci, "A comparative study between HLS and HDL on SoC for image processing applications". *Elektron* (ISSN 2525-0159), vol. 4, num. 2, págs. 100-106, doi. [10.37537/rev.elektron.4.2.117.2020](https://doi.org/10.37537/rev.elektron.4.2.117.2020), 2020.
- [27] R. Millon, F. E. Frati, and E. Rucci, "Implementación de Filtro de Detección de Bordes Sobel en SoC usando Síntesis de Alto Nivel". *Actas del Congreso Argentino de Sistemas Embebidos (CASE2020)*, ISBN: 978-987-46297-7-7, págs. 73-75, 2020.
- [28] R. Millon, E. Rucci, and F. E. Frati. "Análisis Comparativo de Implementaciones HLS de Filtro Sobel en SoC". *Actas del Congreso Argentino de Ciencias de la Computación (CACIC2020)*, En prensa, 2020.
- [29] 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.
- [30] A. E. De Giusti, M. Naiouf, F. Chichizola, E. Rucci, and L. C. De Giusti. *Short Papers of the 8th Conference on Cloud Computing Conference, Big Data & Emerging Topics (JCC-BD&ET 2020)*. Facultad de Informática (UNLP), ISBN: 978-950-34-1927-4, 2020.

Cálculo Paralelo para la Multiplicación de Matrices

Ponce de León Alejo¹, Díaz-Acevedo Karvin¹, Chirino Pamela¹, Galdamez Mariela¹, Caymes-Scutari Paola^{1,2}, Bianchini Germán¹

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

alejo.ponce@alumnos.frm.utn.edu.ar, karvin.diaz@alumnos.frm.utn.edu.ar,
pamela.chirino@alumnos.frm.utn.edu.ar, mariela.galdamez@alumnos.frm.utn.edu.ar
pcaymesscutari@frm.utn.edu.ar, gbianchini@frm.utn.edu.ar

RESUMEN

La utilización de la multiplicación de matrices toma un alto grado de importancia en aplicaciones tales como resolución de sistemas de ecuaciones de muchas variables, cálculo numérico y también actualmente se utiliza mucho en el cálculo de *microarrays*, en el área de bioinformática. En este proyecto nos enfocamos en el uso de algoritmos para realizar dicha operación a fin de aplicar el concepto de paralelismo y así conseguir mayores velocidades en el proceso de resolución. Además, se evaluarán los métodos de balanceo de carga con el objetivo de comparar eficiencia entre diversos códigos y el incremento de velocidad con la utilización del sistema paralelo (*Speedup*).

Palabras clave: Cálculo, Multiplicación, Matrices, Algoritmo, Paralelismo.

CONTEXTO

Este proyecto se encuentra en ejecución desde Noviembre de 2020, llevándose a cabo inicialmente en el marco de la asignatura Computación Paralela y

posteriormente trasladado hacia el LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido) dentro del ámbito de la UTN-FRM.

1. INTRODUCCIÓN

El cómputo paralelo resulta de gran utilidad en varios aspectos de la investigación en la actualidad. Se consideró su potencial uso para la resolución de cálculos complejos, como lo es la multiplicación de matrices de gran volumen. La multiplicación se define dadas dos matrices A y B, se dicen multiplicables si el número de columnas de A coincide con el número de filas de B. En la nueva matriz C, los elementos C_{ij} parten del producto que se obtiene multiplicando cada elemento de la fila i de la matriz A por cada elemento de la columna j de la matriz B y sumándolos. [1] Partiendo de esas premisas se desarrollaron dos algoritmos: uno que aprovechase la capacidad de cómputo de un *cluster* de computadoras parte del LICPaD ubicados en la UTN-FRM, y otro cuya ejecución fuese secuencial. De ésta manera y realizando una serie de pruebas comparativas, lograríamos determinar la

verdadera ventaja (de existir una) de utilizar paralelismo para dichos cálculos, y también lograr discernir los límites operativos del algoritmo y sus posibles optimizaciones.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Esta temática está enfocada directamente en la utilización del concepto de paralelismo. El objetivo es comparar la resolución del cálculo en formato secuencial y en formato paralelo. A su vez, el formato paralelo será puesto a prueba mediante el uso de distintos tipos de balanceo de cargas, es decir, la forma en que el algoritmo divide las tareas al resto de procesadores, para obtener ciertas mediciones de rendimiento del algoritmo y concluir en la selección de la mejor opción. Este proyecto es posible mediante el uso de la librería MPI [2], la interfaz de paso de mensajes que define la sintaxis y la semántica de las funciones diseñadas para ser usada en programas que exploten la existencia de múltiples procesadores. Otro concepto importante es la utilización de un modelo de resolución, una guía que plantee el orden de cómo deberían actuar los procesadores. El modelo utilizado es el Master-Worker, el procesador Master que se encarga de dividir las tareas y los Workers que se encarga de realizar los cálculos. Finalmente los conceptos claros, se comenzó la programación en lenguaje C. [3]

3. RESULTADOS ESPERADOS

El principal aporte que se espera alcanzar es llegar a comprender los beneficios del uso de paralelismo aplicado a la

resolución de problemas que sean de gran importancia y de gran envergadura, en este caso, multiplicación de matrices. Dicho objetivo será alcanzado testeando los algoritmos mediante distintos tamaños de matrices, para analizar su rendimiento en cuanto al *Speedup*, la Eficiencia, y el Balanceo de Carga [4]. Bajo la hipótesis de que cuanto más crezca el problema, más beneficioso es el uso del paralelismo, la información que será de utilidad para realizar el estudio comparativo serán los tiempos de ejecución de cada uno de los experimentos, en función del tamaño de los datos considerado (tamaño de las matrices) y del tamaño de la máquina paralela utilizada en la ejecución (cantidad de unidades computacionales involucradas).

4. FORMACIÓN DE RECURSOS HUMANOS

La temática propuesta permite continuar el desarrollo académico de los integrantes del grupo de trabajo, como también así de forma complementaria su formación en el área de investigación de la computación paralela. Los primeros cuatro autores del artículo actualmente se encuentran cursando la carrera de Ingeniería en Sistemas de Información en la Universidad Tecnológica Nacional, y aspiran a ser becarios de investigación y desarrollo del LICPaD.

5. BIBLIOGRAFÍA

[1] Nykamp, Duane. "Multiplying matrices and vectors". Math Insight.

https://mathinsight.org/matrix_vector_multiplication Accedida en febrero de 2021.

[2] MPI Forum – MPI Documents.
<https://www.mpi-forum.org/docs/>.
Accedida en febrero de 2021.

[3] UTN, FRM - Lenguaje C
<http://www1.frm.utn.edu.ar/informatica1/VIANI/LENGUAJE%20C/LENGUAJE%20C.pdf> Accedida en enero 2021

[4] Wilkinson, B., Allen, M. (2005)
Parallel Programming: Techniques and Applications Using Networked Workstations and Parallel Computers.
Pearson Prentice Hall.

Cloud Computing, IaaS privados y públicos para el análisis de performance.

Diego Encinas^{1,2}, Brian Galarza¹, Román Bond¹, Gonzalo Zaccardi¹, Nicolás Benquerença Mendes¹, Jorge Osio¹, David Duarte¹, Martín Morales^{1,3}

¹Proyecto de Investigación SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - UNAJ

²Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

³Centro CodApli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, bgalarza@unaj.edu.ar, rbond@unaj.edu.ar, gzaccardi@unaj.edu.ar, nicobenquerenca@yahoo.com.ar, 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 y tecnología en Cómputo de Altas Prestaciones (High Performance Computing, HPC) para aplicaciones de interés social” – SimHPC de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 148/18. 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), autenticación (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 hipervisores 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.

Redes definidas por Software

Las redes definidas por software (SDN) son un paradigma de gestión y administración de redes por medio de software que permiten tener un control más flexible con respecto al control del tráfico de datos por medio de hardware ya que permite cambiar en tiempo real las normas y políticas establecidas en la red[11].

Mininet[12] es un emulador de redes SDN open source que permite generar tráfico artificial entre nodos virtuales de la red.

Se ha llevado a cabo el montaje de un laboratorio SDN utilizando Mininet junto con Amazon WebServices con el fin de analizar el comportamiento de este nuevo paradigma de redes de computadoras en un entorno de Cloud Computing.[13]

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.

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 el controller [14].
- Lograr escalabilidad agregando nuevos nodos compute a la infraestructura obteniendo un mayor performance en el sistema.
- 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].

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.

Durante 2020 se han realizado publicaciones nacionales. Además, se encuentran en desarrollo y concluidas varias Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Informática.

En esta línea de I/D existe cooperación a nivel nacional. Hay 4 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. Kreutz, D., Ramos, F. M. V., Esteves Verissimo, P., Esteve Rothenberg, C., Azodolmolky, S., & Uhlig, S. (2015). Software-Defined Networking: A Comprehensive Survey. Proceedings of the IEEE, 103(1), 14–76.
12. Mininet. <http://mininet.org>
13. Montes de Oca, F.; Galarza, B.; Morales, M.; Encinas, D. "Redes Definidas por Software en Entorno de Cloud Computing". 6° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2018). Mar del Plata, Argentina.
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". 8° 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". 8° 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". 8° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.

Cómputo paralelo y distribuido: métricas de rendimiento, aplicaciones de datos masivos e inteligencia artificial

Javier Balladini¹, Marina Morán¹, Claudio Zanellato¹, Claudia Rozas¹, Rodrigo Cañibano¹, Darío Semenzato¹, Agustín Chiarotto¹, Cristina Orlandi², Armando De Giusti³, Remo Suppi⁴, Dolores Rexachs⁴, Emilio Luque⁴

¹ Facultad de Informática, Universidad Nacional del Comahue
{javier.balladini, marina, claudio.zanellato, claudia.rozas, rcanibano}@fi.uncoma.edu.ar
{dario.semenzato, guillermo.chiarotto}@est.fi.uncoma.edu.ar

² Hospital Francisco Lopez Lima - orlandi.mariacristina@gmail.com

³ Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata - degiusti@lidi.info.unlp.edu.ar

⁴ Departamento de Arquitectura de Computadores y Sistemas Operativos, Universidad Autónoma de Barcelona
{remo.suppi, dolores.rexachs, emilio.luque}@uab.es

Resumen

Los grandes avances tecnológicos de los sistemas de cómputo paralelo y distribuido hacen viable nuevas soluciones a problemas. Por un lado, nos enfocamos en la métrica del consumo energético, un tema de enorme relevancia actual dado el gran número de unidades de procesamiento que componen los sistemas. Por otro lado, buscamos aplicar las técnicas de cómputo paralelo y distribuido para solucionar problemas del sector salud. En particular, nos orientamos a soluciones de alertas tempranas de gravedad para Unidades de Cuidados Intensivos, afectadas por un gran volumen de datos y la necesidad de aplicar técnicas de inteligencia artificial. También nos enfocamos en una aplicación para atender y clasificar pacientes con COVID-19 según el riesgo de salud. Los trabajos se desarrollan en colaboración con otras universidades, dos hospitales públicos de Argentina y un centro de investigación. La formación de recursos humanos en estas líneas está orientada al nivel de licenciatura, grado, maestría y doctoral.

Palabras claves: computación de altas prestaciones, eficiencia energética, big data, salud, inteligencia artificial.

1. Contexto

Nuestra investigación es parte de una solicitud de nuevo proyecto de investigación de la Universidad Nacional del Comahue que está en evaluación, denominado "Cómputo paralelo y distribuido: métricas de rendimiento, aplicaciones de big data e inteligencia artificial". La mayoría de los temas surgieron en proyectos anteriores acreditados. El eje de aplicaciones para la salud se desarrolla en colaboración con el Hospital Francisco López Lima de General Roca (Río Negro), el Hospital Italiano y el Instituto de Medicina Traslacional e Ingeniería Biomédica de Buenos Aires. Respecto al eje relacionado con el consumo energético de los sistemas de cómputo paralelo, se desarrolla en colaboración con el Instituto de Investigación en Informática LIDI de la Universidad Nacional de La Plata, y el grupo de investigación "High Performance Computing for Efficient Applications and Simulation" de la Universidad Autónoma de Barcelona, España.

2. Introducción

Numerosas aplicaciones informáticas que requieren el procesamiento de grandes volúmenes de datos están surgiendo gracias al uso de

tecnologías de cómputo paralelo y distribuido, que hasta hace unos años eran excepcionales. La masificación tecnológica y consecuente reducción de costos hizo que se pueda pensar en el desarrollo de nuevas aplicaciones que tengan un fuerte impacto social. Nos enfocamos principalmente en el sector de salud. A su vez, una de las mayores problemáticas que afectan a los sistemas de cómputo paralelo (y más aún a los de grandes dimensiones) es el elevado consumo energético y potencia eléctrica que demandan. Desde hace años no solo interesan las métricas de rendimiento computacional sino también energéticas. Por esta razón, también orientamos nuestra investigación a la gestión del consumo energético de estos sistemas, para reducir su impacto económico, medioambiental y social. A continuación se introduce la motivación y problemática que direccionan ambas temáticas de investigación.

2.1. Aplicaciones para la salud

Unidad de Cuidados Intensivos

Una Unidad de Cuidados Intensivos (UCI) provee cuidados continuos y rigurosos a personas adultas críticamente enfermas que pueden beneficiarse de tratamiento, y da un buen morir a pacientes irrecuperables. Los datos de los pacientes involucran datos clínicos de baja frecuencia y flujos de datos fisiológicos de alta frecuencia generados por el equipamiento médico (como monitores médicos de signos vitales). En una UCI típica, los enfermeros completan manualmente datos en formularios, registrando datos clínicos y fisiológicos. Los datos fisiológicos se obtienen por observación de las pantallas del equipamiento médico (ubicados a los lados de las camas) a intervalos de tiempo (por ejemplo de dos horas o fracciones) que definen los médicos para cada paciente. El equipamiento médico emite alertas cuando hay riesgo en la salud del paciente basándose en mediciones de ciertos parámetros. Luego, los médicos analizan los datos de los formularios y dan a los enfermeros indicaciones de tratamientos a realizar.

Los principales problemas que ocurren en las

UCIs típicas son: (a) gestión de la información es proclive a errores humanos, (b) pérdida de datos entre registros de enfermería, (c) detección tardía del deterioro de la salud de los pacientes. Las causas tienen origen en el tratamiento manual de la información (lento y propenso a errores), y cuyos efectos incluyen diagnósticos imprecisos o incorrectos o retrasados, inconsistencias en la información, menos datos para investigaciones médicas, y mayores recursos humanos y materiales. Esta situación es agravada por déficit de personal intensivista, causada porque es una profesión poco atractiva. Así, en la mayoría de las UCIs del país, no hay médicos intensivistas durante las 24hs, y hay sobrecarga laboral, impactando negativamente en la salud de los pacientes.

Las soluciones se orientan a sistema expertos de detección automática, temprana y progresiva del deterioro de pacientes, tal como los propuestos en [1, 10, 13].

Detección de riesgo en pacientes con COVID-19

El 11 de marzo de 2020 la Organización Mundial de la Salud dictaminó la pandemia de COVID-19. Esta enfermedad es causada por un virus extremadamente contagioso. El sistema sanitario argentino tiene su límite crítico en la cantidad de enfermeros disponibles por habitantes, con un número de 2,6 enfermos cada mil habitantes en contra parte de los 5,7 que tienen España e Italia¹, quienes no han podido contener la situación con ese personal. Esta situación generó que los recursos humanos especializados escaseen.

El “triage” es un método que permite la organización y optimización de los recursos en situaciones críticas, seleccionando y clasificando a los pacientes en diferentes niveles de gravedad para una correcta asignación de la atención. Este método fue aplicado en China [11] durante la emergencia sanitaria de pacientes con COVID-19 por medio de la adaptación de uno de los sistemas de alertas tempranas más

¹Datos relevados en 2017 por el Banco Mundial (<https://datos.bancomundial.org>).

conocido y validados. Esto permite mejorar el rendimiento del personal, reduciendo los controles en pacientes estables y aumentando los mismos en pacientes graves, y reducir la mortalidad hospitalaria inesperada en áreas de internación general.

Si bien el triage puede ser realizado mediante cálculos hechos a mano, no solo se agregaría una nueva tarea al personal de salud sino que sería un método muy propenso a errores humanos y por ende poco fiable. Así, nos hemos propuesto desarrollar un sistema informático inteligente que implemente un sistema de alerta temprana para pacientes en internación general con COVID-19, que realice un triage a través del análisis automático y en tiempo real de datos clínicos, clasificando a los pacientes según el riesgo, emitiendo alertas al personal de salud, y permitiendo organizar y optimizar recursos materiales y humanos. El sistema, a partir de la carga de datos de enfermería, comorbilidades, resultados radiológicos y de laboratorio, predice la gravedad de los pacientes (bajo, moderado, alto, crítico). No conocemos actualmente sistemas de este tipo para COVID-19 o similares en funcionalidad.

2.2. Consumo energético de sistemas de HPC

La computación de alto rendimiento (HPC) sigue aumentando su rendimiento computacional y su eficiencia energética. Por ejemplo, el superordenador Fugaku, la supercomputadora más rápida del mundo en el ranking Top500, presenta 415,5 PFlops frente a los 148,6 PFlops de su predecesor, Summit. Al mismo tiempo, estas supercomputadoras se encuentran entre las diez primeras de las más eficientes energéticamente del ranking Green500, con alrededor de 14,7 GFlops/W. Sin embargo, el consumo de energía sigue aumentando; mientras que la computadora Summit tiene un consumo de 10 MW, Fugaku sube a 28 MW (aproximadamente lo mismo que consume una ciudad de 400.000 habitantes). Como este aumento del consumo de energía no es sostenible, es necesario reducirlo.

La computación ecológica es el estudio y la práctica de la computación ambientalmente sostenible. Ella se ocupa de diferentes aspectos de los sistemas de cómputo: diseño, manufactura, eliminación, y uso. Este último aspecto, el uso ecológico, se refiere al uso de los sistemas de cómputo con conciencia ambiental. Es posible reducir el consumo de energía de los sistemas de cómputo utilizando diferentes estrategias que deben ser consideradas a nivel del software [9, 6], y consisten en realizar cambios en la configuración del sistema (escalado dinámico de frecuencia y tensión, hibernación de recursos) o en las aplicaciones (uso adecuado de la jerarquía de memoria, rediseño de algoritmos, planificación de tareas, y asignación de tareas a recursos hardware).

El aumento significativo de la cantidad de unidades de procesamiento causa el incremento del consumo de energía y la disminución de la confiabilidad del sistema de cómputo. Así, a poco tiempo de arribar a la era exaescala, la tolerancia a fallos y el consumo energético se han identificado como los dos mayores desafíos a los que deberemos enfrentarnos [12, 8]. En consecuencia, proponemos desarrollar metodologías, modelos y software para gestionar el consumo energético, en especial al utilizar mecanismos de tolerancia a fallos en máquinas paralelas de tipo cluster.

3. Líneas de investigación

El eje central de nuestra investigación es desarrollar metodologías, modelos y soluciones informáticas para colaborar en la reducción del consumo energético de sistemas de HPC, y en la resolución de problemas que tengan una alta demanda computacional e impacto social en el campo de la salud.

3.1. Aplicaciones para la salud

UCI: Nuestro sistema intenta emular el comportamiento de un médico intensivista experto, dando recomendaciones para la toma de decisiones clínicas, con el objetivo de reducir la

incertidumbre sobre el diagnóstico, las opciones de tratamiento y el pronóstico. La solución requiere la aplicación de técnicas de cómputo paralelo y distribuido para procesamiento en tiempo real de algoritmos de inteligencia artificial sobre grandes volúmenes de datos. A diferencia de otras alternativas, nuestro objetivo se orienta a la construcción de un sistema multihospitalario (con el fin de incrementar el volumen de datos y consecuente extracción de conocimiento) con soporte de telemedicina, que integre componentes de software libre maduros, que nosotros optimizados para nuestro dominio, que sea seguro, tolerante a fallos y resiliente.

COVID-19: Nuestro objetivo es construir un sistema y modelo computacional de alerta temprana para pacientes con COVID-19, inicialmente basado en el conocimiento de personal experto en salud, y posteriormente mejorado mediante técnicas de aprendizaje automático. La complejidad de la arquitectura del sistema no está en la alta demanda computacional sino en evitar el uso de servidores estándares en los hospitales (a fin de simplificar la administración multihospitalaria), y que el sistema local de un hospital se mantenga operativo ante fallos de conexión con la nube, utilizando únicamente dispositivos móviles del tipo smartphones y tablets.

3.2. Consumo energético de sistemas de HPC

Nos centramos en el desarrollo de metodologías, modelos y construcción de software para administrar y gestionar el consumo de energía y prestaciones de sistemas de cómputo paralelo. Actualmente, nuestro principal objetivo es la gestión energética en mecanismos de tolerancia a fallos basados en checkpoints.

4. Resultados y objetivos

4.1. Aplicaciones para la salud

UCI

En [5] presentamos un análisis del estado general de la UCI del hospital Francisco López Lima, y la propuesta del diseño de alto nivel del sistema. En [2] presentamos el diseño de la infraestructura del sistema de procesamiento de reglas clínicas y un prototipo, y el desarrollo de un dispositivo embebido adquirente de datos que envía datos a la plataforma de procesamiento por medio de una red WiFi. En [3] propusimos la herramienta como mejora de la prestación de servicios integrados de cuidados intensivos de la salud. En [7] presentamos una optimización de la administración de datos de curvas fisiológicas. Próximos objetivos: continuar con el desarrollo de hardware y software para extraer datos del equipamiento médico, desarrollar aplicaciones para el procesamiento eficiente de señales, manejo de errores en datos fisiológicos, optimización de la infraestructura de datos masivos de tiempo real para el procesamiento de reglas clínicas, y el desarrollo de una aplicación para la interacción del sistema con médicos y enfermeros.

COVINDEX

Se construyó un modelo de alertas tempranas basado en el conocimiento de médicos expertos nacionales e internacionales. Se está desarrollando un sistema informático, denominado COVINDEX, que implementa dicho modelo. El sistema se compone de una aplicación web, una aplicación móvil y una aplicación que ejecuta en un servidor en la nube. El sistema se pondrá a disposición de los hospitales en Abril de este año. A partir de los datos históricos que se vayan recolectando de los pacientes, se calibrará periódicamente el sistema de alerta temprana de insuficiencia respiratoria, sobre la base de técnicas de aprendizaje automático, para aumentar la capacidad predictiva del sistema. Aún no se cuenta con publicaciones científicas.

4.2. Consumo energético de sistemas de HPC

Los métodos de tolerancia a fallos tienen fuerte incidencia en el consumo energético de los sistemas de HPC, y resulta de suma importancia conocer, antes de ejecutar una cierta aplicación, el impacto que pueden producir los diferentes métodos y configuraciones del mismo. En [4], presentamos una metodología para predecir el consumo energético producido por el método de checkpoint coordinado remoto, y en [15] expusimos un análisis de los factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. En [14] propusimos un modelo para estimar el consumo energético de operaciones de checkpoint y restart, y un método para su construcción. En [16] presentamos estrategias para checkpoints no coordinados que, al momento de un fallo de un nodo, permiten gestionar y reducir el consumo energético de los nodos que no han fallado; se construyó también un modelo energético y un simulador que permite evaluar las estrategias.

5. Formación de recursos humanos

El equipo de trabajo de la Universidad Nacional del Comahue tiene un Doctor y un Magíster, una estudiante de Doctorado (a graduarse en el presente año) y dos estudiantes de Maestría (a graduarse en 2022). Tres estudiantes de grado están realizando trabajos de tesis, y hay un estudiante de tecnicatura. En 2020 no se ha finalizado ninguna tesis.

Referencias

- [1] S. Balaji, M. Patil, and C. McGregor. A cloud based big data based online health analytics for rural nicus and picus in india: Opportunities and challenges. In *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 385–390, 2017.
- [2] Javier Balladini, Pablo Bruno, Rafael Zurita, and Cristina Orlandi. An automatic and early detection of the deterioration of patients in intensive and intermediate care units. *Journal of Computer Science and Technology*, 18(03):e25, Dec. 2018.
- [3] Javier Balladini, Pablo Bruno, Rafael Zurita, Cristina Orlandi, Remo Suppi, Dolores Rexachs, and Emilio Luque. A tool for improving the delivery of integrated intensive health care performance. *International Journal of Integrated Care*, 19(4), 07/2019 2019.
- [4] Javier Balladini, Marina Morán, Dolores Rexachs, and Emilio Luque. Metodología para predecir el consumo energético de checkpoints en sistemas de hpc. *XX Congreso Argentino de Ciencias de la Computación (CACIC 2014)*, 2014.
- [5] Javier Balladini, Claudia Rozas, Emmanuel Frati, Nestor Vicente, and Cristina Orlandi. Big data analytics in intensive care units: challenges and applicability in an argentinian hospital. *Computer Science and Technology (JCST)*, 2015.
- [6] Aurelien Bouteiller, Franck Cappello, Jack Dongarra, Amine Guermouche, Thomas Héroult, and Yves Robert. Multi-criteria checkpointing strategies: Response-time versus resource utilization. In *European Conference on Parallel Processing*, pages 420–431. Springer, 2013.
- [7] Rodrigo Cañibano, Claudia Rozas, Cristina Orlandi, and Javier Balladini. Data management optimization in a real-time big data analysis system for intensive care. In *Conference on Cloud Computing, Big Data & Emerging Topics*, pages 93–107. Springer, Cham, 2020.
- [8] Franck Cappello, Al Geist, William Gropp, Sanjay Kale, Bill Kramer, and Marc Snir. Toward exascale resilience: 2014 update. *Supercomputing Frontiers and Innovations*, 1(1), 2014.
- [9] Mohammed El Mehdi Diouri, Olivier Glück, Laurent Lefèvre, and Franck Cappello. Ecofit: A framework to estimate energy consumption of fault tolerance protocols for hpc applications. In *Proceedings of the 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing*, pages 522–529, 2013.
- [10] Jagreet Kaur and Dr. Kulwinder Singh Mann. AI based HealthCare platform for real time, predictive and prescriptive analytics using reactive programming. *Journal of Physics: Conference Series*, 933:012010, jan 2018.
- [11] Xuelian Liao, Bo Wang, and Yan Kang. Novel coronavirus infection during the 2019–2020 epidemic: preparing intensive care units-the experience in sichuan province, china. *Intensive care medicine*, 46(2):357–360, 2020.
- [12] Yongpeng Liu and Hong Zhu. A survey of the research on power management techniques for high-performance systems. *Software: Practice and Experience*, 40(11):943–964, 2010.
- [13] Fernando López-Martínez, Edward Rolando Núñez-Valdez, Vicente García-Díaz, and Zoran Bursac. A case study for a big data and machine learning platform to improve medical decision support in population health management. *Algorithms*, 13(4), 2020.
- [14] M. Morán, J. Balladini, D. Rexachs, and E. Luque. Prediction of energy consumption by checkpoint/restart in hpc. *IEEE Access*, 7:71791–71803, 2019.
- [15] Marina Morán, Javier Balladini, Dolores Rexachs, and Emilio Luque. Factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. *XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, 2018.
- [16] Marina Morán, Javier Balladini, Dolores Rexachs, and Enzo Rucci. Towards management of energy consumption in hpc systems with fault tolerance. *arXiv preprint arXiv:2012.11396*, 2020.

Métodos de Inteligencia artificial para la reducción de incertidumbre en modelos de predicción de incendios

Chirino Pamela¹, Galdamez Mariela¹, Diaz-Acevedo Karvin¹, Ponce de León Alejo¹, Caymes-Scutari Paola^{1,2}, Bianchini Germán¹

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional

Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

pamela.chirino@alumnos.frm.utn.edu.ar, mariela.galdamez@alumnos.frm.utn.edu.ar ,
karvindiaz@gmail.com, alejo.poncedleon@gmail.com, pcaymesscutari@frm.utn.edu.ar,
gbianchini@frm.utn.edu.ar

RESUMEN

La inteligencia artificial ha generado una revolución importante en los últimos años de la computación. En esta línea de trabajo se estudiarán dos formas de inteligencia artificial para aplicarlas en la reducción de incertidumbre en modelos de predicción, en este caso, el modelo de predicción de incendios llevado a cabo en el Laboratorio de Investigación en Cómputo Paralelo/Distribuido de la UTN-FRM. En el mismo se trabaja con paralelismo, por lo tanto se analizará la posible paralelización de estos métodos. Los métodos de inteligencia artificial que se estudiarán son: Redes neuronales y Visión Computacional.

Palabras Clave: Inteligencia Artificial, Redes neuronales, Paralelismo, Visión Computacional, Predicción

CONTEXTO

Esta investigación es llevada a cabo en el Laboratorio de Investigación en Cómputo

Paralelo/Distribuido de la Universidad Tecnológica Nacional, Facultad regional de Mendoza.

Está enmarcada en el proyecto PID TEUTIME0007658TC titulado “Formación de docentes y alumnos de grado como Investigadores Científicos Iniciales en las áreas de Informática y Ciencias de la Computación.”

1. INTRODUCCIÓN

La Inteligencia Artificial es la disciplina que trata de crear sistemas capaces de aprender y razonar como un ser humano.

Normalmente, un sistema de inteligencia artificial es capaz de analizar datos en grandes cantidades, identificar patrones y tendencias y, por lo tanto, formular predicciones de forma automática, con rapidez y precisión [1].

Estudiaremos las redes neuronales y la visión computacional, analizando su rendimiento en paralelo y su aplicación al

modelo de predicción de incendios previamente mencionado [2].

A continuación, detallaremos algunos aspectos de relevancia para mayor comprensión a la hora de explicar el desarrollo.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto cuenta con tres ejes temáticos: Redes Neuronales, Visión Computacional y Paralelismo.

2.1. Redes Neuronales

Las redes neuronales artificiales son un modelo computacional inspirado en el comportamiento observado en el cerebro humano. Consiste en un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitir o comunicar señales [3].

Una red neuronal artificial está constituida por neuronas interconectadas y arregladas en capas. Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas de neuronas.

Además de las capas ya mencionadas, una red neuronal cuenta con tres funciones principales.

La función de entrada trata muchos valores de entrada como si fueran uno solo y esto recibe el nombre de entrada global. Además le asigna pesos a las entradas, y estos pesos se suman a los pesos que anteriormente tenían las neuronas de las capas subsiguientes.

La segunda función es la de activación que tiene como objetivo acotar los valores de salida de una red neuronal para mantenerlos en ciertos rangos, es decir, calcula el estado de actividad de una neurona; transformando la entrada global en un valor (estado) de activación.

El último componente que una neurona necesita es la función de salida. El valor resultante de esta función es la salida de la neurona; por ende, la función de salida determina qué valor se transfiere a las neuronas vinculadas. Si la función de activación está por debajo de un umbral determinado, ninguna salida se pasa a la neurona subsiguiente.

En particular, para esta investigación, la red neuronal que nos interesa estudiar es el perceptrón.

El perceptrón es una de las redes neuronales más sencillas, su funcionamiento es muy simple, lee los valores de entrada, suma todas las entradas de acuerdo a unos pesos y el resultado lo introduce en una función de activación que genera el resultado final [1].

Para que el perceptrón nos sea de utilidad tenemos que entrenarlo o enseñarle. Al proceso de enseñarle se le denomina aprendizaje, que es el mecanismo por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas.

Un aspecto importante a tener en cuenta es que el perceptrón utiliza un tipo particular de aprendizaje supervisado, es decir, aprendizaje controlado por un

agente externo, llamado retropropagación de error o *backpropagation* [1].

En las redes de retropropagación primero se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir una salida de la misma. Esta salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida hacia todas las neuronas de las capas intermedias. Cada neurona recibe un error que es proporcional a su contribución sobre el error total de la red. Basándose en este error recibido se ajustan los pesos sinápticos de cada neurona [1].

Se considera que la red ha aprendido cuando el error es 0 o un margen próximo al mismo [5] [6].

2.2. Visión computacional

La visión computacional trata de interpretar las imágenes recibidas por dispositivos y reconocer los objetos, ambiente y posición en el espacio [7]. Debido a que parte de nuestro trabajo se centrará en el tratamiento de imágenes satelitales, hemos considerado estudiar este tema y su posible paralelización para aplicarlo en el modelo de predicción de incendios anteriormente mencionado.

Estudiaremos y aplicaremos la visión computacional a través de las redes convolucionales [7].

Las redes neuronales convolucionales son muy similares a las redes neuronales ordinarias como el perceptrón multicapa que fue descrito anteriormente; se componen de neuronas que tienen pesos y capacidad de aprender.

Lo que diferencia a las redes neuronales convolucionales es que suponen explícitamente que las entradas son imágenes, lo que nos permite codificar ciertas propiedades en la arquitectura de las redes neuronales tradicionales, permitiendo ganar en eficiencia y reducir la cantidad de parámetros en la red.

Las redes neuronales convolucionales trabajan modelando de forma consecutiva pequeñas piezas de información, y luego combinan esta información en las capas más profundas de la red.

En general, las redes neuronales convolucionales van a estar construidas con una estructura que contendrá tres tipos distintos de capas.

La capa convolucional es la que le da nombre a la red. Lo que distingue a las redes neuronales convolucionales de cualquier otra red neuronal es que utilizan una operación llamada convolución en alguna de sus capas.

La operación de convolución recibe como entrada o *input* la imagen y luego aplica sobre ella un filtro o *kernel* que nos devuelve un mapa de las características de la imagen original, y de esta forma logramos reducir el tamaño de los parámetros.

Luego sigue la capa de reducción o *pooling*. Su utilidad principal radica en la reducción de las dimensiones espaciales (ancho x alto) del volumen de entrada para la siguiente capa convolucional.

Al final de las capas mencionadas anteriormente, las redes utilizan generalmente capas completamente conectadas en la que cada pixel se

considera como una neurona separada al igual que en una red neuronal regular.

2.3. Paralelismo

El procesamiento paralelo es un método mediante el cual se dividen grandes problemas en componentes, tareas o cálculos que puedan resolverse en simultáneo. En particular, estamos analizando el modelo Master-Worker para incluir en nuestro estudio.

El modelo Master-Worker es un modelo aplicado a la descomposición de dominio, es decir, el dominio del problema se divide en subconjuntos de datos y los mismos son asignados a nodos diferentes [8].

El proceso principal denominado Master es el proceso encargado de coordinar todo el tratamiento y procesamiento del problema, para lo que genera muchos subprocesos, que son ejecutados como procesos independientes denominados Workers, y en general se ejecutan en procesadores independientes de forma simultánea. La interacción que existe entre ellos es que el Master inicia los procesos Worker, les asigna el trabajo a realizar, y estos devuelven el resultado al proceso Master [8].

3. RESULTADOS OBTENIDOS/ ESPERADOS

Como ya mencionamos anteriormente, en el laboratorio de computo paralelo distribuido se trabaja continuamente con la incertidumbre de variables en el modelo de predicción de incendios.

Nuestra propuesta es trabajar con esas variables donde se presenta incertidumbre haciéndolas variar en las posibilidades de

valores que podrían tomar y presentarle como conjunto de entrenamiento los resultados reales o esperados del incendio, para que la red neuronal aprenda a diferenciar qué variables influyeron más y en qué posibles valores en un incendio.

En cuanto a los avances actuales, se determinó sistematizar la introducción de datos con los que operará la red neuronal mediante archivos a raíz del volumen de datos. También se tuvo en cuenta que los desarrollos previos estaban implementados bajo los lenguajes C/C++ y que una de las pocas redes neuronales que se pueden implementar en este lenguaje es el perceptrón, debido a que es más sencillo de implementar cuando no se cuenta con clases.

Para la primera propuesta de paralelización, se propone determinar al Master como el encargado de almacenar datos realizando el envío y recepción de estos. Además, se encargará de realizar los cálculos de la función de entrada. Proponemos considerar a cada capa oculta y a la capa de salida como un proceso. De esta forma, cada procesador - Worker- quedaría a cargo de cierta cantidad de neuronas y realizaría los cálculos correspondientes en cada neurona que le fuera asignada obteniendo una salida. Luego de ingresar los datos de entrada en la capa de entrada, se obtendrían sus salidas que serán enviadas a cada proceso Worker. El procedimiento explicado anteriormente representa el primer ajuste de pesos. Luego se lleva a cabo el proceso de aprendizaje con backpropagation explicado anteriormente, hasta que se considere que la red neuronal ha aprendido.

En cuanto a la Visión Computacional, la propuesta es trabajar con las redes neuronales convolucionales para el entrenamiento a la hora de analizar imágenes o mapas satelitales. En el modelo de predicción de incendios se trabaja constantemente con mapas, [9] por lo tanto, al aplicar visión computacional a través de redes convolucionales se busca automatizar y hacer lo más rápido posible este análisis.

En cuanto la propuesta de paralelización, es similar a la aplicada en las redes neuronales, la descomposición será de dominio, es decir, se repartirá el dominio de problema (datos), considerando que las imágenes para el procesamiento serán un gran conjunto de píxeles. El nodo master se encargaría de administrar y repartir los píxeles correspondientes a las imágenes satelitales, entre los diferentes nodos de la red convolucional. Los diferentes nodos tendrán su propia red convolucional y manejarán un gran número de neuronas. En un futuro se estudiará la descomposición funcional de la misma, teniendo en cuenta las diferentes capas con las que cuenta la red convolucional.

4. FORMACION DE RECURSOS HUMANOS

Nuestro grupo de trabajo se encuentra conformado por los directores de laboratorio, German Bianchini y Paola Caymes Scutari, y cuatro becarios en el marco de las Becas Bis y las Becas I+D.

La temática del proyecto de investigación aporta a la formación académica de los integrantes del grupo de trabajo, proporcionando cooperación y aprendizaje conjunto entre los estudiantes

en sus respectivos trabajos en el marco del laboratorio.

5. BIBLIOGRAFIA

- [1] Pedro Ponce Cruz, “Inteligencia Artificial Aplicada a la ingeniería”, Alfaomega, 2001.
- [2] Germán Bianchini. Paola Caymes Scutari, Miguel Méndez Garabetti, Evolutionary Statistical System: a Parallel Method for Improving Forest Fire Spread Prediction. Journal of Computational Science (JOCS) Vol 6 pp. 58-66. ISSN: 1877-7503 doi: 10.1016/j.jocs.2014.12.001 Elsevier.
- [3] Carlos Alberto Ruiz, “Redes Neuronales: Conceptos básicos y aplicaciones”, UTN-Facultad Regional de Rosario, 2005.
- [4]<http://www.sc.ehu.es/ccwbytes/docencia/mmcc/docs/t8neuronales.pdf>, accedida el 20/08/19.
- [5] Rivera E., “Entrenamiento de redes neuronales en algoritmos evolutivos”, 2005.
- [6] Matich, D. J., “Introducción a las Redes Neuronales Artificiales”, 2001.
- [7] Sucar Enrique, Giovani Gómez, “Vision Computacional”, 2005.
- [8] Barry Wilkinson, Michael Allen. “Parallel Programming” (2005). Pearson.
- [9] G. Bianchini, P. Caymes Scutari, “Metodos basados en computación de alto rendimiento para predecir el comportamiento de incendios forestales” E-ICES7. ISBN 978-987-1323-27-2. Edit. CNEA. pp. 28-36, 2012.

Optimización en el desarrollo de aplicaciones Serverless en entornos distribuidos

Nelson Rodríguez¹, María Murazzo¹, Diego Medel¹, Daniel Arias Figueroa², Lorena Parra¹, Ana Laura Molina¹, Federico Sánchez¹, Adriana Martín¹, Hernán Atencio³, Martín Gómez³

¹ Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

² Consejo de Investigación, FCE, Universidad Nacional De Salta

⁶ 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, vdiego.unsj@hotmail.com, daaf@cidia.unsa.edu.ar, lorenaparra152@yahoo.com.ar, almm95@gmail.com, fgsanchez@unsj-cuim.edu.ar, adrianamartin1@gmail.com, hernan.atencio.98@gmail.com, martinsj0811@gmail.com

Resumen

Los últimos años han sido testigos de avances sin precedentes en el campo de Cloud Computing. Este modelo se refiere al acceso de red ubicuo, conveniente y bajo demanda a un grupo compartido de recursos informáticos. Como una evolución constante, aparece Serverless como un nuevo paradigma de desarrollo de software en el Cloud, que surge a partir de una arquitectura monolítica y luego pasando por máquinas virtuales y contenedores para finalizar en serverless, que en algunos casos se conoce como función como servicios, aunque también incluye Backend como servicio, lo cual en general son servicios de terceros disponibles en el Cloud. La misma está enfocada en proveer una arquitectura que permite la ejecución de funciones arbitrarias con mínima sobrecarga en la administración del servidor y soportada bajo la programación orientada a eventos. La cantidad y variedad de datos provenientes del Edge, necesitan ser procesados y en algunos casos en tiempo real y por lo tanto deben resolverse varios desafíos y problemas abiertos. Debido a que las investigaciones apuntan a diversas áreas como son: arquitectura y hardware, administración de recursos, modelo de negocios y desarrollo

de aplicaciones. Debido a que las aplicaciones serverless son altamente distribuidas, donde cada función realiza una tarea enfocada y depende de un sinnúmero de otros servicios para el resto., el objetivo de esta línea de investigación es resolver problemáticas vinculadas al desarrollo, que generen alta latencia, dificulten el paralelismo de funciones serverless o no se adapten a los principios DevOps, entre otros.

Palabras clave: *Serverless Computing, Distributed Computing, FaaS, , Cloud Computing, Serverless Programming*

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es una de las líneas de investigación del proyecto: Computación Serverless para el tratamiento de datos provenientes de dispositivos de IoT, cuya propuesta ha sido aprobada y está en desarrollo para el período 2020-2021. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación distribuida y de alta performance desde hace más de 20 años. Como continuación del proyecto

anterior: Orquestación de Servicios para la Continuidad Edge al Cloud, se sigue trabajando con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

La computación Serverless es una tecnología emergente con un impacto creciente en nuestra sociedad y una mayor adopción tanto por parte de la academia como de la industria [1]. Es un paradigma emergente en el que las aplicaciones de software se descomponen en múltiples funciones independientes sin estado [2] [3]. Las funciones se ejecutan en contenedores casi sin estados y solo en respuesta a acciones desencadenantes (como interacciones de usuario, eventos de mensajería o cambios en la base de datos), se pueden escalar de forma independiente y pueden ser efímeras (pueden durar una invocación) y están completamente administrados por el proveedor de Cloud.

Debido a que la implementación se realiza mediante la plataforma de funciones, la computación serverless, también es denominada función como servicio (FaaS). En este enfoque, casi todas las preocupaciones operativas son abstraídas lejos de los desarrolladores. Los cuales en principio simplemente escriben código e implementan sus funciones en una plataforma sin servidor. La plataforma se encarga de la ejecución de la función, el almacenamiento y la infraestructura de contenedor, redes y tolerancia a fallas. Adicionalmente, también se encarga de escalar las funciones según la demanda real.

En la mayoría de los casos, se puede escribir funciones en el lenguaje favorito del programador (Node.js, Python, Go, Java y más) y utilizar herramientas de contenedor y serverless, como AWS SAM o la CLI de Docker, para compilar, probar e implementar las funciones.

La computación serverless ha sido identificada como un enfoque prometedor para varias aplicaciones, como el análisis de datos en el borde de la red [4]. En consecuencia, una plataforma sin servidor maneja el ciclo de vida, la ejecución y escalada de las funciones reales;

estas necesitan correr solo cuando son invocadas o activadas. Por lo tanto, el mayor beneficio de esta computación son las pocas preocupaciones operativas y de gestión y la utilización eficiente de los recursos [5].

Un modelo basado en funciones es particularmente adecuado para ráfagas, uso de CPU intensivo, cargas de trabajo granulares. Actualmente, los casos de uso de FaaS varían ampliamente, incluido el procesamiento de datos, el procesamiento de flujo, la computación de borde (IoT) y la computación científica [6] [7]. Con la continua experimentación en torno a FaaS, es probable que otros casos de uso surjan en un futuro.

En los últimos años, el paradigma de la función como servicio ha revolucionado la forma en que resuelve el procesamiento de eventos distribuidos. Los desarrolladores pueden implementar y ejecutar funciones controladas por eventos en el Cloud que se escalan bajo demanda. La ventaja de la alta disponibilidad y poca administración, son muy valoradas por la conducción de las empresas.

El término "sin servidor" es confuso ya que hay tanto hardware de servidor como procesos de servidor que se ejecutan, pero la diferencia en comparación con los enfoques normales es que la organización que construye y admite una aplicación "sin servidor" no se ocupa de ese hardware o esos procesos, debido a que están subcontratando esta responsabilidad.

Los primeros usos del término aparecieron en 2012, incluido el artículo de Ken Fromm [8]. El término se hizo más popular en 2015, luego del lanzamiento de AWS Lambda, y creció en popularidad después del lanzamiento de API Gateway de Amazon en julio de 2015. Posteriormente el proyecto de código abierto "JavaScript Amazon Web Services (JAWS)" se renombró a Serverless Framework, continuando la tendencia.

A mediados de 2016, Serverless se había convertido en un nombre dominante para esta área, dando paso al nacimiento de la serie de conferencias Serverless, y varios proveedores Cloud adoptaron el término en todo, desde marketing de productos hasta descripciones de puestos de trabajo.

Existen varias definiciones y todas son muy similares. Se toma de referencia la del grupo de investigación The SPEC Cloud:

La computación serverless es una forma de computación en el cloud que permite a los usuarios ejecutar eventos y aplicaciones facturadas de forma granular, sin tener que abordar la lógica operativa [9].

Esta definición coloca al servidor como una abstracción, que se superpone parcialmente con la plataforma como servicio (PaaS). Los desarrolladores se centran en abstracciones de alto nivel (por ejemplo, funciones, consultas y eventos) y en crear aplicaciones que los operadores de infraestructura asignan a recursos concretos y servicios de soporte. Esto efectivamente separa las preocupaciones, con los desarrolladores enfocándose en la lógica empresarial y en las formas de interconectar elementos de la lógica empresarial en flujos de trabajo complejos. Mientras tanto, los proveedores de servicios se aseguran de que las aplicaciones están alojadas en contenedores, desplegadas, aprovisionadas y disponibles bajo demanda, mientras se factura al usuario solo por los recursos utilizados [10].

La computación serverless se puede identificar como resultado de la unión de Cloud y Microservices Architecture. Pero la evolución de Cloud ha pasado por varias etapas.

Se investigaron muchas vías diferentes para simplificar la gestión de aplicaciones a escala, entornos virtuales, hosts virtuales dentro de máquinas virtuales, contenedores, por nombrar algunos. Con la aparición de los contenedores comenzaron a ofrecerse plataformas que permitían desplegar contenedores.

De estos microservicios, surgieron nuevas ofertas de "nivel de servicio" de alto nivel como Backend-as-a-Service (BaaS) y Function-as-a-Service (FaaS). Estos representaban los albores de la computación serverless.

Optimización del desarrollo Serverless

La computación serverless es adecuada para aplicaciones de corta duración sin estado dirigidas por eventos, por ejemplo

microservicios, backends IoT móviles, procesamiento de flujo modesto, bots e integración de servicios. No son indicados para cálculo numéricos de larga duración con estado como son: bases de datos, entrenamiento de aprendizaje profundo, análisis de flujo de servicio pesado, simulación numérica y vídeo transmitido en vivo [10].

La computación serverless se está utilizando para admitir una amplia gama de aplicaciones. Desde una perspectiva funcional, serverless y las arquitecturas más tradicionales pueden usarse indistintamente. La determinación de cuándo usar serverless estará influenciada por requisitos no funcionales como la cantidad de control sobre las operaciones requeridas y el costo y las características de la carga de trabajo de la aplicación, por ejemplo [11].

Si bien las primeras ventajas eran la económica y la sencillez en el desarrollo de aplicaciones, posteriormente surgieron nuevos beneficios como son: la escalabilidad, el código puede ejecutarse desde cualquier lugar y se adecúa al desarrollo de productos ágiles.

Por supuesto que existen desventajas como son: la depuración puede resultar más difícil de llevar a cabo, existe pérdida del control operativo, surgen nuevos problemas al utilizar APIs de terceros, nuevos riesgos de seguridad aparecen y los desarrolladores necesitan de un cambio de mentalidad, entre otras.

El desarrollo de aplicaciones presenta una gran cantidad de problemas a resolver. De los cuales se van a enunciar algunos de ellos y se espera que sean resueltos o minimizados por los trabajos del grupo de investigación.

Si bien los frameworks serverless han permitido a los desarrolladores crear e implementar aplicaciones más rápidamente, también traen un conjunto de desafíos de depuración en comparación con los frameworks tradicionales.

Además, debido a que la arquitectura sin servidor todavía está en sus etapas formativas, hay enfoques y herramientas en evolución que se integran sin problemas con una plataforma FaaS, al tiempo que ayudan con la recopilación constante de datos [12].

En cuanto a la configuración, existen muchas dificultades para realizar la supervisión de

sistemas distribuidos dinámicamente para identificar los posibles cuellos de botella que puedan ocurrir.

También se produce un aumento de la latencia, que es causada por comunicación entre intercomponentes complejos

Aparece la necesidad de integrar rápida y eficientemente los microservicios asociados. Pueden ocurrir errores en el Cloud o en el Edge, pero en el borde, sin embargo, las técnicas de rollback de falla son muy limitadas y deben lograrse soluciones más robustas que para lograr el nivel deseado de resiliencia [13].

Otro problema es la gestión eficiente de arranques en frío, dado que se produce un retraso introducido a medida que la infraestructura subyacente aprovisiona dinámicamente nuevos recursos en tiempo de ejecución y además puede generar, si la cantidad de funciones es elevada, latencias adicionales y errores de sistema.

Por otro lado no existen herramientas de seguimiento y depuración adecuadas, tampoco hay expertización en este tipo de programación para atender la refactorización de los sistemas existentes, la capacidad de integrar y componer, gestionar y mantener, diseñar funciones con estado y sin estado, la gestión de simultaneidad, la granularidad de código de forma optimizada, diseño del sistema de recuperación y adaptación de los principios de DevOps entre otras. [14].

Líneas de Investigación, Desarrollo e Innovación

El tipo de investigación que se está utilizando para llevar a cabo esta línea de trabajo ha comenzado hace un año con una análisis documental que permitió definir las bases de conocimiento sobre la “Computación Serverless” y su convergencia al Edge. Esto posibilitó mediante la confección de revisiones sistemáticas y mapeos sistemáticos de la literatura, encontrar cuales son los problemas científicos y desafíos que se van a tratar de solucionar.

Posteriormente la investigación se conducirá de forma hipotética mixta (experimental y deductiva). Esta forma de trabajo permitirá la

selección de una plataforma de despliegue serverless sobre la cual se realizará el tratamiento de datos provenientes del Edge y la comunicación entre las diferentes funciones, debido a que en entornos altamente distribuidos la comunicación entre los componentes es sumamente importante para evaluar. Posteriormente, se seleccionarán los problemas a resolver, se fijará el orden y prioridad en analizar e investigar cada uno de ellos. El método a utilizar será empírico analítico, pero puede variar en función de cada problema en particular.

Resultados y Objetivos

Resultados Obtenidos

Durante los últimos trece años se trabajó en el área de Computación de Altas Prestaciones y distribuidas, en particular sobre análisis de diversas arquitecturas paralelas y distribuidas, tales como: Cloud Computing, Cluster de commodity, arquitecturas distribuidas y paralelas de bajo costo y fog computing. En el último año se inició un proyecto sobre Srverless Computing, cuya experiencia impulsó la línea de investigación del presente trabajo. El grupo ha realizado varias publicaciones en esta área: seis trabajos de investigación en Congresos y Jornadas, se realizaron tres publicaciones en revistas científicas y se transfirieron los resultados mediante conferencias en eventos científicos. Se han aprobado tres tesinas de grado, se incorporó un becario de investigación categoría alumno y otra beca está en evaluación.

Objetivos

El objetivo del grupo de investigación es analizar y resolver los diferentes problemas y retos que surgen y que dificultan la eficiencia a nivel del desarrollo (en particular programación) de aplicaciones serverless, teniendo en cuenta entorno distribuidos como Edge a Cloud o datos provenientes de dispositivos IoT. De esta forma poder optimizar el desarrollo aplicando variadas estrategias y evaluando cómo se comportan los parámetros de desempeño y eficiencia.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de ocho investigadores que figuran en este trabajo de las universidades Nacional de San Juan y Nacional de Salta y dos alumnos de grado. Además, el proyecto marco donde se está desarrollando esta propuesta incluye a tres investigadores más de la Nacional de San Luis, de la Universidad Champagnat y de la Universidad Nacional de San Juan y tres alumnos de grado.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, seis tesinas de grado en el área de Serverless computing, Concurrencia y Computación distribuida y una tesis de maestría en áreas afines. Además se espera aumentar el número de publicaciones. Por otro lado 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 otros organismos del estado.

Referencias

- [1] IDC, FutureScape: Worldwide IT Industry 2019 Predictions," <https://www.idc.com/getdoc.jsp?containerId=US44403818>, Oct 2018.
- [2] Adzic, G., & Chatley, R. (2017). Serverless computing: economic and architectural impact. In Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering (pp. 884-889). ACM.
- [3] Baldini, I., Castro, P., Chang, K., Cheng, P., Fink, S., Ishakian, V. & Suter, P. (2017). Serverless computing: Current trends and open problems. In Research Advances in Cloud Computing (pp. 1-20). Springer, Singapore.
- [4] Nastic, S., Rausch, T., Scekcic, O., Dustdar, S., Gusev, M., Koteska, B. & Prodan, R. (2017). A serverless real-time data analytics platform for edge computing. IEEE Internet Computing, 21(4), 64-71.
- [5] Mohanty, S. K., Premsankar, G., & Di Francesco, M. (2018). An Evaluation of Open Source Serverless Computing Frameworks. In CloudCom (pp. 115-120).
- [6] Gottlieb, N. (2016). State of the Serverless Community Survey Results. <https://serverless.com/blog/state-of-serverless-community/>.
- [7] Jonas, E., Pu, Q., Venkataraman, S., Stoica, I., & Recht, B. (2017). Occupy the cloud: Distributed computing for the 99%. In Proceedings of the 2017 Symposium on Cloud Computing (pp. 445-451). ACM.
- [8] Fromm, K. (2012). <https://readwrite.com/2012/10/15/why-the-future-of-software-and-apps-is-serverless/>
- [9] Van Eyk, E., Iosup, A., Seif, S., & Thömmes, M. (2017). The SPEC cloud group's research vision on FaaS and serverless architectures. In Proceedings of the 2nd International Workshop on Serverless Computing (pp. 1-4). ACM.
- [10] Van Eyk, E., Toader, L., Talluri, S., Versluis, L., Uță, A., Iosup, A. (2018). Serverless is more: From paas to present cloud computing. IEEE Internet Computing, 22(5), 8-17.
- [11] Castro, P., Ishakian, V., Muthusamy, V., & Slominski, A. (2017). Serverless programming (function as a service). In 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS) (pp. 2658-2659). IEEE. DOI 10.1109/ICDCS.2017.305.
- [12] Rodríguez N., et al (2020). Computación Serverless para tratamiento de datos provenientes de dispositivos de IoT. WICC 2020.
- [13] Aslanpour M. and Toosi A. et al. (2021). Serverless Edge Computing: Vision and Challenges. ACSW '21, February 1–5, 2021, Dunedin, New Zealand.
- [14] Arokia Paul Rajan R (2020) A review on serverless architectures - function as a service (FaaS) in cloud computing. TELKOMNIKA Telecommunication, Computing, Electronics and Control. Vol. 18, No. 1, February 2020, pp. 530~537

Procesamiento de problemas de Mecánica de Sólidos en entornos de Cloud Computing. El caso de estudios paramétricos.

Carlos García Garino⁽¹⁾, Claudio Careglio⁽¹⁾, Elina Pacini^(1,2), Anibal Mirasso⁽¹⁾,
Luc Papeleux⁽³⁾, Jean Philippe Ponthot⁽³⁾.

⁽¹⁾ Universidad Nacional de Cuyo, Facultad de Ingeniería e ITIC,
⁽²⁾ CONICET,

⁽³⁾ Aerospace and Mechanical Engineering Department. LTAS-Computational mechanics.
Liège Université.

cgarcia@itu.uncu.edu.ar, epacini@itu.uncu.edu.ar, ccareglio@uncu.edu.ar,
L.Papeleux@uliege.be, JP.Ponthot@uliege.be

RESUMEN

El proyecto estudia el procesamiento en entornos de Cloud Computing de la simulación numérica de problemas con grandes deformaciones elastoplásticas. En muchos casos de interés se debe estudiar la sensibilidad de los resultados del problema frente a cambios en los datos de entrada y/o en la discretización del problema de interés. Por ejemplo, la simulación numérica del ensayo de tracción simple en el rango de grandes deformaciones presenta estas características. Desde el punto de vista de Cloud Computing se estudiará la eficiencia de algoritmos de planificación de trabajos como es el caso de Ant Colony Optimization (ACO) y Particle Swarm Optimization (PSO) para llevar a cabo la asignación de trabajos propios del procesamiento paramétrico de este tipo de problemas.

Palabras clave: *Cloud Computing, Procesamiento Paramétrico, Mecánica de Sólidos.*

CONTEXTO

El grupo de trabajo de la Universidad Nacional de Cuyo desarrolla sus tareas en el Instituto para las Tecnologías de la Información y las Comunicaciones, ITIC y la Facultad de Ingeniería. Actualmente, en el contexto de la convocatoria de Proyectos de

investigación SIIP 2019, se llevan a cabo dos proyectos de investigación. El primero de ellos está orientado al estudio de problemas de plasticidad con grandes deformaciones [1] y contempla el procesamiento en entornos de Cloud Computing. En este caso se trabaja en cooperación con el grupo del Prof. Ponthot de la Universidad de Lieja. El segundo proyecto estudia planificadores de tareas en el entorno de Cloud Computing [2].

1. INTRODUCCIÓN

El proyecto propone el estudio de problemas de Mecánica de Sólidos [3], así como su procesamiento en entornos de Cloud Computing. Para ello se emplean códigos basados en el Método de Elementos Finitos [4], en particular el código SOGDE3D [5]. Este programa, de índole académica, posee la capacidad de resolver problemas con grandes deformaciones elastoplásticas. De esta manera se pueden resolver complejos problemas de ingeniería estructural, ingeniería mecánica, etc.

La presente propuesta continúa proyectos previos de los autores [6,7] que en general han estudiado problemas de plasticidad con grandes deformaciones y su procesamiento en entornos de computación distribuida. En trabajos previos del grupo [8,9], pueden consultarse algunos avances y resultados obtenidos en el campo de la plasticidad con grandes deformaciones.

De los resultados obtenidos en los trabajos citados [8,9] surge que en muchos casos de interés práctico los resultados son sensibles a cambios en la geometría. Resulta entonces de interés plantear estudios paramétricos que permitan evaluar la sensibilidad de los resultados del problema en función del cambio de algunas variables del problema.

Cloud Computing se ha convertido en una herramienta de mucha potencialidad y valor para para ejecutar flujos de trabajo científicos y de ingeniería.

Los estudios paramétricos consisten en procesar un problema determinado con diferentes valores en los datos del problema, cuya variación sigue una ley previamente definida. Luego, en general, se debe procesar una cantidad de casos, cada una de las cuales conlleva una simulación de elementos finitos. Es conveniente procesar estas simulaciones en un entorno de computación distribuida y Cloud Computing aparece como una tecnología de interés [10,11].

La simulación numérica del ensayo de tracción simple, también conocido *necking* en la literatura debido la estricción que se produce en la probeta en presencia de grandes deformaciones, es un proceso relativamente complejo desde el punto de vista numérico. Para el caso de probetas rectangulares es necesario simular el problema en 3D y el procesamiento es costoso en términos de esfuerzo computacional. Entonces es necesario profundizar el estudio de los planificadores en entornos de Cloud Computing.

Un caso de interés académico es analizar este problema para pequeños espesores de la probeta, reproduciendo así (en el límite) un estado aproximado al problema plano de tensiones. Luego resulta de interés investigar la respuesta del problema frente a cambios en el espesor y la sensibilidad de los resultados obtenidos frente a dichos cambios.

También es importante seleccionar un entorno adecuado de Cloud Computing y estudiar los planificadores para asignar de manera óptima los recursos a los trabajos con el fin de disminuir los tiempos de cálculo.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto plantea el siguiente Objetivo General:

Investigar la sensibilidad de los resultados de la simulación de problemas con grandes deformaciones frente a cambios en la malla de elementos finitos y otras variables de entrada de interés. Para ello se llevarán a cabo estudios paramétricos en entorno de Cloud Computing.

Para alcanzar el Objetivo general se plantean los siguientes Objetivos específicos

- a) Relevar el estado del arte
- b) Estudiar la sensibilidad de los resultados de la simulación del ensayo de tracción simple para el caso de probetas rectangulares, frente a la variación del espesor.
- c) Estudiar y mejorar los planificadores para la ejecución de problemas de Cloud Computing.
- d) Formar Recursos Humanos
- e) Transferir y difundir los resultados obtenidos.

3. RESULTADOS OBTENIDOS/ESPERADOS

Resultados Obtenidos

Los resultados obtenidos al presente se informan desde el punto de vista disciplinar y de Cloud Computing.

Desde el punto de vista de la Mecánica Computacional, cabe destacar, entre otros estudios, la simulación numérica del problema de tracción simple para el caso de probetas de sección rectangular [12], como se observa en la Figura 1.

Se llevaron a cabo dos estudios de elementos finitos con el código SOGDE3D [6], uno de referencia que simuló el caso general 3D, ver

figura 2 superior y otro que modeló una probeta delgada, ver figura 2 inferior.

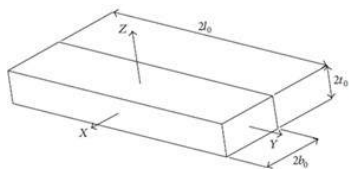


Figura 1: Probeta de sección rectangular, sistema de coordenadas y dimensiones $2b_0$, $2l_0$ y $2t_0$.

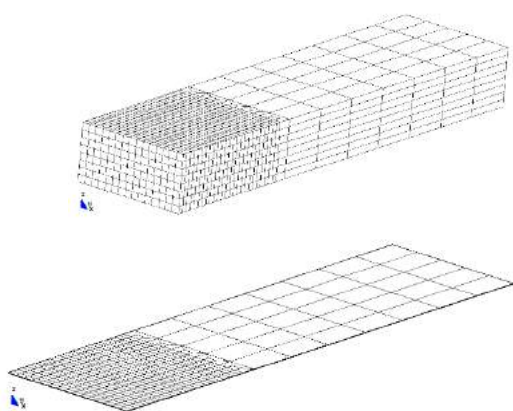


Figura 2: Mallas de Elementos Finitos para el para el caso general 3D (superior) y el caso de la probeta delgada (inferior).

Se ha comprobado que es factible obtener buenos resultados pese a la gran esbeltez de las probetas delgadas, cuyo modelado conlleva grandes relaciones de aspecto en los elementos hexaédricos utilizados en la malla de elementos finitos.

En la figura 4 se muestran de los contornos de la tensión efectiva de Von Mises, que es un indicador del daño que sufre el material. En la misma se observan dos bandas de color rojo dispuestas aproximadamente a 40 grados, que son características de la simulación de probetas delgadas y que no aparecen para los casos generales 3D (ver figura 2 superior).

También se han obtenido resultados valiosos en lo que respecta a la asignación de trabajos en Cloud Computing [13-15].

El procesamiento en la nube ha recibido mucho interés por parte de los investigadores disciplinares. La nube permite el acceso

rápido y elástico a grandes cantidades de recursos informáticos y entonces es posible ejecutar experimentos a gran escala con carga de trabajo variable, como resultan experimentos de barrido de parámetros (PSE).

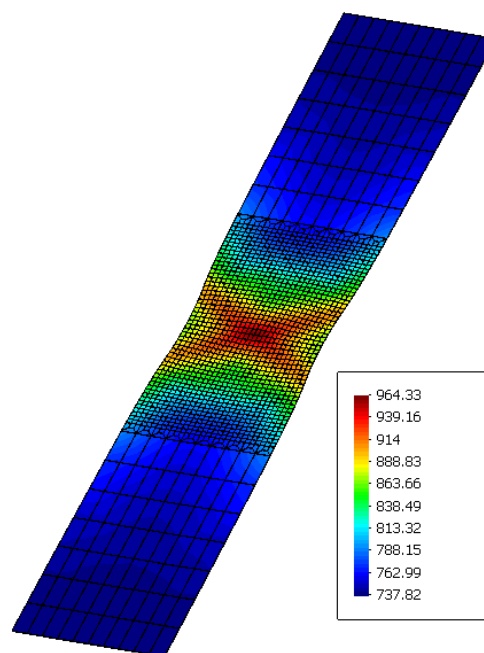


Figura 3: Contornos de tensiones efectivas de Von Mises, superpuestos a la malla deformada para el caso de la probeta delgada.

En este sentido resultan de gran utilidad Los escaladores automáticos o autoscalers que son componentes de software, a nivel de middleware que permiten incrementar o disminuir el tamaño de la plataforma de cálculo mediante la adquisición o finalización de las tareas de las máquinas virtuales (VM), en función de la carga de tareas del flujo de trabajo.

En un trabajo previo [13] se ha propuesto el escalador automático evolutivo multiobjetivo (MOEA), que utiliza un algoritmo de optimización multiobjetivo para determinar el conjunto de posibles configuraciones de infraestructura virtual. El rendimiento de MOEA está muy influenciado por el algoritmo de optimización subyacente utilizado y su ajuste. Luego se analizaron dos algoritmos evolutivos multiobjetivos bien

conocidos (NSGA-II y NSGA-III) y su impacto en el rendimiento del escalador. Los experimentos simulados con tres PSE del mundo real muestran que MOEA mejora en gran medida cuando se usa NSGA-III en lugar de NSGA-II debido a que el primero proporciona una mejor compensación de explotación frente a exploración.

En el trabajo [14] se propuso un escalador automático multiobjetivo en línea para flujos de trabajo denominado Cloud Multi-Object Intelligence (CMI). El objetivo de CMI es la minimización de la duración, el costo monetario y el impacto potencial de los errores derivados de la falta de fiabilidad de las VM. Además, el escalador está sujeto a restricciones presupuestarias monetarias propias del problema. Así CMI es responsable de resolver periódicamente de manera automática los problemas de ajuste de la plataforma simultáneamente con la ejecución del flujo de trabajo. Los experimentos realizados han mostrado muy buenos resultados que se discuten en el trabajo y se comparan con otros disponibles en la literatura.

Un problema clave del ajuste de escala automático del flujo de trabajo bajo restricciones presupuestarias (es decir, con un límite máximo en el costo) es determinar la proporción correcta entre: VM costosas pero confiables llamadas instancias bajo demanda, y VM más baratas pero sujetas a fallas llamadas instancias puntuales. Las instancias puntuales pueden proporcionar enormes posibilidades de paralelismo a bajo costo, pero deben usarse con prudencia, ya que pueden fallar inesperadamente y obstaculizar la fabricación. Dada la imprevisibilidad de las fallas y la variabilidad inherente del rendimiento de las nubes, diseñar una política para asignar el presupuesto para cada tipo de instancia no es una tarea trivial. Por tal razón, se ha formalizado el problema descrito como un proceso de decisión de Markov [15] que permite aprender políticas casi óptimas a partir de la experiencia de otras políticas de base. Los experimentos sobre cuatro flujos de trabajo científicos bien conocidos demuestran

que las políticas aprendidas superan a las políticas de referencia considerando la diferencia porcentual relativa agregada de la fabricación y el costo de ejecución.

Resultados Esperados

Se ´parte de los resultados preliminares informados en el trabajo acerca de la simulación del ensayo de tracción simple sometida de probetas delgadas [12]. A partir de los mismos se plantea diseñar primero y procesar y analizar luego un estudio paramétrico que varíe el espesor de la probeta y permita identificar el cambio de las variables de interés en función del espesor de la probeta. Es difícil estimar a priori la cantidad de casos a considerar. Posiblemente se parta de un número reducido de casos, 10 por ejemplo y una vez que se obtengan los primeros resultados *refinar* el estudio, agregando más casos con distintos espesores, que puedan cambiar de manera uniforme o quizás en un intervalo particular. Será de mucho interés analizar en este caso las dificultades de convergencia de los diferentes casos considerados. En la práctica, si cambia el número de iteraciones de equilibrio requeridas para alcanzar convergencia, o si fuese necesario cambiar el programa de cargas del proceso iterativo incremental propio de los estudios no lineales de elementos finitos, entonces puede cambiar significativamente el tiempo de cómputo de caso del estudio paramétrico.

Luego será de interés procesar el estudio paramétrico con algoritmos ya estudiados por los autores para este tipo de problemas. La novedad frente a los estudios previos [11], es que en los mismos los diferentes casos procesados presentaban mayor uniformidad en los tiempos de cómputo, característica que posiblemente no se cumpla para el estudio de interés y que quizás puede impactar en el rendimiento de los planificadores utilizados previamente [11].

Por otra parte, será de interés analizar la posibilidad de procesar estos problemas en entornos de cloud comerciales, además de cloud privadas, apuntando a instancias spot,

bajo demanda o una combinación de las mismas, siguiendo las recomendaciones de los trabajos previos de los autores [13-15].

4. FORMACIÓN DE RECURSOS HUMANOS

La Lic Yisel Gari, becaria doctoral de CONICET, trabaja en temas relacionados con los proyectos citados y está próxima a terminar su doctorado. El equipo del proyecto a cargo de García Garino incluye a varios alumnos avanzados de Ingeniería.

AGRADECIMIENTOS

Los autores agradecen el financiamiento recibido de la Universidad Nacional de Cuyo a través de los proyectos 06/B369 y B082.

5. BIBLIOGRAFÍA

- [1] Proyecto Bienal Tipo I: 06/B369 Simulación Numérica de Problemas con grandes deformaciones. Estudio de la sensibilidad de los resultados frente a cambios en la malla de elementos finitos y datos del problema. Director: Carlos García Garino. SIIP, UNCuyo, 2019-2021
- [2] Proyecto Bienal Tipo I: B082 Optimización y Autoescalado basado en Metaheurísticas para la ejecución eficiente de Aplicaciones Científicas en Infraestructuras Cloud. Director: Elina Pacini. SIIP, UNCuyo, 2019-2021
- [3] Niels Ottosen y Matii Ristinmaa: *The Mechanics of Constitutive Modelling*, Elsevier, 2005.
- [4] Oleg Zienkiewicz and Robert Taylor: *The Finite Element Method*, 5th Edition, Vol. 1 and Vol. 2, Butterworth-Heinemann, 2000.
- [5] Claudio Careglio: *Respuesta mecánica global y local en problemas elastoplásticos con grandes deformaciones*, Tesis Doctoral, Carrera de Doctorado en Ingeniería, Universidad Nacional de Cuyo, Mendoza, Argentina, 2017.
- [6] Proyecto 06/B308: *Procesamiento de problemas de mecánica de sólidos en entornos de computación distribuida*. Director Carlos García Garino. Proyecto bienal UNCuyo, 2013-2015.
- [7] Proyecto B057: *Simulación de estudios paramétricos de estudios de forjado en caliente sobre entornos de Cloud Computing*. Director Dr. Claudio Careglio. Proyecto Bienal Tipo II. UNCuyo, 2016-2018.
- [8] Carlos García Garino, Felipe Gabaldón y José M^a Goicolea: *Finite Element Simulation of Simple tension test.*, *Finite Element Analysis and Design*, 2, 1187-1197, 2006.
- [9] Claudio Careglio, Diego Celentano, Carlos García Garino y Anibal Mirasso: *Global and Local Mechanical Responses for Necking of Rectangular Bars Using Updated and Total Lagrangian Finite Element Formulations*. *Mathematical Problems in Engineering*. Hindawi Publishing Corporation. 2016 vol.2016 n^o. p1 - 12. issn 1024-123X. eissn 1563-5147
- [10] Carlos García Garino, Elina Pacini, David Monge, Claudio Careglio y Anibal Mirasso: *Computational Mechanics Software as a Service Project*. *Journal of Computer Science & Technology*. La Plata: ISTECS., 2013 vol.13 n^o. p160 - 166. issn 1666-6038.
- [11] Elina Pacini, Cristian Mateos, Carlos García Garino, Claudio Careglio y Anibal Mirasso A: *A bioinspired scheduler for minimizing makespan and flowtime of computational mechanics applications on federated clouds.* *Journal of Intelligent and Fuzzy Systems*. Amsterdam. Ios Press. 2016 vol.31 n^o3. p1731 - 1743. issn 1064-1246.
- [12] Claudio Careglio, Carlos García Garino, Anibal Mirasso, Diego Celentano, Luc Papeleux y Jean-Philippe Ponthot: *simulación numérica del ensayo de tracción simple Para probetas delgadas. Análisis del caso límite 3D. Enviado para su evaluación y eventual publicación al Congreso MACI 2001. En revisión.*
- [13] Virginia Yannibelli, Elina Pacini, David Monge, Cristian Mateos y Guillermo Rodriguez: *A Comparative Analysis of NSGA-II and NSGA-III for Autoscaling Parameter Sweep Experiments in the Cloud*. *Sci. Program*. 2020: 4653204:1-4653204:17 (2020).
- [14]. Monge, D.; Pacini, E.; Mateos, C.; Alba, E., García Garino, C.: *CMI: An Online Multi-objective Genetic Autoscaler for Scientific and Engineering Workflows in Cloud Infrastructures with Unreliable Virtual Machines*. *Journal of Network and Computer Applications*, 149 (2020).
- [15] Yisel Gari, David Monge, Cristian Mateos, Carlos García Garino: *Learning budget assignment policies for autoscaling scientific workflows in the cloud*. *Clust. Comput*. 23(1): 87-105 (2020).

Procesamiento paralelo sobre arquitecturas serverless para tratamiento de datos provenientes del IoT

Nelson Rodríguez¹, María Murazzo¹, Diego Medel¹, Lorena Parra¹, Ana Laura Molina¹, Federico Sánchez¹, Adriana Martín¹, Fabiana Piccoli², Zamira Ale³, Lucas Vargas³

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

⁶ 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, vdiego.unsj@hotmail.com, lorenaparra152@yahoo.com.ar, almm95@gmail.com, fgsanchez@unsj-cuim.edu.ar, adrianamartin1@gmail.com, mpiccoli@unsl.edu.ar, zamiranahirale@gmail.com, lucasgvargas96@gmail.com

Resumen

Según un estudio realizado por la multinacional Ericsson, para 2022 existirán más de 18 mil millones de dispositivos IoT comunicados entre sí. Todos estos dispositivos generan una gran cantidad de datos que deberán ser procesados en el menor tiempo posible, con el objeto de poder realizar una toma de decisiones óptima. En este sentido, el cloud parece ser la solución obvia, pero hay casos en los que la latencia que viene con la transmisión de datos de un lado a otro podría ser una limitación. En determinadas situaciones, tampoco es factible esperar una conexión a Internet constante y confiable, ya sea porque podría no ser viable económicamente o porque puede que no sea posible desde el punto de vista de la infraestructura.

Con el objeto de solucionar estos problemas, se propone trabajar con el paradigma serverless, en articular con FaaS usando lenguajes de programación paralelos que permitan reducir el tiempo del procesamiento de los datos obtenidos desde los dispositivos de IoT.

Palabras Claves: IoT, Serverless Computing, FaaS, Paralelismo, Métricas.

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es una de las líneas de investigación del proyecto: Computación Serverless para el tratamiento de datos provenientes de dispositivos de IoT, cuya propuesta ha sido aprobada y está en desarrollo para el período 2020-2021. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación distribuida y de alta performance desde hace más de 20 años. Como continuación del proyecto anterior: Orquestación de Servicios para la Continuidad Edge al Cloud, se sigue trabajando con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

IoT [1] es una tecnología creada para optimizar procesos. Esta tecnología, que busca interconectar dispositivos a través de internet para facilitar y acelerar tareas, está avanzando a pasos agigantados. Según un estudio realizado por la multinacional Ericsson, para 2022 existirán más de 18 mil millones de dispositivos IoT comunicados entre sí. Pero este crecimiento de dispositivos inteligentes trae aparejado la generación de grandes cantidades de datos, los cuales no pueden ser tratados y procesados con los paradigmas de programación tradicionales [2]. Es por ello que es necesario aplicar algoritmos que permitan aprovechar la escalabilidad de recursos de cómputo y procesamiento de datos [3] [4].

En este sentido se plantea como solución al procesamiento de datos provenientes del IoT, técnicas de computación de alta prestaciones (HPC) con el fin de aumentar la velocidad de procesamiento. HPC es la evolución de los sistemas de cómputo convencional, los cuales permiten realizar operaciones de cómputo intensivo y mejorar la velocidad de procesamiento; involucrando diferentes tecnologías tal como los sistemas distribuidos y los sistemas paralelos.

Todos estos entornos son ideales para resolver aplicaciones científicas, computacionalmente costosas con manejo de grandes cantidades de datos, a fin de lograr resultados en menor tiempo. Dadas estas características, estas arquitecturas son candidatas ideales para procesar datos provenientes del IoT; puesto que estos son generados a una elevada tasa y requieren un tratamiento en tiempo real [5].

Si bien, las arquitecturas HPC han evolucionado en pos de obtener mejores tiempos de respuesta para las aplicaciones, presentan el inconveniente del escalado, tanto vertical como horizontal, de recursos de cómputo. Es por ello que una alternativa es el cloud, el cual permite contar con una cantidad de recursos computacionales virtualmente infinitos, administrados por terceros y accedidos bajo demanda pagando por el uso,

de esta manera es posible lograr calidad de servicio (QoS) garantizado por parte del proveedor de la infraestructura cloud [6] [7].

Cloud Computing se ha convertido en una tecnología centrada en ofrecer cómputo bajo demanda como cualquier otro servicio. Uno de los aspectos clave que afecta la performance de las aplicaciones al montarlas sobre arquitectura cloud, es la capacidad de virtualización de cualquier recurso (bases de datos, red, procesador, etc.) y ofrecerlo como un servicio (AaaS, Anything as a Service). Esto puede tener un impacto negativo en la ejecución de aplicaciones de cálculo intensivo, debido a que los ambientes cloud deben virtualizar la plataforma en la cual se ejecuta y esto conlleva un overhead de a contextualización de las máquinas virtuales [8] [9] [10].

El principal inconveniente que existe con migrar las aplicaciones al cloud, es que es responsabilidad de la organización mantener funcionando de forma correcta la infraestructura (IaaS) que se necesite para el despliegue de las aplicaciones. Esta responsabilidad lleva a cargar costos sobre el presupuesto para el mantenimiento y soporte de la infraestructura.

En este sentido, la aparición del Serverless Computing [11] [12] supone una revolución en el desarrollo de servicios web. Los desarrolladores ya no se tienen que preocupar por el aprovisionamiento y escalado de la infraestructura, por lo que se pueden centrar en la lógica de sus aplicaciones. Serverless Computing trata de abstraer de la gestión de servidores (aprovisionamiento, configuración, escalado, etc.) para que los usuarios, en este caso desarrolladores, puedan enfocarse en la lógica de sus aplicaciones. Este paradigma se relaciona normalmente con el modelo FaaS [13], que consiste en la definición de funciones en un determinado lenguaje de programación que son ejecutadas sobre un proveedor Cloud. Para gestionar este tipo de servicios se utilizan tecnologías de contenedores software, que son los entornos sobre los cuales se ejecutan las funciones, si bien su gestión suele quedar oculta a los desarrolladores.

FaaS Paralelas

Con FaaS se llega a la máxima expresión de abstracción pues no se despliegan microservicios si no funciones que se ejecutan bajo eventos. Es decir, se basa en una arquitectura bajo eventos donde se abstrae completamente de toda la infraestructura subyacente. Se pueden hacer configuraciones básicas, pero olvidándose de la gestión del escalado, disponibilidad, despliegue y configuraciones de recursos. Aquí es donde llega el concepto de arquitectura serverless a tener una capa de abstracción que donde no necesita saber las características de los recursos subyacentes. Esto permite que sea un servicio con auto-escalado completo y solo se pague por el consumo real, ya que no se reservan recursos de computación. Además, las funciones solo se ejecutan cuando ha habido un evento y, en caso contrario, se mantienen “apagadas” sin consumir recursos. Otra característica para tener en cuenta es que deben ser ejecuciones en tiempos mínimos, entre 5 y 10 minutos, ya que las funciones deberán consumir recursos el mínimo tiempo posible. En caso contrario es necesario dividir esa función en funciones más pequeñas para distribuir los procesos.

En función de estas características se puede ver que en FaaS la lógica es, desarrollada dentro de un flujo de trabajo clásico que automáticamente es desplegada en procesos sin estado que son capaces de responder a distintos eventos que se produzcan en la infraestructura, con tiempo de vida limitado y completamente gestionada por el proveedor.

Hay diversas razones para desplegar IoT serverless entre las que se incluyen [14]:

- Al construir un backend basado en Cloud es difícil de monitorear la escalabilidad y disponibilidad. Las redes IoT suelen estar compuestas de millones de nodos, a través de distintas zonas geográficas, utilizando una arquitectura serverless permite eliminar toda la gestión sobre los proveedores Cloud.
- Las redes IoT suelen tener variaciones de carga donde a ciertas horas del día o

debido a determinados eventos se puede desencadenar que se active un gran número de dispositivos. Debido a que los servicios IoT suelen ejecutar un gran número de dispositivos inactivos, estos están esperando a ser activados y, por lo tanto, tener todos los procesos con recursos reservados no es óptimo y causará sobrecostos por infrautilización. Construir estos servicios con una arquitectura serverless permite pagar por lo que realmente consumen.

- El software que ejecuta y habilita los servicios de IoT generalmente comienza de manera simple, pero se van agregando características continuamente. La adopción de metodologías serverless permite a las organizaciones centrarse en su valor comercial principal, agilizar los cambios y el time to market en vez de la orquestación y el mantenimiento.

En base a esto se puede observar como FaaS encaja perfectamente con una solución paralela para procesar datos del IoT [15].

Líneas de Investigación, Desarrollo e Innovación

Este trabajo presenta una línea de investigación dentro del proyecto que comienza a desarrollarse este año. Si bien se ha trabajado anteriormente con paradigmas de HPC para procesamiento de datos masivos provenientes de aplicaciones móviles y de dispositivos de IoT, esto se ha realizado sobre arquitecturas clustes y cloud.

Esta línea de investigación estará orientado a investigar el concepto de arquitectura serverless o FaaS (Functions as a Service) para soluciones IoT. Debido a la disrupción de las tecnologías y su gran velocidad de desarrollo aparece la necesidad de implementar soluciones escalables, flexibles, estables y a bajos costos. Por lo tanto, el concepto de arquitectura serverless encaja perfectamente con estas necesidades. Pudiendo conseguir desplegar soluciones paralelas para IoT sin la necesidad de tener una infraestructura, ni costos de mantenimiento y operación.

Con el objeto de trabajar con serverless computing, las investigaciones comenzarán con un análisis de las plataformas cloud más adecuadas para conducir las investigaciones. Posteriormente y una vez elegida la plataforma cloud se realizarán trabajos de procesamiento paralelo analizando las métricas y midiendo el grado de eficiencia y eficacia de las soluciones planteadas.

Resultados y Objetivos

Resultados Obtenidos

Durante los últimos doce años se trabajó en el área de Computación de Altas Prestaciones y distribuidas. Dicha experiencia impulsó el proyecto de investigación y la línea que en este trabajo se presenta.

Objetivos

El objetivo del grupo de investigación es trabajar con serverless computing sobre cloud públicos mediante FaaS que usen lenguajes de programación paralela. Una vez realizadas las ejecuciones, se realizará el análisis de las métricas con el objetivo de determinar cuál es el lenguaje que mejores resultados arroja en función de los datos procesados.

La finalidad última de esta línea de investigación es determinar el grado de satisfacción alcanzado con el procesamiento paralelo en serverless computing contra los paradigmas tradicionales de procesamiento.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de ocho investigadores que figuran en este trabajo de las universidades, Nacional de Salta y San Juan y dos alumnos de grado. Además, el proyecto marco donde se está desarrollando esta propuesta incluye a tres investigadores más de la Nacional de San Luis, de la Universidad Champagnat y de la Universidad Nacional de San Juan y tres alumnos de grado.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, seis tesis de grado en el área de Serverless computing, Concurrencia y Computación distribuida y una tesis de maestría en áreas afines. Además se espera aumentar el número de publicaciones. Por otro lado 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 otros organismos del estado.

Bibliografía

- [1] Stoyanova, M., Nikoloudakis, Y., Panagiotakis, S., Pallis, E., & Markakis, E. K. (2020). A survey on the internet of things (IoT) forensics: challenges, approaches, and open issues. *IEEE Communications Surveys & Tutorials*, 22(2), 1191-1221.
- [2] Farhan, L., Kharel, R., Kaiwartya, O., Quiroz-Castellanos, M., Alissa, A., & Abdulsalam, M. (2018, July). A concise review on Internet of Things (IoT)-problems, challenges and opportunities. In *2018 11th International Symposium On Communication Systems, Networks & Digital Signal Processing (CSNDSP)* (pp. 1-6). IEEE.
- [3] Medel, D., Murazzo, M. A., Molina, A. L., Sánchez, F., Cornejo, M., Rodríguez, N. R., ... & Piccoli, M. F. (2019). La Computación de Alta Performance como soporte a los sistemas altamente distribuidos. In *XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)*.
- [4] Barrionuevo, M., Escalante, J., Lopresti, M., Lucero, M., Miranda, N. C., Murazzo, M. A., & Piccoli, M. F. (2020). Solución de grandes problemas aplicando HPC multi-tecnología. In *XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz)*.
- [5] de Souza Cimino, L., de Resende, J. E. E., Silva, L. H. M., Rocha, S. Q. S., de Oliveira Correia, M., Monteiro, G. S., ... & de Castro Lima, J. (2017, November). IoT and HPC integration: revision and perspectives. In *2017 VII Brazilian Symposium on Computing Systems Engineering (SBESC)* (pp. 132-139). IEEE.

- [6] Mauch, V., Kunze, M., & Hillenbrand, M. (2013). High performance cloud computing. *Future Generation Computer Systems*, 29(6), 1408-1416.
- [7] Murazzo, M. A., Rodriguez, N. R., Guevara, M. J., & Tinetti, F. G. (2016). Identificación de algoritmos de cómputo Intensivo para big data y su implementación en clouds. In *Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina)* (Vol. 18).
- [8] Rodríguez, N. R., Murazzo, M. A., Medel, D., Chávez, S. B., Martín, A. E., & Valenzuela, A. (2015). Análisis de mejora en la escalabilidad de las infraestructuras de cloud computing. In *II Jornadas de Cloud Computing & Big Data (La Plata, 2015)*.
- [9] Murazzo, M. A., Castiglione, F., Medel, D., Tinetti, F. G., & Rodríguez, N. R. (2016). Análisis de la migración del cómputo intensivo a un cloud privado con OpenStack. In *XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*.
- [10] Murazzo, M. A., & Rodríguez, N. R. (2016). Evaluación del Impacto de Migración al Cloud. *Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação*, 1(4).
- [11] Baldini, I., Castro, P., Chang, K., Cheng, P., Fink, S., Ishakian, V., ... & Suter, P. (2017). Serverless computing: Current trends and open problems. In *Research Advances in Cloud Computing* (pp. 1-20). Springer, Singapore.
- [12] McGrath, G., & Brenner, P. R. (2017, June). Serverless computing: Design, implementation, and performance. In *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)* (pp. 405-410). IEEE.
- [13] Fox, G. C., Ishakian, V., Muthusamy, V., & Slominski, A. (2017). Status of serverless computing and function-as-a-service (faas) in industry and research. *arXiv preprint arXiv:1708.08028*.
- [14] Ciuffoletti, A. (2020). Stateless IoT. *Information*, 11(2), 85.
- [15] Pinto, D., Dias, J. P., & Ferreira, H. S. (2018, October). Dynamic allocation of serverless functions in IoT environments. In *2018 IEEE 16th international conference on embedded and ubiquitous computing (EUC)* (pp. 1-8). IEEE.

Redes neuronales paralelas aplicadas a la visión computacional

Galdamez Mariela¹, Chirino Pamela¹, Diaz-Acevedo Karvin¹, Ponce de León Alejo¹, Caymes-Scutari Paola^{1,2}, Bianchini Germán¹

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
mariela.galdamez.16@gmail.com, pamelaachirino@gmail.com, karvindiaz@gmail.com,
alejo.poncedleon@gmail.com, pcaymesscutari@frm.utn.edu.ar, gbianchini@frm.utn.edu.ar

RESUMEN

La capacidad de permitir que una computadora reconozca en una imagen los objetos, ambiente y posición en el espacio fue el inicio de la visión computacional, un área dentro de la inteligencia artificial. Al utilizar el modelo de redes neuronales en este campo, se consiguen resultados en diversas ramas de la ciencia. Por ejemplo, diagnóstico de ciertas patologías en imágenes provenientes de ecografías o resonancia magnética; realizar vigilancia, reconocimiento dactilar y ocular como mecanismos de seguridad; entre otras tantas aplicaciones. Pero estructurar una red neuronal no es tarea sencilla, pues para conseguir que funcione y aprenda adecuadamente se requiere de un conjunto de datos de entrada que deberá analizar cierta cantidad de veces hasta producir una salida coherente con los datos ingresados. Esto implica la necesidad de una velocidad de cómputo enorme para que sea capaz de aprender en un período de tiempo razonable. Además, si el tamaño de la red es mucho mayor,

los datos de entrada aumentan en cantidad y se complejiza el aprendizaje de la misma, lo que infiere en un aumento considerable de tiempo. Los procesadores actuales no brindan la velocidad suficiente y es aquí donde la programación paralela se presenta como una solución alternativa.

Palabras clave: Redes Neuronales, Programación Paralela, Visión Computacional.

CONTEXTO

La línea de trabajo propuesta está incluida en el proyecto TEUTIME0007658TC [1], titulado “Formación de docentes y alumnos de grado como Investigadores Científicos Iniciales en las áreas de Informática y Ciencias de la Computación”. Este proyecto es financiado por la Universidad Tecnológica Nacional y llevado a cabo en el Laboratorio de Investigación en Cómputo Paralelo/Distribuido ubicado en la Facultad Regional Mendoza. Propone

la iniciación de docentes y estudiantes en el proceso de investigación científica.

1. INTRODUCCIÓN

Una de las disciplinas de la inteligencia artificial es la ampliamente conocida como visión computacional. La misma consiste en adquirir, procesar, analizar y comprender las imágenes del mundo real, recibidas por dispositivos como cámaras, con el fin de producir información numérica o simbólica que pueda ser tratada por una computadora [2]. Existen numerosas situaciones en donde se puede aplicar la visión computacional para la solución de determinados problemas [3], y cabe destacar el uso de redes neuronales como una opción de modelo computacional a emplear. Una red neuronal está caracterizada por su estructura, su aprendizaje y su función de activación. Según Laurene Fausett en [4], “la estructura de una red neuronal consiste en un gran número de elementos simples de procesamiento llamados neuronas o nodos”, las cuales se organizan en una única capa de entrada, una única capa de salida y en capas ocultas, posicionadas entre las dos mencionadas anteriormente. Además, este autor describe en [4] que “cada neurona está conectada a otras neuronas por medio de enlaces de comunicación dirigidos, cada uno con un peso asociado, y cuenta con un estado interno, llamado su nivel de activación, el cual es una función de las entradas que ha recibido. Por lo general, una neurona envía su activación, es decir, su única salida, como una señal a varias otras neuronas”. La clasificación de una red neuronal depende de su estructura, específicamente de la cantidad de capas ocultas y de la forma de conexión entre las neuronas de las mismas [5]. Para este proyecto, se escogió trabajar con redes

neuronales multicapas, que se caracterizan por tener una o varias capas ocultas, y por las conexiones inexistentes entre neuronas de la misma capa. De todas formas, no se descarta el estudio de las otras variantes en el futuro. Lo más importante en una red neuronal es su capacidad de aprendizaje. Una vez elegido un volumen de datos considerable para su entrenamiento y para la validación de este, la red neuronal procede a hacer un ajuste de los pesos de las conexiones entre sus neuronas a medida que recorre los datos de entrada. Una red neuronal multicapa puede resolver problemas más complicados, pero su aprendizaje se vuelve mucho más difícil de concretar [4]. Además, el tiempo de aprendizaje de una red neuronal multicapa es una variable que aumenta a medida que la estructura de la misma se complejiza. Aprovecharemos esta problemática para introducir el cómputo paralelo, una técnica que permite atender a la demanda de gran velocidad computacional para resolver determinados problemas, en ámbitos como la ciencia y la ingeniería. Tales problemas suelen necesitar enormes cantidades de cálculos repetitivos en grandes volúmenes de datos para dar resultados válidos, los cuales deben ser completados dentro de un período de tiempo “razonable” [6]. En el caso particular de las redes neuronales, con solo calcular la función de un número grande de entradas en una neurona, se infiere en un costo computacional importante. Luego, a este costo se le suman los cálculos realizados por las neuronas restantes y el ajuste de todos los pesos en la red. Todo este costo computacional para una única entrada ejemplo, después la red neuronal deberá analizar los ejemplos restantes. Por consiguiente, para las redes neuronales, el paralelismo es sumamente aplicable, y es la solución que se propone a la

complejidad computacional del aprendizaje de las mismas.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Esta temática se enmarca en una propuesta de investigación enfocada en el cómputo paralelo como herramienta para la reducción en tiempo de ejecución en diversas problemáticas. Dentro de este marco, se ha investigado sobre redes neuronales con el objetivo de que las mismas mejoren la toma de decisiones en un modelo de predicción de incendios [7]. Actualmente, el enfoque se orienta a la visión computacional y específicamente al uso de redes neuronales multicapa como modelo computacional. Las redes neuronales multicapa son capaces de hacer tareas muy complejas. Sin embargo, es imprescindible el paralelismo puesto que para conseguir la estructura correcta de la red neuronal es necesaria una cantidad importante de cómputo [8]. Por esto mismo, se pretende aplicar los conocimientos actuales en programación paralela y los algoritmos desarrollados en este ámbito para mejorar la eficiencia del aprendizaje de una red neuronal orientada al reconocimiento y clasificación de objetos en una imagen. A partir de los avances que se obtengan, se pretende encarar a futuro una propuesta de investigación enfocada en reconocer y determinar un delito u accidente vial en tiempo real, con el objetivo de alertar a los organismos correspondientes de forma casi inmediata, quienes se encargarán de tomar las medidas adecuadas.

3. RESULTADOS OBTENIDOS/ ESPERADOS

El principal aporte que se espera alcanzar es brindar una mejora en la velocidad de aprendizaje de las redes neuronales multicapa al utilizar paralelismo. Contribuir con la computación paralela en el ámbito de la inteligencia artificial permitiría cubrir las exigencias de cómputo en los diversos ámbitos que hacen uso de los algoritmos de redes neuronales. Otro aporte que se pretende conseguir es expresar estadísticamente los beneficios del cómputo paralelo en éstas. Para contabilizar estas estadísticas, se busca realizar una comparativa del tiempo de ejecución requerido para conseguir un aprendizaje adecuado de distintas redes neuronales multicapa. Éstas se diferenciarán según el tamaño, la cantidad de ejemplos que requieren aprender para funcionar correctamente, y la cantidad y las características de los recursos utilizados. Finalmente, se espera comprobar que, al paralelizar, se consigue en un tiempo menor una convergencia a algo cercano al óptimo global de los pesos en cada neurona de la red neuronal.

4. FORMACIÓN DE RECURSOS HUMANOS

La temática propuesta por este proyecto permite la formación de los distintos integrantes en relación a la inteligencia artificial y la programación paralela. Los autores Mariela Galdamez, Pamela Chirino, Karvin Diaz Acevedo y Alejo Ponce de León son estudiantes de tercer año en la carrera Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Mendoza. Además, son integrantes del Laboratorio LICPaD, dirigido por el Dr. German Bianchini y la Dra. Paola Caymes Scutari, en el cual realizan actividades de investigación de índole científica complementarias a sus estudios.

5. BIBLIOGRAFÍA

- [1] Bianchini Germán, Caymes Scutari Paola, Ontiveros Patricia, Rotella Carina, Salinas Sergio, Tagarelli Sandra, Chirino Pamela, Galdamez Mariela, Formación de docentes y alumnos de grado como Investigadores Científicos Iniciales en las áreas de Informática y Ciencias de la Computación. Actas del XXII Workshop de Investigadores en Ciencias de la Computación: WICC 2020. ISBN 978-987-3714-82-5. pp.590-594.
- [2] Forsyth D., Ponce J. (2012) Computer Vision - A Modern Approach. Pearson.
- [3] Iván García S., Víctor Caranqui S. (2015) La Visión Artificial y los Campos de Aplicación. Artículo Revista Digital “Tierra Infinita”. URL: <https://revistasdigitales.upec.edu.ec/index.php/tierrainfinita/article/view/76> (fecha de consulta: febrero 2021).
- [4] Fausett L. (1994) Fundamentals of Neural Networks – Architectures, Algorithms, and Applications. Pearson.
- [5] Berzal F. (2018) Redes Neuronales y Deep Learning. Autoedición.
- [6] Wilkinson B., Allen M. (2005) Parallel Programming – Techniques and Applications Using Networked Workstations and Parallel Computers. Pearson.
- [7] Chirino P., Galdamez M., Bianchini G., Caymes-Scutari P., Propuesta de paralelización en redes neuronales. Libro de Actas 7mo CONNAIISI 2019. pp.1620--1630. Edit. UNLaM, San Justo (2020)
- [8] Russell S., Norvig P. (2004) Inteligencia Artificial – Un Enfoque Moderno. Pearson.

Sintonización de un Método de Reducción de Incertidumbre: un proceso de medición-mejora para incrementar la eficiencia

Caymes-Scutari Paola^{1,2}, Bianchini Germán¹, Méndez-Garabetti Miguel^{1,2}

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

pcaymesscutari@frm.utn.edu.ar, gbianchini@frm.utn.edu.ar, miguelmendezgarabetti@gmail.com

RESUMEN

La reducción de incertidumbre constituye un proceso complejo orientado a contrarrestar o paliar los efectos negativos que ocasiona la imprecisión en los datos de entrada de un programa o en los cálculos que se realiza con ellos. ESS (*Evolutionary Statistical System*) constituye un método de reducción de incertidumbre en el proceso de predicción de fenómenos de propagación, y su funcionamiento tiene tres pilares como fundamento: un algoritmo evolutivo para guiar la búsqueda, la estadística para determinar la tendencia, y el paralelismo para potenciar la cantidad de procesamiento computacional que debe realizarse. Si bien ESS obtiene predicciones suficientemente acertadas, en este proyecto se propone incorporar capacidades para la sintonización dinámica y automática. Ello significa que debe desarrollarse un modelo de medición-mejora, con la finalidad de detectar las situaciones de sobrecarga computacional que se susciten durante la ejecución, para ajustar y/o adecuar el comportamiento dinámico de la aplicación y así lograr una

ejecución globalmente más eficiente y precisa.

Palabras clave: Sintonización Automática, Modelo de Rendimiento, Sintonización Dinámica, Aplicaciones Paralelas, Reducción de Incertidumbre.

CONTEXTO

El año 2020 trajo consigo la pandemia de COVID-19, y ello impuso cambios en todos los ámbitos de la vida. La llamada ‘virtualidad’ se volvió la nueva forma de comunicación y de trabajo, demandando una rápida adaptación al cambio de paradigma y un esfuerzo adicional para llevar a la virtualidad tantos aspectos cotidianos y/o académicos que usualmente se transmitían o elaboraban de forma presencial. Sumado a ello, los órganos financiadores también introdujeron modificatorias en cuanto a las normativas relativas a la utilización del presupuesto de los proyectos, restricciones en las compras de

equipamiento, retrasos en las comunicaciones y formalidades atinentes a la ejecución de los proyectos, etc. Todo ello ha ralentizado el avance de los proyectos, pues el personal involucrado se ha visto afectado por la sobrecarga de la virtualidad, y refrenado por las restricciones presupuestarias. En este contexto, en el ámbito de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional (UTN-FRM), más específicamente en el marco del LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido) hemos continuado con el desarrollo del Proyecto de I+D SIUTIME0007840ME financiado por la UTN (que inició su ejecución en enero de 2020), y estamos a la espera del inicio del recientemente aprobado Proyecto de Investigación Científica y Tecnológica Orientado PICTO-UUMM-2019-00042, adjudicado por el FONCyT. Ambos proyectos dan continuidad a la línea de investigación de proyectos previos del LICPaD, pero en esta oportunidad haciendo foco en la necesidad del proceso de medición-mejora de un método de reducción de incertidumbre para la predicción de fenómenos de propagación.

La necesidad y motivación por medir está ligada fundamentalmente a la mejora de calidad continua que requieren los procesos y productos. Para poder asegurar que un proceso o sus productos resultantes son de calidad, o bien para poder realizar comparaciones, es necesario asignar valores, descriptores, indicadores o algún otro mecanismo mediante el cual se pueda llevar a cabo dicha comparación [1]. La predicción de fenómenos de propagación como tema específico de la ciencia computacional no escapa a esta necesidad de medición, evaluación y mejora, y es allí donde emerge la Sintonización como un proceso de medición-mejora permanente que se

desea incorporar al método de predicción ESS [2], con la finalidad de ajustar y/o adecuar el comportamiento dinámico del mismo y lograr una ejecución globalmente más eficiente [2]. En 2021 esperamos poder avanzar en esta dirección, más allá de las dificultades y restricciones operativas ocasionadas por el contexto de pandemia.

1. INTRODUCCIÓN

La predicción de fenómenos de propagación constituye una tarea compleja desde distintos puntos de vista. En primer lugar, ha de considerarse el volumen de información que es necesario manejar y procesar, acompañado de la variedad de técnicas y conceptos provenientes de distintas áreas entrelazados para modelar y resolver la situación [2]. En segundo lugar, y teniendo en cuenta los desastres que podrían ocasionar los distintos fenómenos de propagación, el nivel de precisión alcanzado por los resultados de la predicción constituye un aspecto crítico para asistir a la toma de decisiones. En tercer lugar, puede mencionarse la velocidad de respuesta con la que la predicción es arrojada como otro aspecto crítico para la toma de decisiones paliativas.

El primer aspecto mencionado en el párrafo anterior constituye la principal fuente de incertidumbre, por la imprecisión o discontinuidad de la información y las mediciones o estimaciones de los parámetros y datos, los errores o redondeos en los modelos y sus implementaciones, etc. Todo ello es, de alguna manera, inevitable y por lo tanto es menester trabajar en pos de minimizar el efecto negativo que la incertidumbre proveniente de las imprecisiones ocasiona en la salida del

sistema. El segundo aspecto, conjuntamente con el tercero, hacen necesario un análisis más profundo del método de reducción de incertidumbre, a fin de mejorar su precisión y su velocidad de respuesta. Por ello, ha de introducirse el proceso de sintonización. La sintonización es un paradigma tecnológico que permite que los programas sean adaptables a una variedad de condiciones computacionales [2,4,5]. El proceso de Sintonización involucra varias **fases** sucesivas que se relacionan en pos del objetivo de mejorar la calidad de la aplicación y de su ejecución. La primera etapa es la **Instrumentación**, en la cual la aplicación es anotada a fin de habilitar la medición de valores propios de la aplicación y/o su ejecución (capturar valores de variables, medir tiempos de ocio o ejecución, etc.). La segunda etapa es la **Monitorización**, durante la cual se recolectan, se clasifican y se almacenan los datos obtenidos por medio de la instrumentación. La siguiente etapa es la de **Análisis**, en la que las mediciones y la información monitorizadas son evaluadas y valoradas a fin de detectar posibles cuellos de botella que actúen en detrimento de la calidad de la ejecución [3,5]. La etapa final del proceso se denomina **Sintonización** en sí misma, pues es la que introduce y materializa los cambios necesarios en la aplicación para adaptar su comportamiento y ajustar su rendimiento.

En el caso de ESS, podemos resumir su funcionamiento diciendo que realiza las predicciones en base a múltiples variaciones en los parámetros de entrada para constituir un conjunto de soluciones candidatas (o población), cuya transformación para alcanzar un resultado de calidad está orientada por un algoritmo evolutivo [7]. Es por ello que ESS realiza la predicción en base a la información que

le brinda una población de soluciones [7, 8]. Dicha población constituye una muestra del universo de búsqueda, y está compuesta por elementos del dominio de búsqueda (también denominados **individuos** o **escenarios**, con diferentes características: cada individuo, usualmente, representa un conjunto de parámetros con los que ha de ejecutarse una simulación, y es en el valor de tales parámetros de entrada en donde difieren unos individuos de otros, con el consecuente efecto en la salida que arrojará cada simulación). La calidad de los individuos y la predicción que arrojan se mide a través de una función de aptitud. Los individuos son sometidos a transformaciones evolutivas, mediante los operadores de cruzamiento y mutación, hasta que se converge a un óptimo o a un cierto criterio de parada. Los criterios de parada permiten decidir cuándo finalizar una generación para pasar a la siguiente, y cuándo terminar el algoritmo [6,7]. Frecuentemente el proceso evolutivo se ve degradado por la convergencia prematura o el estancamiento, dos situaciones que van en detrimento de la calidad de los resultados y que sobrecargan y retrasan la toma de decisiones incrementando innecesariamente el tiempo de cómputo [5]. Teniendo como antecedente otros trabajos en el área [9,10], en este trabajo se propone dar tratamiento al estancamiento y convergencia [5] en ESS, a fin de modelar, detectar y actuar en consecuencia para aplicar acciones de sintonización que permitan paliar el efecto negativo que ocasionan en los resultados

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La temática propuesta continúa la línea de trabajo de otros cinco proyectos previos vinculados a la sintonización y al desarrollo automático. Se propone la aplicación de un modelo matemático de rendimiento para su utilización en la sintonización automática y dinámica de un método de reducción de incertidumbre. Que la sintonización sea automática significa que el propio algoritmo o aplicación será dotado con la capacidad de realizar todas las etapas del proceso de sintonización, sin necesidad de que el usuario intervenga. Que sea dinámica significa que el proceso de mejora se llevará a cabo a la vez que el programa o aplicación esté ejecutándose, lo cual permitirá reflejar de modo inmediato las decisiones que se tomen en pos de la mejora del rendimiento, y así lograr ejecuciones más eficientes y adaptadas a las características de los datos particulares de entrada y al estado corriente del entorno de ejecución. [3,5]. Que el proceso de sintonización esté basado en modelos matemáticos de rendimiento significa que el conocimiento experto que se utilizará en la etapa de Análisis estará centrado en la utilización de un conjunto de expresiones que modelan el problema de rendimiento particular, lo cual, a la hora de tomar decisiones se reduce a la evaluación de tales expresiones matemáticas a partir de los datos recopilados durante la monitorización, permitiendo una toma de decisiones concisa en un tiempo de cómputo razonable, muchas veces despreciable en relación a la ejecución global del programa [3].

En resumen, tenemos como hipótesis que las características de dinamicidad y

automatismo antes mencionadas para el proceso de sintonización, contribuirán a lograr predicciones más precisas y veloces para el método predictivo ESS propuesto como caso de estudio.

3. OBJETIVOS

Como objetivos experimentales, se espera lograr la sintonización de ESS en sí misma, mejorando el desempeño del método de reducción de incertidumbre, tanto en calidad de resultados como en velocidad de respuesta. Como objetivo teórico, se espera definir un modelo de rendimiento que permita caracterizar el comportamiento del algoritmo de ESS para anticipar la detección de situaciones de sobrecarga computacional innecesaria y aplicar acciones paliativas que las contrarresten, automáticamente y en sintonía con la propia ejecución.

4. FORMACIÓN DE RECURSOS HUMANOS

La temática propuesta por este proyecto permite continuar con la formación y desarrollo de los distintos integrantes del grupo de trabajo, de forma complementaria a la formación adquirida hasta el momento. En 2020 se defendieron dos tesis doctorales relacionadas con esta línea de investigación, desarrolladas por la Dra. Laura Tardivo (bajo la dirección de la Dra. Paola Caymes Scutari) y el Dr. Miguel Méndez Garabetti (bajo la dirección del Dr. Germán Bianchini y la codirección de la Dra. Paola Caymes Scutari). En el caso del Dr. Ing. Méndez Garabetti, continúa su labor en el marco de la misma línea de investigación, la cual resulta transversal a su plan de beca postdoctoral. Asimismo, el grupo de

trabajo está evaluando la incorporación de un nuevo integrante, y siempre está abierto a la incorporación de nuevos integrantes (de grado o postgrado) que deseen familiarizarse con las temáticas que aquí se describen.

5. BIBLIOGRAFÍA

- [1] Piattini Velthuis, M., García Rubio, F., Garzías Parra, J., Genero Bocco, M. (2008) *Medición y Estimación del Software – Técnicas y métodos para mejorar la calidad y la productividad*. Alfaomega Ra-Ma.
- [2] Bianchini, G., Caymes-Scutari, P., & Méndez-Garabetti, M. (2015). Evolutionary-Statistical System: A parallel method for improving forest fire spread prediction. *Journal of Computational Science*, 6(1), 58–66. <https://doi.org/10.1016/j.jocs.2014.12.001>
- [3] Caymes-Scutari, P., Bianchini, G., Sikora, A., Margalef, T., (2016) “Environment for Automatic Development and Tuning of Parallel Applications” HPCS 2016, International Workshop on Parallel Optimization using/for Multi and Many-core High Performance Computing, Innsbruck, Austria. pp. 743-750. IEEE.
- [4] Wilkinson, B., Allen, M. (2005) *Parallel Programming: Techniques and Applications Using Networked Workstations and Parallel Computers*. Pearson Prentice Hall.
- [5] Naono, K., Teranishi, K., Cavazos, J. y Suda, R. (2010) *Software Automatic Tuning: From Concepts to State-of-the-Art Results*, Springer, New York.
- [6] Lampinen, J. y Zelinka, I. (2000), “On the Stagnation of the Differential Evolution algorithm”, *I.C. Soft Computing*, pp. 76-83.
- [7] Talbi, E. (2009) *Metaheuristics: From Design to Implementation*. John Wiley & Sons, Hoboken, New Jersey
- [8] Bianchini, G., Denham, M., Cortés, A., Margalef, T., Luque, E. (2010) “Wildland Fire Growth Prediction Method Based on Multiple Overlapping Solution”, *Journal of Computational Science Vol 1 Issue 4*, pp. 229-237. Elsevier.
- [9] Tardivo, M.L., Caymes-Scutari, P., Méndez-Garabetti, M. y Bianchini, G. (2017) “Optimization for an Uncertainty Reduction Method Applied to Forest Fires Spread Prediction”, *Computer Science – Cacic 2017*, pp. 13-23. Springer.
- [10] Mendez Garabetti, M. Bianchini, G., Caymes Scutari, P., Tardivo M.L. (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.

Técnicas de modelado y simulación para arquitecturas HPC y salud.

Diego Encinas^{1,2}, Jimena Jara¹, Román Bond¹, Daniel Rosatto¹, Lucas Maccallini, Mauro Gomez¹, Federico Montes de Oca¹, Adriana Gaudiani³, Martín Morales^{1,4}

¹Proyecto de Investigación SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - UNAJ

²Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

³Área de Computación - Instituto de Ciencias - UNGS

⁴Centro CodApli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, elchejime@gmail.com, rbond@unaj.edu.ar, drosatto@unaj.edu.ar, lucas.maccallini@gmail.com, mauro.f.gmz@gmail.com, federicomdo97@gmail.com, agaudiani@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 y tecnología en Cómputo de Altas Prestaciones

(High Performance Computing, HPC) para aplicaciones de interés social” - SimHPC de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 148/18. 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. Una manera de llevar a cabo este trabajo es utilizar técnicas de simulación para evaluar el efecto de los cambios de los factores con mayores influencias en las prestaciones del sistema de E/S paralelo.

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. Dicho espacio de memoria principal está controlado por otro componente que es capaz de procesar instrucciones guardadas en el espacio de memoria sintético denominado RamEntity. Si las instrucciones se guardan lógicamente, la nueva versión de Vm es capaz de administrar procesos en la nueva capa de ejecución. La idea general de la implementación consiste en la creación de una nueva SimEntity y sus agregaciones necesarias para manipular el espacio de memoria proporcionada por RamEntity, en tiempos de simulación. De tal forma que cada celda de RamEntity no sólo guarda datos, sino que también es posible que mantenga objetos interpretables como eventos para la nueva SimEntity que se denomina InitEntity.

Lo destacable es que el entorno de la cola de procesos es en el espacio de memoria de RamEntity y que cada instrucción que compone a un código objeto del proceso es en realidad un evento con todos sus parámetros.

Simulación y Salud

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 Symphony [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, la cantidad de insumos hospitalarios, etc. Al realizar distintas ejecuciones, 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.

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.
- Análisis y modelado de librerías de archivos para aplicaciones que utilizan cómputo de altas prestaciones-HPC. 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 soportes de memoria principal en tiempos de simulación que se activan como componentes de las máquinas 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].
- Contraste de una simulación de un cluster en la nube y uno idéntico desplegado en un sistema de cloud computing público [20].
- Modelado de infraestructuras de sistemas de salud.
- Implementación de un simulador para analizar el contagio de enfermedades intrahospitalarias.
- Paralelización en la ejecución de simuladores.

- Implementación de un simulador para analizar la propagación de enfermedades [21].

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.

Durante 2020 se han realizado publicaciones nacionales e internacionales. Además, se encuentran en desarrollo y concluidas varias Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Informática.

En esta línea de I/D existe cooperación a nivel nacional e internacional. Hay dos investigadores realizando estudios de postgrado, 1 becario EVC CIN, un becario de Iniciación a la Investigación UNAJ y 2 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Second Edition. 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. Tauber et al. (Eds.): ISC High Performance Workshops 2016, LNCS 9945, pp. 591–603, 2016. DOI: 10.1007/978-3-319-46079-6_40.
13. Kushang Parikh, Nagesh 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/simj ava 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 Symphony Frequently Asked Questions, disponible en internet: <https://repast.github.io/docs/RepastFAQ/RepastFAQ.html>. Fecha: 29/03/2020
18. 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.
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. 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
21. J. Baez, A. Barreto, B. Galarza, M. Morales, D. Encinas. Simulación para estimar propagación de enfermedades. 8° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba. 2020

Tecnología HPC como motor de ciencia de la UNdeC

Fernando Emmanuel FRATI, José TEXIER, Paula Cecilia RIVERA, Jonathan ALVAREZ, Fernanda CARMONA, Patricia FIGUEROLA, Francisco FRATI, Sebastián GUIDET, Roberto MILLON, Raul MORALEJO, Matías PEREZ, Emmanuel PORTUGAL, Donna RATTALINO, Alberto RIBA, Daniel ROBINS, Mara ROVERO, Javier RUITTI, Jorge TEJADA, Jusmeidy ZAMBRANO, Carlos Esteban GRAFFIGNA

Universidad Nacional de Chilecito

9 de julio 22, Chilecito, La Rioja, Argentina

{fefrati, jtexier, privera, jalvarez, fbcarmona, pfiguerola, flfrati, sguidet, rmillon, rmoralejo, mperez, eportugal, drattalino, ariba, drobins, mrovero, jruiitti, jtejada, jzambrano, cgraffigna}@undec.edu.ar

RESUMEN

En la actualidad, para abordar problemas de mayor tamaño y complejidad, los estudios de ciencia básica y aplicada utilizan Computación de Altas Prestaciones (HPC - High Performance Computing). El HPC permite mejorar la capacidad, velocidad y precisión en el procesamiento de datos. Con el proyecto que da origen a este trabajo se abordan seis estudios desde la perspectiva del HPC, para explorar los aspectos centrales del paralelismo aplicado desde las Ciencias de la Computación en otras disciplinas.

Algunos de estos estudios se realizan exclusivamente en la Universidad Nacional de Chilecito, mientras que otros son en cooperación con otras instituciones nacionales y extranjeras. Entre estos, tres formalizan trabajos finales de postgrado. En todos los casos, el HPC será abordado a través de un proceso metodológico organizado para:

- Consolidar una infraestructura de experimentación, desarrollo y producción de soluciones a problemas de HPC
- Desarrollar las capacidades científico-tecnológicas del equipo
- Fomentar la vinculación y transferencia con los sectores académico, social y productivo

Cada problema abordado reúne entre sus integrantes investigadores especialistas en la disciplina del estudio, investigadores de Ciencias de la Computación y estudiantes en sus últimos años de formación de grado. Con esto, se está consolidando un grupo de investigación, desarrollo y transferencia que generará

oportunidades de formación de recursos humanos, proveerá de servicios a la comunidad en el área de estudio y potenciará los vínculos de cooperación con otras instituciones.

Palabras clave: HPC, cómputo paralelo, aplicaciones, interdisciplinariedad.

CONTEXTO

La línea de investigación presentada es parte del proyecto “Software y aplicaciones en Computación de Altas Prestaciones” fue aprobado en la convocatoria a proyectos de Investigación y Desarrollo 2018 de la Secretaría de Ciencia y Tecnología de la UNdeC, y se encuentra en ejecución desde junio de 2019. Además, en 2018 la UNdeC destinó fondos de PROMINF para la adquisición de 12 PC con procesadores i7 y 8GB RAM, 5 de las cuales están equipadas con placas de video NVIDIA GTX 1060 para el “laboratorio de sistemas paralelos”. También, con el objeto de desarrollar las capacidades en HPC de la UNdeC, a fines de 2019 se adquirió un servidor Dell PowerEdge R740, equipado con 2 Xeon Platinum 8176 (56 núcleos físicos, 112 threads en total), 256 GB de RAM y 2 GPGPU NVIDIA Quadro P4000 (financiado a través del “Plan de mejoramiento de la función de I+D+i” - MINCyT), actualmente en funcionamiento. Finalmente, se destinaron fondos del PROMINF para financiar parcialmente la formación de postgrado de docentes afines a la disciplina, 20 de los cuales están realizando la *Especialización en inteligencia de datos orientada a big data* (acreditada por CONEAU) de la Facultad de

Informática de la UNLP. Estas iniciativas permitirán consolidar una infraestructura de experimentación, desarrollo y producción de soluciones a problemas de HPC, como así también recursos humanos formados para aprovecharla.

1. INTRODUCCIÓN

La informática tiene su origen en la necesidad de los distintos sectores de la sociedad de conseguir mayor velocidad, confiabilidad y precisión para resolver sus problemas. Sin embargo, la capacidad de solución a un problema dado encuentra su límite en los tiempos requeridos por sus algoritmos. Superar ese límite requiere que el problema sea abordado mediante cómputo paralelo. Normalmente, esto implica estudiar tres aspectos clave: hardware, aplicaciones y software.

Durante décadas, la industria respondió a la creciente demanda de mayor poder computacional incrementando exponencialmente el rendimiento de los procesadores[1]. Sin embargo, esta forma de obtener mayor poder de cómputo encontró barreras físicas, limitando el rendimiento de los microprocesadores y los sistemas en general [2]. Desde el año 2005, el escalado tecnológico se ha venido aprovechando para aumentar el número de cores dentro del chip, dando lugar a una importante variedad de arquitecturas (multicores, commodity clusters, GPGPU y Cloud) [3], [4].

No obstante, reducir los tiempos de procesamiento y obtener la mayor eficiencia de ese hardware requiere el diseño y desarrollo de algoritmos paralelos [5]. Transformar un algoritmo secuencial en uno paralelo no es trivial. En general los procesos concurrentes necesitan algún mecanismo para comunicar resultados parciales entre sí. Dependiendo de la arquitectura de cómputo, se consigue a través del uso de variables compartidas o del paso de mensajes entre procesos. Una transformación ‘implícita’ o transparente es deseable, pero el costo es una pérdida importante de rendimiento [6]. En su lugar, el programador recurre a librerías estándares para expresar explícitamente

el paralelismo: OpenMP, Pthreads, CUDA, OpenCL, MPI [7].

Las suposiciones de orden de ejecución entre instrucciones heredadas del modelo de programación secuencial ya no son válidas, obligando al programador a utilizar algún mecanismo de sincronización para garantizar estados consistentes del programa. En este contexto, la correctitud de los algoritmos es más difícil de garantizar que en la computación serial. Frecuentemente ocurren errores al sincronizar los procesos, dando lugar a deadlocks, condiciones de carrera, violaciones de orden, violaciones de atomicidad simple y violaciones de atomicidad multivariable, requiriendo el uso de herramientas de depuración específicas [8]. A diferencia de soluciones secuenciales donde existe un modelo teórico que permite estimar el desempeño de los programas antes de escribirlos, la evaluación del sistema paralelo (software y hardware) requiere la definición de distintas métricas: tiempo de ejecución, speedup, eficiencia y overhead. Las soluciones paralelas están tan estrechamente vinculadas con el hardware subyacente que dificultan enormemente conseguir portabilidad de rendimiento [9].

Existen muchos aspectos que requieren ser tomados en cuenta al diseñar la solución paralela: tamaño del problema, división de datos o tareas, balance de carga, requerimientos de memoria, precisión de los cálculos, comunicaciones y sincronización entre procesos, errores de concurrencia, detección y tolerancia a fallos entre los más relevantes. La complejidad de los problemas requiere habilidades especiales de los desarrolladores: dominio de múltiples paradigmas de programación y frecuentemente múltiples lenguajes, conocimientos de redes y comprensión de la concurrencia y sus consecuencias. Por todo esto, se considera de gran interés el estudio de estos temas para el desarrollo de capacidades científico-tecnológicas en la UNdeC que favorezcan el trabajo interdisciplinario en la institución.

2. LÍNEAS DE I+D

Se abordaron las siguientes líneas de I+D desde la perspectiva del HPC como eje central:

- Análisis de la diversidad molecular de microorganismos del suelo [10]–[13]. Estudio e implementación de algoritmos que contribuyan a reducir los tiempos de procesamiento y aumentar la capacidad de análisis referidos a este campo de la bioinformática, a fin de profundizar en el estudio de la diversidad molecular de microorganismos del suelo asociados a cultivos regionales.
- Evaluación de enfoques de desarrollo HDL y HLL en FPGA para aplicaciones de procesamiento de imágenes [14]–[16]. Estudio de lenguajes de desarrollo HDL y HLL en FPGA para implementar soluciones de procesamiento de imágenes eficientes.
- Identificación biométrica masiva mediante venas del dedo usando redes de aprendizaje extremo (ELM) [17], [18]. Estudio de técnicas de computación paralela para mejorar la eficiencia y aceleración del preprocesamiento, extracción de características biométricas, búsqueda e identificación de individuos, y el diseño de algoritmos de ELM mejorados que manejen eficientemente lotes de datos de gran tamaño.
- Servicios basados en lingüística computacional para análisis de texto [19]–[25]. Estudio sobre modelos computacionales que reproduzcan aspectos del lenguaje humano, con el fin de realizar análisis lingüísticos como servicios para el Centro de Escritura en la UNdeC.
- Documentos inteligentes a través del Blockchain [26]–[29]. Estudio de la tecnología blockchain para garantizar la integridad de documentos universitarios.
- Nodo de información meteorológica [30], [31]. Estudio, diseño e implementación de algoritmos para reducir los tiempos de procesamiento, aumentar la capacidad de análisis y favorecer la escalabilidad de aplicaciones de análisis y proyección de datos climáticos.

3. RESULTADOS

Se alcanzaron los siguientes resultados:

- Tres publicaciones en revista[32]–[34] y una en revisión.
- Diez presentaciones en congresos y/o workshops [35]–[44].
- Seis charlas científicas con invitados externos expertos en los distintos temas del proyecto.
- Cinco capacitaciones en otros centros de formación del país y del exterior en temas relacionados con el proyecto.
- Tres tesis de maestría en desarrollo, dos codirigidas por doctores especialistas en los temas de la UCM (Chile) y la tercera por un doctor de la UNLP (Argentina).
- Una tesina de grado en desarrollo (Asistente virtual académico de estudiantes de IS de la UNdeC).
- Un documento de recomendaciones para escribir tesinas de Ingeniería en Sistemas (en desarrollo).
- Un becario CIN.
- Puesta en funcionamiento del cluster del laboratorio de Sistemas Paralelos.
- Puesta en funcionamiento del servidor de altas prestaciones recientemente adquirido con fondos del Plan de Mejoramiento de la función de I+D+i del MINCyT (ver contexto).

4. FORMACIÓN DE RECURSOS HUMANOS

Seis miembros del equipo poseen formación de postgrado a nivel de doctorado, uno de ellos es especialista en Cómputo de Altas Prestaciones. Cinco miembros se encuentran en su etapa final para obtener el grado de maestría en Informática, tres de los cuales desarrollan como tesis temas abordados por esta propuesta. Dos de estas tesis de maestría están siendo codirigidas por docentes de la Universidad Católica de Maule (Chile) y una está siendo codirigida con un docente de la Universidad Nacional de La Plata. Nueve miembros están cursando la Especialización en inteligencia de datos orientada a Big Data (UNLP). Cada línea I+D propuesta integra al menos un docente

investigador experto en el campo de cada estudio específico. Todos los temas propuestos se trabajan con estudiantes de grado de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC (ambas acreditadas por CONEAU). Los docentes forman parte de los equipos de diversas asignaturas de estas carreras, entre las que se encuentran programación, arquitecturas de computadoras y arquitecturas paralelas. Nueve docentes se encuentran categorizados en el programa de incentivos.

5. BIBLIOGRAFÍA

- [1] S. Borkar y A. A. Chien, «The Future of Microprocessors», *Commun ACM*, vol. 54, n.º 5, pp. 67–77, may 2011, doi: 10.1145/1941487.1941507.
- [2] K. Ahmed y K. Schuegraf, «Transistor wars», *IEEE Spectr.*, vol. 48, n.º 11, pp. 50–66, nov. 2011, doi: 10.1109/MSPEC.2011.6056626.
- [3] V. V. Kindratenko *et al.*, «GPU clusters for high-performance computing», en *2009 IEEE International Conference on Cluster Computing and Workshops*, 2009, pp. 1–8.
- [4] J. Jeffers, J. Reinders, y A. Sodani, *Intel Xeon Phi Processor High Performance Programming: Knights Landing Edition*. Morgan Kaufmann, 2016.
- [5] A. Grama, A. Gupta, G. Karypis, y V. Kumar, *Introduction to Parallel Computing - Second Edition*. Pearson Education and Addison Wesley, 2003.
- [6] G. Hager y G. Wellein, *Introduction to high performance computing for scientists and engineers*. Boca Raton, FL: CRC Press, 2011.
- [7] J. Dongarra *et al.*, *Sourcebook of parallel computing*, vol. 3003. Morgan Kaufmann Publishers San Francisco, 2003.
- [8] F. E. Frati, «Software para arquitecturas basadas en procesadores de múltiples núcleos», Tesis, Facultad de Informática, 2015.
- [9] S. Ghosh, *Distributed Systems: An Algorithmic Approach*, 1.ª ed. University of Iowa, Iowa City, USA: Chapman and Hall/CRC, 2006.
- [10] M. S. De, M. Prager, R. E. Naranjo, y O. E. Sanclemente, «El suelo, su metabolismo, ciclaje de nutrientes y prácticas agroecológicas», *Agroecología*, vol. 7, n.º 1, pp. 19–34, 2012.
- [11] J. P. Hulsenbeck y F. Ronquist, «MrBayes: Bayesian inference of phylogeny», *Bioinformatics*, vol. 17, pp. 754–755, 2001.
- [12] A. Stamatakis, «RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies», *Bioinformatics*, vol. 30, n.º 9, pp. 1312–1313, may 2014, doi: 10.1093/bioinformatics/btu033.
- [13] M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, y A. Rambaut, «Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10», *Virus Evol.*, vol. 4, n.º 1, ene. 2018, doi: 10.1093/ve/vey016.
- [14] R. Nane *et al.*, «A Survey and Evaluation of FPGA High-Level Synthesis Tools», *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 35, n.º 10, pp. 1591–1604, oct. 2016, doi: 10.1109/TCAD.2015.2513673.
- [15] S. Windh *et al.*, «High-Level Language Tools for Reconfigurable Computing», *Proc. IEEE*, vol. 103, n.º 3, pp. 390–408, mar. 2015, doi: 10.1109/JPROC.2015.2399275.
- [16] R. Tessier, K. Pocek, y A. DeHon, «Reconfigurable Computing Architectures», *Proc. IEEE*, vol. 103, n.º 3, pp. 332–354, mar. 2015, doi: 10.1109/JPROC.2014.2386883.
- [17] D. Ezhilmaran y P. R. B. Joseph, «a Study Of Feature Extraction Techniques And Image Enhancement Algorithms For Finger Vein Recognition», p. 8, 2015.
- [18] A. Akusok, K. Björk, Y. Miche, y A. Lendasse, «High-Performance Extreme Learning Machines: A Complete Toolbox for Big Data Applications», *IEEE Access*, vol. 3, pp. 1011–1025, 2015, doi: 10.1109/ACCESS.2015.2450498.
- [19] AMPLN, «Asociación Mexicana para el Procesamiento del Lenguaje Natural Main/Home Page». <https://www.ampln.org/> (accedido ago. 30, 2018).
- [20] M. Vallez y R. Pedraza, «El Procesamiento del Lenguaje Natural en la Recuperación de Información Textual y áreas afines», *Hipertext Net*, 2007.
- [21] J. Texier, F. E. Frati, F. B. Carmona, A. E. Riba, M. Pérez, y J. Zambrano, «La gestión de la información en abierto, vehículo importante para maximizar la visibilidad web», presentado en XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina), may 2016.
- [22] P. Gamallo Otero, J. C. Pichel Campos, M. García González, J. M. Abuín Mosquera, y T. Fernández Pena, «Análisis morfosintáctico y clasificación de entidades nombradas en un entorno Big Data», 2014.
- [23] S. W. D. Chien, C. P. Sishla, S. Markidis, J. Zhang, I. B. Peng, y E. Laure, «An Evaluation of the TensorFlow Programming Model for Solving Traditional HPC Problems», en *International Conference on Exascale*

- Applications and Software*, 2018, p. 34.
- [24] H. Guan, X. Shen, y H. Krim, «Egeria: A Framework for Automatic Synthesis of HPC Advising Tools Through Multi-layered Natural Language Processing», en *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, New York, NY, USA, 2017, p. 10:1–10:14, doi: 10.1145/3126908.3126961.
- [25] P. Suber, *Ensuring open access for publicly funded research*. British Medical Journal Publishing Group, 2012.
- [26] M. U. Wasim, A. A. Ibrahim, P. Bouvry, y T. Limba, «Law as a service (LaaS): Enabling legal protection over a blockchain network», en *Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT), 2017 14th International Conference on*, 2017.
- [27] A. Preukschat, *Blockchain: la revolución industrial de internet*. Gestión 2000, 2017.
- [28] L.-Y. Yeh, P. J. Lu, y J.-W. Hu, «NCHC blockchain construction platform (NBCP): rapidly constructing blockchain nodes around Taiwan», en *Digital Libraries (JCDL), 2017 ACM/IEEE Joint Conference on*, 2017.
- [29] H. Dai *et al.*, «TrialChain: A Blockchain-Based Platform to Validate Data Integrity in Large, Biomedical Research Studies», *ArXiv Prepr. ArXiv180703662*, 2018.
- [30] A. Botta, W. De Donato, V. Persico, y A. Pescapé, «On the integration of cloud computing and internet of things», en *Future internet of things and cloud (FiCloud), 2014 international conference on*, 2014, pp. 23–30.
- [31] P. Yue, H. Zhou, J. Gong, y L. Hu, «Geoprocessing in cloud computing platforms—a comparative analysis», *Int. J. Digit. Earth*, vol. 6, n.º 4, pp. 404–425, 2013.
- [32] R. Hernández-García *et al.*, «Fast Finger Vein Recognition Based on Sparse Matching Algorithm under a Multicore Platform for Real-Time Individuals Identification», *Symmetry*, vol. 11, n.º 9, p. 1167, sep. 2019, doi: 10.3390/sym11091167.
- [33] S. Guidet, R. J. Barrientos, R. Hernández-García, y F. E. Frati, «Exhaustive similarity search on a many-core architecture for finger-vein massive identification», *J. Phys. Conf. Ser.*, vol. 1702, p. 012012, nov. 2020, doi: 10.1088/1742-6596/1702/1/012012.
- [34] R. Millón, E. Frati, y E. Rucci, «A Comparative Study between HLS and HDL on SoC for Image Processing Applications», *Elektron*, vol. 4, n.º 2, pp. 100-106, dic. 2020, doi: 10.37537/rev.elektron.4.2.117.2020.
- [35] S. Guidet y E. Frati, «Desarrollo de un método para identificación de personas por venas de dedo en grandes bases de datos.», presentado en VIII Jornadas Científicas de Estudiantes Investigadores, Chilecito, La Rioja, Argentina, oct. 2019.
- [36] A. Riperto y J. Texier, «Smart Contract en la generación de título de grado bajo infraestructura HPC», presentado en VII Jornadas Científicas de Estudiantes Investigadores, Chilecito, La Rioja, Argentina, oct. 2019.
- [37] A. Ortiz y I. Flores, «Sistema para el análisis de textos científicos a través de HPC», presentado en VIII Jornadas Científicas de Estudiantes Investigadores, Chilecito, La Rioja, Argentina, oct. 2019.
- [38] J. Zambrano y J. Texier, «Laboratorio de escritura: dispositivos de acompañamiento en escritura para estudiantes de la UNdeC.», presentado en 1er Congreso Internacional de Ingeniería Aplicada de Ibero-American Science & Technology Education Consortium, 2019.
- [39] R. Hernández-García, S. Guidet, R. J. Barrientos, y F. E. Frati, «Massive Finger-vein Identification based on Local Line Binary Pattern under Parallel and Distributed Systems», en *2019 38th International Conference of the Chilean Computer Science Society (SCCC)*, 2019, pp. 1–7.
- [40] R. Millon, E. Rucci, y E. Frati, «Análisis Comparativo de Implementaciones HLS de Filtro Sobel en SoC», en *Libro de actas del XXVI Congreso Argentino de Ciencias de la Computación - CACIC 2020*, Buenos Aires, Argentina, oct. 2020, pp. 639-648.
- [41] R. Millon, F. E. Frati, y E. Rucci, «Implementación de Filtro de Detección de Bordes Sobel en SoC usando Síntesis de Alto Nivel», presentado en Congreso Argentino de Sistemas Embebidos (CASE2020) (Encuentro virtual, 24 al 26 de agosto de 2020), 2020.
- [42] S. Guidet, R. J. Barrientos, F. E. Frati, y R. Hernández-García, «Finger-vein individuals identification on massive databases», presentado en VIII Conference on Cloud Computing, Big Data & Emerging Topics (Modalidad virtual), 2020.
- [43] F. E. Frati *et al.*, «Software y aplicaciones en computación de altas prestaciones para el contexto de la UNdeC», presentado en XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz), 2020.
- [44] F. E. Frati *et al.*, «Software y aplicaciones en computación de altas prestaciones para el contexto de la UNdeC», presentado en XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)., abr. 2019.

Tecnologías HPC Híbridas en la Solución de Problemas Complejos

Mercedes Barrionuevo ⁽¹⁾, Francisco Constantini⁽²⁾, Julian Escalante ⁽²⁾, Mariela Lopresti ⁽¹⁾,

Maximiliano Lucero ⁽¹⁾, Natalia Miranda ⁽¹⁾ y Fabiana Piccoli ^(1 y 2).

⁽¹⁾LIDIC- Univ. Nacional de San Luis, San Luis

⁽²⁾Universidad Autónoma de Entre Ríos, Concepción del Uruguay

⁽³⁾ Univ. Nacional de San Juan, San Juan

Argentina

{mbarrio, omlopres, mlucero, ncmiran, mpiccoli}@unsl.edu.ar

{franciscostantini,escalantejulian18}@gmail.com

RESUMEN

Los problemas a resolver hoy en día mediante soluciones computacionales demandan muchos recursos, esto puede obedecer a dos factores: el tamaño del problema generado por los datos que involucra, o su complejidad. Cualquiera sea el caso, siempre los resultados se esperan en un tiempo razonable. Una de las soluciones propuestas es pensar en sistemas heterogéneos: una computadora formada por procesadores many y multicores con software capaz de tomar ventaja de cada una de las componentes subyacentes.

En este trabajo se exponen dos líneas de trabajo orientadas a distintos tipos de problemas, en las cuales se propone desarrollar técnicas de Computación de Alto Desempeño para el escenario descrito.

Palabras clave: Computación de Alto Desempeño. Modelos Híbridos. Problemas Complejos. Big Data.

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro de los proyectos de investigación: “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” (LIDIC, UNSL) y “Cómputo de Altas Prestaciones

aplicado a la Solución de Grandes Problemas” (UADER).

1. INTRODUCCIÓN

Los datos, las soluciones complejas y los problemas a resolver computacionalmente han crecido exponencialmente en la última década. Por esta razón tanto el hardware como el software tuvieron que evolucionar ajustándose a los nuevos requerimientos.

Los sistemas de computación híbridos permiten combinar las características de las arquitecturas multi-core [Bas16] y many-core [Aam18, Hon10], los sistemas de memoria compartida [Vaj11] y de memoria distribuida [Kau14, Mar17], incrementan la capacidad y poder de cómputo del hardware del sistema de computación resultante. Esto posibilita aprovechar las ventajas ofrecidas: múltiples procesos y threads con distintas administraciones de memoria coexistiendo en el mismo sistema.

Si se considera el software, se han desarrollado diferentes paradigmas teniendo en cuenta las características de estos nuevos sistemas. Si bien los primeros desarrollos se enfocaron en los paradigmas típicos de paralelismo de datos y de tareas [Pac11]; o según la arquitectura: modelo de memoria compartida (OpenMP[Op19]), de memoria

distribuida o pasaje de mensajes (MPI[Gro14]) y para placa GPU (CUDA[NVi19]); hoy ya existen propuestas donde nos permiten combinar todas estas herramientas o independizarnos de los dispositivos utilizado como co-procesador alguna de las placas: GPU Intel/Radeon/NVidia, FPGA[Sk119], arduino, etc. a través de una única herramienta. Esto es posible utilizando tecnologías de alto nivel para desarrollar aplicaciones multiplataforma como son openCL (Open Computing Language)[Kae15] y oneAPI[One20]. Ambas tecnologías permiten desarrollar soluciones computacionales portables independientes del proveedor y del dispositivo, capaces de lograr tiempos de respuestas menores en muchas plataformas de hardware diferentes. Ambas herramientas permiten la programación de aplicaciones paralelas sobre unidades de computación heterogéneas, como GPUs, CPUs y otros dispositivos de aceleración. Su estructura consta de dos partes: Una API para acceder y usar las unidades de computación, y un lenguaje de programación que se ejecutará dentro del dispositivo o unidad de proceso. [Kae15, One20].

Teniendo en cuenta la demanda de cómputo de los problemas actuales y la posibilidad que nos brindan tanto las arquitecturas híbridas como los modelos de programación de alto nivel se puede lograr aplicar una estrategia de paralelización más efectiva mediante múltiples niveles de paralelismo y posibilitar así la reducción del overhead de comunicación. Existen numerosas soluciones a problemas complejos de gran escala, las cuales aplican técnicas HPC multi-tecnología y multiplataforma [Car18, Isu19].

Por todo lo expuesto, nuestra motivación es investigar, verificar y poner en marcha nuevas técnicas y arquitecturas híbridas para mejorar el procesamiento y los tiempos de respuesta de diferentes aplicaciones complejas. En base a ello, nuestro objetivo es aplicar HPC

multi-tecnología para resolver problemas complejos en ambientes híbridos. En la próxima sección presentamos los diferentes casos de estudio considerados

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Varios son los sistemas complejos considerados como benchmark para aplicar técnicas de HPC híbridas. Se destacan:

- Problemas de Simulación de Sistemas Reales: La toma de decisiones en sistemas compuestos por múltiples componentes interrelacionadas, no es una tarea fácil. Una posibilidad es abordarlos analíticamente o mediante ensayos o experimentos, lo cual puede implicar mucho riesgo. Una buena solución suele ser analizarlos a través de técnicas de simulación por computadora, las cuales representan una de las herramientas más poderosas para la resolución de problemas. Su potencial es incalculable, ya que permiten abstraer sistemas del mundo real y tomar decisiones basadas en múltiples experimentos. Los Autómatas Celulares (AC) son modelos matemáticos muy útiles y adecuados, los cuales a través de un conjunto de reglas simples pueden representar sistemas complejos pertenecientes a diferentes áreas científicas, como química, bioquímica, economía, física, etc.[Kau84, Wol84]. Además tienen como característica, su naturaleza paralela. En función de todas estas características, hemos enfocado el estudio de varios sistemas a través de AC paralelos y en distintas tecnologías. Ellos son:
 - El contagio de enfermedades como la Gripe A y Covid-19. Para ambos casos se desarrolló un AC donde se simuló un ambiente social dinámico y las reglas para cada una de las

- enfermedades[Hup13, Joh09, Wor14]. La propuesta fue desarrollada para arquitecturas many-core, tanto Nvidia, AMD e Intel, utilizando las herramientas de programación provistas por CUDA y OpenCL. Los desarrollos nos permiten estudiar, en menos tiempo, el proceso de propagación teniendo en cuenta diferentes realidades: el tipo de población, su distribución y otras características; a fin de tomar decisiones, como las campañas de vacunación, el aislamiento, cuarentena, mecanismo de contención, etc.. Los resultados fueron validados con los reportados por la OMS.
- Difusión de Noticias/Rumores: Este trabajo plantea, mediante un simulador con AC y técnica HPC, analizar el comportamiento de una sociedad frente a la dispersión de una noticia o rumor [Mai13, Yir12]. También se definió el AC, especificando los estados y reglas de transición. En la implementación se aplicaron técnicas de HPC para memoria distribuida. Esto nos habilita a realizar un análisis del comportamiento de una sociedad, frente a una noticia o rumor, siendo de utilidad tanto para fines buenos como programas de concientización o campañas de solidaridad, de política, etc., o no buenos como las fake news, campañas políticas, de desprestigio, etc.. De los resultados preliminares se pudo observar cómo estas últimas tienen mayor difusión.
- La versión actual está siendo ajustada y extendida a fin de incorporar, no sólo otras características relacionadas al problema cómo la difusión virtual, sino también la inclusión de otras tecnologías HPC.

- Otros problemas: En este caso, se están analizando distintos problemas a fin de simular su comportamiento, para ello estamos en la etapa de investigación de la dinámica del comportamiento de los siguientes fenómenos: Desarrollo de la roya en el trigo (Agricultura); Desplazamiento de los bancos arenas en el río Uruguay (Hidrología) y Crecimiento demográfico de una región para su buena planificación urbana (Sociedad). La solución presentada tendrá la característica de aplicar técnicas HPC híbridas utilizando las nuevas herramientas como oneAPI.
 - Problemas con Grandes Volúmenes de Datos: Muchas aplicaciones de Internet generan un tráfico de datos con características específicas susceptibles de clasificación utilizando Aprendizaje de Máquina. Las técnicas de aprendizaje automático guiados sin supervisión son un camino prometedor para hacer frente a los constantes cambios en el tráfico de red.
- El principal obstáculo al que nos enfrentamos al momento de avanzar en la clasificación del tráfico es un problema persistente de la investigación de Internet en general: la falta de una variedad de huellas que se puedan compartir para servir como datos de prueba, así como la ausencia de objetos de flujo utilizados como referencia para su validación. Estamos desarrollando un sistema paralelo de detección de ataques y anomalías circulantes en una red de computadoras, buscando llevar a cabo la tarea de identificación de ataques a nivel de red y transporte del modelo TCP/IP.
- La tarea de detectar posibles paquetes de datos anómalos en una red de computadoras es muy costosa, combinar

técnicas de clasificación y visualización de datos implementadas con técnicas HPC en su solución parece una buena alternativa. Para ello se busca identificar patrones que se desvían del comportamiento normal a partir del tráfico circulante en la red (gran volumen de datos) en un tiempo cercano a su ocurrencia, a fin de poder tomar decisiones rápidas [Bar16, Bar17, Bar18]. El trabajo está siendo desarrollado para GPU Nvidia.

Todas estas líneas de investigación tienen en cuenta la escalabilidad del problema y la portabilidad de los desarrollos, no sólo respecto a las arquitecturas sino también al tipo de problema.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como objetivos de las líneas de investigación nos planteamos facilitar el desarrollo de soluciones paralelas portables, de costo predecible, capaces de explotar las ventajas de modernos ambientes HPC a través de herramientas y “frameworks de computación” de alto nivel. Los resultados obtenidos hasta el momento son muy satisfactorios.

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados esperados respecto a la formación de recursos humanos son hasta el momento el desarrollo de 3 tesis de posgrado: 1 de doctor y 2 de maestría; y varias tesinas de grado en las distintas universidades intervinientes.

5. BIBLIOGRAFÍA

[Aam18] Aamodt, T., Lun Fung, W. and Rogers, T. - General-Purpose Graphics

Processor Architectures Synthesis Lectures on Computer Architecture. Editor: Morgan & Claypool Publishers. ISBN: 1627056181, 9781627056182. 2018.

[Bas16] Basu, S. “*Parallel and Distributed Computing Architectures and Algorithms*”. ISBN: 8120352122, 9788120352124. PHI Learning Pvt. Ltd., 2016.

[Bar16] Barrionuevo, M. Lopresti, M., Miranda, N. Piccoli, F.. “*Un enfoque para la detección de anomalías en el tráfico de red usando imágenes y técnicas de Computación de Alto Desempeño*”. XXII Congreso Argentino De Ciencias de la Computación. CACIC 2016. Pp. 1166-1175. Octubre 2016, San Luis, Argentina.

[Bar17] Barrionuevo, M. Lopresti, M. Miranda, N. Piccoli, F. “*Un Modelo de Detección de Anomalías en una LAN usando K-NN y Técnicas de Computación de Alto Desempeño*”. Communications in Computer and Information Science” (CCIS). ISBN: 978-3-319-75213-6. Editorial: Springer. Edición: Extranjera (inglés). Año: 2017. Volumen 790. Pp 219-228.

[Bar18] Barrionuevo, M., Lopresti, M., Miranda, N. and Piccoli, F. . “*Secure Computer Network: Strategies and Challenges in Big Data Era*”. Journal of Computer Science & Technologies. Editorial Red de Universidades Nacionales con Carreras de Informática (RedUNCI) Iberoamerican Science & Technology Education Consortium (ISTEC). Vol 3, pp 248-257. ISSN: 1666-6038 (electrónico), 1666-6046 (impreso). Diciembre 2018.

[Car18] CARLA 2018, *Selected Papers High of Performance Computing 5th Latin American Conference*. Communications in Computer and Information Science, Vol. 979. ISBN: 3030162052, 9783030162054. Colombia, 2018. Springer, 2019.

[Gro14] Gropp, W., Hoefler, T., Thakur, R. and Lusk, E. “*Using Advanced MPI: Modern Features of the Message-Passing Interface*”. Computer science & intelligent systems Scientific and Engineering Computation. ISBN: 0262527634 - 9780262527637. MIT Press, 2014.

- [Hon10] Hong, S., Kim, H., Hong, S. and Kim, H.. “*An integrated GPU power and performance model*” in Proceedings of the 37th annual international symposium on Computer architecture - ISCA '10, Vol. 38, N° 3. Pp. 280-289. 2010.
- [Hup13] Huppert, A. and Katriel, G.. “*Mathematical modelling and prediction*” in infectious disease epidemiology. Clinical Microbiology and Infection. Vol 19 N°11. Pp 999-1005, 2013.
- [Isu19] ISUM 2019, *Selected Papers of Supercomputing: 10th International Conference on Supercomputing. Communications in Computer and Information Science*, Vol. 1151. ISBN: 3030380432, 9783030380434. Springer Nature. México 2019.
- [Joh09] Johnson, T. and McQuarrie, B.. “*Mathematical modeling of diseases: Susceptible-infected-recovered (sir) model*”. In University of Minnesota, Morris, Math 4901 Senior Seminar, 2009.
- [Kae15] Kaeli, D. Mistry, P. Schaa, D. Zhang, D. P “*Heterogeneous Computing with OpenCL 2.0*” Third Edition, Morgan Kaufmann, ISBN: 978-0-12-801414-1, 2015
- [Kau14] Kaur, K. and Rai, A.. “*A Comparative Analysis: Grid, Cluster and Cloud Computing*” Int. J. Adv. Res. Comput. Commun. Eng., vol. 3, no. 3, pp. 2278–1021, 2014.
- [Kau84] Kaufman, S.. “*Emergent properties in random complex automata*”. Physica D: Nonlinear Phenomena. Vol 10 N°1 Pp 145-156, 1984.
- [Mai13] Maitanmi, O., Adekunle, Y. and Agbaje, M.. “*Model-Based Cellular Automata on Spread of Rumours*”. Computer Science, Babcock University. Nigeria. 2013.
- [Mar17] Marinescu, D.. “*Cloud Computing: Theory and Practice*”. 2nd Ed. ISBN: 0128128119, 9780128128114. Morgan Kaufmann, 2017.
- [NVi19] Nvidia. “*CUDA C++ Programming Guide, Design Guide*”. https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf. 2019.
- [One20] OneAPI Specification, release 1.0-rev 3. <https://spec.oneapi.com/versions/1.0-rev-3/>. 2020
- [Op19] OpenMP Architecture Review Board. “*OpenMP Application Programming Interface Specification Version 5.0*”. ISBN: 1795759887 - 978-1795759885. 2019.
- [Pac11] Pacheco, P. “*An Introduction to Parallel Programming*”, 1st ed., San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.
- [Sk119] Skliarova, I. and Sklyarov, V. - FPGA-BASED Hardware Accelerators. Lecture Notes in Electrical Engineering, Vol. 566 ISBN 3030207218, 9783030207212. Springer, 2019
- [Vaj11] Vajda, A. “*Multi-core and Many-core Processor Architectures. In: Programming Many-Core Chips*”. Springer, Boston, MA. 2011.
- [Wol84] Wolfram, S.. “*Universality and complexity in cellular automata*”. Physica D: Non-linear Phenomena, Vol 10 N° 1. Pp 1-35, 1984.
- [Wor14] World Health Organization. *Influenza (seasonal)*. Fact sheet N211. 2014.
- [Yir12] Yiran, G. and Jinzhu, D. “*Research on Rumors Spread Based on Cellular Automata*”, in Proceedings of the 2nd International Conference on Green Communications and Networks. Vol 2. ISBN 978-3-642-35418-2. Pp 235-244 Londres, Inglaterra. Springer Heidelberg. 2012.

Procesamiento de señales y Sistemas de Tiempo Real

Análisis de la revolución industrial 4.0 y Big Data. Proyecto de software: Oxímetro fotográfico.

Caffetti, Yanina Andrea.

Facultad de Exactas Químicas y Naturales. Universidad Nacional de Misiones. Posadas, Misiones.
yanina007@gmail.com

Resumen:

El entorno tecnológico actual, la competencia global y la posibilidad única de generar un valor agregado a los productos y servicios que se presentan en el mercado, implican necesariamente una transformación de los procesos productivos hacia una industria 4.0. Las TIC crean sistemas de manufactura, gestión, formas de hacer negocios y marketing que impacta en una mayor eficiencia y eficacia a la hora de responder a la demanda existente de la sociedad. Pero si queremos entender el fenómeno global que interactúa en la productividad del mercado potenciándola, debemos entender los alcances de Big Data en la vida cotidiana. El presente artículo es una revisión sistemática de la literatura que intenta contextualizar las transformaciones del mercado hacia una industrial 4.0 inmersas dentro del entorno tecnológico actual, partiendo de los conceptos básicos y proyectando las perspectivas futuras para una tesis de Doctorado en Informática.

Palabras claves: Industria 4.0, Big Data, Comercio, Marketing Digital, Mercadotecnia.

Contexto:

Big Data constituye un fenómeno global que puede llegar a tener un impacto económico real y potencial, que beneficie tanto al sector público como al privado en

el aumento de la productividad, la competitividad sectorial y la calidad de vida de la ciudadanía[1]. Si bien los datos siempre han sido parte del impacto de las TIC, los cambios que Big Data provoca impactan en el panorama económico, lo que genera nuevas oportunidades de negocios y mejoras en las tomas de decisiones a partir de la disponibilidad de datos en tiempo real que permiten cambios en la productividad[2]. Aún tenemos que comprender plenamente la velocidad y la amplitud que se encuentra presente en esta nueva revolución industrial. Consideremos las posibilidades ilimitadas de tener miles de millones de personas conectadas mediante dispositivos móviles, lo que da lugar a un poder de campos, como la inteligencia artificial (IA), la robótica, el internet de las cosas (IoT), los vehículos autónomos, la impresión 3D, la nanotecnología, la biotecnología, la ciencia de materiales, el almacenamiento de energía y la computación cuántica, por nombrar unos pocos. Muchas de estas innovaciones están en sus albores, pero ya están llegando a un punto de inflexión en su desarrollo a medida que se construyen y amplifican mutuamente en una fusión de tecnologías a través de los mundos físico, digital y biológico. [3]

En Alemania se debate sobre la «industria 4.0», un término acuñado en la Feria de Hannover de 2011 para describir cómo esta revolucionará la organización de las cadenas de valor globales. Mediante la

creación de «fábricas inteligentes», la cuarta revolución industrial genera un mundo en el que sistemas de fabricación virtuales y físicos cooperan entre sí de una manera flexible en todo el planeta. Esto permite la absoluta personalización de los productos y la creación de nuevos modelos de operación[3].

En definitiva, cuarta revolución industrial está creando un enfoque centrado en ecosistemas digitales, generando modelos de negocios innovadores basados en la interconexión de millones de consumidores, máquinas, productos y servicios, siendo Big Data una fase del paradigma intensivo en comunicación e información. Big Data emerge a partir de la revolución tecnológica iniciada en la década del setenta[1], es decir la tercera revolución industrial, y esto nos lleva a evaluar las perspectivas a futuro acerca del impacto de la disrupción y la inevitabilidad de los efectos que tendrá sobre nosotros su continuidad y transformación en la industria 4.0.

Introducción:

Las empresas, administraciones e individuos tienen cada día más datos disponibles y mejores herramientas para analizarlos, pero la velocidad, el volumen y la fuerza con la que se están dando los cambios que agitan la tecnología configuran además lo que muchos denominan “La Cuarta Revolución Industrial”. Esas bases de datos gigantes a las que ya hoy podemos acceder, y con las que podemos trabajar se denominan “Big Data”. Sacarle partido es uno de los retos de los próximos años, y parte de la clave para hacerlo está en la capacidad de estudiarlos de manera eficaz y en la capacidad de comunicación e interoperar

con la que podamos dotar a los objetos conectados a la Red[4].

Las tecnologías que impulsan la cuarta revolución industrial son variadas, los descubrimientos científicos y las creaciones que estos generan parecen ilimitados. La bibliografía al respecto puede contabilizar ciertas nuevas tecnologías que se encuentran en auge, sin embargo, estamos frente a una oportunidad única de innovación global.

Todos los nuevos desarrollos y tecnologías tienen una característica clave en común: aprovechan el poder de penetración que tienen la digitalización y las tecnologías de la información[3].

Líneas de Investigación y Desarrollo:

Con el fin de identificar las mega tendencias y poner de relieve el amplio panorama de los impulsores tecnológicos de la cuarta revolución industrial, Klaus Schwab por ejemplo, ha organizado la lista en tres grupos: físicos, digitales y biológicos. Los tres están profundamente interrelacionados y las diferentes tecnologías se benefician entre sí gracias a los descubrimientos y los avances que cada grupo va logrando[3].

Físicos:

Hay cuatro manifestaciones físicas principales de las mega tendencias tecnológicas, que son las más fáciles de detectar debido a su carácter tangible:

- Vehículos autónomos: a medida que progresen tecnologías como los sensores y la inteligencia artificial, las capacidades de todas estas máquinas autónomas se incrementarán a un ritmo rápido[3]. Google, en el año 2011, obtuvo la patente para desarrollar una tecnología que permita la conducción autónoma. Ahora el desafío es que los vehículos autónomos en cierta forma se vuelvan más “sensibles” y

puedan interpretar los datos del entorno en tiempo real haciendo la conducción más precisa y eficiente.

- **Impresión 3D:** la impresión en 3D comienza con material suelto y luego construye un objeto de forma tridimensional utilizando una plantilla digital. A diferencia de los bienes manufacturados producidos en serie, los productos impresos en 3D se pueden personalizar fácilmente. A medida que se superen las limitaciones actuales de tamaño, costos y velocidad, la impresión 3D se difundirá aún más e incluirá componentes electrónicos como circuitos impresos e incluso células y órganos humanos. Los investigadores ya están trabajando en 4D, un proceso que podría crear una nueva generación de productos que se modifican a sí mismos, capaces de responder a cambios ambientales como el calor y la humedad[3].

- **Robótica avanzada:** los avances en materia de sensores permiten a los robots comprender y responder mejor a su entorno y dedicarse a una variedad más amplia de tareas, como los trabajos del hogar. Al contrario de lo que sucedía en el pasado, cuando tenían que ser programados por medio de una unidad autónoma, los robots pueden ahora tener acceso a la información de manera remota mediante la nube y conectarse así con una red de otros robots[3]. Muchos de éstos avances corresponden a la amplia formación profesional a la que apuestan las sociedades visionarias, un caso de ejemplo es La Escuela de Robótica de la Provincia de Misiones, Argentina[5].

- **Nuevos materiales:** en general, son más ligeros, sólidos, reciclables y adaptables. En la actualidad existen aplicaciones para materiales inteligentes que se auto reparan o se limpian a sí mismos, metales con

memoria que vuelven a sus formas originales, cerámicas y cristales que convierten la presión en energía, y así sucesivamente[3]. Un ejemplo es el nanomaterial denominado grafeno. Las propiedades del grafeno son asombrosas. Un millón de veces más delgado que un cabello humano, pero más fuerte que el diamante y 200 veces más fuerte que el acero, el grafeno es el material más delgado y resistente del mundo; también es un conductor eficiente de calor y electricidad, es muy flexible y muy ligero. No es sorprendente que sea aclamada como una tecnología potencialmente disruptiva en muchas industrias, incluida la automotriz[6].

Digital

Una de las principales conexiones entre las aplicaciones físicas y digitales que ha sido habilitada por la cuarta revolución industrial es el internet de las cosas (IoT, por sus siglas en inglés), a veces llamado el «internet de todas las cosas». En su forma más simple, se puede describir como una relación entre las cosas (productos, servicios, lugares, etc.) y la gente, que resulta posible mediante tecnologías conectadas y plataformas varias[3].

Consideremos el monitoreo remoto, una aplicación muy conocida del IoT. Cualquier paquete, palé o contenedor ahora puede estar equipado con un sensor, un transmisor o una etiqueta de identificación por radiofrecuencia (RFID) que permite a una empresa rastrear sus movimientos a través de la cadena de suministro, cómo se comporta el objeto, cómo se está utilizando y así sucesivamente[3]. Enfoquemos ésta idea a la telemedicina. La telemedicina se refiere a todas las formas que posee el intercambio de información médica,

incluyendo una variedad de tecnologías de comunicación. Las aplicaciones referidas a la medicina y al cuidado de la salud incluyen telecomunicaciones, y datos que son usados para transferir información médica[7]. Hoy en día, el significado de telemedicina se relaciona con el uso tecnológico de los sistemas de información proveyendo información médica y servicios para miles de propósitos, como diagnósticos de enfermedades, transferencia de datos y registros médicos, monitoreo de rehabilitación de pacientes o procesos de terapias[8]. Todavía no comprendemos los alcances tecnológicos y desarrollos actuales que ha impulsado esta pandemia en la que vivimos. Una de las motivaciones del presente trabajo es poder realizar un sistema de monitoreo remoto para pacientes covid19 positivos en su etapa crítica, incluyendo la posibilidad de implementar IA para diagnósticos más precisos que puedan generar tratamientos particulares.

Biológicas

Las innovaciones en el campo biológico y la genética en particular son, cuando menos, impresionantes. En los últimos años se han logrado considerables progresos en reducir costos y aumentar la facilidad para la secuenciación genética, y más recientemente para activar o modificar genes. Muchos de nuestros insuperables retos en salud, desde las cardiopatías hasta el cáncer, tienen un componente genético. Debido a ello, la capacidad para determinar la constitución genética individual de una manera eficaz y rentable (por medio de la secuenciación en máquinas utilizadas en el diagnóstico rutinario) revolucionará la sanidad de forma personalizada y eficaz[3].

Estamos desarrollando nuevos métodos para integrar y utilizar dispositivos que monitoreen nuestros niveles de actividad y de química sanguínea, y cómo enlazar todo esto con el bienestar, la salud mental y la productividad en casa y en el trabajo. También estamos aprendiendo mucho más acerca de cómo funciona el cerebro humano y estamos viendo desarrollos increíbles en el campo de la neurotecnología. De ello da fe el hecho de que, en los últimos años, dos de los programas de investigación más financiados del mundo pertenecen al ámbito del cerebro[3].

En definitiva, Big Data toma un rol fundamental en los cambios que estamos viviendo. Lo crucial no es el gran aumento de los datos en sí mismo, sino su análisis para la toma de decisiones inteligentes lo que potenciará cada una de las tendencias de la industria 4.0.

Resultados obtenidos/esperados:

La escala y la amplitud de la creciente revolución tecnológica producirán cambios económicos, sociales y culturales de proporciones tan fenomenales que son casi imposibles de prever[3].

Sin lugar a dudas, la revisión de la literatura acerca de la temática presentada en este artículo nos determina las bases para describir y analizar el impacto potencial de la cuarta revolución industrial en la economía, los negocios, los gobiernos y países, la sociedad y los individuos visualizando el punto de entramado y protagonismo que adquiere Big Data. Nos deja la posibilidad de continuar investigando para prever las nuevas tendencias y generar soluciones frente a la disrupción que se presenta con la cuarta revolución industrial.

Los resultados obtenidos son la base bibliográfica para el desarrollo de un software de medición del ritmo respiratorio a través de la cámara fotográfica de un smartphone, procesando las imágenes en tiempo real para controlar la función respiratoria en pacientes covid19 positivos críticos emulando la funcionalidad de un oxímetro y utilizando inteligencia artificial para los registros médicos del paciente a fin de proyectar un diagnóstico a corto plazo de la evolución del sistema respiratorio.

Formación de Recursos Humanos:

Este proyecto de investigación forma parte del desarrollo de una tesis de posgrado, correspondiente a la carrera de Doctorado en Tecnologías de la Información dictada por la Universidad Nacional del Nordeste (UNNE), la Universidad Nacional de Misiones (UNaM) y la Universidad Tecnológica Nacional (UTN)[9].

Bibliografía:

- [1] F. Malvicino and G. Yoguel, "Big Data : Avances Recientes a Nivel Internacional y Perspectivas para el Desarrollo Local Autores : Facundo Malvicino a y Gabriel Yoguel b Coordinación : Gabriel Yoguel Centro Interdisciplinario de Estudios en Ciencia Tecnología e Innovación Ministerio d," 2015. Accessed: Feb. 10, 2021. [Online]. Available: www.ciecti.org.ar.
- [2] E. Brynjolfsson, L. Hitt, and H. Kim, "Strength in numbers: How does data-driven decision-making affect firm performance?," in *International Conference on Information Systems 2011, ICIS 2011*, Apr. 2011, vol. 1, pp. 541–558, doi: 10.2139/ssrn.1819486.
- [3] K. Schwab, *La cuarta revolución industrial*, Debate. 2016.
- [4] M. Tascón, *Big Data y el Internet de las cosas: Qué hay detrás y cómo nos va a cambiar.*, Los Libros. 2020.
- [5] Gobierno de la Provincia de Misiones Argentina., "Escuela de Robotica Misiones." <https://www.escueladeroticamisiones.com/node/35>.
- [6] D. Isaiah, "Grafeno de grado automotriz: el tiempo corre | Mundo automotriz," *Agosto 2015*. <https://www.automotiveworld.com/articles/automotive-grade-graphene-clock-ticking/> (accessed Feb. 13, 2021).
- [7] M. Berggren, "Wireless communication in telemedicine using Bluetooth and IEEE 802.11b," 2001, Accessed: Feb. 13, 2021. [Online]. Available: <http://www.it.uu.se/research/reports/2001-028/>.
- [8] D. Andrada, P. M. Sparhagl, H. M. Novillo, and J. Ierache, "Arquitectura para el Monitoreo Remoto de Funciones Vitales en Pacientes Ambulatorios," 2006. Accessed: Feb. 13, 2021. [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/22020>.
- [9] "Doctorado en Informática: inscripciones desde el 1/11 – FCEQyN." <https://www.fceqyn.unam.edu.ar/en-marzo-2021-comienza-el-doctorado-en-informatica/> (accessed Feb. 13, 2021).

Modelos Matemáticos y Métodos Computacionales en Ingeniería

Javier Giacomantone - Oscar Bria

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – UNLP

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
La Plata, Buenos Aires

{jog, onb}@lidi.info.unlp.edu.ar

Resumen

Este trabajo describe brevemente el contexto y los objetivos generales de un tema principal y sus líneas de investigación particulares. Su objetivo principal es estudiar y evaluar modelos matemáticos y métodos numéricos que permitan abordar problemas específicos en ingeniería. También se estudian sistemas y procesos, que por su nivel de complejidad, requieren abordajes y trabajos multidisciplinarios para operar en el dominio del problema de interés. Los modelos abordados son dependientes del tipo de sistema estudiado, del fenómeno analizado y del área particular de ingeniería que originó el requerimiento. Determinar el tipo de sistema, el método para evaluar su rendimiento y las soluciones numéricas óptimas o sub-óptimas forma parte de los objetivos generales.

Palabras Clave: modelos matemáticos aplicados, métodos computacionales, modelos probabilísticos, desempeño de sistemas, confiabilidad de sistemas.

Contexto

Las líneas de investigación y desarrollo (I/D) forman parte del proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real”. En particular del sub-proyecto “Mode-

los y métodos computacionales. Procesamiento de señales y reconocimiento de patrones”.

1. Introducción

En ingeniería y en ciencias básicas se utilizan distinto tipo de modelo para estudiar y caracterizar objetos y fenómenos de interés. Los modelos matemáticos nos permiten predecir fenómenos naturales y el comportamiento de estructuras, dispositivos, procesos y sistemas diseñados por el hombre. Los métodos computacionales asociados a modelos determinísticos y probabilísticos permiten estudiar el dominio del problema considerando distintas condiciones estructurales y funcionales. El análisis del comportamiento a medida que el objeto o fenómeno real se aleja de las hipótesis fundamentales del modelo permite determinar su alcance, su validez y su utilidad respecto a los objetivos iniciales [1][2]. Este abordaje permite implementaciones en *hardware* o *software* que ofrezcan soluciones viables. Los sistemas estudiados pueden ser lineales o no lineales requiriendo modelos sofisticados [3][4]. Determinar si una solución es viable con fundamento científico, es una tarea compleja y dependiente del problema particular analizado [5]. Involucra estudiar el dominio del problema, los objetos o los fenómenos modelados, y por lo tanto eventualmente requiere soporte multidisciplinar. Este proyecto tiene como primer objetivo analizar y proponer modelos computacionales, métodos y las soluciones

particulares derivadas de los mismos. Alcanzar los objetivos anteriores requiere estudiar los fundamentos que subyacen a cada modelo, evitando soluciones, que por su nivel de encapsulamiento, limiten una verdadera comprensión y abordaje científico de los mismos [6]. El tipo de problema de interés en esta línea de investigación requiere la integración de soluciones de tres áreas, ciencias de la computación, matemáticas aplicadas y un área de ingeniería o ciencia básica en particular. La evaluación de rendimiento es un aspecto fundamental para poder validar las soluciones propuestas o los modelos analizados [7]. Por lo tanto, otro aspecto fundamental es el estudio de las métricas y paradigmas de desempeño en sistemas específicos.

En la sección 2 se presenta un breve resumen de los temas de I/D específicos en el período actual. La sección 3 enumera resultados obtenidos y esperados. Finalmente, la sección 4 resume los objetivos con respecto a la formación de recursos humanos.

2. Líneas de Investigación

2.1 Modelos Probabilísticos

El primer objetivo de esta línea de investigación es estudiar si un determinado modelo es viable en el contexto de un sistema particular. El segundo objetivo general es determinar si es robusto cuando compromisos de diseño implican apartarse de las hipótesis iniciales del modelo.

2.1.1 Imágenes de Tiempo de Vuelo (TOF)

Las cámaras de TOF permiten obtener imágenes de rango, también denominadas $2 \frac{1}{2} D$. El ruido y los artefactos en este tipo de imagen requieren de modelos probabilísticos adecuados para caracterizar, filtrar y eventualmente reducir efectos indeseados. Segmentar imágenes de tiempo de vuelo requiere modelos de segmentación con características específicas [8][9][10]. En particular métodos estadísticamente robustos [11][12].

2.1.2 Datos Fuertemente Desbalanceados

Cuando el conjunto de patrones de entrenamiento de un clasificador probabilístico es marcadamente asimétrico, esta línea de investigación estudia el dominio específico del problema y la capacidad de generalización del sistema de clasificación. Se analizan paradigmas de aprendizaje automático estadístico supervisado y semi-supervisado [13][14][15].

2.1.3 Generación de Descriptores

El objetivo de esta línea de trabajo es mejorar la calidad de los descriptores obtenidos a partir de modelos probabilísticos y modelos espectrales, considerando: unicidad, invariancia, sensibilidad y su impacto en el sistema de clasificación.

2.2 Desempeño de Sistemas de Posicionamiento, Navegación y Vigilancia.

En los sistemas de posicionamiento, de navegación y de localización [16][17], el concepto de desempeño excede al habitual que está limitado a la calidad nominal de la estimación de ubicación y eventualmente a la confiabilidad [18][19]. En estos sistemas deben considerarse además los parámetros de integridad y continuidad que le garanticen al usuario que la información proporcionada por el sistema es correcta para que una operación crítica pueda realizarse en forma segura [20][21].

Un tema relacionado con el desempeño de los sistemas vigilancia es el volumen de transacciones con características aleatorias [22].

Esta línea de trabajo se avoca al estudio de problemas puntuales de desempeño en los sistemas mencionados, utilizando criterios y métodos diversos de modelado y procesamiento de señales [23][24].

3. Resultados y Objetivos

3.1 Resultados publicados

- Se estudiaron y propusieron métodos para detección en series temporales de fMRI [25][26].
- Se desarrollaron métodos de segmentación de imágenes de rango y supresión del plano de fondo [27][28][29][30].
- Se analizaron y propusieron alternativas para el agrupamiento de objetos de interés en video [31].
- Se estudió el desempeño de un método de exclusión de satélites en un sistema de ayuda a la aeronavegación basado en GNSS [32].
- Se presentaron resultados experimentales de un método de aprendizaje en aritmética computacional [33].
- Se propuso un método de segmentación espectral para imágenes TOF [34].
- Se estudió el comportamiento bajo carga de un algoritmo para programar transacciones de radares aeroportuarios [35].

3.2 Objetivos generales

- Desarrollar modelos y optimizar algoritmos particulares de clasificación supervisada y no supervisada.
- Evaluar métodos de análisis de desempeño y su aplicación sobre los clasificadores y conjuntos de datos particulares.
- Evaluar la monitorización de la integridad de los sistemas de ayuda a la navegación aérea basados en sistemas GNSS.
- Estudiar métodos de selección y extracción de características.
- Promover la interacción con otros grupos y líneas de I/D resultando en un mecanismo de permanente consulta y transferencia.

4. Formación de Recursos Humanos

La formación de recursos humanos en primer lugar implica la transferencia de los resultados

de cada línea de investigación a las asignaturas de grado y cursos de postgrado que los integrantes dictan. Los alumnos también tienen la posibilidad de realizar trabajos supervisados de investigación, resultantes en tesinas y tesis en el área. Debido al carácter trans-disciplinar de las líneas de I/D expuestas, se espera orientar y brindar apoyo a investigadores y alumnos de otras líneas de investigación relacionadas.

BIBLIOGRAFIA

1. Zalizniak V. Essentials of Scientific Computing – Numerical Methods for Science and Engineering. Woodhead Publishing, 2008.
2. Juergen G. Coupled Systems: Theory, Models, and Applications in Engineering. CRC, 2014.
3. Torokhti A., Howlett P. Computational Methods for Modelling of Nonlinear Systems. Elsevier, 2007.
4. Canuto C., et al. Spectral Methods. Evolution of Complex Geometries and Applications to Fluids Dynamics Scientific Computation. Springer, 2007.
5. Ciurpina G. Scientific Computing in Electrical Engineering Springer, 2007.
6. Gustafsson B. Fundamentals of Scientific Computing Springer, 2011.
7. Aslak T., et al. Elements of Scientific Computing. Springer, 2010.
8. Von Luxburg U. A Tutorial on Spectral Clustering. Statistics and Computing, v.17(4), 2007.
9. Kim H.Y., Giacomantone J. O., Cho, Z. H. Robust Anisotropic Diffusion to Produce Enhanced Statistical Parametric Map, Computer Vision and Image Understanding, v.99, pp.435-452, 2005.
10. Han Y., Feng X., Baci G. Variational and PCA based natural image segmentation. Pattern Recognition 46, pp. 1971-1984, 2013.
11. Fukunaga K. Introduction to Statistical Pattern Recognition. Second Edition. Academic Press, 1990.
12. Devijer P, Kittler, J. Pattern Recognition: theory and applications. Springer, 1986.
13. Cortes C, Vapnik V, Support vector networks. Machine Learning v.20, pp.273-297, 1995.

14. Vapnik, V. *The Nature of Statistical Learning Theory*. N. Y. Springer, 1995.
15. Aytug H. Feature selection for support vector machines using Generalized Benders Decomposition. *European Journal of Operational Research*, v. 244(1), pp. 210-218, 2015.
16. Partap Misra, Per Enge. *Global Positioning System: Signals, Measurements and Performance*, Ganga-Jamuna Press, 2010.
17. Hakan Koyuncu, Shuang Hua Yang. *A Survey of Indoor Positioning and Object Locating Systems Indoor Positioning System*, International Journal of Computer Science, 2010.
18. Petevelo Mark. *Quantifying the performance of Navigation Systems and Standards for assisted-GNSS, Inside GNSS*, 2008.
19. Morurikis A., Roumeliotis S. *Performance Analysis of Multirobot Cooperative Localization*, IEEE, 2005.
20. Murphy T., et. al., *Fault Modeling for GBAS Airworthiness Assessments*, Navigation, 2012.
21. Cosmen-Schortmann J., Azaola-Sáenz, Martínez-Olagüe M. A., Toledo-López M., *Integrity in Urban and Road Environments and its use in Liability Critical Applications*, IEEE, 2008.
22. Pengfei Duan, Maarten Uijy De Haa. *Flight Test Results of a Measurement-Based ADS-B System for Separation Assurance*, Navigation, 2013.
23. Shuo-Ju Yeh1, Shau-Shiun, *GBAS airport availability simulation tool*, GPS Solutions, 2015.
24. Sam Pullen, Todd Walter, Per Enge. *SBAS and GBAS Integrity for Non-Aviation Users: Moving Away from "Specific Risk," International Technical Meeting of The Institute of Navigation*, 2011.
25. Giacomantone J., Tarutina T. *Diffuse Outlier Detection Technique for Functional Magnetic Resonance Imaging*. Computer Science and Technology Series. XVI Argentine Congress of Computer Science Selected Papers. pp. 255-265, 2011.
26. Giacomantone J., De Giusti A. *Detección de áreas de interés bajo la hipótesis de relación espacial de voxels activados en fMRI*. XX Congreso Argentino de Ciencias de la Computación. San Justo. Argentina, 2014.
27. Lorenti L., Giacomantone J. *Segmentación espectral de imágenes utilizando cámaras de tiempo de vuelo*. XI Workshop Computación Gráfica, Imágenes y Visualización. pp. 430-439. Mar del Plata, Argentina, 2013.
28. Lorenti L., Giacomantone J. *Time of flight image segmentation through co-regularized spectral clustering*. Computer Science & Technology Series. XX Argentine Congress of Computer Science. Selected papers. La Plata, Editorial de la Universidad Nacional de La Plata, 2015.
29. Giacomantone J., et al. *Supresión del plano de fondo en imágenes de tiempo de vuelo*. VII Workshop Procesamiento de Señales y Sistemas de Tiempo Real, 2016.
30. Lorenti, L., Giacomantone, J., Bria, O. N., De Giusti, A. E. (2017). *Fusión de información de geometría e intensidad para segmentación de imágenes TOF*. XXIII CACIC. La Plata, Argentina, 2017.
31. Lorenti L., Giacomantone J., De Giusti A. *Agrupamiento de trayectorias vía clustering espectral incremental*. XXII CACIC, pp. 222-231, 2016.
32. Bria, O., Giacomantone, J., Lorenti, L., *Excluding Ionospherically Unsafe Satellite Geometries in GBAS CAT-I*. XXII CACIC. CCIS 790: Revised CACIC Selected Papers, 790, pp. 243-252, 2018.
33. Giacomantone, J., Bria, O., *Proactive Independent Learning Approach: A case study in computer arithmetic*. XXII CACIC. La Plata, Argentina, 2017.
34. Lorenti L., Giacomantone J., Bria O., *Unsupervised TOF Image Segmentation through Spectral Clustering and Region Merging*. *Journal of Computer Science & Technology*, v.18(2), pp. 97-104, 2018.
35. Bria O., Giacomantone J., Villagarcía H., *Compound Interleaving Scheduling Algorithm for SLM Transactions in Mode S Surveillance Radar*. *Communications in Computer and Information Science* 995, pp. 297-312. Springer, 2019.

Redes de Sensores, Robots Móviles y Simulación en Sistemas de Tiempo Real

Fernando Romero¹, Diego Encinas¹, Armando De Giusti^{1,2}, Santiago Medina¹, Martín Pi Puig¹, Horacio Villagarcía^{1,3}, Juan Manuel Paniego¹, Fernando G. Tinetti^{1,3}

¹Instituto de Investigación en Informática LIDI (III-LIDI)³

Facultad de Informática – Universidad Nacional de La Plata – Centro Asociado CIC

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{fromero, dencinas, degiusti, smedina, mpipuig, hvw, jmpaniego, fernando}@lidi.info.unlp.edu.ar

Resumen

Esta línea de investigación trata sobre Sistemas de Tiempo Real. Dentro de esta temática se desarrollan tareas en el contexto de tres temas: 1) Redes de sensores inalámbricas 2) Modelado y Simulación 3) Robótica móvil.

Dentro de las redes de sensores inalámbricas y robótica móvil se estudia la interacción que provee dos aspectos: los sensores que están sobre el móvil que proveen información, y la posibilidad de a través de información intercambiada entre la red y la generada por el propio robot provee la ubicación del mismo. Tanto las redes como los robots incluyen productos elaborados en el laboratorio como comprados y terminados. Respecto de las simulaciones se remiten tanto a sistemas de hardware y cloud computing como a situaciones de evacuaciones en casos de emergencia y transmisión de enfermedades.

Contexto

Esta línea de Investigación forma parte del proyecto 11/F024 – Computación de Alto Desempeño: Arquitecturas,

Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real SubProyecto CAD-3. Procesamiento para problemas de Tiempo Real / Robótica del Instituto de Investigación en Informática LIDI acreditado por la UNLP.

Palabras Claves: Tiempo Real, Simulación, Sistemas Embebidos, Comunicaciones, Redes de Sensores, Robots, drones, Microcontroladores, Cloud Computing.

1. Introducción

Son Sistemas de Tiempo Real (STR) aquellos que requieren la existencia de plazos de tiempo para llevar a cabo sus acciones [5] [6] [11] [12] [16] [17]. Una característica es que deben interactuar con el mundo físico. Ello determina los plazos en que ante una entrada o cambio en el sistema físico debe elaborar la respuesta y sus acciones. Por ello es necesario que el procesamiento debe estar sincronizado con un sistema de tiempo que esté sincronizado y en una escala estándar, o sea deben tener reloj de tiempo real. Las entradas desde el mundo físico son a través de sensores y se utilizan actuadores para

las respuestas. Estos sensores pueden adquirir configuraciones complejas en red cuando la cantidad de variables a controlar sea grande y sobre todo estén situadas en forma remota. Muchas veces se utilizan robots móviles, terrestres y aéreos [4] [9] [10] [15] [22] [23] [18], en combinación con los sensores y actuadores. Estos sensores sobre los móviles permiten doble funcionalidad: proveer datos del ambiente que recorren y en una interacción con el ambiente y las redes de sensores fijos en el mismo, proveer información sobre la ubicación del móvil. Para ello se recurre a sistemas GPS en exteriores y técnicas de posicionamiento en interiores (indoor) [3] [13] utilizando para ello diferentes tipos de sensores de tecnología de ultrasonido, infrarrojo como también los sistemas de comunicaciones wifi y bluetooth. En el desarrollo de robots propios y redes de sensores se utilizan placas de desarrollo basadas en microcontroladores (como, Arduino, NodeMCU, CIAA [19] [27]) y Computadoras de Placa Simple (como, Raspberry Pi), utilizando diferentes SOTR (Linux RT-Preempt, Free RTOS, MQX, OSEK-OS, etc.) [8]. Se realizan pruebas de alcance, integridad y funcionalidad de redes de sensores inalámbricas [28] [29] [30] [21] principalmente utilizando módulos WiFi y LoRa [26]. Por otra parte, se trabaja en el campo del modelado y simulación [7] [14] [24] [25] con el fin de obtener simuladores que permitan predecir el comportamiento y la eficiencia de distintos sistemas ante diferentes escenarios. Además, se ajustan estas simulaciones con datos reales, lo cual permite luego realizar ensayos sobre la simulación: ejemplos de ellos son las simulaciones de robots, placas, incendios, evacuación de edificios en catástrofes y procesos industriales.

2. Resultados y Objetivos

Se han desarrollado tareas sobre los temas antes expuestos tales como:

- Desarrollo de un robot con encoders para odometría, combinados con sensores de ultrasonido e infrarrojo.
- Medición de consumo energético de diferentes microcontroladores con cámara de video, variando su ancho de banda y velocidad de transmisión.
- Construcción y estudio de redes de sensores inalámbricas basadas en WiFi y LoRa [1].
- Desarrollo de interfaces para el control de drones.
- Modelado y simulación de arquitecturas de Cloud Computing para comparar con arquitecturas de HPC [21] [2].
- Modelado y simulación de evacuaciones en caso de catástrofe en diversos edificios y transmisión de enfermedades intrahospitalarias.

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. Además, se encuentran en desarrollo y concluidas tesinas de grado de alumnos de Licenciaturas de Informática y Sistemas, como así también Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Computación y Analista en TICs. De postgrado, investigadores del grupo están desarrollando un trabajo final de especialización, tres tesis de Maestría y una tesis de Doctorado.

5. Referencias

[1] Integration of Sensor Networks with Cloud Computing, S. 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] 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.
- [3] Azizi, F., N. Houshangi. "Mobile robot position determination using data from gyro and odometry." In Electrical and Computer Engineering, 2004. Canadian Conference on, vol. 2, pp. 719-722. IEEE, 2004.
- [4] Bekey, George A. Robotics: state of the art and future challenges. Imperial College Press, 2008.
- [5] 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.
- [6] Buttazzo, G. C., "Hard RealTime Computing Systems", Third edition, Springer, 2011.
- [7] Eickhoff, J., Simulating Spacecraft Systems, Springer, 2009.
- [8] "FreeRTOS - market leading RTOS (real time operating system) for embedded systems supporting 34 microcontroller architectures". [http:// www.freertos.org/](http://www.freertos.org/).
- [9] Jenkins, T., I. Bogost. "Designing for the internet of things: prototyping material interactions." In CHI'14 Extended Abstracts on Human Factors in Computing Systems, pp. 731-740. ACM, 2014.
- [10] Jones, J. L., A. M. Flynn, Bruce A. Seiger. Mobile robots: inspiration to implementation. Vol. 2. Wellesley MA: AK peters, 1999.
- [11] Kopetz. H., "Real-Time Systems, Design Principles for Distributed Embedded Applications". Second Edition. Springer. 2011.
- [12] Liu, J. W. S. Liu, "Real Time Systems", Integre Technical Publishing Co., Inc., 2000
- [13] Rekleitis, I. M., G. Dudek, E. E. Milios. "Multi-robot exploration of an unknown environment, efficiently reducing the odometry error". In International Joint Conference on Artificial Intelligence, vol. 15, pp. 1340-1345. Lawrence Erlbaum Associates Ltd, 1997.
- [14] 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.
- [15] Unmanned Vehicles: Real Time Problems in Drone Receivers F. G. Tinetti, O. C. Valderrama Riveros, and F. L. Romero, Proceedings of the 2019 International Conference on Computational Science and Computational Intelligence (CSCI), págs. 1080-1085, doi. 10.1109/CSCI49370.2019.00205, 2019.
- [16] Silberschatz, A., P. B. Galvin, G. Gagne, Operating System Concepts, 8th Edition, ISBN: 978-0-470-12872-5, Wiley, 2009.
- [17] PHILLIP A. LAPLANTE, SEPO J. OVASKA. REAL-TIME SYSTEMS DESIGN AND ANALYSIS Tools for the Practitioner Fourth Edition. A JOHN WILEY & SONS, INC., PUBLICATION. IEEE PRESS. 2012.
- [18] <http://ftp.k-team.com/khepera/documentation/Kh2IRAN.pdf>
- [19] <http://www.proyecto-ciaa.com.ar/devwiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp>
- [21] 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

[22] 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.

[23] 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.

[24] D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.

[25] Proteus. <https://www.labcenter.com>. 2017

[26] LoRa <https://www.lora-alliance.org/> 2017

[27] NodeMcu <http://www.nodemcu.com/> 2017

[28] Akyildiz, Ian F., and Mehmet Can Vuran. "Wireless sensor networks" Vol. 4. John Wiley & Sons, 2010.

[29] Lewis, Franck L. "Wireless sensor networks." Smart environments: technologies, protocols, and applications 11 (2004): 46.

[30] Raghavendra, Cauligi S., Krishna M. Sivalingam, and Taieb Znati, eds. "Wireless sensor networks" Springer, 2006.

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

Martín Pi Puig: 0000-0002-7202-7638

Horacio Villagarcía:

Juan Manuel Paniego: 0000-0001-6721-9822

SEGUIMIENTO DEL ESTADO DE UNA CÁMARA DE GERMINACIÓN HIDROPÓNICA UTILIZANDO EMAIL BASADO EN NODE-RED.

Urquijo, Rubén R., Marinelli, Marcelo J.

Departamento de Informática, Facultad de Ciencias Exactas, Químicas y Naturales.

Universidad Nacional de Misiones.

Félix de Azara 1552, Posadas, Misiones

chinourquijo@gmail.com; marcelomarinelli@gmail.com

RESUMEN

El objetivo de este trabajo es implementar sobre un sistema de cultivo hidropónico, funcionalidades de IoT (*Internet of things*, Internet de las cosas). Con esto, se pretende sacar provecho del auge que está teniendo dicha tecnología en distintos ámbitos, y aplicarlo a la cámara de germinación hidropónica desarrollada en [1].

La idea principal es poder interactuar, mediante el uso de un software especial para IoT, con el controlador que obtiene información de los sensores de temperatura, humedad y nivel de nutrientes de la cámara, utilizando la comunicación con la cámara vía Twitter y E-mail [2] para la obtención de los valores al instante de la medición de estas variables. De la misma manera, el controlador enviará alertas a las cuentas de Twitter y E-mail configuradas cuando los valores de temperatura, humedad y nivel de nutrientes superen o estén por debajo de los umbrales establecidos en [1].

Palabras clave: *Hidroponía, Internet de las Cosas, Web Services, E-mail, Twitter, Raspberry Pi.*

CONTEXTO

Este trabajo se encuentra enmarcado en el “Programa de Investigación en Computación” del Instituto de Investigación, Desarrollo e Innovación en Informática de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones; también posee vínculos con el Doctorado en Ciencias Aplicadas y la Maestría en Tecnologías de la Información de la misma casa de estudios.

Puntualmente, este trabajo es un avance de la Tesis de Maestría denominada “Internet de las

Cosas (IoT) aplicada a un sistema de monitoreo de una cámara de germinación hidropónica.”

En el presente trabajo, lo que se quiere realizar es la elaboración de un prototipo para poder comunicarse con la plataforma IoT. Para ello se necesitará una cuenta de E-mail y una cuenta de Twitter previamente creadas para los fines del trabajo.

Las personas que interactúen con la plataforma, podrán recibir información mediante un mensaje personalizado que llegue a sus cuentas, ya sea de mail o twitter con el objetivo de conocer los valores que posean las variables de Temperatura, Humedad y nivel de Nutrientes de la cámara hidropónica realizada en [1]. Básicamente, la idea es agregar a lo realizado en [3] la posibilidad que las personas puedan interactuar con la plataforma IoT mediante el envío de un mail o un tweet; los cuales serán creados para tal fin. La plataforma, de acuerdo a unas palabras claves definidas previamente, deberá armar la respuesta para enviársela nuevamente a la persona que haya realizado la consulta. Si la consulta llega a la plataforma vía mail, la respuesta vuelve por la misma vía. Si la misma es recibida a través de Twitter, la respuesta llegará como respuesta al tweet, “*arrobando*” al usuario que realiza la consulta.

De la misma manera, cuando por motivos desconocidos para la plataforma se detecten valores anormales que pongan el peligro el cultivo, debe generarse una alarma automáticamente con el objetivo de avisar dicha anomalía. Esta alarma se disparará hacia las cuentas de mail y twitter avisando que la cámara requiere su atención.

Hidroponía.

La hidroponía puede considerarse como un conjunto de técnicas que permitirá el cultivo de plantas sin suelo. Puede permitir en estructuras simples o complejas, la producción de plantas (principalmente, de tipo herbáceo) aprovechando sitios o lugares como azoteas, invernaderos climatizados, suelo infértiles, etc. [4] Por medio de una solución de nutrientes aportan el alimento que se necesario para el desarrollo de las plantas. Los elementos químicos esenciales se diluyen en agua y son transportados por diversas técnicas a los soportes de las plantas. El término hidroponía deriva del griego Hydro (agua) y Ponos (labor o trabajo) [5]. A partir de este concepto, se han desarrollado técnicas que se sustentan en sustratos (medios para sostener a las plantas) o sistemas que aportan soluciones de nutrientes, ya sea de manera estática o circulante, sin perder de vista las necesidades básicas de la planta, como ser humedad, temperatura, agua y nutrientes.

Cuando se emplea esta técnica de cultivo, las raíces absorben una solución rica en nutrientes, necesarios para su crecimiento, y concentraciones equilibradas diluida en agua. Dicha solución dispone de todos los elementos químicos necesarios para que la planta tenga un correcto desarrollo. De esta manera, la planta podrá crecer en una solución mineral únicamente, o en un medio inerte, como ser arena, fibra de coco o lana de roca, entre otros [6].

En condiciones normales, el suelo es quien actúa como reserva de nutrientes minerales, pero el suelo en sí no es esencial para el crecimiento de la planta. Cuando los nutrientes minerales que posee la tierra se diluyen en agua, las raíces son capaces de absorberlo. Cuando estos nutrientes son incorporados dentro del suministro de agua de la planta, ya no es necesario el suelo para que la planta se desarrolle y crezca. Con esto, cualquier planta que se cultiva por el método de siembra directa (Figura 1), puede cultivarse también mediante la utilización de la hidroponía, aunque haya algunas que puedan crecer con mayor celeridad y mejores resultados que otras.



Figura 1. Método de siembra directa.

En la actualidad, esta actividad está teniendo un gran auge en países donde las condiciones para la agricultura resultan desfavorables. Combinando ésta técnica con un buen manejo de invernadero, es posible alcanzar rendimientos muy óptimos con relación a los que se pueden obtener en cultivos a cielo abierto.

Es una manera simple, sencilla y de un costo muy bajo para producir vegetales de rápido desarrollo y, por lo general, ricos en elementos nutritivos. Con esta técnica de agricultura a pequeña escala, se usan los recursos y medios que las personas tienen a mano, como ser materiales de desecho, espacios inutilizables y tiempo libre.

La hidroponía ha conseguido estándares comerciales y algunos alimentos, plantas decorativas y jóvenes plantas de tabaco se cultivan de esta manera, debido en gran parte a la falta de suelos adecuados, suelos contaminados que pueden producir enfermedades a las plantas o la utilización de aguas subterráneas que degradan la calidad de los suelos.

Cuando no se utiliza el suelo como lugar de crecimiento de la planta, no se cuenta con el efecto amortiguador que brinda el suelo agrícola. Sin embargo, también se tiene diversos problemas con la oxigenación de las raíces y no es algo que pueda considerarse limpio a escalas comerciales.

La hidroponía es ideal para personas con tiempo libre que desean divertirse, para trabajos de investigación, para que docentes realicen demostraciones a sus alumnos sobre la esencialidad de distintos elementos químicos, también para quien desee cultivar en un contenedor o tina pequeña, para cultivar en aeronaves espaciales o cultivos en gran

escala; aunque se presentarán distintos niveles de complejidad, sobre todo si se desea que sea una actividad económica y tenga un impacto ambiental bajo [6].

Cámara de germinación hidropónica.

Para construir la cámara de germinación, se ha utilizado un contenedor de policloruro de vinilo (PVC), en la cual se han realizado los siguientes montajes:

1. En la parte superior: el hardware que hace de unidad de control, los sensores de humedad y temperatura exterior.
2. En la parte inferior: están los sensores de humedad y temperatura interior, un turboventilador, sistema de iluminación artificial con una lámpara de bajo consumo de 75w y una lámpara incandescente, la cual será utilizada como una fuente de calor para los días de baja temperatura (Figura 2).

Además, en la parte inferior se ha montado un sensor de nivel de nutrientes que, en caso de que el nivel descienda de un mínimo preestablecido, acciona una bomba de impulsión sumergida en un depósito de nutrientes ubicada al lado de la misma.

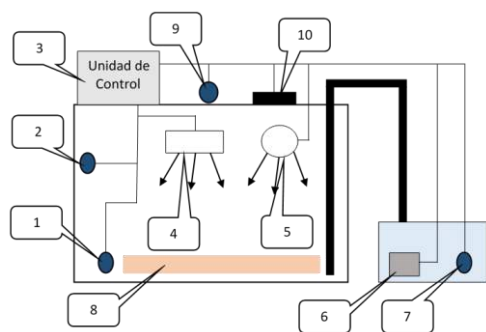


Figura 2 – Cámara de Germinación [1]. 1) Sensor de nivel de nutrientes; 2) Sensor Humedad y Temperatura; 3) Unidad de Control; 4) Lámpara Iluminación artificial; 5) Lámpara incandescente; 6) Bomba de impulsión de nutrientes; 7) Sensor nivel de nutrientes; 8) Espuma de germinación; 9) Sensor Humedad y Temperatura Ambiente; 10) Ventilador.

Placa Raspberry Pi y Sistema Operativo Raspbian.

La Raspberry es considerada una maravilla en miniatura, en cuyo interior existe un importante poder de cómputo en un tamaño no mayor al de una tarjeta de crédito. [7]

Raspberry Pi es un computador de placa reducida o SBC (*Single Board Computer*) de bajo costo, desarrollado por la Fundación

Raspberry Pi en el Reino Unido, con el objetivo de poder estimular a enseñanza de las ciencias de la computación en las escuelas y colegios.

Posee un reducido tamaño (85mm de largo por 56mm de ancho), además posee una salida HDMI, una interfaz Ethernet y 4 USB; por lo que si se conecta un monitor, teclado y mouse, se convierte en una computadora. Está equipado con un procesador ARM1176JZF-S (armv6k) a 700MHz, una GPU Broadcom VideoCore IV3, memoria de 512Mb además de permitir almacenamiento por tarjetas SD o SDHC [8].

Aunque no se indica si es hardware libre o con derechos de marca, en su web oficial [9] explican que se disponen de contratos de distribución y ventas con varias empresas, pero también se pueden convertir en revendedores o redistribuidores de las Raspberry Pi. En cambio, su software sí es *Open Source*, siendo su sistema operativo oficial una versión de Debian adaptada, la cual se denomina Raspbian, aunque también es posible utilizarlo con otros sistemas operativos, entre ellos una versión de Windows 10.

Raspbian [10] es un sistema operativo libre basado en Debian, optimizado para el hardware de Raspberry Pi. Dicho sistema operativo presenta un conjunto de programas básicos y utilidades que hacen funcionar a la Raspberry. Sin embargo, Raspbian no solamente es un sistema operativo puro ya que viene con más de 30.000 paquetes de software pre compilados, en un formato agradable para una fácil instalación y configuración de la Raspberry Pi.

La idea principal de que el sistema Raspbian venga por defecto con más de 30.000 paquetes preinstalados y optimizados para la obtención de un mayor rendimiento en los Raspberry Pi, se completó en junio de 2012. De todas maneras, Raspbian todavía se encuentra en etapa de desarrollo activo, haciendo mayor énfasis en mejorar la estabilidad y el rendimiento del mayor número de paquetes Debian posibles [11].

Cabe mencionar que el sistema Raspbian no se encuentra afiliado a la Fundación Raspberry

Pi. Raspbian fue creado por un equipo pequeño y dedicado de desarrolladores fanáticos del hardware Raspberry Pi, los objetivos educativos de la Fundación Raspberry Pi y el Proyecto Debian. El desarrollo de este sistema operativo es financiado y respaldado por la comunidad, aunque los costos asociados no son gratuitos [10].

Internet de las Cosas.

Al realizar la combinación de Internet con las tecnologías actuales, denominadas emergentes, tales como localización en tiempo real, comunicación de corto alcance, red de sensores integrados, permite transformar los objetos cotidianos como ser heladeras, lavarropas entre otros, en objetos inteligentes que puedan entender y reaccionar ante ciertos eventos de su entorno [12]. ¿Es posible imaginar que una heladera le avise acerca de las fechas de vencimiento de los productos que posee? ¿Qué sucedería si el cepillo de dientes lo alertara sobre alguna carie y solicitase un turno al dentista, sin su intervención? Estos objetos son los elementos fundamentales sobre los que se ha formulado, lo que hoy en día se conoce como, Internet de las Cosas (IoT, *Internet of Things*). Es un tema sobre el cual se presta mayor atención en la actualidad, en ambientes informáticos y de alta tecnología. Para muchos es considerada como una nueva generación de intercomunicación, entre objetos de la vida real o cotidiana, el cual constituye una tendencia hacia donde el desarrollo de la tecnología de las telecomunicaciones se mueve de manera acelerada [13].

El término IoT, como tal, comenzó a utilizarse a finales de los años noventa. Más precisamente, fue utilizado por primera vez en el año 1999 por Kevin Ashton (1968 -), pionero de la tecnología británica, en el cual describe un sistema en el que los objetos en el mundo físico podrían conectarse a Internet por medio de sensores. El IoT puede considerarse como un paradigma que sienta las bases y modelos para resolver el inconveniente de tener interconectados todos los objetos que nos rodean [13].

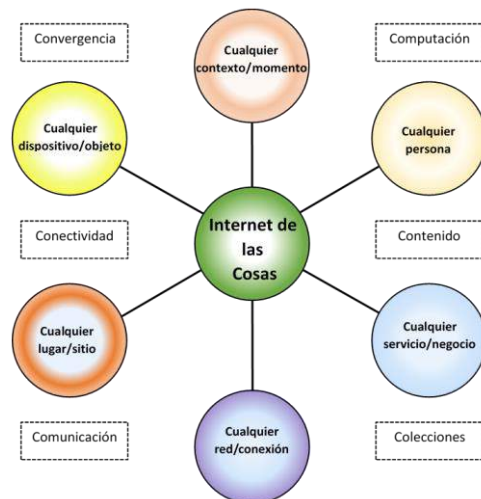


Figura 3 – Internet de las cosas.

La ventaja que tiene IoT es que utiliza la gran mayoría de los estándares existentes de Internet, para la transferencia de información, recolección y análisis de datos, así como también el desarrollo de aplicaciones que permitan la interacción y comunicación con los usuarios. La evolución que ha tenido Internet, ha permitido la interconexión e interrelación de las personas por sí mismas a través de aplicaciones como correo electrónico, redes sociales, entre otros. En este momento, nos encontramos en la era de la interconexión con los objetos y cosas cotidianas, con el propósito de crear y promover un ambiente informado y confortable para perfeccionar la toma de decisiones y obtener una mejor calidad de vida.

1 INTRODUCCION

1.1 Antecedentes

La cámara de germinación consiste en un recipiente plástico con una solución de nutrientes con una placa de poliestireno expandido con rectangulares perforaciones flotando. Si bien, actualmente se está trabajando en la realización de cultivos hidropónicos para los invernaderos de la provincia de Misiones [14], muchos productores de hortalizas aún utilizan el método de siembra directa para la obtención de los plantines y realizar su posterior trasplante en canteros.

Las nuevas tecnologías de Internet tienden a la integración de servicios de la web, redes

sociales, redes de sensores que puedan obtener información de variables ambientales o de cualquier otro tipo y además controlar actuadores, cámaras, etc.; la integración de estos dispositivos y servicios dan como resultado la IoT [15] [16]. Cuando se trabaja con un conjunto de sensores inalámbricos distribuidos espacialmente e interconectados por distintos concentradores, se denomina WSN (*Wireless Sensor Network*) utilizado para monitoreo ambiental, sistemas interconectados de energía, etc., [17].

A la hora de desarrollar aplicaciones del tipo IoT, debemos tener en cuenta que se deben considerar la adquisición de datos proveniente de sensores, tomar decisiones para realizar acciones sobre actuadores, cámaras u otros dispositivos y en tiempo real interactuar con servicios en línea. Si tuviéramos que programar una aplicación IoT de este tipo, se necesitan conocer varios lenguajes y protocolos, así como desarrollar APIs específicas y vincularlas con todos los componentes del sistema. Por esto, surgen varias aplicaciones que integran los componentes que permiten desarrollar aplicaciones IoT. Estas se componen de nodos con funciones específicas y se interconectan por conexiones que intercambian flujo de datos.

El trabajo propone el desarrollo de un prototipo de sistema de monitoreo para sistemas hidropónicos automatizados mediante la evaluación y control de las variables intervinientes. Dicho monitoreo se debe realizar a través de una cuenta de correo electrónico y de la red social Twitter, creadas para tal fin.

Del relevamiento bibliográfico surgió, habiendo hecho una comparación entre las distintas herramientas disponibles en el mercado para la realización del sistema IoT, además se busca que sea compatible con la placa SBC Raspberry Pi 2 modelo B, que es el hardware utilizado en la cámara de germinación, se opta por el desarrollo utilizando Node-RED.

1.2 Diseño de sistemas de control con hardware

Como se ha mencionado anteriormente, la placa utilizada es la Raspberry Pi 2 modelo B (Figura 4); la cual tiene un tamaño aproximado de 85mm de largo por 56mm de ancho; posee una salida HDMI, un conector de red Ethernet y 4 puertos USB, con lo que si conectamos un monitor, teclado y mouse tenemos un ordenador de bajo costo. Además, posee una CPU ARM1176JZF-S (armv6k) a 700 MHz3, GPU Broadcom VideoCore IV3, memoria 512 Mb, capacidad de almacenamiento tarjeta SD o SDHC.



Figura 4 – Raspberry Pi 2 modelo B.

1.3 Puertos GPIO.

Los puertos GPIO son, como su nombre lo indica, puertos de entrada/salida de propósito general, es decir, un conjunto de conexiones que pueden utilizarse como entradas o salidas para diversos usos. Estos puertos representan la interfaz entre la Raspberry Pi y el medio exterior [17].

Posee un total de 17 puertos configurables como entrada o salida. Por defecto están todos configurados como entradas excepto los GPIO 14 y 15 que operan como salidas (Figura 5).

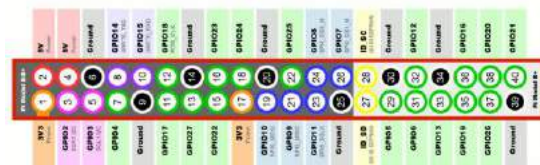


Figura 5 – Puertos GPIO de Raspberry Pi.

1.4 Sensores de humedad y temperatura DHT22

Para la obtención de los valores de las variables de humedad relativa y temperatura se utilizan sensores DHT22 (Figura 6 y Tabla 1), los cuales poseen un rango de operación de 0 a 100% HR y de -40 a 80° C. Además cuentan con una precisión de humedad de 2%

HR y de temperatura 0,5%. Por otro lado, poseen la característica de trabajar con protocolo serial enviando los datos por una única vía, por el pin de datos (Figura 6).



Figura 6 – Sensor DHT22

Tabla 1. Pines de salida de DHT22

Pin	Función
1	5 v
2	Datos
3	NC
4	GND

De acuerdo a la Tabla 1, el pin 1 se corresponde a la alimentación, el 4 a masa (GND) y el pin 2 posee una resistencia *pull up* de 10 K por donde se envían los datos de humedad y temperatura en forma serial hacia el puerto GPIO de la Raspberry Pi (Figura 7).

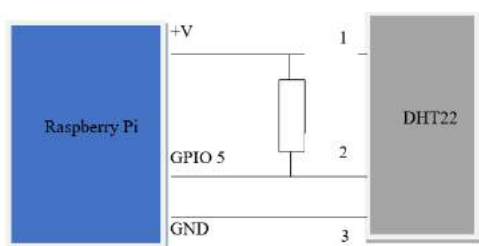


Figura 7 – Conexiones del sensor DHT22 con el puerto GPIO.

1.4 Node-RED.

Es un editor basado en web para generar aplicaciones de IoT. De código abierto, creado por el equipo de *IBM Emerging Technology*. Posee una interfaz amigable que, mediante la función de arrastrar y soltar, permite colocar en el área de trabajo nodos que representan APIs de web, servicios en línea o dispositivos

de hardware. Estos nodos se pueden unir con un cableado que representan el flujo de datos [18], también se pueden programar funciones en JavaScript; de esta manera es posible programar nodos con determinadas funcionalidades. El motor de tiempo real se basa en Node.js que es un entorno para ejecutar JavaScript realizado con el motor V8 de Chrome. Dado que Node.js trabaja con arquitectura basada en eventos no bloqueantes que se ejecutan del lado del servidor, esto hace que consuma poco costo de hardware lo que lo hace ideal para correr en placas Arduino (www.arduino.cc) o Raspberry Pi [19].

2 LÍNEAS DE INVESTIGACION, DESARROLLO E INNOVACIÓN

A continuación se detallan los avances más importantes con respecto a la realización del sistema de monitoreo.

- Actualmente se está trabajando en el desarrollo del sistema de monitoreo a través de Node-RED. Se posee un bosquejo inicial que recibe un correo electrónico y lo responde (Figura 8).
- Se trabajó en la creación y seguridad de la cuenta de correo electrónico, este paso permitió poder realizar el bosquejo antes mencionado.
- Se habilitó el acceso de aplicaciones externas en Gmail para poder obtener los correos electrónicos a través del sistema propuesto.
- Se realizó la petición para poder utilizar la API de Twitter con el objetivo de poder obtener y responder los tweets que se reciban.

3 RESULTADOS Y OBJETIVOS

Actualmente, el sistema desarrollado recibe correos electrónicos y los responde al remitente, con un mensaje personalizado del estado de las variables que se consultan a través del mismo.

Al no obtener todavía, acceso a la API de Twitter, no es posible consultar los datos de las variables utilizando esta red social.

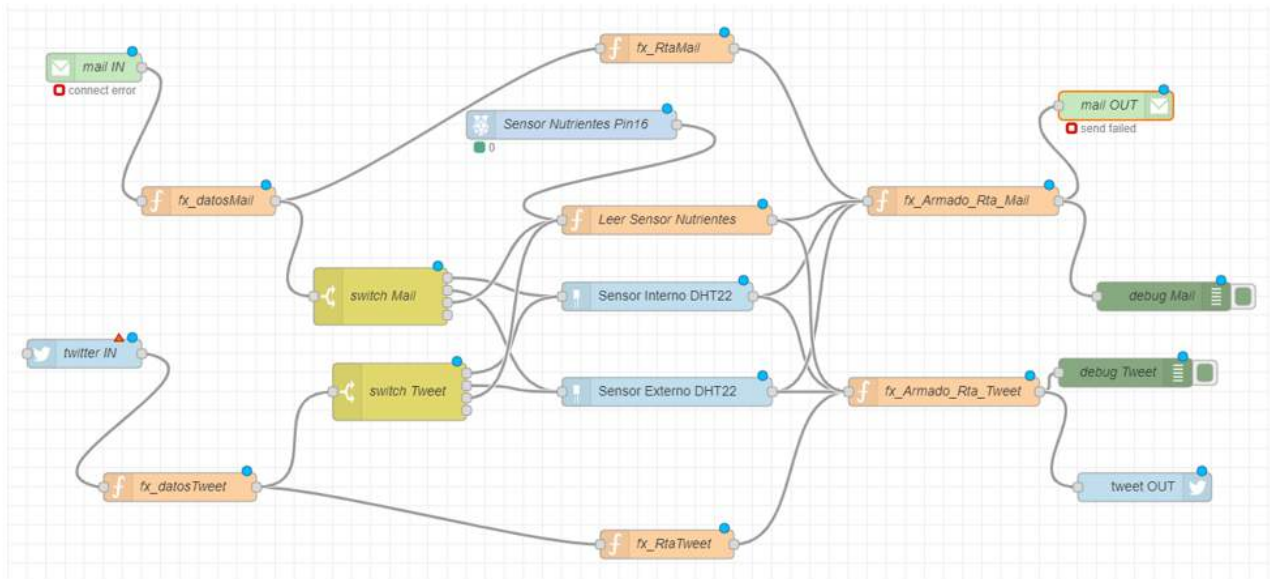


Figura 8 – Diseño del sistema de Monitoreo de la cámara hidropónica

Este trabajo cuenta con publicaciones en distintos congresos y eventos académicos, los cuales se mencionan a continuación:

- M. Marinelli & R. Urquijo, “Sistema de Control de una cámara de germinación hidropónica con IoT”. Libro de Actas del Congreso Argentino de Ciencias de la Computación (CACIC) 2017, (pp. 1101-1107) ISBN 978-950-34-1539-9.
- Urquijo, R. & Marinelli, M., “Sistema de monitoreo de una cámara de germinación hidropónica con IoT basado en Raspberry Pi”, Anales de la 47ª Jornadas Argentinas de Informática (47JAIIO), Simposio Argentino de Grandes Datos AGRANDA, ISSN: 2451-7569, pp. 64-73, Argentina, 2018.
- Marinelli, Marcelo; Lombardo, Graciela; Wurm, Guillermo & Urquijo Rubén R., “Implementación de sistemas de control automático para cultivos hidropónicos en invernaderos de la provincia De Misiones.”, Libro de resúmenes: Jornadas Científico-Tecnológicas UNaM, 1º Edición, Editorial Universitaria, ISBN: 978-950-579-495-4, pág. 277, 2018, Posadas Misiones, Argentina.
- Urquijo, Rubén R. & Marinelli, Marcelo, “Monitoreo de una cámara de germinación hidropónica utilizando Internet de las Cosas”, Libro de resúmenes: Jornadas Científico-Tecnológicas UNaM, 1º Edición, Editorial Universitaria, ISBN: 978-950-579-495-4, pág. 289, 2018, Posadas Misiones, Argentina.
- Marcelo Marinelli, Graciela Lombardo, Kuna Horacio, Guillermo Wurm, Rubén Urquijo, Verónica González, “Implementación de sistemas de control automático para cultivos hidropónicos en invernaderos de la provincia De Misiones”, Libro de Actas XX Workshop de Investigadores de Ciencias de la Computación (WICC 2018), ISBN: 978-987-3619-27-4, pp. 128 – 132.
- Urquijo, Rubén R. & Marinelli, Marcelo J., “Monitoreo de una cámara de germinación hidropónica basada en Raspberry Pi y Node-RED”, 4º Congreso de Ingeniería y Ciencias Aplicadas de las Tres Fronteras, 2019, Foz do Iguaçu, Brasil. (Póster).
- Marcelo Marinelli, Myriam Kurtz, Rubén R. Urquijo, Guillermo Wurm,

“Control de cámaras de germinación hidropónicas mediante Internet de las Cosas”, IEEE ARGENCON 2020, V Congreso Bienal de la Sección Argentina del IEEE (Modo virtual).

Como resultado, se espera poder desarrollar el sistema completo, contemplando el uso de correo electrónico y la red social Twitter (en este caso) para poder consultar el estado de las variables de la cámara de germinación hidropónica.

4 BIBLIOGRAFIA

- [1] M. Marinelli, “Control automatizado de cultivos hidropónicos mediante Lógica Difusa”, Tesis Doctoral, Universidad Nacional de Misiones, 2015.
- [2] Sarthak Jain, Anant Vaibhav & Lovely Goyal, "Raspberry Pi based Interactive Home Automation System through E-mail", ICROIT 2014 - Proceedings of the 2014 International Conference on Reliability, Optimization and Information Technology, ISBN: 978-1-4799-2995-5, 2014.
- [3] M. Marinelli & R. Urquijo, “Sistema de Control de una cámara de germinación hidropónica con IoT”. Libro de Actas del Congreso Argentino de Ciencias de la Computación (CACIC) 2017, (pp. 1101-1107) ISBN 978-950-34-1539-9.
- [4] Beltrano, J. & Gimenez, D. O., “Cultivo en hidroponía”, Libro de Cátedra, Editorial de la Universidad Nacional de La Plata, ISBN: 978-950-34-1258-9, 2015. Consultado el 10-04-2018. Disponible en URL: <http://sedici.unlp.edu.ar/handle/10915/46752>
- [5] Errecart, N., “Contenidos de nitratos, oxalatos, vitamina C y grado de aceptación en Cultivo sin suelo de espinaca vs. Cultivo en suelo”, Tesis de Grado, Universidad Fasta, Mar del Plata, Argentina, 2011. Consultado el 11-06-2018. Disponible en URL: <http://redi.ufasta.edu.ar:8080/xmlui/handle/123456789/355>
- [6] Portillo, G., “La hidroponía”, Blog RenovablesVerdes, 2017. Consultado el 19-07-2018. Disponible en URL: <https://www.renovablesverdes.com/hidroponia/>
- [7] E. Upton & G. Halfacree “Raspberry Pi: Guía del Usuario 2º Edición en Español”, Ed. Wiley, ISBN: 978-1-118-46446-5, 2012.
- [8] "Ultimate guide to Raspberry Pi", No. 324, Computer Shopper. Consultado el 21-09-2017. Disponible en URL: <http://micklord.com/foru/Raspberry%20Pi%20Pages%20from%20Computer%20Shopper%202015-02.pdf>
- [9] Sitio Oficial de Raspberry Pi, <https://www.raspberrypi.org/>. Consultado el 06-09-2017.
- [10] Sitio Oficial de Raspbian, <http://www.raspbian.org/>. Consultado el 06-09-2017.
- [11] William Harrington, "Learning Raspbian", Pack Publishing, 2015, ISBN: 978-1-78439-219-2. Consultado el 09-05-2017. Disponible en URL: <http://1.droppdf.com/files/VIFV2/learning-raspbian-william-harrington-2015.pdf>
- [12] Kortuem, G., Kawsar, F., Sundramoorthy, V., & Fitton, D., “Smart Objects as Buildings Blocks for the Internet of Things”, Journal IEEE Internet Computing, Vol. 14 (1), pp. 44-51, 2010.
- [13] Gonzalez, S. M., “Entendiendo el Internet de las Cosas”, Revista Investiga.TEC, Num. 24, ISSN: 1659-3383, pp. 22-23, 2015.
- [14] 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.

[15] Kelly, S. D. T., Suryadevara, N. K., & Mukhopadhyay, S. C. (2013). Towards the implementation of IoT for environmental condition monitoring in homes. *IEEE Sensors Journal*, 13(10), 3846-3853.

[16] Mainetti, L., Patrono, L., & Vilei, A. (2011, September). Evolution of wireless sensor networks towards the internet of things: A survey. In *Software, Telecommunications and Computer Networks (SoftCOM)*, 2011. 19th International Conference on (pp. 1-6). IEEE.

[17] Jain, S., Vaibhav, A., & Goyal, L. (2014, February). Raspberry Pi based interactive home automation system through E-mail. In *Optimization, Reliability, and Information Technology (ICROIT)*, 2014 International Conference on (pp. 277-280). IEEE.

[18] Kleinfeld, R., Steglich, S., Radziwonowicz, L., & Doukas, C. (2014, October). glue. things: a Mashup Platform for wiring the Internet of Things with the Internet of Services. In *Proceedings of the 5th International Workshop on Web of Things* (pp. 16-21). ACM.

[19] Salihbegovic, A., Eterovic, T., Kaljic, E., & Ribic, S. (2015, May). Design of a domain specific language and IDE for Internet of things applications. In *Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2015 38th International Convention on (pp. 996-1001). IEEE.

Técnicas de Optimización de Soluciones en Sistemas Embebidos

Marcelo Tosini, Lucas Leiva, Martín Vázquez, Oscar Goñi, Juan Toloza

LabSET - INTIA - UNICEN, Tandil, Argentina

{mtosini, lleiva, mvazquez, oegoni, jmtoloza}@labset.exa.unicen.edu.ar

RESUMEN

Se presentan los aspectos relevantes de un proyecto de investigación y desarrollo abordado en el Laboratorio de Sistemas Embebidos -LabSET- del Instituto INTIA de la UNCPBA. Este proyecto fue oportunamente aprobado en el marco de los incentivos a la investigación y busca como objetivos generales el trabajo conjunto tendiente a desarrollar técnicas para optimizar soluciones de hardware y software aplicadas a sistemas embebidos.

Conjuntamente se realiza la formación de parte de los integrantes del laboratorio quienes se encuentran realizando sus estudios de posgrado en temas afines al proyecto.

El proyecto busca desarrollar tareas de investigación y desarrollo en temas relacionados a los sistemas embebidos, tanto en aspectos de software como de hardware. De este modo se busca desarrollar soluciones a problemas en áreas como visualización, aritmética de computadoras, sensores, etc. no solo desde el punto de vista algorítmico sino a partir de su implementación en sistemas de hardware programable como FPGAs.

Palabras clave: sistemas embebidos, tiempo real, procesamiento de imágenes, aritmética decimal.

CONTEXTO

La presente propuesta continúa la línea de trabajo del proyecto de Incentivos “Diseño de sistemas Embebidos de alto Rendimiento” (03/C257) llevado adelante por el grupo de Sistemas Digitales del INTIA durante el período 2015 a 2017 y financiado por SeCAT de UNICEN. La presente propuesta busca extender y especializar a la original con una permanente filosofía de trabajo orientada a la

generación de soluciones informáticas para la resolución de problemas en sistemas digitales en general, y embebidos en particular; fuertemente relacionados con la lógica programable.

El contexto tecnológico de esta propuesta es la generación de soluciones informáticas que resuelvan problemas en el ámbito de los llamados sistemas embebidos, en general, y relacionados en particular con la lógica programable basada tanto en FPGAs como en SoC FPGAs.

El objetivo principal de la nueva propuesta es continuar con el estudio de soluciones de hardware (sensores, operaciones aritméticas, procesadores dedicados y coprocesadores) y software (aplicaciones de control y procesamiento digital de señales – imágenes, audio, video) orientados al desarrollo de sistemas dedicados de alta performance para su uso en sistemas de control, procesamiento e IoT (Internet of Things).

1. INTRODUCCIÓN

Se define como sistema embebido [1] a todo equipo o dispositivo electrónico que permite realizar algún procesamiento de datos, pero que, a diferencia de una computadora de propósito general, se diseña para cumplir una función determinada. El mercado global de sistemas embebidos ha evolucionado considerablemente en los últimos años. Esto incluye la tecnología y las industrias atendidas. Con el advenimiento de IoT y el IoT industrial (IIoT), la tecnología de sistemas embebidos se ha convertido en un facilitador para la rápida expansión mundial de ecosistemas IoT inteligentes y conectados [2]. El mercado de sistemas embebidos

incluye comunicaciones, industria automotriz, aeroespacial, electrónica de consumo, sistemas militares, controles industriales y otros sectores, como las ciudades inteligentes (Smart Cities).

Los sistemas embebidos ofrecen soluciones transparentes en situaciones comunes, resolviendo problemáticas y tomando decisiones de acuerdo a las condiciones de entrada. Un sistema embebido es una máquina que emplea una combinación de Hardware y Software para realizar una función específica. Es parte de un sistema más grande y trabaja en un ambiente reactivo con restricciones temporales. En estos sistemas el software proporciona flexibilidad y funcionalidad; y el hardware proporciona desempeño y seguridad fomentando la transformación digital de la industria 4.0.

Componentes fundamentales de un sistema embebido son la arquitectura de hardware subyacente (FPGAs [3], SoC FPGAs [4], IPs, ICs), los dispositivos de acceso (fundamentalmente sensores) y el software de control, en general un sistema operativo de tiempo real de altas prestaciones y optimizado para funcionar en sistemas de tiempos óptimos de respuesta, bajo consumo y alta fiabilidad.

De este modo, un objetivo general de la propuesta es el estudio de soluciones de hardware (sensores, aritmética, procesadores dedicados y coprocesadores) y software (sistemas de control acordes para sistemas embebidos y aplicaciones de control) orientados al desarrollo de sistemas dedicados de alta performance. La aplicación de los desarrollos logrados se realiza, preferentemente, en tareas de análisis, detección y control a partir del procesamiento de imágenes como eje motivador de preferencia, sin perjuicio de otros ejes de estudio en los que la inclusión de un sistema embebido sea apropiada y superadora.

2. LÍNEAS DE INVESTIGACIÓN Y DE DESARROLLO

Las líneas de investigación y desarrollo se enmarcan en dos ejes principales: referente a plataformas Hardware/Software y respecto a las aplicaciones.

2.1 Plataformas Hardware/Software

Las plataformas utilizadas son placas genéricas con microcontroladores, dispositivos programables FPGAs y SoC (System on a Chip) programables. Este eje comprende:

- *Diseño de cores aritméticos*: estudio, diseño e implementación de unidades aritméticas en dispositivos FPGAs, a través del diseño e implementación hardware eficiente de funciones elementales, tales como logaritmo, raíz cuadrada, exponenciación, potencia, etc.
- *Técnicas y mecanismos de aceleración en SoC programables*: Elaboración de técnicas y mecanismos de identificación de rutinas críticas en cuanto a tiempo de cómputo para posterior aceleración. Profundización del estudio y aplicación de técnicas conocidas en síntesis de circuitos digitales descritos en lenguajes de alto nivel (HLS - High Level Synthesis) para sistemas digitales basados en SoC programables. Desarrollo de núcleos genéricos y específicos descritos en lenguajes de descripción de hardware (HDLs).

2.2 Aplicaciones:

- *Procesamiento de señales*: Utilización y desarrollo de técnicas para el procesamiento digital de señales, imágenes, audio, video, entre otras; en sistemas embebidos sobre las plataformas seleccionadas. Particularmente se propone el estudio de técnicas de *Machine Learning*.
- Internet de las cosas o IoT (*Internet of Things*): estudio, adaptación e implementación de protocolos estándares industrializados. El estudio abarca tanto a los sistemas microprocesados y microcontrolados, así como aquellos basados en lógica programable. Además, se pretende aprovechar la capacidad de procesamiento paralelo de las FPGAs para desarrollar e implementar módulos de encriptación de datos.

3. RESULTADOS OBTENIDOS

En el área de aritmética de computadoras los esfuerzos se centraron en la implementación de soluciones optimizantes para la realización de operaciones aritméticas decimales conformes al estándar IEEE 754-2008[5]. En este sentido, se trabajó en el desarrollo de alternativas para operaciones decimales básicas de suma, resta y multiplicación en punto flotante decimal en formato DPD (Densely Packed Decimal), logrando resultados muy competitivos en la implementación de soluciones en dispositivos programables[6].

También se trabajó en soluciones a operaciones más complejas como funciones logaritmos y métodos de cálculo rápido de raíces cuadradas, siempre en formato decimal DPD [7].

Además, se trabaja actualmente en el análisis de comportamiento de soluciones para el redondeo de resultados en punto flotante decimal codificados en binario, Binary

Integer Decimal (BID). Esta línea de trabajo permitió la obtención de un grado doctoral de uno de los integrantes del proyecto y otro trabajo similar se encuentra actualmente en proceso.

Por otra parte, se lograron avances significativos en el área de IoT, llevando adelante un proyecto que implementa monitoreo de parámetros biométricos y la geolocalización de operarios en instalaciones mineras. El mayor desafío de este proyecto es utilizar la incapacidad que poseen las señales de radio para penetrar muros o rocas como entradas de métodos de estimación de distancia y posición. Como resultado, se implementó un método de mejora de estimación de distancia basado en redes neuronales [8].

Se fortaleció la vinculación con otros centros de investigación, como CIVETAN (UNICEN) en donde actualmente se encuentra en desarrollo un sistema embebido para el conteo automático de huevos de parásitos para su aplicación en veterinaria ganadera. El sistema está basado en un sistema de procesamiento de imágenes desarrollado con herramientas de síntesis de alto nivel (HLS), y participa un alumno de la Universidad Nacional de Tres de Febrero.

Además, se generaron vínculos con el Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense (InFo-Lab) de la Universidad FASTA, investigando sobre el uso de FPGAs aplicadas a la inteligencia artificial, particularmente en SVM y Deep Learning.

Por otra parte, se establecieron vínculos con los responsables del desarrollo de la plataforma EDU-CIAA-FPGA pertenecientes a la Universidad Tecnológica Nacional (UTN-Haedo), a fin de aportar a este proyecto.

Respecto de la transferencia, se logró una vinculación con empresas interesadas en el desarrollo de microemprendimientos

productivos que involucran la generación de soluciones con soporte de sistemas digitales. En este aspecto, se ejecutó un proyecto referente a la convocatoria Universidades Agregando Valor 2018, financiado por SPU. El proyecto fue ejecutado con Satellogic participando como contraparte, con el objetivo de investigar el consumo en algoritmos de compresión embebidos, específicamente sobre FPGAs. El proyecto involucró el análisis de algoritmos de compresión de imágenes y el desarrollo de su arquitecturas eficientes. Algunos de los resultados logrados se encuentran publicados en [9]. Además, se ejecutó un proyecto financiado por la fundación Dr. Manuel Sadosky en su convocatoria Fase Cero, a través de un convenio tripartito entre la Fundación, Redimec SRL y UNICEN. El objetivo del proyecto fue el desarrollo de un dispositivo que permitiera la detección de malezas en tiempo real, para aplicación selectiva, y los resultados se encuentran publicados en [10].

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por docentes-investigadores, egresados y alumnos de la Universidad Nacional de Centro de la Provincia de Buenos Aires. Cuenta con cinco docentes investigadores (dos doctores en ciencias informáticas, un doctor en matemática computacional e industrial y dos magísteres en sistemas), 5 alumnos de la carrera de ingeniería de sistemas, y un maestrando de la maestría en ingeniería de sistemas. Hasta la fecha, uno de los investigadores ha concluido su doctorado. Además se han dirigido 5 becas de estímulo a la vocación científica (EVC-CIN) y actualmente dos alumnos se encuentran en proceso de admisión. En el transcurso 6 alumnos finalizaron sus tesis de grado. Se

prevé la finalización de 3 tesis de grado más durante este año.

5. BIBLIOGRAFÍA

- [1] Lee, Edward Ashford, and Sanjit A. Seshia. *Introduction to embedded systems: A cyber-physical systems approach*. Mit Press, 2017.
- [2] Kim, J. H. (2017). A review of cyber-physical system research relevant to the emerging IT trends: industry 4.0, IoT, big data, and cloud computing. *Journal of industrial integration and management*, 2(03), 1750011.
- [3] Kilts, Steve. *Advanced FPGA design: architecture, implementation, and optimization*. John Wiley & Sons, 2007.
- [4] Crockett, Louise Helen, et al. *The Zynq Book: Embedded Processing with the Arm Cortex-A9 on the Xilinx Zynq-7000 All Programmable Soc*. Strathclyde Academic Media, 2014.
- [5] IEEE Working Group of the Microprocessor Standards Subcommittee. *IEEE Standard for Floating-Point Arithmetic*. New York: The Institute of Electrical and Electronics Engineers, 2008.
- [6] Vázquez, Martín. *Análisis e implementación de operaciones aritméticas en base diez sobre dispositivos de lógica programable*. Tesis de Doctorado en Matemática Computacional e Industrial, Universidad Nacional del Centro de la Provincia de Buenos Aires, 2018.
- [7] Vázquez, Martín & Tosini, Marcelo, *Design and Implementation of Decimal Fixed-Point Square Root in LUT-6 FPGAs*, IX Southern Conference on Programmable Logic (SPL), IEEE, Argentina, 78-1-4799-6848-0, 2014.
- [8] Gerez, Agustín, Oscar Enrique Goñi, Lucas Leiva. "Aumento de Precisión en

- Localización Indoor basado en Redes Neuronales." *Elektron* 4.2 (2020): 74-80.
- [9] Leiva, L., Vázquez, M., Tosini, M., Goñi, O., & Noguera, J. (2020). FPGA based implementation of imagezero compression algorithm. *IEEE Latin America Transactions*, 18(02), 344-350.
- [10] Jose Noguera, Lucas Leiva, Oscar Goñi, Martín Vázquez y Marcelo Tosini, "Prototipo de Sistema de Detección de Malezas en Cultivos Basado en SoC", CASE2020

Redes de Cooperación Científica Internacionales

Consortio para la colaboración en I+D+I en Temas de Cloud Computing, Big Data y Emerging Topics (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)

LCG – Laboratorio de Computación Gráfica (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)

& 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, el análisis de datos masivo y tópicos emergentes, como la visión por computadora, el aprendizaje automático y los sistemas inteligentes, entre otros. Estas temáticas, y su integración, han adquirido creciente importancia por su aplicación en dominios de alto impacto como las ciudades inteligentes, la internet de las cosas, los sistemas de e-health y los basados en tecnologías de block-chain.

Los integrantes del consorcio, provenientes mayoritariamente de Argentina, Chile y España, han tenido a lo largo de los años distintas 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, HPC, Data Analytics, Sistemas Inteligentes, Emerging Technologies.*

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 **JCC-BD&ET** surgieron como *Jornadas de Cloud Computing* realizándose por primera vez del 17 al 19 de junio de 2013 en La Plata, organizadas por el III-LIDI, Facultad de Informática, de la UNLP. 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**. Estas Jornadas se llevan a cabo anualmente

en la Facultad de Informática de la UNLP y constituyen un foro 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. Desde sus inicios, en el contexto de las Jornadas se desarrollan conferencias, paneles, cursos de posgrado, y también se integran ponencias científicas con experiencias de desarrollos y aplicaciones, fomentando la interacción entre la academia y los sectores productivos/industriales. 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.

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. En este contexto de colaboración, las JCC-BD&ET seguirán 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.

Objetivos

El Consorcio se propone realizar acciones de cooperación académica vinculadas con la formación de recursos humanos, la formulación y ejecución de proyectos de investigación conjuntos, la publicación de trabajos científicos, y la coordinación de

acciones de vinculación con empresas y organismos relacionados con la industria informática de las regiones y países a los que pertenecen estos grupos de I+D+I que constituyen el Consorcio.

II. Áreas de Colaboración en el CCC-BD&ET

En el año 2020 se constituye el Consorcio y, a fin de lograr los objetivos propuestos, se establecen diversas áreas de colaboración proponiéndose:

- Fomentar y facilitar la movilidad de investigadores y alumnos entre los grupos/Universidades miembros del Consorcio.
- Organizar de manera colaborativa eventos científicos y de divulgación de las tecnologías y aplicaciones de *Cloud Computing*, *Big Data* y Tecnologías Emergentes.
- Dirigir proyectos de Tesis y/o Cursos y Carreras compartidas en los temas de interés del Consorcio.
- Realizar tareas de investigación colaborativas.
- Formular proyectos conjuntos y aplicar en presentaciones internacionales/nacionales para su financiamiento.
- Compartir un repositorio de publicaciones relacionadas con los temas de interés del Consorcio y una base de datos de referencia a recursos digitales propios y/o publicados por otros y que sean de interés para los temas del Consorcio.
- Organizar al menos dos actividades anuales que se sumen a las JCC-BD&ET: 1) un workshop de Investigadores del Consorcio para discutir líneas de Investigación y posibles trabajos conjuntos y 2) una

Jornada de discusión de Proyectos de Tesis de Postgrado, para potenciar la generación de conocimiento en el área y también fomentar las co-direcciones de Tesis.

- Otras actividades en áreas de interés colectivo.

III. Líneas de Interés de GRUPOS e INVESTIGADORES del CCC-BD&ET

1- Temas Transversales (tratados por varios grupos/investigadores)

- 1.1- HPC (High Performance Computing)
- 1.2- Inteligencia Artificial. Sistemas Inteligentes.
- 1.3- Big Data
- 1.4- Energía. Eficiencia energética.
- 1.5- Simulación. Modelos.
- 1.6- IoT (Internet of Things). Edge Computing. Fog Computing.

2- Temas Verticales de interés para el Consorcio

- 2.1- Seguridad.
- 2.2- Tolerancia a fallos.
- 2.3- Gobierno Digital.
- 2.4- Visualización.
- 2.5- Realidad Virtual, Realidad Aumentada, Realidad Extendida.
- 2.6- Biometría.
- 2.7- BioInformática.

3- Aplicaciones generales de interés en el Consorcio

- 3.1- Ciudades Inteligentes.
- 3.2- Aplicaciones en Salud.
- 3.3- Aplicaciones en Educación.
- 3.4- Aplicaciones en Economía.
- 3.5- Aplicaciones Industriales / Agrícolas
- 3.6- Ingeniería de Software en escenarios híbridos

Se anexa una matriz con las líneas abordadas por cada uno de los grupos e investigadores asociados.

IV Resultados obtenidos y esperados

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.

V. Formación de RRHH

En lo concerniente a la formación de recursos humanos los distintos integrantes están abocados a la formación de recursos humanos tanto a nivel de grado como de posgrado, en temáticas afines al Consorcio. Poner que se realizan individualmente dirigidas por los miembros y en colaboración entre los miembros.

VI. Integrantes del Consorcio

El Consorcio está integrado tanto por Grupos como por Investigadores Asociados. A continuación se detallan los mismos.

Grupos que conforman el CCC- BD&ET

- **III-LIDI** – Instituto de Investigación en Informática LIDI (Universidad Nacional de La Plata – Argentina) [2]
Director: *Ing. Armando De Giusti*
- **LISSI** – Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información (Universidad Nacional del Sur – Argentina) [3]
Director: *Dr. Pablo Fillotrani*

- **VyGLab** – Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (Universidad Nacional del Sur – Argentina) [4]
Directora: *Dr. Silvia Castro*
- **LIDIC** – Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (Universidad Nacional de San Luis – Argentina) [5]
Director: *Dr. Marcelo Errecalde*
- **LCG** – Laboratorio de Computación Gráfica (Universidad Nacional de San Luis – Argentina) [6]
Director: *Mg. Roberto Guerrero*
- **HPC4EAS** – High Performance Computing for Efficient Applications and Simulation (Universidad Autónoma de Barcelona – España) [7]
Director: *Dr. Emilio Luque*
- **SMILe** – Soft Management of Internet and Learning (Universidad de Castilla-La Mancha – España) [8]
Director: *Dr. José A. Olivas Varela*
- **ArTeCS** – Group of Architecture and Technology of Computing Systems (Universidad Complutense de Madrid – España) [9]
Director: *Dr. Francisco Tirado*
- **LITRP** – Laboratorio de Investigaciones Tecnológicas en Reconocimiento de Patrones (Universidad Católica de Maule – Chile) [10]
Representante para el Consorcio: *Dr. Ricardo Barrientos*

Investigadores Asociados al CCC-BD&ET

- Dr. Aurelio Fernández (Universidad Rovira i Virgili – España).
- Dr. Carlos García Garino (Universidad Nacional de Cuyo – Argentina).

- 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] <http://weblidi.info.unlp.edu.ar>
- [3] <https://lissi.cs.uns.edu.ar>
- [4] <http://vyglab.cs.uns.edu.ar>
- [5] <http://lidic.unsl.edu.ar>
- [6] <http://www.lcg.unsl.edu.ar/>
- [7] <https://grupsderecerca.uab.cat/hpc4eas/>
- [8] <http://smile.esi.uclm.es>
- [9] <https://artecs.dacya.ucm.es>
- [10] <http://www.litrp.cl>

Investigadores participantes.

Balladini Javier	Guerrero Roberto
Barrientos Ricardo	Hasperue Waldo
Boracchia Marcos	Lanzarini Laura
Castro Silvia	Luque Emilio
Chichizola Franco	Naiouf Marcelo
De Giusti Armando	Olcoz Katzalin
De Giusti Laura	Olivas Varela Jose A.
Denham Mónica	Pasini Ariel

Errecalde Marcelo	Pesado Patricia
Esponda Silvia	Piccoli Fabiana
Estevez Elsa	Piñuel Luis
Fernandez Aurelio	Printista Marcela
Fillottrani Pablo	Rodriguez Nelson
Frati Emmanuel	Ronchetti Franco
Ganuzza María Lujan	Rucci Enzo
García Garino Carlos	Sanz Cecilia
Gaudiani Adriana	Suppi Boldrito Remo

ANEXO - MATRIZ DE TEMAS EN LOS GRUPOS DEL **C**CC-BD&ET

Grupo	Universidad	Temas de Interés mencionados		
G.1	UAB HPC4eas	1.1 – 1.2 – 1.4 – 1.5	2.2	3.2
G.2	UCM ARTECS	1.1 - 1.2 – 1.4 – 1.6		3.5
G.3	UCLM SMILE	1.2 – 1.3		3.4 – 3.5
G.4	LIPTRP (UC Maule)	1.1 – 1.2 – 1.3	2.1 - 2.6	3.1
G.5	UN Sur VGyLAB		2.4 – 2.5	3.3 – 3.5
G.6	UN Sur LISSI	1.3	2.3	3.1 – 3.6
G.7	UN San Luis LIDIC	1.1 – 1.2 – 1.3 – 1.5		3.1 – 3.2
G.8	UN San Luis LCGrafica		2.4 – 2.5	3.1 – 3.3
G.9	UNLP – III-LIDI P1	1.1 – 1.3 - 1.4 – 1.5 – 1.6	2.2- 2.7	3.3 - 3.5
G.9	UNLP – III-LIDI P2	1.2 – 1.3	2.2- 2.7	3.3 – 3.4
G.9	UNLP – III-LIDI P3	1.3	2.3- 2.5	3.1 – 3.6
I.1	URV Aurelio Fernández	1.2 – 1.3		3.4
I.2	UN Cuyo ITIC Carlos G. Garino	1.1 – 1.6	2.7	3.5
I.3	UNCOMA Javier Balladini	1.1 – 1.4	2.7	3.2
I.4	UN Chilecito Emmanuel Frati	1.1 – 1.3 – 1.4 – 1.5		3.1 - 3.5
I.5	UNGS Adriana Gaudiani	1.1 – 1.5		3.5
I.6	UN San Juan Nelson Rodríguez	1.1 – 1.6		3.5
I.7	UN Río Negro Mónica Denham	1.1 – 1.5	2.4	3.5
OTROS		Grupos I+D+I /Organismos y Universidades que NO están en el Consorcio		

El Uso de Técnicas Computacionales para Mejorar el Cumplimiento de Técnicas del Recuerdo en Entornos Inteligentes (REMIND)

Gramajo Sergio^a, Espinilla Macarena^b, Medina Javier^b Bernal Edna^c

^aCentro de Investigación Aplicada en Tecnologías de la Información y la Comunicación / Universidad Tecnológica Nacional, Facultad Regional Resistencia (Argentina)
^bGrupo de Investigación Avances en Sistemas Inteligentes y Aplicaciones (ASIA) / Departamento de Informática / Universidad de Jaén (España)

^cGrupo de investigación Davinci / Universidad Nacional Abierta y a Distancia UNAD (Colombia)

sergiogramajo@gfe.fre.utn.edu.ar, {mestevez, jmquero}@ujaen.es, edna.bernal@unad.edu.co

RESUMEN

El proyecto “El Uso de Técnicas Computacionales para Mejorar el Cumplimiento de Técnicas del Recuerdo en Entornos Inteligentes” o por su título original “The Use of Computational Techniques to Improve Compliance to Reminders within Smarts Environments (REMIND)” tiene por objetivo general crear una red internacional e intersectorial para desarrollar el marco de intercambio de conocimiento y masa crítica necesario para avanzar en el desarrollo de técnicas del recuerdo que se implementarán en entornos inteligentes orientado a resolver problemas de recuerdo o personas con demencia.

Para lograr este objetivo, el enfoque se centra en el desarrollo de tecnología basada en la ciencia del comportamiento y mejores técnicas computacionales orientadas a soluciones apropiadas del recuerdo.

Palabras Clave: Marie Skłodowska-Curie Actions, Horizonte 2020, Técnicas del recuerdo, Ambientes Inteligentes.

CONTEXTO

Este proyecto internacional se enmarca en el financiamiento de Marie Skłodowska-Curie Actions. Research and Innovation Staff Exchange (RISE). CALL: H2020-MSCA-RISE-2016.

ID de Proyecto: 734355. Período: 01/2017 al 12/2020. Extendido actualmente por COVID-19.

El proyecto consta de 16 participantes beneficiarios distribuidos en 10 países. 7 universidades europeas y 5 socios del sector industrial europeo. Además 3 socios de países fuera de la comunidad y una empresa. La clasificación puede verse en la siguiente tabla y figura:

Nombre Legal	País
University of Ulster	United Kingdom
Universidad de Jaén	Spain
National University of Ireland, Galway	Ireland
Lulea Tekniska Universitet	Sweden
Hogskolan I Halmstad	Sweden
Universita Degli Studi di Firenze	Italy
I+Srl	Italy
Karde As	Norway
F.Ageinglab	Spain
Kyung Hee University	Rep. of Korea
Corporation University de la Costa CUC	Colombia
Universidad Nacional Abierta y a Distancia	Colombia
Universidad Tecnológica Nacional	Argentina
Academisch Ziekenhuis Croningen	Netherlands
Associazione Novilunio Onlus	Italy
Swedish Adrenaline AB	Sweden

Tabla 1. Listado de Participantes



Fig. 1. Ubicación de los Participantes

1. INTRODUCCIÓN

De acuerdo con los objetivos mencionados REMIND [1] centra la atención en desarrollar las habilidades del personal y los socios en áreas de diseño centradas en el usuario y la ciencia del comportamiento, junto con técnicas computacionales mejoradas que a su vez ofrecen soluciones de recuerdo más apropiadas y eficaces.

Se ha establecido un programa de trabajo para maximizar la transferencia de conocimientos entre los diferentes sectores ofreciendo una gama de oportunidades de desarrollo y formación para el personal.

El personal industrial se beneficia a su vez de los intercambios bilaterales con dominios técnicos de las tecnologías tratadas específicamente.

Los miembros europeos académicos se beneficiarán de la adquisición de experiencia en el desarrollo de software estándar industrial conforme a las normas ISO y médicas, el compromiso con las partes interesadas a través de un proceso de diseño centrado en el usuario y el trabajo con organizaciones que brindan atención a ancianos y personas con demencia.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación que se abordan en el proyecto están vinculadas a la experiencia de cada beneficiario. Y se han dividido en 8 Working Parties que tratan dichos temas. En la Figura se observan los WPs.

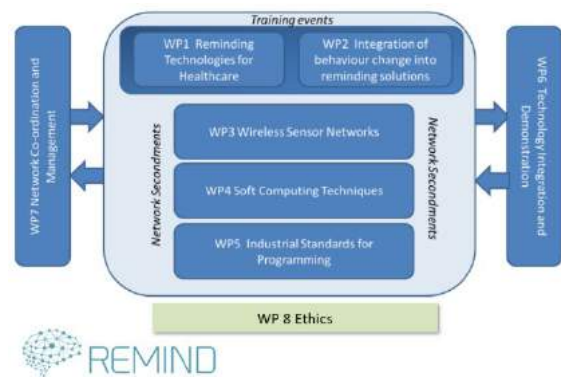


Fig 2. Working Parties y Técnicas Metodológicas de Trabajo.

El líder beneficiario por cada WP es el siguiente:

- WP1. KARDE
- WP2. NUI Galway
- WP3. LTU
- WP4. UJAEN
- WP5. I+
- WP6. HALMSTAD
- WP7. ULSTER
- WP8. ULSTER

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto que promueve la interacción entre la industria y la academia para lograr los objetivos de transferencia de conocimiento y creación de soluciones a través de redes interdisciplinarias. El desarrollo de las soluciones se enfoca en el cuidado de la salud general para personas con problemas de demencia y generar conocimiento tecnológico compartido en este sentido. Específicamente se han desarrollado soluciones para mejorar condiciones que promuevan técnicas de recuerdo en medicamentos, por ejemplo, y asistencia a personas mayores usando para ello sistemas y ambientes inteligentes.

Si bien los resultados obtenidos han sido amplios por cada WP, en lo que respecta a los autores, se ha trabajado en WP4 (Técnicas de Soft Computing) logrando las diversas publicaciones [1] [2] [3].

Los resultados buscados son:

- La formación del personal
- Nuevas colaboraciones
- Colaboraciones extendidas
- Oportunidades de desarrollo profesional

- Entregables del Proyectos
- Oportunidades de difusión
- Intercambio de conocimiento
- Construcción de capacidades
- El resultado final tangible será una solución de las técnicas del recuerdo.

4. FORMACIÓN DE RECURSOS HUMANOS

Con el proyecto se realizaron numerosas estancias que totalizan más de 200 meses distribuidos entre los investigadores de todos los participantes. Las estancias realizadas por los equipos de los autores a Europa son:

- **España:**
 - Sergio Gramajo (UTN)
 - Sixto Campaña (UNAD)
- **Italia:**
 - Luis Lezcano Airaldi (UTN)
 - Edna Rocio Bernal (UNAD)
- **Irlanda:**
 - Raimundo Vázquez (UTN)

- **Suecia:**
 - Edna Rocio Bernal (UNAD)

5. BIBLIOGRAFÍA

- [1] REMIND Research website <https://remind-research.com/index.html>
<https://cordis.europa.eu/project/id/734355/es>
- [2] Javier Medina Quero, Carmen Martinez-Cruz, Macarena Espinilla Estevez and Sergio Gramajo Analyzing daily behaviours from wearable trackers using linguistic protoforms and fuzzy clustering. 24th European Conference on Artificial Intelligence-ECAI, Santiago de Compostella, España. 29/08-08/09 de 2020.
- [3] Carmen Martinez-Cruz, Javier Medina Quero, Jose Maria Serrano and Sergio Gramajo. Monwatch: A fuzzy application to monitorize the user behavior using wearable trackers. IEEE World Congress on Computational Intelligence (WCCI). FUZZ-IEEE 2020. Glasgow (UK). 19 – 24th July, 2020

Ciudades Inteligentes Sostenibles en América Latina – Proyecto CAP4CITY

Armando De Giusti , Patricia Pesado , Ariel Pasini , Pablo Thomas ,
Rocío Muñoz¹  Juan Santiago Preisegger¹ 
Elsa Estevez , Pablo Fillotrani , Sonia Rueda, Karina Cenci, Gabriela A. Diaz

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata

¹ Becario postgrado UNLP

50 y 120 - La Plata, Buenos Aires

Centro Asociado CIC

526 e/ 10 y 11 - La Plata, Buenos Aires

(degiusti, ppesado, apasini, pthomas, rmunoz, jspreisegger) @lidi.info.unlp.edu.ar

Laboratorio de Ingeniería de Software y Sistemas de Información (LISSI)

Departamento de Ciencias e Ingeniería de la Computación – Universidad Nacional del Sur

Av. San Andrés 800 – Campus de Palihue - Bahía Blanca, Buenos Aires

Centro Asociado CIC

526 e/ 10 y 11 - La Plata, Buenos Aires

(ece, prf, svr, kmc, gabriela.diaz) @cs.uns.edu.ar

RESUMEN

El proyecto CAP4CITY tiene como objetivo fortalecer y desarrollar la capacidad de instituciones académicas en América Latina y Europa para mejorar la calidad de la educación superior en el campo de ciudades inteligentes sostenibles. El Proyecto es co-financiado en el marco del Programa Erasmus+ de la Unión Europea, para la construcción de capacidades en la educación superior. Se presenta el avance de los primeros años del proyecto y las actividades a realizar en el siguiente periodo.

Palabras Claves

Ciudades Inteligentes - Capacidades de Gobernanza - Gobernanza Digital - Colaboración Internacional

CONTEXTO

La línea de investigación y desarrollo aquí descrita es parte de las tareas planificadas por el proyecto “Strengthening Governance Capacity for Smart Sustainable Cities (CAP4CITY)”. El Proyecto, financiado como parte del programa Erasmus+ de la Unión Europea

es ejecutado por un consorcio integrado por 12 universidades; cuatro de ellas europeas – Donau Universität für Weiterbildung (DUK) en Austria, Tallinn University of Technology (TUT) en Estonia, Delft University of Technology (TU Delft) en los Países Bajos, y Gdańsk University of Technology (GUT) en Polonia; y ocho universidades en la región de América Latina – Universidad Nacional de La Plata (UNLP) y Universidad Nacional del Sur (UNS) en Argentina; Pontificia Universidade Católica do Rio Grande do Sul (PUCRS) y Faculdade Meridional (IMED) en Brazil; Universidad Técnica Federico Santa María (UTFSM) y Universidad Católica del Norte (UCN) en Chile; y Universidad Externado de Colombia (UEC) y Escuela Colombiana de Ingeniería (ECI) en Colombia. El Proyecto lleva el número 598273 y se ejecuta bajo el acuerdo 598273-EPP-1-2018-1-AT-EPPK A2-CBHE-JP. Debido a la pandemia de COVID-19 el proyecto se ha extendido por otros 12 meses.

1. INTRODUCCION

El proyecto CAP4CITY tiene como principal objetivo utilizar la gran atención

que el concepto de Ciudades Inteligentes Sostenibles (CIS) ha alcanzado en América Latina e integrarlo en varios cursos universitarios utilizando nuevas herramientas de enseñanza y aprendizaje, así como desarrollar nuevos planes de estudio en todos los niveles del proceso educativo.

Se define una SSC como una ciudad innovadora que utiliza Tecnologías de Información y Comunicación (TIC) y otros medios para mejorar la calidad de vida, la eficiencia de la operación y los servicios urbanos, y la competitividad, al tiempo que se garantiza que satisfaga las necesidades de las generaciones presentes y futuras con respecto a los aspectos económicos, sociales, ambientales y culturales.

Uno de los modelos de SSC propone cinco dimensiones para su desarrollo: 1) Social, 2) Económica, 3) Ambiental, 4) Gobernanza, y 5) Infraestructura Urbana. La Figura 1 muestra el modelo.

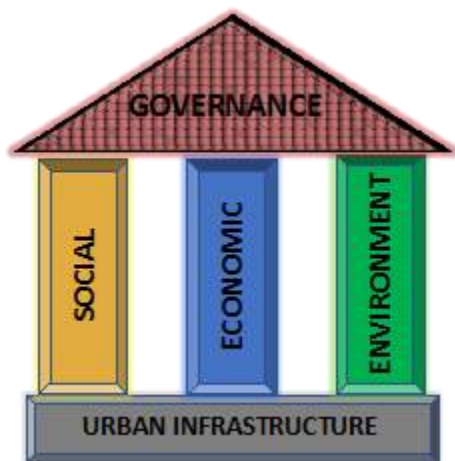


Figura 1. Dimensiones de una CIS

La dimensión Social cubre los aspectos relacionados con las personas y comunidades para garantizar la calidad de vida, como, por ejemplo: salud, seguridad, educación, entre otros. La Económica cubre los aspectos relacionados con el crecimiento económico responsable y sustentable, y la generación de oportunidades laborales. Por su parte, la Ambiental se relaciona con la utilización

de prácticas ecológicas, la protección y restauración del medio ambiente. La dimensión de Gobernanza se refiere a la capacidad de administrar recursos, políticas e involucrar a diferentes partes interesadas, proponiendo mecanismos y procesos regulatorios y de cumplimiento bien equilibrados de manera estandarizada y continua. Por último, la dimensión de Infraestructura Urbana se refiere a la infraestructura física (carreteras, transporte, etc.) y a la infraestructura digital (tecnología de la información y comunicación), que son herramientas esenciales para permitir ciudades inteligentes y sostenibles.

Dadas estas cinco dimensiones, la complejidad de las mismas, así como la característica multidisciplinar de los problemas a resolver en SSC, hace que la construcción de capacidades humanas, particularmente, la formación de líderes que se dediquen al liderazgo y gobernanza de estas iniciativas sea un gran desafío.

Debido al creciente número de competencias necesarias y su característica interdisciplinaria, los planes de estudio de SSC se implementarán en áreas como Administración de Empresas, Informática, Ingeniería, Arquitectura y Urbanismo, Planificación Urbana, Ciencias Políticas, entre otras, a través de una red de colaboración internacional de instituciones académicas en América Latina y Europa, que apoyan la modernización e internacionalización del campo de la educación superior en los países socios.

A fin de definir las competencias necesarias para el desarrollo de SSC en América Latina, se hizo un relevamiento de datos en la región. A tal efecto, se realizaron dos talleres en las ciudades de Bogotá y Medellín, Colombia, como experiencia piloto. Luego cada una de las universidades de Latino América replicó el “Taller de Relevamiento de Competencias

para el Desarrollo de Ciudades Inteligentes y Sostenibles” con representante de su comunidad, incluyendo ciudadanos y representantes de la industria, la academia y el gobierno.

Con los resultados obtenidos de todas las universidades se estableció una lista de competencias en base a las cuales se definieron 31 cursos. Estos cursos fueron validados, durante el transcurso del 2020, mediante entrevistas online, encuestas y workshops con los interesados locales y se contó, además, con un asesoramiento de 11 expertos internacionales externos al proyecto, obteniendo así una retroalimentación de los cursos en general, de cada una de las áreas establecidas y de los contenidos que serán impartidos.

Los beneficiarios del Proyecto CAP4CITY incluyen instituciones de educación superior a nivel local, provincial, nacional e internacional. Estas podrán utilizar los cursos diseñados, compartir conocimientos e intercambiar experiencias sobre el desarrollo y la entrega de programas educativos relacionados con SSC.

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.

1. OBJETIVOS DE INVESTIGACIÓN Y ACTIVIDADES

- Mejorar la calidad de la educación superior en el campo de SSC, aumentando su relevancia para el mercado laboral y la sociedad en general.
- Aumentar las competencias de los recursos humanos en los países en desarrollo de América Latina para

enfrentar los desafíos del mundo digital.

- Proveer el desarrollo continuo de competencias en SSC a través de programas de capacitación y educación.
- Asegurar una enseñanza multidisciplinaria orientada a la resolución de problemas en SSC.
- Desarrollar una red de cooperación entre los socios de diferentes regiones del mundo.
- Promover la colaboración entre los socios del Consorcio, las entidades públicas, las empresas y otros interesados en SSC.
- Facilitar el intercambio de conocimientos, experiencias y buenas prácticas e iniciativas conjuntas en materia de SSC entre socios académicos ubicados en diferentes ciudades y países.

La Universidad Nacional de La Plata y la Universidad Nacional del Sur colaborarán con las siguientes actividades:

- Relevamiento de iniciativas de SSC que puedan identificarse como buenas prácticas y de programas de postgrado relacionados con temas de SSC.
- Capacitación de formadores en SSC.
- Implementación de nuevos programas educativos en SSC.
- Difusión y explotación del concepto de SSC en Argentina.
- Fortalecimiento de los vínculos laborales con sectores públicos y privados de la región.

2. RESULTADOS ESPERADOS/OBTENIDOS

Al primer trimestre de 2021, los resultados obtenidos por la UNLP y la UNS son:

- Se ha desarrollado un relevamiento de programas de postgrado relacionados con temas de SSC.

- Se organizaron talleres para identificar competencias y validar los módulos de formación propuestos.
- Se han diseñado y validado 5 de los 31 cursos (de diferentes niveles – pregrado, posgrado y formación continua) relacionados a SSC.
- Se han estudiado y desarrollado dos casos de estudio relacionados con SSC en Argentina.
- Se ha desarrollado e implementado un curso a distancia de “capacitación de formadores” sobre la construcción de MOOC para los cursos del proyecto.
- Se dictó un curso de posgrado de Fundamentos de Ciudades Inteligentes Sostenibles en la UNLP.
- Se está trabajando para la definición de una nueva Maestría conjunta en Ciudades Inteligentes Sostenibles a ser ofrecido y dictado en colaboración entre la UNLP y la UNS.
- Se organizó un panel para explicar los alcances del proyecto y sus resultados preliminares en el ERASMUS-DAY, organizado por el programa ERASMUS.
- Participación de los integrantes de esta línea de investigación en el dictado de asignaturas/cursos de grado/postgrado en la Facultad de Informática de la UNLP y en el Departamento de Ciencias e Ingeniería de la Computación de la UNS.

4. AGRADECIMIENTOS

Project co-funded by the Erasmus+ Programme of the European Union. Grant no: 598273-EPP-1-2018-1-AT-EPPKA2-CBHE-JP.

5. BIBLIOGRAFIA

- [1] Akande, A., Cabral, P., Gomes, P., Casteleyn, S. 2019. “The Lisbon ranking for smart sustainable cities in Europe.” *Sustainable Cities and Society*, 44, pp. 475-487.
- [2] Caragliu, Andrea, Chiara Del Bo, and Peter Nijkamp. 2011. “Smart Cities in Europe.” *Journal of Urban Technology* 18(2):65–82.
- [3] Craglia, Massimo, Lila Leontidou, Giampaolo Nuvolati, and Jürgen Schweikart. 2004. “Towards the Development of Quality of Life Indicators in the ‘Digital’ City.” *Environment and Planning B: Planning and Design* 31(1):51–64.
- [4] Elsa Estevez, Nuno Lopes, Tomasz Janowski. 2016. “Smart sustainable cities: Reconnaissance Study”, 1-330.
- [5] Huovila, A., Bosch, P., Airaksinen, M. 2019. “Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when?” *Cities*, 89, pp. 141-153.
- [6] Keshvardoost, S., Renukappa, S., Suresh, S. 2019. “Developments of policies related to smart cities: A critical review”. *Proceedings - 11th IEEE/ACM International Conference on Utility and Cloud Computing Companion, UCC Companion 2018*, art. no. 8605807, pp. 365-369.

A futuro se espera lograr:

- Incremento de la matrícula de alumnos interesados en los postgrados de SSC.
- Incremento de publicaciones relacionadas con SSC por parte de las universidades involucradas.
- Inclusión de temas de SSC en currícula de programas de pregrado y posgrado.
- Fortalecimiento de la colaboración en diferentes áreas entre los miembros del Consorcio.

3. FORMACION DE RECURSOS HUMANOS

Los esfuerzos dedicados a la formación de recursos humanos incluyen:

- Capacitación de los miembros del proyecto en SSC.
- Desarrollo de tesis de postgrado y tesinas de grado en el área.

- [7] Kurebayashi, Toshihiko, Yoshihiro Masuyama, Kiyonori Morita, Naoyuki Taniguchi, and Fumio Mizuki. 2011. "Global Initiatives for Smart Urban Development." *Hitachi Review* 60(2):89–93.
- [8] Martin, C., Evans, J., Karvonen, A., Paskaleva, K., Yang, D., Linjordet, T. 2019. "Smart-sustainability: A new urban fix?" *Sustainable Cities and Society*, 45, pp. 640-648.
- [9] Paroutis, Sotirios, Mark Bennett, and Loizos Heracleous. 2014. "A Strategic View on Smart City Technology: The Case of IBM Smarter Cities During a Recession." *Technological Forecasting and Social Change* 89:262– 72.

Seguridad Informática

Aplicación de los Contratos Inteligentes en Internet de las Cosas

Jorge Eterovic; Marcelo Cipriano; Luis Torres; Dalma Agostina Lomoro

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{jorge.eterovic; cipriano1.618 }@gmail.com; torreslu@ar.ibm.com; agostina.lomoro@usal.edu.ar

RESUMEN

Blockchain es una de las tecnologías más innovadoras de nuestro tiempo gracias a su capacidad para asegurar la integridad de las transacciones y la autenticidad entre cualquier entidad conectada a Internet, de manera descentralizada. Entre las ventajas que ofrece Blockchain, se incluyen la condición de permanente e inmutable del registro en la cadena de bloques y la capacidad de ejecutar contratos inteligentes.

Internet de las Cosas (IoT) es un concepto que se refiere a interconectar distintos dispositivos a través de Internet, cuestión que puede traer muchos beneficios a la sociedad de diferentes maneras, pero a la vez, es muy importante investigar y proponer la mejor solución para proteger la seguridad de los datos y de las comunicaciones entre todos los dispositivos interconectados. Esto constituye un gran desafío.

Este proyecto de investigación se centra en la búsqueda y análisis de distintas plataformas Blockchain donde se pueden desarrollar contratos inteligentes y mediante ello permiten dar soporte a la interacción entre dispositivos que nos propone el IoT; la recopilación y el estudio de las vulnerabilidades detectadas y cómo es su comportamiento con respecto a la escalabilidad, la complejidad del sistema y los factores del protocolo de consenso.

El resultado esperado es, en el contexto de la integración de contratos inteligentes entre Blockchain e Internet de las Cosas, encontrar

las oportunidades y resolver los desafíos de esta integración para proteger la seguridad de los datos y de las comunicaciones garantizando la integridad de las transacciones y un ecosistema seguro para los dispositivos interconectados.

Palabras Clave:

Contratos Inteligentes. Ethereum. Blockchain. Internet de las Cosas.

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 “Integración de Blockchain e Internet de las Cosas usando Contratos Inteligentes.”, con una duración de 2 años (2021-2022) y que ya ha sido evaluado y aprobado para su realización.

1. INTRODUCCIÓN

Si buscamos una tecnología que impactará y beneficiará nuestras vidas en los próximos años, es el Internet de las cosas. Los automóviles, electrodomésticos, teléfonos inteligentes, medidores de servicios públicos, sensores incorporados al cuerpo, indumentaria y casi cualquier cosa que podamos imaginar estarán conectados a Internet y serán accesibles desde cualquier parte del mundo [1]. La revolución traída por IoT será inigualable, algunos autores dicen que será similar a la construcción de carreteras y ferrocarriles que impulsaron la Revolución Industrial de los siglos XVIII al XIX [2], y será transversal a todos los sectores de la sociedad y todas las industrias, desde educación, salud, hogar y ciudad inteligente, hasta manufactura, minería, comercio, logística y vigilancia, solo por mencionar algunas [3].

En los últimos años, los investigadores han centrado su atención principalmente en abordar los problemas de escalabilidad de las capacidades computacionales y de comunicación de IoT [4]. Si bien estos temas son primordiales para el éxito de IoT y deben investigarse a fondo, la comunidad científica ahora ha reconocido que deben considerarse los problemas de seguridad y privacidad de los datos en IoT, que no tienen precedentes en cuanto a alcance y magnitud y requerirá un considerable esfuerzo de investigación para ser solucionados [5]. Es fácil imaginar que una vez que las personas, los sensores, los automóviles, los robots y los drones puedan interactuar entre sí desde cualquier lugar del mundo, se encontrarán nuevas amenazas y vulnerabilidades que hoy no podemos imaginar.

IoT implementa una arquitectura de acceso centralizado basado en el modelo cliente-

servidor en el que las transacciones de IoT (datos, dinero, información en general) entre dispositivos de IoT (sensores, redes de comunicaciones, servidores, aplicaciones, etc.) se realiza con proveedores de servicios monolíticos y centralizados [6]. Este modelo simplifica claramente las interacciones entre los dispositivos de IoT y facilita el proceso de recopilación de datos. Sin embargo, hace que IoT sea vulnerable a una serie de problemas de seguridad y privacidad.

Los proveedores de servicios centralizados pueden hacer un uso ilegítimo de los datos de IoT, como por ejemplo de los sistemas de vigilancia masiva [7]. Aún más importante, los modelos centralizados de recolección de datos pueden exponer a todo el sistema a piratería por actividades maliciosas, con consecuencias nefastas para los ciudadanos, como se dio a conocer en el trabajo de Dan Goodin [8]. Otro desafío importante es la autenticación de los dispositivos de IoT que se conectan a la red [9]. Si no se abordan los problemas de autenticación de IoT, los atacantes pueden generar botnets maliciosas, como ya ocurrió con el ataque del malware Mirai [10] y ataques Sybil [11], muy difíciles de detectar.

En un ataque Sybil, un atacante puede contaminar un sistema distribuido creando un gran número de identidades que aparenten ser independientes y usarlas para obtener una influencia desproporcionada, alterar rutas o modificar contenido almacenado de forma redundante. De esta forma, ciertos nodos legítimos pueden sufrir una usurpación de identidad al estar solo conectados a los del atacante. La vulnerabilidad del sistema depende de la facilidad para crear nuevas identidades y la falta de una cadena de confianza bien constituida, que pueda hacer que todas las identidades sean tratadas por igual.

Una manera de abordar los desafíos antes descritos es organizar el IoT de manera descentralizada, de modo que ningún dispositivo individual tenga control sobre las transacciones de la red de IoT. La descentralización no solo proporciona seguridad y privacidad por diseño, sino que

también les dará a los usuarios la opción de compartir o vender sus datos de sensores con dispositivos de terceros sin intermediarios. El control descentralizado también implica escalabilidad, lo que ha afectado a Internet de las cosas desde su inicio [12]. El desafío es investigar modelos descentralizados de acceso a datos para el IoT, lo que asegurará que los datos de los usuarios no se confíen a dispositivos o empresas centralizadas, sino que se conviertan en propiedad de los propios usuarios.

Las tecnologías y sistemas basados en el concepto de Blockchain están irrumpiendo en el mercado mundial de criptomonedas y pueden resultar adecuados para lograr los objetivos de seguridad y privacidad de IoT mediante un modelo descentralizado [13]. Aunque los algoritmos y principios de Blockchain se conocen desde los años 70, cuando se desarrollaron los árboles de Merkle [14] y los algoritmos de consenso [15], la primera aplicación práctica de Blockchain se propuso en 2008 como parte de la criptomoneda Bitcoin [16]. Desde entonces, se ha aplicado también a una amplia gama de desarrollos no monetarios, que incluyen logística, gestión de energía, ciudades inteligentes, drones, robots y manufactura industrial.

Una cadena de bloques mantiene una colección (o libro mayor) de transacciones de manera descentralizada y distribuida. El libro mayor es inmutable o irreversible, lo que significa que las transacciones pasadas no pueden ser modificadas por ninguna entidad que registre transacciones en la Blockchain, y se comparte y sincroniza en todos los nodos participantes. De esta manera, la cadena de bloques garantiza que el libro mayor no puede ser manipulado, y que todos los datos que posee la Blockchain son confiables.

Un algoritmo de consenso, que implica un desafío difícil de resolver y que requiere de importantes recursos informáticos, pero fácil de verificar, llamado prueba de trabajo (PoW: Proof of Work), se utiliza para agregar nuevos bloques en la cadena de bloques (minar), y así establecer una red segura y confiable entre

entidades no confiables. Para fines de identificación, los nodos de blockchain pueden optar por emplear claves públicas cambiables para evitar el seguimiento. Se fusionan varias transacciones para formar un bloque que se agrega al libro mayor siguiendo el algoritmo de consenso. Cada bloque incluye el hash del bloque anterior en el libro mayor. Cualquier modificación de una transacción de un bloque puede detectarse fácilmente ya que el hash del bloque posterior no coincidirá.

La combinación de Blockchain e IoT tiene un potencial disruptivo. La cadena de bloques puede ayudar a la consolidación de IoT al proporcionar las siguientes ventajas:

Anonimato. Los dispositivos de IoT pueden participar en la cadena de bloques con una clave pública / privada, que no revela en sí misma la identidad real de su propietario;

Descentralización. Los sistemas centralizados tradicionales requieren que cada transacción se valide a través de una autoridad centralizada, por ejemplo, un banco central, lo que produce un cuello de botella en el rendimiento. Por el contrario, la validación de terceros ya no es necesaria en la cadena de bloques, ya que los algoritmos de consenso mantienen la consistencia de los datos;

No repudio. La cadena de bloques asegura que las transacciones se puedan validar fácilmente; y no se admiten transacciones no válidas: es casi imposible eliminar o revertir las transacciones una vez incluidas en la cadena de bloques.

Aunque la cadena de bloques puede parecer la solución para los problemas de seguridad y privacidad de IoT, todavía hay muchos aspectos a resolver que impiden su aplicación inmediata en la mayoría de las redes actuales de IoT. La mayoría de los algoritmos de consenso utilizados por los sistemas basados en Blockchain no fueron diseñados para ejecutarse en dispositivos con restricciones de capacidad de cómputo, de consumo de energía y con ancho de banda extremadamente estricto, como ocurre en Internet de las Cosas. Es necesario abordar varios desafíos clave, que incluyen: problemas de escalabilidad que

surgen de la necesidad de lograr un consenso entre miles de millones de mineros; altas demandas de cómputo debido al uso de algoritmos de prueba de trabajo; y grandes demoras debido a mecanismos contra el doble gasto, problema que no siempre aplica al IoT.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Si analizamos los sistemas de IoT basados en Blockchain investigados hasta ahora en las publicaciones científicas, podemos dividir los documentos por categorías, cada una con el nombre de las aplicaciones de IoT más comunes disponibles en la actualidad: energía inteligente, ambientes inteligentes, robótica, transporte y cadena de suministro.

Energía inteligente. Este campo ha atraído una atención significativa de la comunidad de IoT en los últimos años. La mayoría de los sistemas de IoT propuestos aprovechan la cadena de bloques para preservar la privacidad de los usuarios junto con su información personal; y proteger el sistema de transacciones maliciosas, como los usuarios que intentan vender o comprar una cantidad irrazonable de energía. Se proponen sistemas de subasta donde los usuarios pueden vender al mejor postor su exceso de energía basada en una subasta definida en un contrato inteligente, eliminando así la necesidad de un moderador externo. También se exploró el uso de Blockchain para reconstruir los patrones actuales de transacciones de energía distribuida para permitir transacciones descentralizadas en tiempo real y contratos inteligentes de comercio de energía utilizando un mecanismo de confianza automático.

Ambientes inteligentes. Los ambientes inteligentes han sido estudiados para entornos industriales, para la asistencia sanitaria inteligente, ciudades y hogares inteligentes. En este contexto, la cadena de bloques se utiliza para garantizar la disponibilidad y la no confiabilidad de los datos detectados recopilados en la naturaleza, por ejemplo, una granja.

Robótica. El trabajo existente en esta área aprovecha la cadena de bloques como un sistema para soportar comunicaciones seguras y confiables de vehículos aéreos no tripulados (UAV: Unmanned Aerial Vehicle). Los UAV necesitan coordinar de manera confiable sus acciones, intercambiar datos y tomar decisiones. Se ha presentado un sistema donde los drones están programados para usar Blockchain para transmitir información de forma segura. Se está investigando el uso de Blockchain para proporcionar seguridad, autonomía y toma de decisiones colectivas en sistemas robóticos haciendo uso de una combinación de Blockchain y almacenamiento en la nube para proteger la integridad de los datos recopilados por los drones.

Transporte. En los últimos años, muchos conceptos de IoT se han utilizado para diseñar sistemas de transporte de próxima generación. El aspecto más prometedor es que los vehículos inteligentes probablemente no estarán tan limitados computacionalmente como otros dispositivos IoT, como las plataformas de sensores. Por lo tanto, Blockchain podría convertirse en un sistema para el intercambio de datos a prueba de manipulaciones entre vehículos inteligentes. Del mismo modo, supervise los datos relacionados con el vehículo, por ejemplo, información de mantenimiento e informes de diagnóstico del vehículo, utilizando Blockchain. Se podría usar Blockchain para diseñar una arquitectura de sistema de transporte inteligente completa, que incluya capas de aplicación, contrato, incentivo, consenso, datos, capa física y de red. La cadena de bloques también se ha aprovechado para implementar sistemas para manejar las claves públicas de los vehículos, y en general compartir datos sin una gestión centralizada de terceros. Se han propuesto sistemas de preservación de la privacidad para compartir información relevante, por ejemplo, accidentes, tráfico entre vehículos, donde los participantes son recompensados con tokens monetarios. También se propusieron sistemas de reputación basados en Blockchain que estima la confiabilidad de los mensajes recibidos.

Cadena de suministro y otros. Se han propuesto algunos sistemas para mejorar la funcionalidad de la fabricación basada en la nube y bajo demanda. Se han presentado marcos de distribución basados en Blockchain para compartir conocimientos y servicios entre empresas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como se ve en los puntos anteriores, una de las aplicaciones más importantes de Blockchain que se pueden usar en IoT son los Contratos Inteligentes.

Una de las principales ventajas de IoT es su capacidad para permitir comunicaciones autónomas y autodirigidas de máquina a máquina (M2M: Machine to Machine). En este contexto específico, es de suma importancia diseñar mecanismos de gestión de manera que las interacciones se inicien automáticamente de acuerdo a condiciones previamente acordadas; y que no haya necesidad de controlar y verificar individualmente la confiabilidad de cada interacción / comunicación. Con este fin, se ha demostrado que los contratos inteligentes [17] son efectivos para resolver los desafíos anteriores.

Si bien la aplicación de contratos inteligentes para IoT aún se está investigando, los resultados preliminares muestran que varias aplicaciones de IoT se beneficiarían de las tecnologías Blockchain como los contratos inteligentes. Por ejemplo, se han investigado los mecanismos de control de acceso que se basan en contratos inteligentes para regular el acceso a la red IoT. Estos trabajos aprovechan la inmutabilidad de Blockchain para generar una lista de control de acceso en tiempo real que también regula y describe las políticas de acceso a los recursos del dispositivo. Blockchain servirá como la columna vertebral de la confianza digital y la seguridad para las interacciones entre dispositivos.

También se discute la posibilidad de lograr un monitoreo inteligente de la cadena de suministro por medio de contratos inteligentes. No solo se pueden usar contratos inteligentes

para regular las transacciones y tarifas relacionadas con los procesos de producción y envío de mercancías, sino que también se pueden usar para realizar un seguimiento de su posición.

4. 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 una alumna que se encuentra promediando la carrera de Ingeniería en Informática.

Esto redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes investigadores, como así también sembrará las bases para la investigación a futuro, a través de la participación de alumnos de la Facultad de Ingeniería.

5. BIBLIOGRAFÍA.

- [1] A. Whitmore, A. Agarwal, and L. Da Xu. "The Internet of Things – A survey of topics and trends". *Information Systems Frontiers*, vol. 17, nro. 2, pp. 261–274. 2015.
- [2] Glen Martin (Forbes). "How the Internet of Things Is More Like the Industrial Revolution than the Digital Revolution". <https://www.forbes.com/sites/oreillymedia/2014/02/10/more-1876-than-1995/#674c4e0b66d2>. Última consulta: enero 2021.
- [3] L. Da Xu, W. He, and S. Li. "Internet of things in industries: A survey". *IEEE Transactions on industrial informatics*, vol. 10, nro. 4, pp. 2233–2243. 2014.
- [4] S. Tayeb, S. Latifi, and Y. Kim. "A survey on IoT communication and computation frameworks: An industrial perspective".

Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual. IEEE, pp. 1–6. 2017.

[5] F. Restuccia, S. D’Oro, and T. Melodia. “Securing the internet of things in the age of machine learning and software-defined networking”. *IEEE Internet of Things Journal*. 2018.

[6] M. S. Ali, K. Dolui, and F. Antonelli. “IoT data privacy via blockchains and IPFS”. *Proceedings of the Seventh International Conference on the Internet of Things*. ACM, p. 14. 2017.

[7] Julia Powles (The Guardian). “Internet of things: the greatest mass surveillance infrastructure ever?” <https://www.theguardian.com/technology/2015/jul/15/internet-of-things-mass-surveillance>. Ultima consulta: enero 2013.

[8] Dan Goodin, *Ars Technica*. “9 Baby Monitors Wide Open to Hacks that Expose Users’ Most Private Moments”. <http://tinyurl.com/ya7w43e9>. 2015.

[9] R. Amin, N. Kumar, G. Biswas, R. Iqbal, and V. Chang. “A lightweight authentication protocol for IoT-enabled devices in distributed Cloud Computing environment”. *Future Generation Computer Systems*, vol. 78, pp. 1005–1019. 2018.

[10] C. Koliass, G. Kambourakis, A. Stavrou, and J. Voas. “DDoS in the IoT: Mirai and other botnets”. *Computer*, vol. 50, nro. 7, pp. 80–84. 2017.

[11] K. Zhang, X. Liang, R. Lu, and X. Shen. “Sybil attacks and their defenses in the internet of things”. *IEEE Internet of Things Journal*, vol. 1, nro. 5, pp. 372–383. 2014.

[12] M. Gharbieh, H. ElSawy, A. Bader, and M.-S. Alouini. “Spatiotemporal stochastic modeling of iot enabled cellular networks: Scalability and stability analysis”. *IEEE Transactions on Communications*, vol. 65, nro. 8, pp. 3585–3600. 2017.

[13] M. Conoscenti, A. Vetro, and J. C. De Martin. “Blockchain for the internet of things: A systematic literature review”. *Computer*

Systems and Applications (AICCSA), IEEE/ACS 13th International Conference of. IEEE, 2016, pp. 1–6. 2016.

[14] R. C. Merkle. “Protocols for public key cryptosystems”. *Security and Privacy, IEEE Symposium*. IEEE, pp. 122–122. 1980.

[15] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang. “An overview of blockchain technology: Architecture, consensus, and future trends”. *2017 IEEE International Congress on Big Data (BigData Congress)*, pp. 557–564. June 2017.

[16] S. Nakamoto. “Bitcoin: A peer-to-peer electronic cash system”. 2008.

[17] K. Christidis and M. Devetsikiotis. “Blockchains and smart contracts for the internet of things”. *IEEE Access*, vol. 4, pp. 2292–2303. 2016.

Aplicación de redes neuronales profundas para la detección automática de Nombres de Dominio generados de manera algorítmica.

Carlos A. Catania, Jorge Guerra, Martin Marchetta, Gabriel Caffaratti, Lucia Cortez, Alfredo Rezinovsky, Franco Palau y Juan Manuel Romero.

Universidad Nacional de Cuyo, Facultad de Ingeniería, LABSIN
Campus Universitario, Mendoza, Argentina.

{harpo, martin.marchetta, gabriel.caffaratti, jorge.guerra, franco.palau, alfredo.rezinovsky, lucia.cortez,juan.romero}@ingenieria.uncuyo.edu.ar

RESUMEN

En el contexto de la seguridad de redes de datos, un nombre de dominio generado de manera algorítmica (DGA, de sus siglas en inglés) es utilizado por el software malicioso (malware) para generar de manera dinámica un gran número de nombres de dominios de manera pseudo aleatoria, y luego utilizar un subconjunto de estos como parte del canal de Comando y Control (C&C). El presente proyecto se enfoca en el desarrollo de algoritmos de detección de DGA mediante la utilización de algoritmos de aprendizaje de máquinas en general y las redes neuronales profundas en particular. Durante el último periodo del proyecto, se ha puesto especial énfasis en la puesta a punto de los modelos obtenidos con vista a su despliegue en ambientes de producción. En particular lo referido a la evaluación de los distintos aspectos necesarios para la estimación del error de generalización, más allá de la división aleatoria entre conjuntos de entrenamiento y prueba.

Palabras Claves: Seguridad de Redes, Detección de Anomalías, Aprendizaje Automático

CONTEXTO

El presente proyecto se desarrolla en el marco de Facultad de Ingeniería dentro Laboratorio de sistemas inteligentes (LABSIN). Este trabajo es parte del proyecto de investigación

que dio inicio en septiembre de 2019 en el marco de los proyectos bienales de secretaria de Investigación, Internacionales y Posgrados (SIIP) de la Universidad Nacional de Cuyo.

1 INTRODUCCIÓN

La detección temprana de secuencias DGA y su posterior bloqueo por parte de los administradores de sistemas resulta fundamental a fin de evitar o al menos mitigar la propagación de las acciones maliciosas dentro de la red. Es por ello que es de vital importancia desarrollar herramientas de detección, no solo con una baja tasa de falsos positivos sino también con la capacidad de operar en tiempo real. La utilización de técnicas de aprendizaje de máquinas surge como la alternativa más interesante para la construcción de una herramienta de detección de DGA. La metodología más común para la construcción de modelos de detección basados en técnicas de aprendizaje de máquinas consiste en utilizar un conjunto de datos conteniendo información sobre nombres de dominios normales y DGA. Este conjunto de datos es utilizado para entrenar un algoritmo que da como resultado un modelo de detección capaz de discriminar entre dominios normales y DGA. La principal ventaja de las técnicas de aprendizaje de máquinas es su capacidad de generalizar a casos similares nunca antes vistos. Lo que facilita considerablemente la tarea del administrador de redes, ya que este no tiene

que mantener actualizada la base de datos con nombres de dominios maliciosos. Sin embargo estas presentan algunos inconvenientes para su aplicación en escenarios reales. Entre los inconvenientes principales se destacan: (a) la tasa de falsos positivos, (b) el diseño del grupo correcto de atributos de entrada y (c) la necesidad de reentrenar periódicamente.

La capacidad de reconocer casos nunca antes vistos trae como consecuencia que algunos nombres de dominios normales, puedan ser detectados como maliciosos. Esto se conoce como falsos positivos. En algunos dominios de aplicación, el número de falsos positivos puede llegar a ser considerablemente alto sin perjuicio de la viabilidad en la aplicación de la técnica de aprendizaje de máquina elegida. Sin embargo, en el caso de la detección de DGA, un nombre de dominio bloqueado de manera incorrecta puede perjudicar seriamente la usabilidad del servicio de nombres de dominios (DNS). Basta que un usuario que no pueda acceder a un sitio con un nombre de dominio erróneamente clasificado como DGA para que genere inconvenientes en las tareas cotidianas de los administradores de sistemas. Es por esto último que mantener un número de falsos positivos muy bajo, resulta fundamental para realizar la detección en escenarios reales.

El número de falsos positivos está directamente relacionado no solo con el tipo de algoritmo de aprendizaje de máquina elegido, sino también con los atributos (variables predictoras) utilizados como entrada del algoritmo. Dentro del área de aprendizaje de máquinas, la construcción de los atributos de entrada adecuados al problema se denomina normalmente ingeniería de atributos. La ingeniería de atributos es una de las tareas que muchas veces demanda no solo los mayores recursos computacionales sino también humanos.

Por otro lado, los modelos de detección construidos a partir de técnicas de aprendizaje de máquinas requieren en muchos casos ajustes periódicos a fin de lidiar con casos significativamente diferentes respecto a

los vistos durante la construcción del modelo de detección. Este proceso normalmente implica la incorporación al conjunto de entrenamiento de los nuevos casos observados y la posterior re-ejecución del algoritmo de aprendizaje de máquinas. Lamentablemente, el tiempo de entrenamiento de algunas de las técnicas de aprendizaje puede resultar excesivamente alto para los requerimientos de una aplicación en un escenario real. La realidad es que si se quiere entrenar un modelo de detección en un tiempo adecuado se debe considerar una alta demanda de recursos computacionales. Sobre todo cuando el volumen de los datos del conjunto de entrenamiento comienza a ser significativo. Es por ello que hay que considerar aquellas técnicas que sean capaces de minimizar el tiempo requerido para el ajuste de los modelos.

En los últimos 10 años han sido desarrolladas técnicas de aprendizaje de máquinas conocidas bajo el nombre de Aprendizaje Profundo (DL). La utilización de estas técnicas ha sido la causa detrás de los mayores avances en el reconocimiento automático de imágenes, audio, video y análisis de texto. En particular las redes neuronales profundas ofrecen ventajas que permiten lidiar con los inconvenientes (a), (b) y (c) mencionados anteriormente. La principal ventaja radica en no tener que lidiar con el diseño del grupo correcto de atributos de entrada. A través de sus múltiples capas, las redes neuronales profundas son capaces de ir aprendiendo los atributos de entrada más adecuados para el problema. Además, esta selección automática de los atributos de entrada puede mejorar significativamente la eficiencia en términos de su tasa de detección de casos tanto positivos como negativos. Por último, recientes avances en la tecnología de procesamiento de los algoritmos de redes neuronales profundas han permitido disminuir considerablemente los tiempos de entrenamiento con grandes conjuntos de datos. En particular, la utilización de la capacidad de cómputo provista por las placas gráficas (GPU, del inglés Graphics Processing

Unit) ofrece una ventaja significativa frente al procesamiento en computadoras con CPU (del inglés Central Processing Unit).

El presente proyecto propone analizar la aplicación de redes neuronales profundas para el aprendizaje de los patrones comunes a los DGA de tal manera que permita desarrollar herramientas de detección no solo con una baja tasa de falsos positivos sino también con la capacidad de operar en tiempo real. Esto último resulta fundamental para lidiar con las amenazas de seguridad de hoy.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO.

El presente proyecto se enmarca en una línea de investigación que se viene desarrollando en el LABSIN desde 2017, la cual se centra en la aplicación de técnicas de aprendizaje automático a la seguridad informática.

Para el desarrollo del presente proyecto pueden diferenciarse 3 etapas principales:

A) Análisis preliminar del problema. Las tareas asociadas a esta etapa tienen por objetivo conocimiento de los problemas asociados a la detección de DGA como así también los modelos probabilísticos actualmente implementados. Una tarea fundamental consiste en realizar una breve revisión de la literatura sobre la aplicación de las técnicas de aprendizaje profundo a problemas de detección de DGA. Dicha tarea tiene por objetivo tratar de determinar cuáles han sido los beneficios e inconvenientes al aplicar este tipo de algoritmos.

B) Desarrollo de un algoritmo para la detección de DGA basado en técnicas de aprendizaje profundo. Esta etapa tiene por objetivo el desarrollo, evaluación y puesta a punto de un primer prototipo funcional para la detección de anomalías en el tráfico de red. Este primer prototipo es desarrollado utilizando alguna de las bibliotecas disponibles que permitan la implementación

de modelos basados en aprendizaje profundo de manera simple y eficiente. En esta etapa se definen los diferentes aspectos de la red como ser: el tipo de red a utilizar, la topología de la red y la secuencia de entrada entre otras. Luego se evalúan los resultados utilizando un conjunto de datos conteniendo tráfico etiquetado (DGA y normal).

C) Experimentación. Finalmente en la última etapa se centra en la evaluación del algoritmo propuesto sobre distintos conjuntos de datos: En particular, se evalúa el algoritmo propuesto con diferentes conjuntos de datos previamente etiquetados. Durante este proceso se consideran las métricas habituales en el área utilizando mecanismos para validar la generalidad del modelo obtenido como validación cruzada. Durante esta etapa se realizan también estudios comparativos con otros modelos de reconocimiento de anomalías de la literatura.

3 RESULTADOS OBTENIDOS

Durante los últimos 6 meses del proyecto se ha puesto el foco en la experimentación y evaluación de los modelos para detección de DGA (Etapa 3 del proyecto). En particular se han explorado distintas técnicas de evaluación con el fin de obtener una estimación de su capacidad de generalización. La estrategia común para evaluar la capacidad de generalización de un modelo consiste en separar el conjunto de datos disponible en un conjunto de entrenamiento y otro de prueba. Si bien esta separación aleatoria del conjunto de datos es una práctica habitual, podría no ser siempre el mejor enfoque para estimar la generalización del rendimiento en algunos escenarios. El hecho es que esta metodología habitual dentro del área de aprendizaje automático puede a veces sobreestimar el error de generalización cuando un conjunto de datos no es representativo o cuando los ejemplos raros y esquivos son un aspecto fundamental del problema de detección. Es por ello que durante los últimos 6 meses del

presente proyecto se puso especial énfasis en el estudio de diferentes estrategias de muestreo a fin de obtener una estimación del error de generalización de los modelos obtenidos.

Durante esta etapa del proyecto se realizaron las siguientes actividades:

- Recopilación de información bibliográfica sobre el tema poniendo especial énfasis en la aplicación de diferentes estrategias para muestreo de datos sobre la cual evaluar los diferentes modelos de detección.
- Construcción de diferentes conjuntos de datos de entrada aplicando las diferentes estrategias de muestreo (Ver Figura 1)
- Evaluación de los resultados de los diferentes modelos implementados sobre los distintos conjuntos de datos.
- Desarrollo de un informe técnico y presentación de los resultados preliminares en CICCASI 2020 [10].

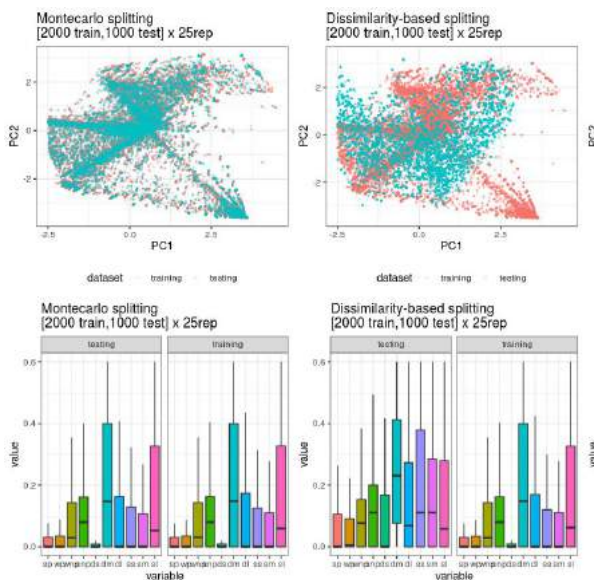


Figura 1: Dos de las técnicas de muestreo evaluadas. Montecarlo y Dissimilarity-based Splitting.

3 RESULTADOS ESPERADOS

Durante los últimos 6 meses del proyecto se espera:

- Continuar fortaleciendo la línea de investigación en la aplicación de aprendizaje de máquinas aplicados a la seguridad informática.
- Obtener una implementación funcional del modelo para la detección de DGA basado en redes neuronales con aprendizaje profundo. Sobre todo conocer los procedimientos necesarios para el despliegue de los modelos obtenidos en el contexto de una infraestructura de red de computadoras
- Incrementar la experiencia para la posterior aplicación de modelos probabilísticos basados en aprendizaje profundo a nuevas líneas de investigación relacionadas con problemas de ciencia y tecnología.

4 FORMACIÓN DE RECURSOS HUMANOS

El proyecto ha permitido capacitar en el ámbito de la investigación a profesores y alumnos interesados en participar en un entorno académico y tecnológico innovador.

Sobre la temática de este proyecto se está trabajando en:

- La tesis doctoral de Jorge Guerra, en el doctorado en Ciencias Informáticas de la Universidad Nacional del Centro de la provincia de Buenos Aires.

Además el desarrollo del proyecto a permitido desarrollar:

- La capacitación del Sr. Franco Palau, alumno de la carrera de Ingeniería en Mecatrónica de la Facultad de Ingeniería.
- La capacitación del Sr. Juan Manuel Romero, alumno de la carrera de Licenciatura en Ciencias de la Computación de la Facultad de Ingeniería.

5 BIBLIOGRAFÍA

- [1] Plohmann, D., Yakdan, K., Klatt, M., Bader, J., Gerhards-Padilla, E. (2016): A comprehensive measurement study of domain generating malware. In: 25th USENIX Security Symposium (USENIX Security 16). pp. 263–278. USENIX Association, Austin, TX.
- [2] Kühner, M., Rossow, C., Holz, T. (2014): Paint it black: Evaluating the effectiveness of malware blacklists. In: Stavrou, A., Bos, H., Portokalidis, G. (eds.) *Research in Attacks, Intrusions and Defenses*. Springer International Publishing.
- [3] Antonakakis, M., & Perdisci, R. (2012). From throw-away traffic to bots: detecting the rise of DGA-based malware. *Proceedings of the 21st USENIX Security Symposium*, 16. Retrieved from <https://www.usenix.org/system/files/conference/usenixsecurity12/sec12-final127.pdf>
- [4] Jason Reed, Adam J Aviv, Daniel Wagner, Andreas Haeberlen, Benjamin C Pierce, and Jonathan M Smith (2010). Differential privacy for collaborative security. In *Proceedings of the Third European Workshop on System Security*, pages 1–7. ACM.
- [5] Grill, M., Nikolaev, I., Valeros, V., & Rehak, M. (2015). Detecting DGA malware using NetFlow.
- [6] M. J. Erquiaga, C. Catania and S. García, "Detecting DGA malware traffic through behavioral models," 2016 IEEE Biennial Congress of Argentina (ARGENCON), Buenos Aires, Argentina, 2016, pp. 1-6.
- [7] Yadav, S., Reddy, A.K.K., Narasimha Reddy, A.L., Ranjan, S.: Detecting algorithmically generated domain-flux attacks with DNS traffic analysis. *IEEE/ACM Transactions on Networking* 20(5), 1663–1677 (2012)
- [8] Schiavoni, S., Maggi, F., Cavallaro, L., & Zanero, S. (2014). Phoenix: DGA-based botnet tracking and intelligence. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8550 LNCS, 192–211. https://doi.org/10.1007/978-3-319-08509-8_11.
- [9] Ahluwalia, A., Traore, I., Ganame, K., Agarwal, N.: Detecting broad length algorithmically generated domains. In: Traore, I., Woungang, I., Awad, A. (eds.) *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*. pp. 19–34. Springer International Publishing, Cham (2017)
- [10] Catania, Guerra, Romero, Caffaratti, Marchetta. Beyond Random Split for Assessing Statistical Model Performance. *Anales 3er Congreso Internacional de Ciencias de la Computación y Sistemas de Información CICCSI 2020*. (en prensa)

Avances en Aspectos de Seguridad en el Sistemas de Voto Electrónico OTP-Vote

Silvia Bast¹ Germán Montejano² MarioBerón²

¹Departamento 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

²Departamento 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

Los sistemas de voto electrónico no son ampliamente aceptados en la sociedad actual. Esta situación se debe principalmente a experiencias fallidas que han tenido lugar en elecciones recientes y conducen a que se acreciente la desconfianza de los electores.

En 2016 se presentaron las bases del sistema de votación denominado OTP Vote. El modelo usa múltiples claves One Time Pad que se combinan para formar una sola y son quienes dan el nombre al sistema. En este último tiempo se ha trabajado sobre el modelo original, con el objetivo de incrementar la confiabilidad e integridad del sistema en las diferentes etapas que incluye el proceso electoral. Las mejoras se orientan a: la configuración de los datos electorales y el proceso de generación y recuperación de votos, además de propuestas de auditoría y de verificabilidad end to end.

En este trabajo se exponen los avances que se llevaron a cabo para cada una de etapas del proceso.

Palabras clave: *Sistemas de Voto Electrónico, Anonimato, Transparencia,*

Auditoría, One Time Pad, Verificabilidad End to End.

CONTEXTO

El presente trabajo surge de una de las líneas de investigación del proyecto "Aspectos de Seguridad en Proyectos de Software", que avanza en el desarrollo de un modelo de voto electrónico basado en criptografía one time pad. El proyecto de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa del (Resolución N° 488/14 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales) es dirigido por el Doctor Germán Antonio Montejano (Universidad Nacional de San Luis) y codirigido por el Magister Pablo Marcelo García (FCEyN - UNLPam) e incluyó como investigadores a la Magister Silvia Gabriela Bast, al Magister Daniel Vidoret, al Analista Programador Adrián García y al Programador Superior Claudio Ponzio.

El proyecto se desprendió de la línea de Investigación "Ingeniería de Software y Defensa Cibernética", presentada en [1], que a su vez se enmarca en el Proyecto "Ingeniería de Software: Aspectos de alta sensibilidad en

el ejercicio de la Profesión de Ingeniero de Software” de la Facultad de Ciencias Físico - Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL) (<http://www.sel.unsl.edu.ar/pro/proyec/2012/index.html>) y que incluye acciones de cooperación con la Universidad Federal de Minas Gerais (UFMG, Brasil).

1. INTRODUCCIÓN

Opiniones contrapuestas que van más allá del ámbito académico, se presentan en la sociedad actual respecto del tema voto electrónico. Quienes se oponen dudan acerca de la transparencia de los datos y presentan como pruebas excluyentes las experiencias fallidas que se han llevado a cabo con sistemas que no verifican los requisitos mínimos. Quienes están a favor, afirman en cambio que produce importantes mejoras en cuanto a precisión y rapidez en la divulgación de los resultados.

La opinión del equipo de trabajo es que, debido a que se trata de sistemas críticos, la confianza del electorado es el aspecto clave a tener en cuenta para lograr su aceptación. Así también lo afirman McGaley y Gibson [2] “Un sistema de votación es tan bueno como el público piensa que es”.

Se asume también, que el sistema manual que se usa actualmente ofrece prestaciones aceptables. Sin embargo, si se tiene en cuenta que un amplio abanico de servicios sensibles y cotidianos se realizan virtualmente, se plantea el desafío de analizar y evaluar las condiciones de seguridad que deben cumplir los sistemas de voto electrónico y las soluciones propuestas por otros autores, con el objetivo de generar un modelo que permita el desarrollo de un sistema robusto y confiable.

En cuanto al concepto de Sistema de Voto Electrónico, Odrisek [3] afirma que es un componente de software que mapea electrónicamente el procedimiento de votación, para McGaley y Gibson [2] el voto electrónico es cualquier forma de recolección de votos que involucre dispositivos electrónicos (generalmente computadoras).

Epstein [4], Kazi, Alam y Tamura [5], Prince [6] y van de Graaf, Henrich y Müller-Quade [7], Hao, Ryan [8], Rivest [9], Ryan P., Schneider S., Teague V. [10], Rabin M y , Rivest R [11] y Awad M. y Leiss E [12], se explican sobre las características y los requisitos que deben cumplir estos sistemas y sobre antecedentes de los mismos.

En relación a la seguridad de los datos, los sistemas de votación electrónica deben proteger: el anonimato del elector por tiempo indefinido y los datos de los votos durante el proceso electoral, ya que luego la información se hace pública.

El Modelo OTP-Vote

El modelo OTP- Vote se describe en [13] se enfoca especialmente en la confidencialidad e integridad de los datos de un sistema de voto electrónico.

El modelo hace uso de los siguientes elementos de datos:

1. Claves One Time Pad (OTP), aleatorias y tan largas como el mensaje que cifran. Cumplen con las hipótesis y condiciones del “Secreto Perfecto” de Shannon [14].
2. Archivos de Datos que Almacenan Bits, son elementos centrales en el modelo propuesto y se van modificando durante el proceso.
 - Archivo Binario de Votos (ABV) surge en base al modelo de almacenamiento Múltiples Canales Dato Único (MCDU) propuesto por García en [15], que se analiza en profundidad en [16] y [17] y surge como una propuesta de resolución a las limitaciones de Birthday Paradox [18].
 - Clave de Descifrado (CD): se genera a partir de operaciones XOR (\oplus) [19] de claves OTP.
3. Tablas del modelo relacional: mantienen los datos básicos de la configuración de la elección: cargos,

candidatos e identificadores de votos y de los votos planos resultantes del proceso electoral.

El modelo propone:

- Anonimato incondicional.
- Seguridad computacional que puede llevarse a cualquier nivel exigible durante el proceso electoral.

El proceso incluye las etapas de:

- Configuración de la elección que incluye:
 - La definición de las dimensiones de ABV y CD, configuración de los atributos de la tupla.
 - Generación de los códigos para cada uno de los atributos identificadores (Cargos, Candidatos e Identificadores de Votos).
 - Generación de las tablas: Identificadores de Votos, Cargos, Candidatos, con los identificadores producidos anteriormente.
 - La inicialización del ABV y la CD con la intervención de las autoridades electorales (CA), mediante el aporte de claves.
- Desarrollo de la elección que involucra:
 - Autenticación: consiste en verificar que el elector se encuentre registrado en el padrón de votantes.
 - Emisión del voto: una vez que el usuario es habilitado, pasa a elegir el candidato de su preferencia para cada cargo y emite su voto.
- Cierre de la elección y recuento de votos que incluye:
 - Inicio del proceso de descifrado mediante la intervención de las autoridades electorales con sus claves.

- Obtención del archivo binario de votos descifrado.
- Generación de la tabla de votos planos.

El modelo teórico presentado supone, para cada una de las etapas mencionadas, el cumplimiento de algunas condiciones que resultan imprescindibles para alcanzar el normal funcionamiento del sistema.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se trabaja en el mejoramiento del modelo original con el objetivo de convertirlo en un sistema de voto electrónico robusto, para ello, es necesario profundizar en aspectos de seguridad, tales como:

1. Uso de atributos de control, de encriptación y variaciones en la configuración de los datos de los votos.
2. Análisis y refinamiento de protocolos antifraude.
3. Análisis y selección de un método criptográfico que asegure la transmisión de datos entre estaciones y servidor.
4. Diseño de un modelo para la automatización del proceso de configuración de parámetros y generación de tablas relacionales del sistema de e-voting.
5. Análisis de la información intermedia que puede ser expuesta a los auditores para su control.
6. Verificabilidad End to End.

El modelo requiere entonces de la especificación de los puntos mencionados, para demostrar que el sistema resultante de la investigación es confiable y seguro.

3. RESULTADOS Y OBJETIVOS

Los avances en la investigación están dados por:

- Mejoras introducidas en las tres etapas del proceso en cuanto al uso de atributos de control y de encriptación, variaciones en la configuración de los datos en cada proceso eleccionario.
- Desarrollo del diseño de un modelo para la automatización del proceso de configuración de parámetros y generación de tablas relacionales del sistema de e-voting, que involucran la primera y última etapas del modelo.
- Análisis de las propuestas que se llevaron a cabo acerca de la información para el control de auditoría
- Avances en la propuesta de Verificabilidad End to End.

Como trabajo futuro, debe focalizarse en los siguientes aspectos:

- Análisis y refinamiento de protocolos antifraude en todas las etapas del modelo.
- Análisis y selección de un método criptográfico que asegure la transmisión de datos entre estaciones y servidor.
- La información intermedia que puede ser expuesta a los auditores para su control en la etapa de Desarrollo de la elección.
- Evaluación y profundización de los avances ya realizados sobre el modelo original.

4. FORMACIÓN DE RECURSOS HUMANOS

En cuanto a la formación de recursos humanos de esta línea de trabajo:

- Silvia Bast completó el cursado de la totalidad de los créditos exigidos en el Doctorado en Ingeniería Informática en la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL). Actualmente se encuentra desarrollando en su tesis doctoral Silvia Bast.

5. BIBLIOGRAFÍA

[1] Uzal R., van de Graaf J., Montejano G., Riesco D., García P., “Inicio de la Línea de Investigación: Ingeniería de Software y Defensa Cibernética”, en Memorias del XV WICC. Ps 769-773. ISBN: 9789872817961. . 2013. <http://sedici.unlp.edu.ar/handle/10915/27537>

[2] M. McGaley, J. Gibson, “A critical analysis of the council of Europe recommendations on e-voting”. In EVT'06: Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop 2006 on Electronic Voting Technology Workshop. 2006.

[3] B. Odrisek, “E-Voting Security Study”, Communications- Electronics Security Group, X/8833/4600/6/21, (Copyright The Crown) Issue 1.2 31 United Kingdom, 2002

[4] J. Epstein, “Electronic Voting” in Computer, vol. 40, no. 8, pp. 92-95, Aug 2007. doi: 10.1109/MC.2007.271.

[5] 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.

[6] A. Prince, “Consideraciones, aportes y experiencias para el Voto electrónico en Argentina”, Editorial Dunken, 2006.

[7] J.van de Graaf, C. Henrich, J. Müller-Quade, “Requirements for secure voting”, Work Notes 2011.

[8] F. Hao, P. Ryan, “Real -World Electronic Voting. Design, Analysis and Deployment”. CRC Press. ISBN-13: 978- 1498714693. ISBN-10: 1498714692. 2017.

[9] R. Rivest, “On the notion of ‘software independence’ in voting systems”.

Philosophical Transactions of The Royal Society A, 366(1881):3759–3767. 2008.

[10] P. Ryan, S. Schneider, V. Teague, “End-to-End Verifiability in Voting Systems, from Theory to Practice”. Voting Systems, from Theory to Practice. IEEE Security & Privacy, 13(3):59–62, 2015.

[11] M. Rabin, R. Rivest, “Efficient End to End Verifiable Electronic Voting Employing Split Value Representations” Bregenz, Austria. Proceedings of EVOTE 2014. ISBN 978-9949-23-688-6. 2014.

[12] M. Awad, E. Leiss, “End-to-End Cryptography: Spreading Democracy”. International Journal of Applied Engineering Research. Volume 11, Issue 11. Ps. 7391-7394. 2016

[13] S. Bast, “Confidencialidad e Integridad de Datos en Sistemas de E-Voting – Un Modelo para la Implementación Segura de un sistema de Voto Presencial”, Editorial Académica Española. ISBN 978-3-639-53793-2. 2017.

[14] 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..

[15] P. García, “Una Optimización para el Protocolo Non Interactive Dining Cryptographers” - Editorial Académica Española (<https://www.eae-publishing.com/> - ISBN-13: 978-3-639-85270-7. ISBN-10: 3639852702. EAN: 9783639852707 – 2017.

[16] J. van de Graaf, G. Montejano, P. García, “Manejo de Colisiones en un Protocolo Non Interactive Dining Cryptographers”. Anales de las 42° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). Páginas 29 a 43. 2013 Disponible en:

<http://42jaiio.sadio.org.ar/proceedings/simposios/Trabajos/WSegI/03.pdf>.

[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.

[18] P. García, J. van de Graaf, A. Hevia, A. Viola, “Beating the Birthday Paradox in Dining Cryptographers Networks”. En “Progress in Cryptology – Latincrypt 2014”. Springer International Publishing. ISSN: 0302-9743. ISSN (electrónico): 1611-3349. ISBN: 978-3-319-16294-2. ISBN (eBook): 978-3-319-16295-9. Ps. 179 – 198. Octubre, 2014.

[19] M. Murdoch, V. Heuring, “Principles of Computer Architecture. Appendix A: Digital Logic”. Editor: Addison Wesley; Edición: US ed (29 de noviembre de 1999) Idioma: Inglés - ISBN-10: 0201436647 - ISBN-13: 978-0201436648

AVANCES EN ROBUSTECIMIENTO ANTE ATAQUES DE PRESENTACIÓN Y FALSIFICACIÓN PARA SISTEMAS BASADOS EN EL ANÁLISIS DE PATRONES DE TECLEO

Nahuel González¹ , Jorge Ierache¹ , Waldo Hasperué^{2,3} , Enrique P. Calot¹ 
, Hernán Merlino¹

Laboratorio de Sistemas de Información Avanzados, Departamento Computación
Facultad de Ingeniería, Universidad de Buenos Aires¹

Instituto de Investigación en Informática (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata²
Investigador Asociado - Comisión de Investigaciones Científicas (CIC)²

{ngonzalez, jierache, ecalot, hmerlino}@lsia.fi.uba.ar
whasperue@lidi.info.unlp.edu.ar

Resumen

El Laboratorio de Sistemas de Información Avanzados (LSIA) de la Facultad de Ingeniería de la Universidad de Buenos Aires (FIUBA) lleva adelante una línea de investigación sobre dinámica de tecleo, que se enmarca en un Proyecto de desarrollo Estratégico (PDE-44 UBA 2019-2020). Los aportes relevantes que se han publicado en la materia incluyen la formulación del método de modelado por contextos finitos, el análisis de la robustez de la técnica ante variaciones emocionales del usuario, la generación de conjuntos de datos para evaluación, y la transferencia de los resultados teóricos anteriores a la industria. El objeto de investigación actual se centra en los tipos de ataques a los que un usuario malintencionado puede someter a este tipo de sistemas, y la formulación de técnicas de defensa eficaces contra ellos.

En particular, se ha demostrado que las implementaciones actuales son vulnerables a los ataques de repetición, de presentación, y a la inyección de falsificaciones sintéticas. Aquí introducimos tanto una nueva modalidad sofisticada de ataque utilizando falsificaciones sintéticas como una defensa eficaz contra el mismo.

Palabras clave: *Seguridad Informática, Ataques de Presentación, Falsificación, Dinámica de tecleo, Contextos finitos.*

Contexto

El Laboratorio de Sistemas de Información Avanzados (LSIA) de la Facultad de Ingeniería de la Universidad de Buenos Aires fue creado en el año 2011 [LSIA, 2020]. Dentro de sus líneas

de investigación se cuenta el estudio de la dinámica de tecleo. Entre sus resultados publicados se cuentan el método de modelado por contextos finitos [González et. al., 2015], la generación de un extenso dataset realista y la replicación con él de experimentos previos destacados [González et. al., 2016], el análisis de la robustez de la técnica ante variaciones emocionales del usuario [Calot et. al., 2019], y la transferencia de los resultados anteriores a la industria, en el marco del Proyecto de desarrollo Estratégico PDE-44-2019, iniciado en el año 2019 [Ierache et. al., 2020]. El antedicho proyecto se basa en los resultados del proyecto UBACYT 20020130200140BA, titulado “Métodos de educación de cadencias de tecleo centrado en el contexto emocional de un individuo aplicando Interfaces Cerebro-Máquina (BMI)”.

Los avances aquí presentados tienen como objetivo fortalecer la robustez de los métodos antedichos frente al estado del arte en ataques contra sistemas de autenticación continua basados en cadencias de tecleo.

Introducción

La dinámica de tecleo (*keystroke dynamics*) es un subcampo de la biometría del comportamiento y la interacción hombre-máquina, que estudia cómo se pueden utilizar los ritmos de escritura para descubrir o verificar la identidad de los usuarios, tanto en el momento de iniciar sesión como de forma

continua durante el trabajo diario [Joyce & Gupta, 1990]. Multitud de técnicas, desde las más sencillas basadas en distancias hasta los más sofisticados clasificadores, han sido ensayadas para la verificación de identidad con dinámica de tecleo. Una comparación exhaustiva puede encontrarse en [Killourhy & Maxion, 2009]. El análisis de la dinámica de tecleo puede ser utilizado tanto para claves como para textos libres [Gunetti et. al., 2005].

Además de ser utilizada para autenticación, el análisis de dinámica de tecleo ha permitido correlacionar el estado emocional detectado por un clasificador con el reporte subjetivo de los usuarios [Epp et al., 2011], aunque alcanzando menor precisión. Otro trabajo realizado en contextos multimodales aplicando EEGs obtenidos por interfase cerebro-maquina registra el patrón de tecleo, bajo un enfoque dimensional; este fue desarrollado en el LSIA FIUBA en colaboración con el ISIER-UM. [Calot et. al., 2019].

En la última década se ha enfatizado la debilidad de los sistemas de autenticación basados en cadencia de tecleo [Rahman et. al., 2011] incluso ante ataques de gran simplicidad [Stefan et al, 2012] como la repetición exacta de una cadencia observada. Sin embargo, estos ataques presuponen acceso físico o lógico con privilegios elevados (suficientes para inyectar eventos de teclado) en los sistemas objetivo y un conocimiento muy detallado del patrón de escritura del usuario a falsificar. Una forma más

sofisticada de este ataque, pero con menos eficacia general, utiliza estadísticas generales para una población a los fines de suplir la ausencia de información de un usuario específico [Stanciu et. al., 2016].

En este artículo se enuncian los avances en materia de robustecimiento ante ataques de presentación y falsificación para sistemas basados en el análisis de patrones de tecleo. Aquí introducimos una nueva modalidad sofisticada de ataque utilizando falsificaciones sintéticas y una defensa eficaz contra el mismo, que asimismo ofrece adecuada protección contra los ataques que se han reseñado en el párrafo anterior.

Resultados y Objetivos

Con el objetivo de perseguir la línea de investigación presentada, se llevó a cabo el diseño y la implementación práctica de un sistema de autenticación continua donde se utiliza la dinámica de tecleo en texto libre. Se integró un modelo de contextos finitos con un detector de falsificaciones sintéticas para brindar protección contra ataques de presentación y repetición, y la implementación se evaluó con varios conjuntos de datos disponibles públicamente para evaluar la generalización de los resultados. Idénticamente, se evaluó en condiciones del mundo real fuera de un ambiente de laboratorio. De manera consistente para todos los conjuntos de datos y respaldando la generalización de estos

resultados, el error de clasificación se estabilizó asintóticamente a un valor entre 1% y 3%, con tasas de falsos negativos muy cercanas al 2% y tasas de falsos positivos que mostraron fluctuaciones más fuertes. Las figuras 1 y 2 muestran el detalle de tasas de error para la cantidad de teclas de entrenamiento en los datasets LSIA y KM.

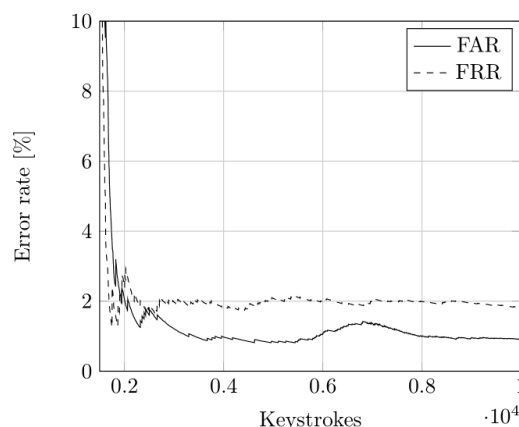


Figura 1. Tasas de error (FRR = falsos negativos, FAR = falsos positivos) de clasificación para el dataset LSIA.

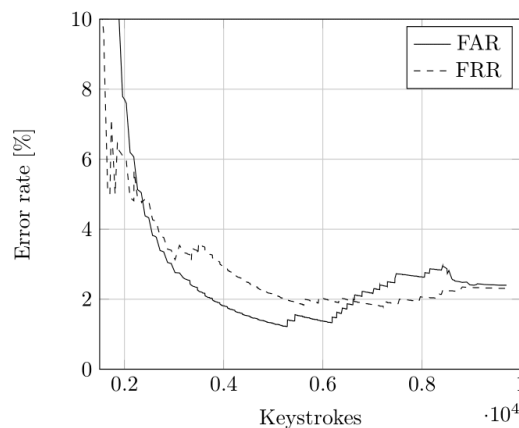


Figura 2. Tasas de error (FRR = falsos negativos, FAR = falsos positivos) de clasificación para el dataset KM.

El valor óptimo se alcanza con un entrenamiento de alrededor 4000

pulsaciones de teclas. Las tasas de error para el clasificador de falsificaciones sintéticas, que detecta ataques conocidos en sesiones individuales de aproximadamente 150 pulsaciones de teclas, se agruparon alrededor del 1% - 2%. Se observaron tasas de falsos negativos ligeramente más bajas que las tasas de falsos positivos, y se demostró que estas no se agregan a los errores del clasificador de usuarios legítimos debido a que ambos clasificadores actúan de manera complementaria.

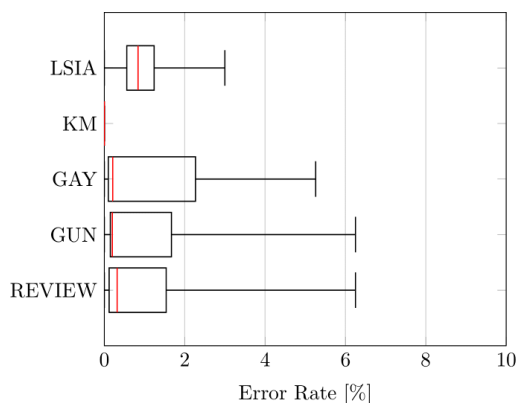


Figura 3. Distribución de falsos positivos utilizando falsificaciones sintéticas con un modelo de Markov de orden uno con desvío gaussiano, para cinco conjuntos de datos.

Para evaluar el rendimiento de la etapa de prevención de ataques, se evaluaron falsificaciones sintéticas generadas con un modelo de Markov de orden uno, utilizando desvío gaussiano, contra el clasificador de usuarios legítimos. Los resultados de rendimiento se muestran en la figura 3. Todos los conjuntos de datos muestran una tasa de error constantemente reducida, lo que demuestra la validez del supuesto. LSIA

es el que tiene el efecto menos pronunciado, probablemente debido a la dificultad de este conjunto de datos. El modelado por contextos finitos puede ser utilizado también para la reconstrucción de un vector patrón que haga las veces de falsificación sintética, si esta puede ser inyectada en el sistema a ser atacado. Las mejoras respecto de un modelo de Markov involucran la variación en el tamaño óptimo del contexto, la selección y ponderación de modelos de distintos órdenes, y la posibilidad de utilizar distribuciones empíricas [González et. al., 2015].

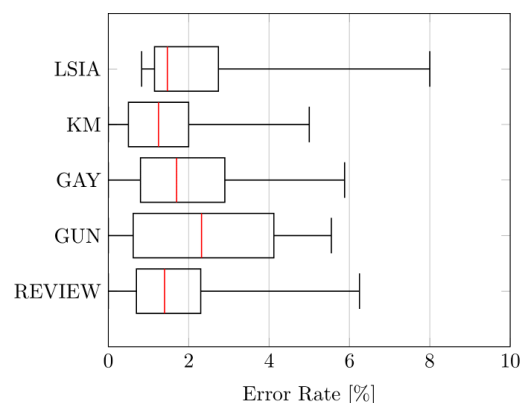


Figura 4. Distribución de falsos positivos utilizando falsificaciones sintéticas con un vector patrón generado con contextos finitos.

En la figura 4 se observa que, para todos los conjuntos de datos evaluados, la tasa de falsos positivos es mayor. La significación de este resultado es doble. En primer lugar, muestra que las falsificaciones sintéticas generadas con este método, más sofisticado, tienen una posibilidad más elevada de ser confundidas con el usuario legítimo. En segundo lugar, muestran que la etapa de

detección sigue siendo robusta aún frente a este tipo de ataques.

Formación de Recursos Humanos

En el marco de este proyecto se están desarrollando dos tesis doctorales. El equipo de investigación se conforma de tres investigadores formados y dos investigadores en formación.

Referencias

- CALOT, E., IERACHE, J., HASPERUÉ, W. ROBUSTNESS OF KEYSTROKE DYNAMICS IDENTIFICATION ALGORITHMS AGAINST BRAIN-WAVE VARIATIONS ASSOCIATED WITH EMOTIONAL VARIATIONS. EN ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING, PÁGINAS 194-211, SPRINGER, CHAM, 2019 .
- EPP, CLAYTON; LIPPOLD, MICHAEL; MANDRYK, REAGAN L. 2011. "IDENTIFYING EMOTIONAL STATES USING KEYSTROKE DYNAMICS", PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, ACM, PP. 715-724.
- GONZÁLEZ, N.; CALOT, E.; IERACHE, J.; FINITE CONTEXT MODELLING OF KEYSTROKE DYNAMICS IN FREE TEXT. EN 2015 INTERNATIONAL CONFERENCE OF THE BIOMETRICS SPECIAL INTEREST GROUP (BIOSIG), PÁGINAS 1-5, IEEE, 2015.
- GONZÁLEZ, N.; CALOT, E.; IERACHE, J.; A REPLICATION OF TWO FREE TEXT KEYSTROKE DYNAMICS EXPERIMENTS UNDER HARSHER CONDITIONS. EN 2016 INTERNATIONAL CONFERENCE OF THE BIOMETRICS SPECIAL INTEREST GROUP (BIOSIG), PÁGINAS 1-6, SEP. 2016. DOI: 10.1109/BIOSIG.2016.7736905
- GUNETTI, D.; PICARDI, C.; KEYSTROKE ANALYSIS OF FREE TEXT, EN ACM TRANSACTIONS ON INFORMATION AND SYSTEM SECURITY (TISSEC) 8.3, PÁGINAS. 312-347, 2005.
- IERACHE, J.; MERLINO, H.; CONCILIO, G.; CALOT E.; GONZÁLEZ, N. AVANCES EN RECONOCIMIENTO DE PATRONES DE TECLEO PARA LA IDENTIFICACIÓN DE PERSONAS EN AMBIENTES WEB, 2020. XXII WORKSHOP DE INVESTIGADORES EN CIENCIAS DE LA COMPUTACIÓN (WICC 2020), ISBN 978-987-3714-82-5, PP. 828-832.
- KILLOURHY, K. S.; MAXION, R. A. 2009. COMPARING ANOMALY-DETECTION ALGORITHMS FOR KEYSTROKE DYNAMICS. IN INTERNATIONAL CONFERENCE ON DEPENDABLE SYSTEMS & NETWORKS (DSN-09), PP. 125-134, ESTORIL, LISBON, PORTUGAL, 29 JUNE TO 02 JULY 2009. IEEE COMPUTER SOCIETY PRESS, LOS ALAMITOS, CALIFORNIA.
- JOYCE, R.; GUPTA, G. 1990. IDENTITY AUTHENTICATION BASED ON KEYSTROKE LATENCIES. COMMUN. ACM 33, 2 (FEBRUARY 1990), PÁGINAS 168-176. [HTTP://DOI.ACM.ORG/10.1145/75577.75582](http://doi.acm.org/10.1145/75577.75582)
- LSIA: [HTTP://LSIA.FI.UBA.AR](http://lsia.fi.uba.ar), VIGENTE AL 22 DE FEBRERO DE 2021
- RAHMAN, K.A., BALAGANI, K.S., PHOHA, V.V., 2011, June. Making impostor pass rates meaningless: A case of snoop-forge-replay attack on continuous cyber-behavioral verification with keystrokes. In *CVPR 2011 workshops* (pp. 31-38). IEEE.
- STANCIU, V.D., SPOLAOR, R., CONTI, M. and Giuffrida, C., 2016, March. On the effectiveness of sensor-enhanced keystroke dynamics against statistical attacks. In *proceedings of the sixth ACM conference on data and application security and privacy* (pp. 105-112).
- STEFAN, D., SHU, X., YAO, D.D., 2012. ROBUSTNESS OF KEYSTROKE-DYNAMICS BASED BIOMETRICS AGAINST SYNTHETIC FORGERIES. *COMPUTERS & SECURITY*, 31(1), PP.109-121.

Criptografía Liviana para Internet de las Cosas e Internet de las Cosas Industrial

Eterovic, Jorge; Cipriano, Marcelo; García, Edith; Torres, Luis.

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{jorge.eterovic; marcelo.cipriano; edith.garcia}@usal.edu.ar
torreslu@ar.ibm.com

RESUMEN

El proyecto tiene por objetivo general el estudio y análisis de algoritmos criptográficos livianos como así también, los protocolos en los que sean incluidos en el contexto de IoT e IIoT en la llamada Industria 4.0.

Particularmente persigue estudiar y analizar:

- algoritmos criptográficos livianos y protocolos en IoT e IIoT, observando fortalezas y debilidades matemáticas y criptográficas.
- mecanismos de intercambio de claves, autenticación, resumen (hash) y protocolos en dispositivos IoT e IIoT.
- test y pruebas de seguridad para ser aplicados a algoritmos criptográficos.
- nuevos estándares criptográficos en el área de IoT e IIoT.
- ataques criptoanalíticos convencionales o recientes y su incidencia sobre los algoritmos.

Otros objetivos del proyecto, enmarcados en el ámbito de las actividades de Difusión y Transferencia, son:

- Explicar y difundir la existencia de nuevos algoritmos criptográficos, como así también sus características de seguridad, su ámbito de aplicación, los criterios de diseño y seguridad dentro de su ámbito de aplicación
- Transferir a la comunidad científica nacional o internacional, docentes e ingenieros del ámbito IT (Tecnologías de la Información, por sus siglas en inglés) y OT (Tecnologías de la Operación, por sus siglas en inglés) la información y resultados obtenidos. En procura de lograr un nexo entre la investigación científico/académica y el mundo de la producción en el marco de la Industria 4.0. La amplia variedad de dispositivos IoT en general, sus perfiles de hardware y software,

como así también el tipo de aplicaciones, no existe al día de hoy una única primitiva criptográfica que pueda aplicarse por igual en todos ellos.

El estudio de cada una de las primitivas y algoritmos criptográficos propuestos por los diferentes autores resulta entonces, de gran importancia. Este es el objetivo que se persigue en el proyecto [9].

Palabras Clave:

Criptografía Ligera, Internet de las Cosas, Internet de las Cosas Industrial, IoT, IIoT.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad Nacional del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndolas como un servicio a la comunidad, 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, asociándose 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 con una duración de 2 años (2021-2023).

1. INTRODUCCIÓN

Desde el año 2015 un tercio de las calles de la Ciudad Autónoma de Buenos Aires cuenta con un Sistema de Telegestión que puede controlar individualmente las luminarias led que se están instalando, también conocidas como “luminarias inteligentes” [1]. Estos y otros dispositivos pertenecen a la llamada IoT, a un sub-campo conocido con el nombre de Smart Cities. Estos mecanismos ofrecen un uso racional de los recursos al poder configurarse para brindar servicio en determinados horarios o condiciones particulares. Ahorran energía y bajan los tiempos de mantenimiento. Son controlados remotamente y se conectan mediante una red de datos (inalámbrica usualmente). Se comunican entre sí y con el Centro de Control informando su estado de funcionamiento y su ubicación satelital exacta (mediante un receptor GPS), además de otros datos considerados relevantes.

1. Otros dispositivos como sensores de contaminación, ruido y tránsito, cámaras de seguridad con reconocimiento facial y lectura de patentes extienden las aplicaciones de IoT en nuestras ciudades, como Salta, Mendoza y Bahía Blanca [2] entre otras.
2. Estos y otros dispositivos pueden presentar vulnerabilidades susceptibles de ser explotadas. Es por ello que dotar a las comunicaciones de confidencialidad y/o autenticación mediante Criptografía Liviana o Ligera [3-4] ofrece solución a una parte del problema.
3. Este proyecto persigue el estudio y análisis de los algoritmos criptográficos que pueden ejecutarse sobre estos dispositivos IoT restringidos en recursos.

El crecimiento exponencial de la cantidad de dispositivos IoT que se ha observado en los últimos años¹ es un claro indicador de cómo esta tecnología está modificando nuestro mundo [5-6]. Sin embargo y de manera sorprendente la mayoría de los dispositivos IoT no presentan mecanismos de seguridad o los mismos son rudimentarios. Se

¹ No hay cifras exactas, pero se presume que en 2020 se habrán instalado en hogares, empresas, industrias y ciudades alrededor del mundo entre 30 y 50 mil millones de dispositivos IoT.

resentaría de esa manera la confidencialidad, autenticación, integridad y privacidad de la información que procesan y transmiten.

La comunidad científica ha expuesto este problema en forma extensa y ha propuesto diferentes medidas y soluciones para reducir su incidencia [7-9].

Una dificultad persistente en estos dispositivos IoT está en la naturaleza misma de su diseño e implementación. Para cumplir con sus funciones, fueron diseñados restringidos en recursos de energía, memoria, capacidad de cómputo, entre otros.

La Criptografía Liviana o Ligera estudia cómo ofrecer confidencialidad, integridad, y autenticación mediante primitivas criptográficas que empleen la menor cantidad de recursos, sin que por ello se resienta la seguridad y la privacidad de información.

Hasta aquí se han expuesto los argumentos acerca de los riesgos que estos dispositivos IoT conllevan para los usuarios que no contemplan mecanismos de protección adecuados.

Sin embargo, cabe aplicar una mirada de mayor amplitud frente a esta situación: se observa cómo los dispositivos IoT han aumentado la llamada “*Superficie de Ataque*”. Esta situación desnuda una vulnerabilidad de mayor escala al poner en riesgo infraestructuras de la red mundial y no a un usuario en particular, como si la vulnerabilidad individual de tal o cual dispositivo IoT pudiera “sumarse” a la de los demás equipos y escalar así hasta poner en riesgo activos críticos de objetivos impensados.

Tal escenario no es teórico. El ataque fue llevado a la práctica y ocurrió el 21 de Octubre de 2016. Y aunque no fue el primero de su tipo, sí fue el más grande registrado. Millones de dispositivos a lo largo del planeta -la mayoría de ellos de tecnología IoT- generaron un tráfico falso de alrededor de 620 Gbps apuntado contra un proveedor de servicio de DNS llamado Dyn. Este ataque Distribuido de Denegación de Servicio (DDoS) afectó un servicio esencial de la infraestructura misma de Internet.

Este exitoso ataque (que se repitió 3 veces a lo largo del día) dejó inaccesibles a plataformas y servicios de la talla de empresas como Amazon, BBC, CNN, Fox News, Github, HBO, Netflix, New York Times, PayPal, Spotify, Starbucks, Twitter, Visa, Wall Street Journal, entre otras, provocándoles pérdidas multimillonarias. Los equipos vulnerados (zombis) fueron explotados por medio de una BotNet y recibieron la orden de

realizar peticiones maliciosas al servicio DNS al que apuntaban.

La autenticación de usuarios y equipos, el uso de una VPN que permita constituir un túnel cifrando las comunicaciones desde y hacia estos dispositivos, como así también la correcta administración de contraseñas, hubieran podido evitar el ataque.

El mundo no puede permitirse conservar un nivel de riesgo semejante que pueda afectar globalmente a Internet, afectando las bases mismas de su funcionamiento.

La Criptografía Liviana ofrece una variedad de algoritmos de clave pública y clave privada, Block Ciphers [11-13] y Stream Ciphers [14-17] como así también algoritmos para la Gestión de Claves, Firma Digital y funciones Hash [18-20]. Lo mismo que puede ofrecer una criptografía convencional pero con la posibilidad de correr sobre los dispositivos Internet de las Cosas (IoT) y dispositivos Internet de las Cosas Industrial (IIoT) y sin sufrir por ello una pérdida de robustez y performance.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Ya se cuenta con una gran cantidad de algoritmos criptográficos que han sido creados para satisfacer las necesidades de IoT. Una parte de ellos fueron desarrollados por empresas que requerían dotar de seguridad dispositivos de su creación. Siguiendo con la línea de investigación propuesta para el proyecto se ha llevado adelante un relevamiento de tales algoritmos, en particular aquellos que se han propuesto para aplicaciones de la llamada Internet de las Cosas Industrial.

Otra línea de investigación busca analizar los protocolos de comunicaciones que se utilizan en los dispositivos IoT que empleen algún tipo de criptografía y mecanismo de seguridad. Se persigue la búsqueda de debilidades y vulnerabilidades de los mismos, a fin de poder realizar las propuestas correspondientes que minimicen o eliminen las mismas.

Por último, el equipo de investigadores sigue de cerca las vicisitudes y etapas por las que transcurre el concurso que en este momento mantiene el NIST [21] el cual persigue la búsqueda del nuevo estándar criptográfico de cifrado autenticado de Criptografía Liviana para dispositivos reducidos en recursos. Se han revisado los algoritmos más prometedores (de acuerdo al criterio vertido del propio NIST) y se está a la espera del algoritmo finalista. Una vez

acaecido este acontecimiento, se podrá analizar y estudiar en profundidad el algoritmo y sus propiedades criptográficas.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Se ha hallado que muchos de los algoritmos existentes para ser usados en dispositivos de la llamada Internet de las Cosas Industrial poseen vulnerabilidades que han permitido debilitar o romper la seguridad que ofrecían [22-23].

Es por ello que el continuo relevamiento, estudio, análisis y evaluación de los algoritmos criptográficos empleados en IoT y en IIoT es de suma importancia para que las fábricas, instalaciones industriales e infraestructuras críticas, preserven la seguridad en la realización de sus tareas y operaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas en la Facultad de Ingeniería, el área de la Seguridad Informática, de la Universidad del Salvador.

El año pasado se incorporaron al mismo dos docentes investigadores y algunos alumnos que se encuentran promediando la carrera de Ingeniería en Informática.

Se espera que en el presente año el equipo pueda crecer con la incorporación de más docentes investigadores y alumnos. Ya que redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes participantes, como así también sembrando las bases para la investigación del futuro, a través de la participación de alumnos de la Facultad de Ingeniería.

5. REFERENCIAS

- [1] Autor no informado. “Buenos Aires, una ciudad con iluminación inteligente”. Portal de información y trámites de la Ciudad Autónoma de Buenos Aires. Mayo, 2015. <https://www.buenosaires.gob.ar/noticias/buenos-aires-una-ciudad-con-luminacion-inteligente>
- [2] Mármol, H. “Cuáles son y qué hacen las ciudades argentinas que quieren parecerse a Japón” Portal del Diario Clarín, Septiembre 2019. <https://www.clarin.com/tecnologia/smart-cities->

hacen-ciudades-argentinas-quieren-parecerse-japon_0_i0n7KiJ5K.html.

- [3] ISO/IEC 29192. Information Technology - Security Techniques - Lightweight Cryptography. 2012.
- [4] Panasenko, S.; Smagin, S. "Lightweight Cryptography: Underlying Principles and Approaches". International Journal of Computer Theory and Engineering, Vol. 3, No. 4, August 2011.
- [5] Manyika, J.; Chui, M.; Bughin, J.; Dobbs, R.; Bisson, P.; Marrs, A. "Disruptive technologies: Advances that will transform life, business, and the global economy". McKinsey Global Institute. 2013.
- [6] Evans, D. "Internet of Things La próxima evolución de Internet lo está cambiando todo". Cisco IBSG. 2012.
- [7] Román R., Nájera P., López J. "Los Desafíos De Seguridad En La Internet De Los Objetos" University.
- [8] Fei Hu, Security and Privacy in Internet of Things (IoTs) : Models, Algorithms, and Implementations. Taylor & Francis Inc. Portland, United States. 2016.
- [9] Shancang Li , Lida Xu, Securing the Internet of Things. Syngress Media, U.S. Rockland, MA, United States. 2017.
- [10] Masanobu Katagi; Shiho Moriai, Lightweight Cryptography for the Internet of Things; Sony Corporation; 2016.
- [11] Satoh, A.; Morioka, S. "Hardware-Focused Performance Comparison for the Standard Block Ciphers AES, Camellia, and Triple-DES". Conference: Information Security, 6th International Conference, ISC 2003, Bristol, UK, October 1-3, 2003, Proceedings.
- [12] Beaulieu, R.; Shors, R.; Smith, J.; Treatman-Clark, S.; Weeks, B.; Wingers, L. "The SIMON and SPECK Families of Lightweight Block Ciphers." Cryptology EPrint Archive. International Association for Cryptologic Research, 19 June 2013.
- [13] Dworkin, M. "NIST SP 800-38B, Recommendation for Block Cipher Modes of Operation: The CMAC Mode for Authentication." NIST Computer Security Resource Center. National Institute of Standards and Technology, Spring (2005).
- [14] Daniel J. Bernstein. "The Salsa20 family of stream ciphers" . URL: <http://cr.yp.to/papers.html#salsafamily>. (2007).
- [15] Babbage, S.; Dodd, M. "The MICKEY stream ciphers". In New Stream Cipher Designs. Pp. 191-209. Springer Berlin Heidelberg. (2008).
- [16] Hell, M.; Johansson, T.; Meier, W. "Grain: a stream cipher for constrained environments". International Journal of Wireless and Mobile Computing, 2, pp. 86-93 (2007).
- [17] De Canniere, C.; Preneel, B. "Trivium. New Stream Cipher Designs (pp. 244-266). Springer Berlin Heidelberg. (2008).
- [18] Kavun, E. B., & Yalcin, T. "On the suitability of SHA-3 finalists for lightweight applications". The Third SHA-3 Candidate Conference. (2012).
- [19] Hirose, S., Ideguchi, K., Kuwakado, H., Owada, T., Preneel, B., & Yoshida, H. "A lightweight 256-bit hash function for hardware and low-end devices: Lesamnta-LW". International Conference on Information Security and Cryptology. Pp. 151-168. Springer Berlin Heidelberg (2010).
- [20] Guo, J.; Peyrin, T.; Poschmann, A. "The PHOTON family of lightweight hash functions". Advances in Cryptology—CRYPTO 2011 (pp. 222-239). Springer Berlin Heidelberg (2011).
- [21] <https://csrc.nist.gov/projects/lightweight-cryptography> (consultada el 13/03/20).
- [22] Eterovic, J.; Cipriano, M.; García, E.; Torres, L. Criptografía Ligera en Internet de las Cosas para la Industria. Congreso Argentino de Ciencias de la Computación. CACIC 2019. Libro de Actas. Pág. 1228-1240. UniRío. ISBN 978-987-688-377-1. 2019.
- [23] Eterovic, J. Cipriano, M. García, E. Torres, L. Lightweight Cryptography in IIoT The Internet of Things in the Industrial field. Computer Science Cacic 2019. Revised Selected Papers. Springer. ISBN 978-3-030-48324-1.

Criptografía Maliciosa y Ciberdefensa.

Cipriano, Marcelo^{1,2}; García, Edith¹, Maiorano, Ariel¹
Malvacio, Eduardo¹, Pazo Robles, María Eugenia¹

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Facultad de Ingeniería del Ejército (FIE), Universidad de la Defensa Nacional - UNDEF
²Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes UNQ.

{marcelocipriano; egarcia; maiorano; emalvacio; mepazorobles}@fie.undef.edu.ar

RESUMEN

El proyecto tiene por finalidad estudiar y analizar la aplicación de paradigmas y herramientas criptológicas modernas para la creación de software malicioso y puertas traseras, como así también indagar técnicas de prevención, detección y protección para ser consideradas en el ámbito de la Ciberdefensa Nacional.

Aunque comúnmente se entiende a la criptografía y a sus aplicaciones como herramientas de carácter defensivo, también pueden emplearse para usos ofensivos y maliciosos; a saber, el secuestro, extorsión y pérdida de información producidos mediante software malicioso denominado *ransomware*, en sus distintas variantes.

Por otra parte, la literatura también da cuenta de ataques en las etapas de diseño e implementación de algoritmos criptográficos, comúnmente llamados *backdoors* o puertas traseras, que pueden vulnerar la confidencialidad, integridad y disponibilidad. Es fácil observar las consecuencias directas sobre la ciberseguridad de usuarios particulares, empresas y organismos no gubernamentales de tales ataques.

Pero un aspecto de este tipo de ataques pasa desapercibido y sin embargo tiene un impacto mayor pues amenaza directamente a la población de una ciudad, provincia o llanamente, un país completo pues es capaz de afectar a sus organismos públicos, fuerzas de seguridad, estructura militar, política y

diplomática, como así también sus activos de información. Puede contribuir a la realización exitosa de ataques a las Infraestructuras Críticas, es decir, aquellas organizaciones relacionadas con la generación y distribución de energía, sistema financiero y bancario, organismos de salud como hospitales, servicio de potabilización y distribución de agua, saneamiento de desechos, entre otras.

Es decir, los ataques basados en puertas traseras o *backdoors* podrían menoscabar la ciberdefensa de una nación.

Palabras Clave

Criptología, Criptovirología, Kleptografía, Puertas Traseras Criptográficas. Ciberdefensa.

CONTEXTO

“MAC: Criptografía Maliciosa para la Ciberdefensa” es un proyecto perteneciente a la Facultad de Ingeniería del Ejército (FIE) “Gral. Div. Manuel N. Savio”, perteneciente a la Universidad de la Defensa Nacional (UNDEF).

Se encuentra enmarcado en el contexto de la carrera de grado de Ingeniería en Informática, la Especialización en Criptografía y Seguridad Teleinformática y, la Maestría en Ciberdefensa, que se dictan en la citada unidad académica.

Allí los investigadores conforman el Grupo de Investigación en Criptología y Seguridad Informática (GICSI) que

depende del Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática y lleva adelante tareas de I+D+i. El equipo está conformado por docentes investigadores categorizados en distintos regímenes científicos, profesionales técnicos, becarios y alumnos de la carrera de grado de Ingeniería en Informática, Especialización en Criptografía y Seguridad Teleinformática y de la Maestría en Ciberdefensa.

1. INTRODUCCIÓN

Comúnmente se aprecia a la criptografía y a sus aplicaciones como instrumentos de carácter defensivo que proporcionan confidencialidad en sistemas de comunicaciones, redes y bases de datos, entre otros.

En el año 1996 en un trabajo fundacional, *Adam Young* y *Moti Yung* [1] presentan lo que han dado en llamar *Criptovirología*. Los autores previeron y mostraron la posibilidad de llevar adelante ataques mediante virus informáticos, que cifran la información de sus víctimas a través de criptografía de clave pública, pidiendo luego rescate para su recuperación. En la actualidad este tipo de malware se conoce como *ransomware*.

Al año siguiente, los mismos autores presentan la llamada "*Kleptografía*": esto es el diseño e implementación de backdoors o puertas traseras en algoritmos criptográficos [2-4]. En particular los autores presentan el mecanismo criptográfico "*Secretly Embedded Trapdoor with Universal Protection*", conocido por sus siglas en inglés por el acrónimo de *SETUP*. Este *kleptograma* es una modificación a nivel matemático del algoritmo de intercambio de llaves *Diffie-Hellman*.

Con las modificaciones pertinentes, estas técnicas *kleptográficas* se podrían implementar en el corazón de otros algoritmos, podrían ser embebidos en

otros mecanismos criptográficos como son los esquemas de cifrado y de firma digital *ElGamal*, *DSA*, el algoritmo de firma de *Schnorr*, y el *PKCS* de *Menezes-Vanstone* y finalmente el reconocido algoritmo *RSA* [5-6, 8,13]. Es importante destacar que estos diseños no se limitan a criptografía de llave pública. Se cuenta en la literatura con publicaciones que describen ejemplos aplicados a funciones de hash, en donde se presentan colisiones para una versión de *SHA-1* modificando sus correspondientes parámetros [7], como también se presentan alternativas para protegerse de funciones de hash comprometidas en algoritmos de nivel superior, como *HMAC* y *HKDF* [14].

Por otro lado, tampoco los generadores de números de pseudo-aleatorios conocidos en la bibliografía como *Pseudo Random Numbers Generators* o *PRNG* por sus siglas en inglés. Estos algoritmos no serían inmunes a este tipo de ataques [10-13]. Básicamente la vulnerabilidad insertada afecta las propiedades estadísticas de un generador haciéndolo muy sensible a la entropía de la entrada. Por ejemplo, cuando los inputs tienen una distribución correcta, este mecanismo no tiene efecto, pero cuando están afectados por algún sesgo, el generador malicioso empeora considerablemente. En síntesis, la seguridad de los esquemas criptográficos se mide tradicionalmente como la incapacidad de un adversario, que cuenta con recursos limitados, de violar un objetivo de seguridad deseado [14]. Sin embargo, este argumento de seguridad generalmente se basa en un diseño sólido de los componentes subyacentes. Podría decirse que uno de los fracasos más devastadores de este enfoque se puede observar al considerar adversarios con la capacidad de influir en el diseño, implementación y estandarización de primitivas criptográficas.

Es entonces que considerando el impacto y la relevancia actual [16,17] de las técnicas y mecanismos mencionados se justificaría esta línea de investigación.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto está conformado por distintas líneas de acción:

- Estudio de material actualizado, asistencia a Cursos, Congresos y Workshops específicos, profundización en el estado del arte tanto de la criptovirología como de la kleptografía, aunque el esfuerzo principal estará dirigido a esta última.
- Estudio y análisis de las diferentes variantes de criptovirología.
- Estudio y análisis de ataques kleptografía en la literatura aplicados a diferentes algoritmos o primitivas criptográficas.
- Profundización en el estudio y análisis de técnicas kleptográficas para el algoritmo RSA y algoritmos de generación de números pseudo-aleatorios específicamente.
- Implementación experimental, conceptual y de referencia de alguna o algunas de las técnicas analizadas.
- Análisis y conclusiones de los resultados obtenidos.

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto propone estudiar y analizar la aplicación de los paradigmas y herramientas en la creación de software malicioso y puertas traseras criptográficas. En procura de desarrollar técnicas de prevención, detección y protección para ser considerados en el ámbito de la Ciberdefensa Nacional.

Por ejemplo someter a análisis a un algoritmo criptográfico o una parte de él, en procura de hallar indicios o técnicas de manipulación Kleptográficas, la inclusión de Backdoors Cryptography. Como así también el análisis del malware de tipo ransomware para el desarrollo de mecanismos de detección y prevención. De esta forma, procurar el desarrollo de herramientas de detección de vulnerabilidades criptográficas susceptibles de ser explotadas, afectando la ciberdefensa.

4. FORMACIÓN DE RECURSOS HUMANOS

Los investigadores que llevan adelante el proyecto dictan las asignaturas Criptografía y Seguridad Teleinformática, Matemática Discreta y Paradigmas de Programación. Desde esas cátedras se invita de forma permanente a los alumnos para participar como colaboradores.

Asimismo y por primera vez, varios alumnos de la especialización en Criptografía y Seguridad Teleinformática que están llevando adelante su Trabajo Final Integrador, como así también maestrandos pertenecientes a la Maestría en Ciberdefensa, que se encuentran trabajando en el desarrollo de sus respectivas tesis; y que además abordan temáticas relacionadas a la de este proyecto, han sido invitados a participar del mismo. Se espera que la contribución mutua entre el equipo de investigadores, especializandos y maestrandos permita alcanzar niveles sinérgicos de avance en la investigación, la formación de recursos humanos.

La Formación de Recursos Humanos permite incrementar el Know-How que tendrá el grupo de investigadores a lo largo de la vida del proyecto. Será un importante beneficio de sus integrantes

y de la institución en la cual desarrollan sus actividades científico-docentes. Por último y atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

- [1] Young, Adam L. and Moti Yung. "Cryptovirology: extortion-based security threats and countermeasures." Proceedings 1996 IEEE Symposium on Security and Privacy (1996): 129-140.
- [2] Young, Adam L. and Moti Yung. "The Prevalence of Kleptographic Attacks on Discrete-Log Based Cryptosystems." CRYPTO (1997).
- [3] Young, Adam L. and Moti Yung. "Kleptography: Using Cryptography Against Cryptography." EUROCRYPT (1997).
- [4] Young, Adam L. and Moti Yung. "Malicious cryptography - exposing cryptovirology." (2004).
- [5] Young, Adam L. and Moti Yung. "A Space Efficient Backdoor in RSA and Its Applications." Selected Areas in Cryptography (2005).
- [6] Young, Adam L. and Moti Yung. "An Elliptic Curve Backdoor Algorithm for RSASSA." Information Hiding (2006).
- [7] Albertini, Ange, Jean-Philippe Aumasson, Maria Eichlseder, Florian Mendel and Martin Schl affer. "Malicious Hashing: Eve's Variant of SHA-1." Selected Areas in Cryptography (2014).
- [8] Young, Adam L. and Moti Yung. "Cryptography as an Attack Technology: Proving the RSA/Factoring Kleptographic Attack." The New Codebreakers (2015).
- [9] Russell, Alexander, Qiang Tang, Moti Yung and Hong-Sheng Zhou. "Ciphertext Clipping: The Power of Kleptographic Attacks." ASIACRYPT (2015).
- [10] Indarjani, Santi. Sugeng, Kiki. Widjaja, Belawati. "Modification Attack Effects on PRNGs: Empirical Studies and Theoretical Proofs." (2015).
- [11] Young, Adam L. and Moti Yung. "Cryptovirology: the birth, neglect, and explosion of ransomware" Commun. ACM 60 (2017): 24-26.
- [12] Teseleanu, George. "Random Number Generators Can Be Fooled to Behave Badly." IACR Cryptology ePrint Archive (2018).
- [13] Markelova, A. V. "Vulnerability of RSA Algorithm." (2018).
- [14] Fischlin, Marc. Janson, Christian. Mazaheri, Sogol. "Backdoored Hash Functions: Immunizing HMAC and HKDF." (2018): 105-118.
- [15] Xiao, Dianyan and Yang Yu. "Klepto for Ring-LWE Encryption." Comput. J. 61 (2018): 1228-1239.
- [16] Yogi, Manas. Aparna, S.. "Novel insights into Cryptovirology A Comprehensive Study." International Journal of Computer Sciences and Engineering. 6. (2018): 1252-1255.
- [17] Zimba, Aaron. Chishimba, Mumbi. "On the Economic Impact of Crypto-ransomware Attacks: The State of the Art on Enterprise Systems." European Journal for Security Research. (2019).

Desarrollo de una DApp académica en la red Blockchain Federal Argentina

Jorge Eterovic; Jonatan Uran Acevedo; Alejandro Rusticcini; Nora Gigante

Programa PROINCE / Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903 (B1754JEC), San Justo, (5411) 4480-8900

eterovic@unlam.edu.ar; juran@unlam.edu.ar; arusticcini@unlam.edu.ar; ngigante@unlam.edu.ar

RESUMEN

Blockchain Federal Argentina (BFA) es la primera plataforma multiservicios sólida, transparente, confiable, abierta y participativa de Argentina, desarrollada para integrar servicios y aplicaciones sobre la Blockchain de Ethereum.

Este trabajo se desarrolló en el marco de un proyecto de investigación que consiste en implementar un nodo Minero (en adelante nodo Sellador) de la Universidad Nacional de La Matanza (UNLaM), dentro de la red Blockchain Federal Argentina.

La UNLaM ha firmado un contrato de colaboración público-privado con BFA, y actualmente forma parte de este consorcio. El mismo permitirá montar un nodo Sellador dentro de la infraestructura de la universidad que formará parte de la red de nodos de BFA.

Uno de los objetivos del proyecto de investigación es desarrollar e implementar una Aplicación Distribuida (DApp) que estará disponible para su uso libre y gratuito para toda la comunidad académica. Esta DApp está siendo desarrollada para las emisiones seguras de actas de examen, certificados de materias aprobadas y títulos académicos, entre otras.

La relevancia de este trabajo radica en la importancia que tiene para una institución académica formar parte de una red con las características de BFA que le permitirá emitir en

formato digital distintos tipos de documentos académicos de manera confiable.

Palabras clave:

Blockchain. DApp. Contrato Inteligente. Blockchain Federal Argentina.

CONTEXTO

Este proyecto de investigación fue presentado como un Programa de Incentivos a Docentes Investigadores de la Secretaría de Políticas Universitarias del Ministerio de Educación - PROINCE en el Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza.

El presente proyecto es del tipo “investigación aplicada” y consiste en el desarrollo e instalación de un nodo sellador en la UNLaM dentro de la red Blockchain Federal Argentina y de la implementación de una DApp para uso académico.

1. INTRODUCCIÓN

Blockchain se están integrando cada vez más en la sociedad. Su gran potencial y su tecnología revolucionaria se están abriendo cada día más hacia las personas. Su modelo seguro, descentralizado y público permite el funcionamiento independiente de autoridades bancarias u otras instituciones.

Una de las criptodivisas con mayor potencial que usan la tecnología blockchain es Ethereum. El objetivo de Ethereum es crear una blockchain programable que cambie el modelo de internet, donde los programas almacenados en su interior sirvan como columna vertebral a futuras aplicaciones descentralizadas (dApps). Estos programas se conocen como smart contracts.

Una Blockchain, contiene un conjunto de tecnologías informáticas y criptográficas que permiten crear una estructura de datos en forma de cadena de bloques cifrados y enlazados entre ellos que a su vez forman una base de datos distribuida y sincronizada.

Una Blockchain, dada su arquitectura y funcionamiento, permite almacenar información de forma verificable, que no puede ser modificada.

Una Blockchain no consiste en una única base de datos, ya que cada nodo tiene una copia de esta y están en constante sincronización. En lugar de confiar en una entidad central, la confianza en la integridad de esa base de datos se consigue mediante las interacciones de los participantes, es decir, el sistema está basado en una confianza descentralizada. Las plataformas que más utilizan este tipo de tecnología actualmente son las criptomonedas.

El ejemplo más famoso de una plataforma de criptomonedas que utiliza la cadena de bloques es Bitcoin, pero hay otras muy populares, como Ethereum, que utiliza su blockchain de manera distinta. La diferencia más significativa entre ambas es que Bitcoin pretende generar un sistema de economía digital y Ethereum, aparte de servir también como red de pago utilizando su propia moneda (Ether), es una plataforma cuyo propósito general es crear un blockchain programable.

Ethereum pretende crear una cadena de bloques donde los programadores puedan crear y subir código para que los participantes de la red puedan usarlo a través de los smart contracts. Estos contratos inteligentes son piezas de código

que viven dentro del blockchain y que pueden ser utilizados de forma autónoma por los usuarios.

La existencia de contratos inteligentes permite realizar tareas autoejecutables que responden a una condición pre-programada en el contrato, es decir, en un acuerdo registrado previamente entre partes donde se cumpla una condición existente en el contrato, se ejecutaría la cláusula correspondiente a esa condición.

Estos contratos tienen gran cantidad de usos en la industria y en la vida diaria, como por ejemplo en la banca, Internet de las Cosas (IoT), autoría, derecho, etc.

Las DApps rompen los esquemas tradicionales al eliminar la necesidad de intermediarios en los servicios que ofrecen, ya que permiten a los proveedores interactuar directamente con los usuarios finales, lo cual brinda mayor flexibilidad y satisfacción para todos [1].

Las DApps se construyen sobre una cadena de bloques determinada, que cuenta con su respectivo protocolo. Las aplicaciones descentralizadas están compuestas por uno o varios contratos inteligentes (smart contract) que operan en la cadena de bloques y una plataforma front-end, que puede ser un sitio web, una aplicación web o móvil, entre otras posibilidades.

La comunicación entre el smart contract y el front-end se realiza mediante una interfaz de programación de aplicaciones (API, por sus siglas en inglés). Parte del procedimiento al crear una DApp es agregarla a un directorio o biblioteca de aplicaciones [2].

Blockchain, en español cadena de bloques, es una tecnología que permite administrar un registro de datos en la nube. Tiene como característica la transparencia y es prácticamente incorruptible.

A pesar de que Bitcoin fue la primera aplicación descentralizada que nació con el propósito de crear una alternativa a los medios de pagos tradicionales, las características de su blockchain

no facilitan la creación de DApps. Años más tarde surge Ethereum como un proyecto que busca superar algunas de las dificultades de Bitcoin.

Ethereum es una plataforma descentralizada de código abierto (open source), que permite la creación de contratos inteligentes sobre una blockchain. En diciembre de 2013, Vitalik Buterin comenzó el desarrollo de Ethereum, con la primera prueba de concepto (PdC) [3].

El enfoque de Ethereum es contar con un mecanismo de desarrollo más eficiente en cuanto a tiempo, seguridad y escalabilidad. Ethereum funciona de manera descentralizada a través de una máquina virtual llamada Ethereum Virtual Machine (EVM). Esta máquina ejecuta un código intermedio o bytecode el cual es una mezcla de lenguaje de programación LISP, un ensamblador y bitcoin script [4].

Los programas en Ethereum se escriben en lenguajes de programación de alto nivel, como Solidity, que es un lenguaje de alto nivel orientado a objetos que permite a los nodos de Ethereum almacenar y procesar datos. Su sintaxis es similar a la de JavaScript y está enfocado específicamente a la EVM para crear los contratos inteligentes [5].

Un contrato inteligente es un programa informático que ejecuta un flujo de trabajo que generalmente representa acuerdos registrados en una Blockchain, entre dos o más partes, por ejemplo, personas u organizaciones [6]. Dichos contratos se ejecutarán como resultado de que se cumplan una serie de condiciones especificadas previamente.

Solidity es un lenguaje de programación orientado a objetos utilizado para escribir contratos inteligentes en la plataforma Ethereum. Fue desarrollado por Gavin Wood y otros programadores.

Es un lenguaje de scripting tipado estáticamente. Esto quiere decir que las variables deben ser declaradas junto con su tipo antes de ser utilizadas. Se debe realizar el proceso de

verificación y hacer cumplir las restricciones en tiempo de compilación, antes de que se ejecute el programa.

Cuenta con un IDE oficial llamado Remix. Un IDE (Integrated Development Environment, entorno de desarrollo integrado), es una aplicación que proporciona servicios para facilitarle al programador el desarrollo de software. Remix es un entorno de desarrollo, compilación y despliegue de contratos inteligentes basado en un navegador Web [7].

Una DApp es una aplicación distribuida sobre una Blockchain. Esta tiene múltiples capas y componentes y no depende de un sistema centralizado. Puede ser Web o Mobile. Una DApp es una aplicación que tiene su Back-end construido sobre contratos inteligentes, en contraposición con los Back-end tradicionales [8].

Finalmente, el desarrollo se hará en una plataforma multiservicios abierta y participativa pensada para integrar servicios y aplicaciones sobre la blockchain Ethereum, la Blockchain Federal Argentina [9]. Esta es una iniciativa confiable y completamente auditable que permite optimizar procesos y funciona como herramienta de empoderamiento para toda la comunidad.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el presente proyecto de investigación, se estudiaron y analizaron los derechos, obligaciones y posibilidades emanados de la firma del contrato de colaboración público-privada celebrado con Blockchain Federal Argentina.

Luego de firmado el acuerdo, se procederá a instalar el hardware necesario para montar el nodo sellador. Seguido a esto, se implementará el software para el correcto funcionamiento del nodo.

Asimismo, se desarrollará e implementará una Aplicación Distribuida en la Blockchain de

BFA. Esto se hará mediante el desarrollo de un Contrato Inteligente, el desarrollo de una API, la implementación de la DApp sobre la Blockchain Ethereum de BFA y por último el diseño, desarrollo e implementación de una aplicación Front-end para entorno Web, y Mobile.

Se escribirán y presentarán informes de avances que incluyan el progreso del proyecto y las conclusiones de cada una de las actividades que forman parte del mismo.

Finalmente, se redactará un informe integral final con el contrato y el software implementado y desarrollado acompañado de recomendaciones y buenas prácticas de uso como conclusión del trabajo de investigación realizado.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo principal de este proyecto de investigación es implementar un nodo Sellador dentro de Blockchain Federal Argentina (BFA).

El objetivo secundario es desarrollar e implementar una DApp (Aplicación Distribuida) perteneciente a la UNLaM. Dicha DApp funcionará sobre la Blockchain de BFA.

Una DApp es una aplicación distribuida. Ésta se desarrollará sobre la blockchain Ethereum. Tendrá múltiples capas y componentes y no dependerá de un sistema centralizado. Podrá ser usada desde un front-end Web o Mobile.

Una DApp es una aplicación que tiene su Back-end construido sobre contratos inteligentes, en contraposición con los Back-end tradicionales [10] [11].

Blockchain Federal Argentina es una plataforma multiservicios abierta y participativa pensada para integrar servicios y aplicaciones sobre blockchain. Una iniciativa confiable y completamente auditable que permite optimizar procesos y funciona como herramienta de empoderamiento para toda la comunidad.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está integrado por docentes-investigadores que pertenecen a distintas cátedras de la carrera de Ingeniería Informática y de la Tecnicatura de Aplicaciones Web de la UNLaM, alguno de los cuales está haciendo sus primeras experiencias en investigación.

Uno de los miembros del equipo de investigación se encuentra desarrollando su trabajo de tesis de posgrado de la Maestría en Ciberdefensa y Ciberseguridad de la Universidad de Buenos Aires y de la Escuela Nacional de Inteligencia (ENI). El tutor de esta tesis es el Mg. Jorge Eterovic, integrante de este proyecto de investigación.

5. BIBLIOGRAFÍA

- [1] Álvaro Santos García; Caracterización de Smart Contracts en Ethereum; Universidad Carlos III de Madrid; Leganés, España. 2019
- [2] Mohammad Dabbagh, Mehdi Sookhak, Nader Sohrabi Safa; The Evolution of Blockchain: A Bibliometric Study; IEEE Access PP (99):1-1. 2019
- [3] Vitalik Buterin; A Next Generation Smart Contract & Decentralized Application Platform; https://blockchainlab.com/pdf/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf. Última visita: diciembre de 2020.
- [4] Gavin Wood; Ethereum: A secure decentralized generalized transaction ledger; Ethereum project yellow paper. 2014.
- [5] Chris Dannen; Introducing Ethereum and Solidity; ISBN-13 (pbk): 978-1-4842-2534-9; Ed. Springer Science; New York, USA. 2017.
- [6] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, Aquinas Hobor; Making Smart Contracts Smarter; CCS '16: Proceedings of the

2016 ACM SIGSAC Conference on Computer and Communications Security; Pages 254–269. 2016.

[7] Susan Elliott Sim, Rosalva E. Gallardo-Valencia; Finding Source Code on the Web for Remix and Reuse; ISBN 978-1-4614-6595-9; Ed. Springer Science; New York, USA. 2013.

[8] Andrea Pinna, Simona Ibba, Gavina Baralla, Roberto Tonelli, Michele Marchesi, A Massive Analysis of Ethereum Smart Contracts. Empirical study and code metrics. DOI: 10.1109/ACCESS.2019.2921936. IEEE Access. 2019.

[9] Blockchain Federal Argentina; 2020; <https://www.bfa.org>. Última visita: febrero de 2021.

[10] Cai, Wei; Wang, Zehua; Ernst, Jason B.; Hong, Zhen; Feng, Chen; Leung, Victor C. M.; Decentralized Applications: The Blockchain-Empowered Software System. IEEE Access. 6: 53019–53033. DOI: 10.1109/ ACCESS.2018. 2870644. ISSN 2169-3536. 2018.

[11] Corbyn, Zoë. "Decentralization: the next big step for the world wide web". The Observer Internet. United Kingdom. <https://www.theguardian.com/technology/2018/sep/08/decentralisation-next-big-step-for-the-world-wide-web-dweb-data-internet-censorship-brewster-kahle>. 2018.

Desarrollo de una Guía para el abordaje de Incidentes de Ciberseguridad en Infraestructuras Críticas Industriales

Jorge Kamlofsky¹, Gerardo Gonzalez² y Santiago Trigo³

¹ CAETI – Universidad Abierta Interamericana
Av. Montes de Oca 725 – Buenos Aires – Argentina
Jorge.Kamlofsky@uai.edu.ar

² Universidad Nacional de la Defensa, Facultad de Ingeniería del Ejército Argentino.
Av. Cabildo 15, Buenos Aires, Argentina
Gerardo.Gonzalez@undef.fie.edu.ar

³ Facultad de Ingeniería, Universidad FASTA.
Gascón 3145, Mar del Plata, Argentina
SantiagoTrigo@ufasta.edu.ar

Resumen

La producción masiva de bienes se automatiza mediante los sistemas de control industrial. Estos sistemas son muy robustos y son aptos para funcionamiento continuo. Gracias a ello, estos sistemas automatizan también, plantas de potabilización de agua, producción y distribución de energía, siderúrgicas, y sistemas de semaforización, entre otros. Es decir, automatizan muchas infraestructuras críticas. Estos sistemas no se pensaron para ser seguros. Por ello, su seguridad se basó en el aislamiento físico.

Nuevas necesidades de mayor eficiencia y productividad requieren de la integración de estos sistemas con los sistemas de administración, con nuevas tecnologías como ser: big data, internet de las cosas, inteligencia artificial, entre otras, e incluso con Internet, dejándolos expuestos a una gran cantidad de riesgos y amenazas para los que no están preparados.

En este proyecto se estudian los riesgos de ciberseguridad a los que están expuestas las infraestructuras críticas industriales. Se plantea la elaboración de

una guía para mitigarlos y remediar los efectos frente a estos incidentes.

Palabras clave: seguridad en sistemas industriales, ciberseguridad en scada, seguridad de tecnologías operacionales, ciberdefensa en infraestructuras críticas.

Contexto

El proyecto presentado en este trabajo es un PDTS (Programa de Desarrollo Tecnológico-Social) aprobado por resolución del Poder Ejecutivo Nacional RS-2021-05630389. Se inició en 07/2020 y tiene una duración de tres años.

Los proyectos PDTS son considerados proyectos de interés o relevancia nacional, local o regional, y deben poseer instituciones financiadoras, adquirientes y demandantes [1].

Las instituciones que desarrollan y financian el proyecto son: la Facultad de Ingeniería de la Universidad Fasta (UFASTA), la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI) y la Facultad de Ingeniería del Ejército de la Universidad Nacional de la Defensa (FIE). Las instituciones adoptantes del proyecto son:

la empresa Trend Ingeniería, el Comando Conjunto de Ciberdefensa, la Dirección de Ciberdefensa del Ejército y la Facultad de Ingeniería del Ejército. Las instituciones demandantes son: La empresa Trend Ingeniería y la Dirección Nacional de Ciberseguridad.

Introducción

Desde mediados del siglo XVIII hasta la fecha las revoluciones industriales han producido una explosión demográfica. Creció fuertemente la esperanza de vida y se redujo notoriamente la pobreza. Estos cambios se lograron gracias al incremento de la disponibilidad de bienes y de alimentos, el incremento de la necesidad de mano de obra y mejoras permanentes en las condiciones sanitarias [2]. La primera revolución industrial se basó en la mecanización de la producción. La segunda (estimada desde 1870 a 1970), se caracterizó por el uso intensivo de energía (eléctrica y petróleo).

La tercer revolución industrial se basó en la incorporación de dispositivos electrónicos, informáticos y redes de comunicaciones para la automatización de la producción.

La automatización de la producción a gran escala se realiza con los ICS (del inglés: Industrial Control Systems). Los ICS consisten en sistemas de tele-mando y tele-control de procesos compuestos por autómatas industriales (según sus siglas en inglés): RTU (Remote Terminal Unit), PLC (Programmable Logic Controller), DCS (Distributed Control System) y/o PAC (Programmable Automation Controller) que pueden interconectarse [3]. A ellos se les conectan entradas y salidas, discretas y/o analógicas como ser: micro-switches, sensores de temperatura, actuadores para encendido de motores, llaves, etc. Poseen procesadores de pequeño porte. Su lógica es determinista,

lo cual favorece a la alta disponibilidad, esencial en el ambiente industrial [4]. Los ICS se supervisan y controlan en tiempo real desde sistemas informáticos llamados SCADA (del inglés: Supervisory Control and Data Acquisition).

Por tratarse de tecnología que impacta físicamente sobre la operación, se denomina OT (del inglés: Operation Technology). Gracias a su robustez, automatiza procesos que requieren uso continuo: plantas de producción de vestimenta, equipos, pero también: plantas alimenticias, de potabilización de agua, de producción y distribución de energía, transporte, siderúrgicas, entre otras. Es decir, está presente en las infraestructuras críticas de naciones. Este último sub-conjunto es el objeto de análisis de este proyecto: las Infraestructuras Críticas Industriales (ICI).

Enormes diferencias tecnológicas entre IT (del inglés: Information Technology) y OT y el aislamiento físico de los SCADA les dieron a los ICS una falsa sensación de seguridad por ocultamiento [5, 6]. Desde el ataque a la central nuclear de Irán de Natanz con el malware Stuxnet en 2010 [7], la comunidad internacional mostró gran preocupación: hasta ese momento, se pensaba imposible que un malware (IT) pudiera afectar la seguridad de las infraestructuras basadas en OT. Se trabaja en soluciones [8–10].

En IT se posee gran experiencia en seguridad. Y diferentes normas (entre ellas: [11,12]) basan la seguridad en tres pilares (en orden según su importancia): confidencialidad, integridad y disponibilidad. Pero en OT, el orden de prioridades es el opuesto. Quizás por ello, no es tan sencillo implementar las soluciones de IT en el mundo OT a pesar de existir normas y estándares de seguridad (entre ellos: [13,14]). Para mostrar ello, en [15] se plantea que una

gran parte de los incidentes de ciberseguridad en sistemas OT podrían haberse evitado si se hubiera implementado algún sistema de gestión de Seguridad Informática.

Hoy, desde hace una década estamos ingresando en la cuarta revolución industrial, caracterizada por la integración de los ICS con las nuevas tecnologías: inteligencia artificial, internet de las cosas, big data, realidad aumentada y con los sistemas corporativos (IT). A esta integración se la denomina Industria 4.0 [16–18]. Esta nueva necesidad de mayor integración tecnológica expone aún más a las ICI a amenazas y riesgos: un problema de seguridad en estas instalaciones puede significar el colapso de servicios vitales para la población.

En este proyecto se propone el desarrollo de una guía para el abordaje de incidentes de ciberseguridad en las ICI que atienda a la prevención de incidentes, su remediación y análisis forense.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se desarrolla íntegramente dentro de la línea de Seguridad Informática o Ciber-Seguridad. Pretende unir los esfuerzos realizados en el tema por cada uno de los grupos de investigación.

El desarrollo del proyecto se inició con presentaciones de lo realizado en el tema por cada uno de los grupos de investigación. También realizaron presentaciones tanto demandantes como adquirientes. Continúa con reuniones de investigación donde se presentan tanto lineamientos de trabajo para los equipos de investigación, como resultados parciales obtenidos. Cada grupo por separado desarrolla las tareas asignadas.

Se considera importante que en todas las reuniones en las que participan todos

los grupos de investigación, se invite tanto a demandantes como a adquirientes a conocer los pasos que los equipos de investigación realizan, de modo de corregir tempranamente algún desvío o acentuar el trabajo sobre un acierto. Además facilita el proceso de transferencias y capacitación hacia adquirientes y demandantes.

Se puede dividir al proyecto en las siguientes etapas: la evaluación de riesgos de seguridad informática en los sistemas de automatización industrial, la elaboración de recomendaciones para mitigar los riesgos en los sistemas de automatización industrial y sugerencias para la actuación para dar respuesta a incidentes y análisis forense en sistemas de automatización industrial.

Resultados y Objetivos

Este proyecto es desarrollado por tres grupos de investigación de distintas instituciones académicas. Cada grupo de investigación posee resultados previos en el tema:

El proyecto de la UAI se denomina “Ciberseguridad en los Sistemas de Control Industrial: Clave para la Ciberdefensa de las Infraestructuras Críticas” y está dirigido por el Lic. Jorge Kamlofsky desde 2015. Los resultados de este proyecto pueden consultarse en [19].

Por parte de la FIE, el proyecto de investigación denominado Infoscopia: es una interesante propuesta que trata el tema de estudio. Está dirigido por el CN Ing. Cesar Cicerchia. Más detalles del proyecto pueden obtenerse en [20].

Por el lado de UFASTA, el laboratorio Infolab posee varios resultados en temas de Ciberseguridad e Informática Forense. Se pueden consultar resultados en el sitio del Info-lab¹.

¹Sitio del Info-lab: <https://info-lab.org.ar>

La presentación de este proyecto generó la firma de convenios de colaboración en temas de investigación entre las entidades académicas, cuyo alcance se amplía más allá del PDTS y comienza a plasmarse en acciones concretas como ser: intercambio docente, interconsultas en investigación, dirección conjunta de tesis, entre otros.

El proyecto busca desarrollar una guía para el abordaje de incidentes de ciberseguridad en infraestructuras críticas industriales.

El proyecto tiene el objetivo principal de desarrollar una guía basada en tres aspectos básicos: los riesgos de seguridad en las infraestructuras críticas industriales, su mitigación y respuesta utilizando para ello el análisis forense.

Esta guía permitirá trabajar tanto ex ante (prevención) como ex post (actuación, remediación, análisis forense) en el abordaje de incidentes de ciberseguridad en infraestructuras que requieren una gestión de extrema seguridad, por su condición de criticidad para la propia organización y la población en general; especialmente, en instalaciones industriales del Estado o de empresas que brindan servicios esenciales (agua, energía, comunicaciones, combustibles, etc).

Formación de Recursos Humanos

El proyecto está dirigido por el Ing. Santiago Trigo. Los co-directores del proyecto son: el Esp. Lic. Jorge Kamlofsky por parte de la UAI y el Ing. Gerardo González por parte de la de la FIE.

Para la FIE, este proyecto se apuntala en la gran importancia que tiene la *concientización y formación* en ciberseguridad para todas las personas involucradas, en especial para los profesionales, en procesos industriales,

Tecnología de la operación e Internet de las cosas (IoT). Además, la aplicación directa en el escenario de Infraestructuras Críticas para todos los sectores esenciales del país y sobre todas las cosas en la reducción de riesgos cibernéticos y mejora en la seguridad en general de los procesos del mundo IT, OT e IoT. Para UFASTA, este proyecto le permite que su equipo de trabajo continúe con el desarrollo de servicios y productos sobre esta temática en proyectos futuros que den continuidad a la Línea de Investigación y Desarrollo en Ciberseguridad. Para la UAI, por su lado, este proyecto permite avanzar en las investigaciones en curso acerca de la ciberseguridad en entornos industriales. En particular, Jorge Kamlofsky se encuentra cursando un Doctorado y se espera que los resultados obtenidos del proyecto colaborarán con el desarrollo de su Tesis Doctoral.

El grupo de investigadores está conformado por: Bruno Constanzo, Hugo Curti, Juan Alberdi, Gonzalo Ruiz de Angeli y Leandro Ferrari por UFASTA; Enrique Belaustegui, Pedro Hecht, Claudio Milio y Oscar Romero por UAI y Pablo Croci, Matias Luzuriaga, Nicolás Díaz País, Rafael Olivieri, Juan Ignacio Raffo Triacca e Ignacio Omaechevarría por FIE. Cada uno de ellos son docentes de las instituciones mencionadas y adquirirán conocimientos específicos en el tema los cuales podrán ser transmitidos a los alumnos.

La participación de alumnos de las tres instituciones se considera de gran importancia y está prevista mediante el desarrollo de tesinas de grado, de especialización y tesis de maestría y doctorado, además está prevista la participación de alumnos en prácticas de laboratorio, y pasantías, entre otros.

Referencias

- [1] Argentina.gob.ar. "Banco Nacional de Proyectos de Desarrollo Tecnológico y Social", (2021). Disponible en: <https://www.argentina.gob.ar/ciencia/banco-pdts/criterios-de-los-proyectos>. [Consultado: 20/02/2020].
- [2] Montagut Contreras, Eduardo. "La transición demográfica en la Revolución Industrial", Los ojos de hipatía, ISSN: 2341-0612, (2017)
- [3] Miguel. "¿DCS, PLC, PAC o RTU?," Control Real Español, (2015). Disponible en: <https://controlreal.com/es/dcs-o-plc-o-pac-o-rtu/>. [Consultado: 8/03/2017].
- [4] Romero Mestre, H. "Ciberseguridad en sistemas de control industrial o ICSs." Trabajo Final de Master. Incibe, UOC, URB, Universitat Autònoma de Barcelona, (2018).
- [5] Courtois, N. "The dark side of security by obscurity, and Cloning MiFare Classic Rail and Building Passes Anywhere, Anytime." IACR Cryptology ePrint. 137, (2009).
- [6] Menezes, A., Van Oorschot, P., Vanstone, S. "Handbook of applied cryptography". CRC press, (1996).
- [7] Englert, M. "Cyber meets nuclear Stuxnet and the cyberattacks on Iranian centrifuges." Deutschen Physikalischen Gesellschaft, (2013).
- [8] Corvalan, F. "Seguridad de Infraestructuras Críticas: Visión desde la Ciberdefensa." III Conferencia Internacional y Taller de Ciberseguridad e Infraestructuras Críticas de Información, Buenos Aires, (2015).
- [9] Andreeva, O., Gordeychik, S., Gritsai, G., Kochetova, O., Potseluevskaya, E., Sidorov, S. and Timorin, A. Industrial control systems vulnerabilities statistics." Kaspersky Labs, (2016).
- [10] Sajid, N., Patel, S. and Patel, D. "Assessing and augmenting SCADA cybersecurity: A survey of techniques." Computers and Security 70, (2017): 436-454.
- [11] ISOTools. "ISO 27001." (2015). Disponible en: <https://www.isotools.org/normas/riesgos-y-seguridad/iso-27001> [Consultado: 20/02/2021].
- [12] NIST. "Special Publication 800 - 30, revision 1." Information Security. National Institute of Standards and Technology, U.S. Department of Commerce, (2012).
- [13] Stouffer, K., Lightman, S., Pillitteri, V., Abrams, M. and Hahn, A. "Guide to Industrial Control Systems (ICS) Security." NIST. Special Publication 800 / 82, revision 2. U.S. Department of Commerce, (2015).
- [14] ISA. "New ISA/IEC 62443 standard specifies security capabilities for control system components", (2018). Disponible en: <https://www.isa.org/intech-home/2018/september-october/departments/new-standard-specifies-security-capabilities-for-c>. [Consultado: 20/02/2021].
- [15] Kamlofsky, J., Colombo, H., Sliafertas, M. y Pedermera, 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).
- [16] Lasi, Heiner, et al. "Industry 4.0." *Business & information systems engineering* 6.4 (2014): 239-242.
- [17] Bartodziej, C.J. "The Concept Industry 4.0", BestMasters, DOI 10.1007/978-3-658-16502-4, (2017)
- [18] Xu, Li Da, Eric L. Xu, and Ling Li. "Industry 4.0: state of the art and future trends." *International Journal of Production Research* 56.8 (2018): 2941-2962.
- [19] Kamlofsky, Jorge, et al. "Ciberseguridad en los sistemas de control industrial: clave para la ciberdefensa de las infraestructuras críticas." XXI Workshop de Investigadores en Ciencias de la Computación WICC, (2019).
- [20] 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).

Detección de vulnerabilidades en especificaciones de contratos inteligentes de la plataforma Ethereum

Mauro C. Argañaraz⁽¹⁾, Mario M. Berón⁽¹⁾, Maria J. Varanda Pereira⁽²⁾, Pedro Rangel Henriques⁽³⁾ & Daniel Riesco⁽¹⁾

⁽¹⁾Departamento de Informática - Facultad de Ciencias Físicas Matemáticas y Naturales (FCFMyN) – Universidad Nacional de San Luis

⁽²⁾ Research Centre in Digitalization and Intelligent Robotics (CeDRI) - Instituto Politécnico de Bragança, Portugal

⁽³⁾Universidade do Minho - Braga, Portugal

marganaraz@gmail.com⁽¹⁾, {mberon, driesco}@unsl.edu.ar⁽¹⁾, mjoao@ipb.pt⁽²⁾, pedrorangelhenriques@gmail.com⁽³⁾

RESUMEN

Ethereum es el principal ecosistema basado en blockchain que proporciona un entorno para codificar y ejecutar contratos inteligentes. El proceso de escritura de contratos seguros y de buen desempeño es un gran desafío para los desarrolladores. Implica la aplicación de paradigmas de programación no convencionales debido a las características inherentes de la ejecución de programas de computación distribuida. Además, los errores en los contratos desplegados pueden tener graves consecuencias debido al acoplamiento inmediato del código del contrato y las transacciones financieras. El manejo directo de los activos significa que las fallas tienen más probabilidades de ser relevantes para la seguridad y tienen mayores consecuencias económicas que los errores en las aplicaciones típicas.

En este artículo, se describe una línea de investigación que se enfoca en diseñar una técnica activa de seguridad que permita comprobar el cumplimiento de una serie de principios, buenas prácticas, pautas y patrones de diseño establecidos en la especificación de contratos inteligentes de la plataforma Ethereum.

Palabras Claves: blockchain, Ethereum, smart contract, Solidity, Vyper, static analysis tool, verification, security patterns.

CONTEXTO

La presente línea de investigación se enmarca en el Proyecto de Investigación: *“Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de*

Sistemas en la Nube”. De la Facultad de Ciencias Físicas Matemáticas y Naturales, Universidad Nacional de San Luis. 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

A principios de la década de 1990 el criptógrafo estadounidense Nick Szabo acuñó el término contrato inteligente. Su planteo consistía en llevar las prácticas de la ley de contratos y las prácticas comerciales relacionadas hacia el diseño de protocolos de comercio electrónico entre extraños en internet. No tuvo éxito debido a las limitaciones tecnológicas de ese momento. Sin embargo, en 2008 se concibió Bitcoin y dio inicio a una nueva generación de protocolos, plataformas, aplicaciones y casos de uso que aprovechaban la descentralización y la computación distribuida para prescindir de los intermediarios tradicionales. A pesar de que Bitcoin sólo estaba pensado para transferir valor entre partes sin un vínculo de confianza, sentaba las bases de la tecnología blockchain.

Blockchain es una tecnología que se basa en una combinación de criptografía, redes y mecanismos de incentivos para respaldar la verificación, ejecución y registro de transacciones entre diferentes partes. En términos simples, las plataformas de blockchain se pueden ver como bases de datos descentralizadas que ofrecen propiedades muy atractivas, incluyendo la

inmutabilidad de las transacciones almacenadas y la creación de confianza entre los participantes sin la intervención de un tercero. Esto hace que esta arquitectura sea adecuada como un libro mayor abierto y distribuido que pueda almacenar transacciones entre partes de manera verificable y permanente.

Las criptomonedas fueron la primera aplicación de la tecnología blockchain y son una nueva forma de moneda virtual basada en una red distribuida de nodos anónimos. Esta estructura descentralizada les permite existir fuera del control de los gobiernos y las autoridades centrales. Las criptomonedas más conocidas y transadas son Bitcoin [1] y Ethereum [2] aunque en la actualidad existen cientos de plataformas blockchain con sus respectivas criptomonedas. Más allá de la transferencia de valor, algunas ofrecen otras propiedades como la privacidad o la ejecución de contratos inteligentes.

Los contratos inteligentes son esencialmente dos conceptos combinados. Uno es la noción de software. Código frío y austero que hace lo que está escrito y se ejecuta para que el mundo lo vea. La otra es la idea de un acuerdo entre las partes. Son programas informáticos que facilitan, verifican y hacen cumplir la negociación y ejecución de contratos legales. Se ejecutan a través de transacciones de blockchain, interactúan con las criptomonedas y tienen interfaces para manejar la información de los participantes del contrato. Cuando se ejecuta en la cadena de bloques, un contrato inteligente se convierte en una entidad autónoma que automáticamente lleva a cabo acciones específicas cuando se cumplen ciertas condiciones.

Ethereum es la principal plataforma de computación distribuida pública basada en blockchain que proporciona un entorno de ejecución de contratos inteligentes dentro del contexto de una máquina virtual descentralizada, conocida como Ethereum Virtual Machine (EVM) [2, 3].

La máquina virtual maneja el cómputo y el estado de los contratos y utiliza un lenguaje basado en una estructura de pila con un conjunto predefinido de instrucciones (códigos de operación) [3]. En esencia, un contrato es simplemente una serie de declaraciones de código de operación, que son ejecutadas secuencialmente por la máquina virtual. Ésta puede considerarse como una computadora global descentralizada en la que se ejecutan todos los contratos inteligentes. Aunque se comporta como una computadora gigante, es más bien una red de máquinas discretas más pequeñas en comunicación constante.

Los contratos inteligentes en Ethereum generalmente se escriben en lenguajes de alto nivel y luego se compilan en bytecode EVM. El lenguaje más prominente y ampliamente adoptado es Solidity [4], inclusive se utiliza en otras plataformas blockchain. Sin embargo, en las primeras etapas de la plataforma se desarrollaron otros lenguajes de alto nivel como LLL y Serpent, mientras que Vyper y Bamboo se encuentran actualmente en desarrollo por la fundación Ethereum.

Solidity es un lenguaje de programación de alto nivel orientado a contratos. Su sintaxis es similar a JavaScript, está tipado de manera estática y admite herencia y polimorfismo, así como bibliotecas y tipos complejos definidos por el usuario.

El proceso de escritura de contratos seguros y de buen desempeño en Ethereum es una tarea difícil para los desarrolladores. Implica la aplicación de paradigmas de programación no convencionales, debido a las características inherentes de la ejecución de programas basados en blockchain. Además, los errores en los contratos desplegados pueden tener graves consecuencias, debido al acoplamiento inmediato del código del contrato y las transacciones financieras. Por lo tanto, es beneficioso contar con una base sólida de diseño, patrones de código establecidos y

probados que faciliten el proceso de escritura de código funcional y libre de errores, y herramientas que permitan analizar y detectar falencias de manera automática.

Un análisis de los contratos inteligentes existentes realizado por Bartoletti y Pompianu [5] muestra que las plataformas Bitcoin y Ethereum se centran principalmente en los contratos financieros. En otras palabras, la mayor parte del código del programa define cómo se mueven los activos (dinero). Por lo tanto, es crucial que la ejecución del contrato se realice correctamente. El manejo directo de los activos significa que las fallas tienen más probabilidades de ser relevantes para la seguridad y tienen mayores consecuencias financieras que los errores en las aplicaciones típicas.

Los incidentes, como el desbordamiento de valor en Bitcoin, o el ataque al proyecto The DAO en Ethereum, causaron que una bifurcación dura de la blockchain anulara las transacciones. Estos incidentes muestran que los problemas de seguridad se han utilizado con propósitos fraudulentos. Muchas de estas vulnerabilidades se pueden clasificar en tres grupos: lenguajes de programación de alto nivel, máquina virtual (EVM) y las peculiaridades propias de la blockchain, y pueden abordarse siguiendo las mejores prácticas para escribir contratos inteligentes seguros, que se encuentran dispersos en toda la comunidad Ethereum. Estas prácticas contienen principalmente información sobre los errores típicos que deben evitarse y la descripción de los enfoques de diseño y problemas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En la búsqueda de asegurar los contratos inteligentes en la plataforma Ethereum se han adoptado diferentes enfoques, cubriendo aspectos como la semántica formal, patrones de seguridad y herramientas de verificación. De acuerdo

con el análisis abordado, se distingue entre verificación y diseño.

El objetivo de los enfoques de verificación es comprobar que los contratos inteligentes escritos en lenguajes de alto nivel existentes (como Solidity por mencionar alguno) o en bytecode EVM cumplan con una política o especificación de seguridad. Algunos ejemplos de investigaciones, técnicas y herramientas en esa dirección son:

- Herramientas de análisis estático para la búsqueda automática de bugs [6, 7, 8, 9]
- Herramientas de análisis estático para la verificación automática de propiedades genéricas [10, 11, 12, 13]
- Frameworks para pruebas semi-automatizadas de propiedades específicas del contrato [14, 15, 16, 17]
- Monitoreo dinámico de propiedades de seguridad predefinidas [18, 19]

En contraste, los enfoques de diseño apuntan a facilitar la creación de contratos inteligentes seguros al proporcionar frameworks para su desarrollo: abarcan nuevos lenguajes que son más susceptibles de verificación, proporcionan una semántica clara y sencilla que es entendible por los desarrolladores de contratos inteligentes o que permiten una codificación directa de las políticas de seguridad deseadas. Además, se incluyen trabajos que tienen como objetivo proporcionar patrones de diseño para contratos inteligentes seguros. Las propuestas vinculadas al diseño se pueden clasificar en:

- Lenguajes de alto nivel [20, 21, 22, 23]
- Lenguajes intermedios [24]
- Patrones de seguridad [25]
- Herramientas [26]

La línea de investigación presentada en este artículo busca plantear una estrategia para la detección de vulnerabilidades, nutriéndose de ambos

campos de estudio. Por un lado, utiliza un enfoque de verificación con las siguientes características:

- Lenguaje destino: lenguajes de alto nivel de la plataforma Ethereum (se selecciona Solidity como caso de estudio de mercado y Vyper como caso de estudio experimental)
- Método de análisis: estático
- Garantías: búsqueda de bugs
- Grado de automatización: verificación automatizada

Por el lado de los enfoques de diseño, se toman como base los patrones de seguridad para lenguajes existentes [25] y la idea de construir una representación intermedia para analizar los aspectos de seguridad.

3. RESULTADOS OBTENIDOS/ESPERADOS

El trabajo de investigación apunta a responder qué especificaciones tienen vulnerabilidades a través de la ejecución de análisis léxicos y sintácticos. El análisis estático garantiza una cobertura completa sin ejecutar el programa y lo suficientemente rápido en un código de tamaño razonable.

Para llevar adelante la propuesta, se creó el proyecto open source denominado OpenBalthazar. Este proyecto consiste en una herramienta web de análisis estático para los contratos inteligentes de la plataforma Ethereum implementada con Microsoft .NET Core.

Actualmente está implementado Solidity y se inició el desarrollo de las reglas de Vyper, si bien es una herramienta extensible y se pueden incorporar nuevos lenguajes como Bamboo a través de los mecanismos previstos.

En el proyecto se utiliza la librería ANTLR 4 y las gramáticas de Solidity y Vyper para generar el árbol de análisis (AST). Este árbol se puede enriquecer con información adicional utilizando algoritmos

y técnicas de procesamiento de lenguajes. Las reglas de verificación construidas utilizan un repositorio de patrones que definen los criterios en términos del árbol.

Los investigadores de esta línea continuarán con estudios en este campo con el objetivo de perfeccionarse en el área y realizar un seguimiento de nuevas amenazas, vulnerabilidades y ciberataques en materia de despliegue y ejecución de contratos inteligentes. También se planea seguir con los siguientes trabajos futuros:

- Generalización para otras plataformas de blockchain que soporten contratos inteligentes, tales como NEM, NEO, Cardano o Hyperledger.
- Obtención de código fuente a partir de bytecode EVM para aplicar el análisis que se propone en esta línea de investigación.
- Análisis dinámico de código fuente, lo cual implica correr el contrato y considerar sólo un conjunto de todas las ejecuciones posibles en base a datos de entrada arbitrarios.
- Análisis de los aspectos de seguridad que surjan de la interoperabilidad con otras plataformas blockchain.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de profesionales de la UNSL que forman parte de la línea de investigación de este trabajo llevan adelante diferentes trabajos finales integradores de *Ingeniería en Informática, Ingeniería en Computación, Licenciatura en Ciencias de la Computación*, y en un futuro próximo trabajos finales de especialización, tesis de maestría y doctorado. En particular, las investigaciones desarrolladas en este trabajo forman parte del lineamiento inicial como trabajo final de uno de los autores para optar al grado de Doctor en Ingeniería Informática en la UNSL.

5. BIBLIOGRAFÍA

- [1] Nakamoto, S., (2008), Bitcoin: Un Sistema de Efectivo Electrónico Usuario-a-Usuario, <http://bitcoin.org/bitcoin.pdf>
- [2] Ethereum White Paper: <https://github.com/ethereum/wiki/wiki/White-Paper>
- [3] Ethereum Yellow Paper: <https://ethereum.github.io/yellowpaper/paper.pdf>
- [4] Solidity - Solidity Documentation. Available: <https://media.readthedocs.org/pdf/solidity/develop/solidity.pdf>
- [5] Bartoletti, M. & Pompianu, L. (2017) *An empirical analysis of smart contracts: platforms, applications, and design patterns*. arXiv preprint arXiv:1703.06322.
- [6] Luu, L., Chu, D.H., Olickel, H., Saxena, P., Hobor, A.: Making smart contracts smarter. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, ACM (2016) 254–269
- [7] Grishchenko, I., Maffei, M., Schneidewind, C.: A semantic framework for the security analysis of ethereum smart contracts. In: Proceedings of the 7th International Conference on Principles of Security and Trust (POST), Springer (2018)
- [8] Zhou, E., Hua, S., Pi, B., Sun, J., Nomura, Y., Yamashita, K., Kurihara, H.: Security assurance for smart contract. In: New Technologies, Mobility and Security (NTMS), 2018 9th IFIP International Conference on, IEEE (2018) 1–5
- [9] Nikolic, I., Kolluri, A., Sergey, I., Saxena, P., Hobor, A.: Finding the greedy, prodigal, and suicidal contracts at scale. arXiv preprint arXiv:1802.06038 (2018)
- [10] Kalra, S., Goel, S., Dhawan, M., Sharma, S.: Zeus: Analyzing safety of smart contracts, NDSS (2018)
- [11] Buenzli, F., Dan, A., Drachsler-Cohen, D., Gervais, A., Tsankov, P., Vechev, M.: Securify (2017) Available at <http://securify.ch>.
- [12] Mythril Available at <https://github.com/ConsenSys/mythril>.
- [13] Manticore Available at <https://github.com/trailofbits/manticore>.
- [14] Hirai, Y.: Defining the ethereum virtual machine for interactive theorem provers. In: International Conference on Financial Cryptography and Data Security, Springer (2017) 520–535
- [15] Amani, S., B´egel, M., Bortin, M., Staples, M.: Towards verifying ethereum smart contract bytecode in isabelle/hol. CPP. ACM. To appear (2018)
- [16] Hildenbrandt, E., Saxena, M., Zhu, X., Rodrigues, N., Daian, P., Guth, D., Rosu, G.: Kevm: A complete semantics of the ethereum virtual machine. Technical report (2017)
- [17] Bhargavan, K., Delignat-Lavaud, A., Fournet, C., Gollamudi, A., Gonthier, G., Kobeissi, N., Kulatova, N., Rastogi, A., Sibut-Pinote, T., Swamy, N., et al.: Formal verification of smart contracts: Short paper. In: Proceedings of the 2016 ACM Workshop on Programming Languages and Analysis for Security, ACM (2016) 91–96
- [18] Grossman, S., Abraham, I., Golan-Gueta, G., Michalevsky, Y., Rinetzky, N., Sagiv, M., Zohar, Y.: Online detection of effectively callback free objects with applications to smart contracts. Proceedings of the ACM on Programming Languages 2(POPL) (2017) 48
- [19] Cook, T., Latham, A., Lee, J.H.: Dappguard: Active monitoring and defense for solidity smart contracts
- [20] OConnor, R.: Simplicity: A new language for blockchains. arXiv preprint arXiv:1711.03028 (2017)
- [21] Pettersson, J., Edström, R.: Safer smart contracts through type-driven development
- [22] Coblenz, M.: Obsidian: A safer blockchain programming language. In: Software Engineering Companion (ICSE-C), 2017 IEEE/ACM 39th International Conference on, IEEE (2017) 97–99
- [23] Schrans, F., Eisenbach, S., Drossopoulou, S.: Writing safe smart contracts in flint
- [24] Sergey, I., Kumar, A., Hobor, A.: Scilla: a smart contract intermediate-level language. arXiv preprint arXiv:1801.00687 (2018)
- [25] Wöhler, M., Zdun, U.: Smart contracts: Security patterns in the ethereum ecosystem and solidity. (2018)
- [26] Mavridou, A., Laszka, A.: Designing secure ethereum smart contracts: A finite state machine based approach. arXiv preprint arXiv:1711.09327 (2017)

Indicadores de ciberseguridad de una red informática en un Laboratorio EDI

Fabián A. Gibellini, Roberto Muñoz, Analía L. Ruhl, Cecilia Sanchez, Juliana Notreni, Milagros N. Zea Cárdenas, Ignacio Sánchez Balzaretti.

Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas de Información/
Universidad Tecnológica Nacional / Facultad Regional Córdoba
Cruz Roja Argentina S/N, 5016

fgibellini@bbs.frc.utn.edu.ar, robertmunioz@gmail.com, analialorenaruhl@gmail.com,
csanchezjuriol@hotmail.com, julinotreni@gmail.com, milyzc@gmail.com,
ignacio@bbs.frc.utn.edu.ar.

RESUMEN

El presente trabajo muestra los avances de la implementación del modelo de defensa de ciberseguridad en la infraestructura de red informática en un Laboratorio EDI.

Además, se presentan los riesgos relacionados a ciberseguridad y sus indicadores asociados. Dichos indicadores permitirán mitigar, contener o tomar acciones en un Laboratorio EDI.

Palabras claves: seguridad, redes, infraestructura, métricas, riesgo, indicadores.

CONTEXTO

Este estudio está inserto y forma parte de un proyecto denominado “Determinación de Indicadores, técnicas y herramientas que evidencian buenas prácticas en la ciberseguridad de la infraestructura tecnológica en un Laboratorio de Educación, Investigación y Desarrollo de la UTN - FRC”, homologado por la Secretaría de Ciencia y Tecnología bajo el código SIUTNCO0005366.

1. INTRODUCCIÓN

Un Laboratorio de Educación, Investigación y Desarrollo (EDI) es un laboratorio donde diversas actividades deben convivir e incluyen necesidades y exigencias del día a día de actividades académicas, estudiantes,

docentes, profesionales e investigadores. Estas exigencias son cada vez mayores en lo que concierne a software, aplicativos y hardware, más ahora, en momentos de pandemia que, el laboratorio bajo estudio, brinda servicios a usuarios.

ITIL v3 define la seguridad de información como la protección activa de información, tanto almacenada como transportada, para asegurar que la misma está disponible sólo a los usuarios autorizados en el momento en que ellos la requieren, con los niveles apropiados de integridad [1].

El laboratorio objeto de estudio es el Laboratorio de Ingeniería en Sistemas de Información (LabSis), de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC) y responde al concepto de un Laboratorio EDI.

LabSis considera que sus servicios deben ser seguros, por lo que basándose en la disponibilidad, confidencialidad e integridad de los datos genera y busca generar mecanismos y procesos que ayuden a dar más seguridad a los datos.

Entre los objetivos diarios que tiene que cumplir LabSis, se encuentran dar soporte académico a cátedras, soporte a la toma de exámenes y backup de los mismos [2], como así también dar soporte de infraestructura de red a proyectos de investigación, además de

generar nuevo software/configuración de software para suplir distintas tareas.

Según la ISO/IEC 270001 es esencial llevar a cabo ciertas acciones que den continuidad a los servicios, protegiendo la confidencialidad, integridad (autenticidad y no repudio de los datos) y disponibilidad de los mismos [3].

En cuanto a la ciberseguridad, Cisco [26] la define como la práctica de proteger sistemas, redes y programas de ataques digitales. Estos ataques son dirigidos generalmente para acceder, cambiar o destruir información sensible; extorsionar por dinero a usuarios; o interrumpir el proceso normal de negocio. Por lo que implementar medidas efectivas de ciberseguridad es un desafío particular hoy en día porque hay más dispositivos que personas, y los atacantes se han vuelto más innovadores [4].

En base a la cantidad de datos que maneja el Laboratorio EDI en cuestión, es crítico considerar la ciberseguridad de su red informática, en parte también por la diversidad de datos sensibles que manipula, teniendo en cuenta que los servicios que debe prestar esta entidad a docentes y la protección de datos sensibles (entre ellos, los parciales y exámenes finales) cuya incumbencia concierne únicamente a la Universidad en que se realiza [5] [6].

Sumado a ello, se tiene que tener en cuenta que los ataques no sólo provienen de atacantes externos, sino también puede ser de atacantes internos, los cuales puede interrumpir procesos cruciales referentes a la información, robar información, manipularla o incriminar al titular de la información en actos que éste desconoce [7].

Por lo que sin una estrategia de ciberseguridad planteada, no se podría medir ni controlar los incidentes relacionados a estos. Un estudio de IBM plantea que muchas organizaciones carecen de una clara y adecuada estrategia de seguridad alineada, tienen una visión limitada de su madurez de ciberseguridad y poseen prácticas insuficientemente para responder a un incidente de ciberseguridad [8].

En el presente trabajo se describen los riesgos relacionados a ciberseguridad y sus indicadores asociados. Este trabajo está relacionado con un modelo de defensa presentado durante el año 2019.

2. LÍNEAS DE INVESTIGACIÓN

La línea de investigación estudiada es la ciberseguridad en redes de información y específicamente aplicada a la infraestructura de red de un Laboratorio de Educación, Investigación y Desarrollo (Laboratorio EDI). En el desarrollo de este proyecto se utiliza el método empírico [9] [10], ya que los indicadores seleccionados, para su posterior recolección, fueron elegidos de forma empírica, teniendo en cuenta la historia del laboratorio y experiencias previas.

El estudio llevado a cabo en el LabSis, confluye sobre la seguridad informática en redes de información que operan sobre infraestructuras tecnológicas de un ámbito público, lo que podría aplicarse a otro ambiente de similares características.

Se llevó a cabo un análisis de riesgos [11], haciendo hincapié en los riesgos de ciberseguridad pero sin olvidar considerar otros tipos de riesgos. Luego se definieron indicadores o métricas [12] que nos permiten tener monitoreo y seguimiento de dichos riesgos

Los indicadores presentados en este trabajo son algunos del total identificado hasta este momento. Estos indicadores fortalecerán la seguridad de la red informática del Laboratorio EDI.

3. RESULTADOS OBTENIDOS/ESPERADOS

Para tener una visión global de todos los procesos/procedimientos de un Laboratorio EDI, procedimos a listar los procesos identificados en el LabSis:

1. Administración de aulas
 - a. Asignación anual/cuatrimstral de aula para cátedra
 - b. Asignación y preparación de aula para parcial o examen

- c. Asignación de aula para práctica libre
- d. Asignación de aula por única vez para cátedra
- 2. Administración de usuarios
- 3. Preparación de nuevo equipamiento
- 4. Administración de software
 - a. Mantenimiento de Software
 - b. Instalación de software para cátedras
- 5. Administración de bases de datos
 - a. Gestión de bases de datos para cátedras
 - b. Gestión de bases de datos del Labsis
- 6. Administración de backups
 - a. Gestión de backups del LabSis
 - b. Gestión de backups para cátedras
- 7. Mantenimiento de hardware
- 8. Mantenimiento preventivo
- 9. Control de inventarios
- 10. Desarrollo de sistemas para el Labsis
- 11. Transferencia al medio
- 12. Investigación
- 13. Tutorías

Para estos procesos se identificaron los riesgos asociados, a través de un análisis de riesgos. El total de riesgos identificados fue de 35, podemos observar un resumen de dicho análisis (Figura 1).

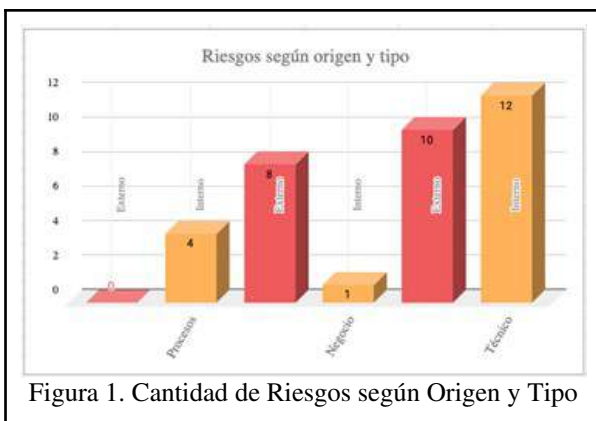


Figura 1. Cantidad de Riesgos según Origen y Tipo

Como podemos observar la mayoría de los riesgos identificados son técnicos. Por otro lado, todos los riesgos técnicos atentan contra la ciberseguridad de la red del LabSis ya que

son consideradas vulnerabilidades que tienen que ser controladas y subsanadas. Cada uno de los riesgos identificados tiene un nivel de exposición (Figura 2), cuyo criterio fue definido antes del análisis de riesgos.

Exposición	Es necesario tomar medidas de control?
Bajo	Riesgo entre 0.00 y 0.19. No es necesario tomar acción para abordar el riesgo
Tolerable	Riesgo entre 0.2 y 0.39. Evaluar si hay controles establecidos que hayan llevado el riesgo a tolerable y asegurar que se mantengan. No son necesarias medidas adicionales.
Significativo	Riesgo entre 0.4 y 0.6. Deben establecerse medidas de control para reducir el riesgo a Tolerable o Bajo. En caso que las medidas de control no sean inmediatas, se establecerán medidas transitorias.
Intolerable	Riesgo entre 0.61 y 1. La implementación de medidas de control debe ser inmediata. En caso no deberse comenzar a ser en subsiguientes hasta implementar las medidas de control.

Figura 2. Criterios de exposición de riesgos

La exposición de un riesgo es el producto entre las ponderaciones dadas al impacto y su frecuencia de ocurrencia. El impacto indica el grado de afectación que tendría el laboratorio si ocurre el riesgo, en donde:

- “0” significa que no afecta
- “1” significa que el impacto afecta completamente a un servicio.

En las Tabla 1 y Tabla 2 se pueden ver algunos riesgos técnicos con una exposición significativa o intolerable, es decir que si ocurren su impacto puede ser catastrófico. Por lo tanto si o si deben poseer medidas de control planificadas

ID	R.30
Amenaza	No realizar pruebas de restauración de los backups almacenados
Impacto	Posibilita la falla de backups a la hora de recuperarlos y pérdida de datos
Frecuencia	0,8
Impacto	0,5
Exposición	0,4
Medidas de	- Realizar capacitación

control	de restauración de cada uno de los backup. - Realizar un instructivo con paso a paso de como realizar la restauración y posibles problemas que se pueden presentar.
----------------	--

Tabla 1. Descripción de riesgo R.30

ID	R.34
Amenaza	No se cambian las contraseñas de los servidores periódicamente
Impacto	Posibilita la filtración de contraseñas que comprometan la seguridad de la infraestructura.
Frecuencia	0,9
Impacto	0,8
Exposición	0,72
Medidas de control	- Cambiar la contraseña cada vez que exista rotación de operadores.

Tabla 2. Descripción de riesgo R.34

Además, se identificaron:

- Riesgos que pueden ser mitigados con procedimientos.
- Riesgos que pueden ser mitigados mediante acciones administrativas que corrijan o mejoren la infraestructura.
- Riesgos que no se pueden mitigar ya que interfieren con el funcionamiento del Laboratorio EDI pero pueden ser monitoreados para reducir los posibles abusos.

- Riesgos que no pueden ser mitigados debido al costo o a la infraestructura y deben ser asumidos.

Como mencionamos antes, los riesgos técnicos son también riesgos de ciberseguridad, para estos riesgos se identificaron que permitan dar seguimiento a los riesgos, para poder tomar acciones preventivas o en el peor de los casos correctivas.

Los indicadores son valores muy disímiles entre sí ya que deben poder ser utilizados para evaluar la incidencia de los riesgos y los riesgos son por naturaleza distintos entre sí. Para esta selección de indicadores, se prioriza que sus valores se generen de forma automática y no dependan de una persona para confeccionarlos, es decir que los indicadores tengan una recolección automática.

Teniendo en cuenta todos los riesgos y principalmente los técnicos que atentan contra la ciberseguridad de la red del LabSis, un ejemplo de riesgo es el uso indebido de equipo, y su indicador asociado es el acceso a internet fuera de los horarios habilitados, para lo cual pueden generarse distintos eventos según sea el origen, el destino y el tipo de tráfico y en el caso en particular que se trata los eventos reflejan una cantidad en el tiempo por lo que pasado el tiempo de muestreo si el incidente persiste se generarán más eventos.

Los indicadores seleccionados se encuentran categorizados:

1. *Indicadores de Monitoreo de suministro eléctrico:* Permitirán autonomía eléctrica durante los cortes de luz
 - a. Cantidad de cortes
 - b. Duración de los cortes
2. *Monitoreo de las condiciones ambientales de los servidores:* Funcionamiento normal o anormal que requiere mantenimiento.
 - a. Temperatura ambiente
 - b. Temperatura de los servidores

3. *Indicadores de Monitoreo del tráfico de red:* Para detectar tráfico de red anómalo
 - a. Cantidad de conexiones /seg
 - b. Cantidad de tráfico /seg
 - c. Tráfico fuera de horario
 - d. Tráfico a puertos o destinos inusuales
4. *Indicadores de Monitoreo de red:* Para detectar equipos posiblemente no autorizados dentro de la red. También para detectar equipos posiblemente no autorizados dentro de la red que estén extrayendo o ingresando datos o aplicaciones no autorizadas.
 - a. Conexión de equipamiento de terceros
 - b. Suplantación de IP de un equipo
 - c. Detección de VMs con acceso directo a la red fuera de horario
5. *Indicadores de Monitoreo de equipos:*
 - a. Uso de disco: Es necesario asegurar el uso de disco para los backup, toma de exámenes y correcto funcionamiento de equipo.
 - b. Uso de memoria: Para asegurar el correcto funcionamiento de aplicaciones y VMs
 - c. Programas sospechosos: Para evitar posibles fugas o daños de datos.
 - d. Intento de elevación de privilegios: Para alertar de actividades sospechosas.
6. *Indicadores de Monitoreo de Usuarios:*
 - a. Intentos de session con clave incorrecta: Para detectar comportamiento anómalo o ataques.
 - b. Uso de usuarios distintos durante exámenes: Para detectar intentos de copia o uso indebido del sistema.

- c. Intentos de conexión remota: Para detectar intentos de copia o uso indebido del sistema

Por heterogeneidad se entiende que la unidad de medición de cada indicador es distinta y depende de lo que este cuantificando dicho indicador, podemos encontrar indicadores de cantidad en un tiempo dado, cantidad de conexiones por minuto, otros, tienen su unidad de medida del tipo absoluto como temperatura.

Además de los indicadores de recolección automática también existen indicadores de recolección manual, como por ejemplo el indicador asociado al riesgo “No se cambian las contraseñas de los servidores periódicamente” (R. 34) requiere un control manual para la verificación del cambio de contraseña para todos los involucrados.

Se determinó que es alta la complejidad de generar indicadores automáticos para los riesgos que se pueden mitigar con procedimientos.

Como conclusiones, respecto a los riesgos identificados en el análisis de riesgos, si bien su implementación fue en el Laboratorio de Sistemas, son extrapolables a cualquier Laboratorio EDI. Siempre considerando un análisis de revisión y/o adaptación previo.

Si bien los indicadores presentados tienen potencialidad de cubrir varios riesgos puede que algunos sean considerados rechazados o refutados durante la evaluación de los mismos, posterior a la etapa de recolección.

Dentro de los pasos a seguir están los de recolección de datos de los indicadores planteados lo que permitirá determinar su aceptación o rechazo como punto de referencia para mantener la red informática contenida o alerta ante cualquier anomalía o ataque.

4. FORMACIÓN DE RECURSOS

El grupo está compuesto por un Director, Co-Director, por investigadores de apoyo, profesores aspirantes a incorporarse a la carrera de investigador, técnicos de soporte, un estudiante investigador de la carrera de Ingeniería Electrónica y becarios que forman

parte del equipo. Este proyecto contribuirá a la formación y crecimiento de la carrera de los integrantes del mismo. En el caso de los estudiantes y algunos integrantes se iniciarán en la línea de seguridad informática.

Australia. Newtown Square, PA: Project Management Institute.

[12] Hauser, John & Katz, Gerry. (1998). Metrics: You are what you measure!.

European Management Journal. 16. 517-528. 10.1016/S0263-2373(98)00029-2.

5. BIBLIOGRAFÍA

- [1] ITIL. Information Technology Infrastructure Library. Vs. 3 - 2001.
- [2] Dharma, R., Sake, S., Manuel, M. (2013). Backup and Recovery in a SAN. Versión 1.2. EMC2 Techbooks
- [3] ISO/IEC 27001. “Tecnología de la información”. Técnicas de seguridad. Sistemas de gestión de la seguridad de la información (SGSI). Requisitos. ISO Ginebra, Suiza 2013.
- [4] What is cybersecurity? CISCO. Recuperado de <https://www.cisco.com/c/en/us/products/security/what-is-cybersecurity.html>.
- [5] (2008). UIT-T X. 1205. Serie X: Redes de Datos, Comunicaciones de Sistemas abiertos y seguridad. Ciberseguridad en el ciberespacio - Ciberseguridad. Aspectos generales de la ciberseguridad.
- [6] What is cybersecurity?. CISCO. Recuperado de <https://www.cisco.com/c/en/us/products/security/what-is-cybersecurity.html>.
- [7] Harmantzis, F., Malek, M. (2004). Security Risk Analysis and Evaluation, IEEE Communications Society, pp. 1897-1901
- [8] (2020) Strategies for managing cybersecurity risk: Assess and advance your security and compliance posture. IBM Security. IBM Global Services.
- [9] Bunge, M. (1998). La ciencia su Método y su Filosofía. Editorial Siglo Veinte. Buenos Aires
- [10] Barchini (2005). G. Métodos “I+D” de la Informática. Universidad Nacional de Santiago del Estero, Argentina.
- [11] Lavanya, N. & Malarvizhi, T. (2008). Risk analysis and management: a vital key to effective project management. Paper presented at PMI® Global Congress 2008—Asia Pacific, Sydney, New South Wales,

Investigación en ciberseguridad en un año de pandemia

Javier Díaz, Paula Venosa, Nicolás Macia, Einar Lanfranco, Alejandro Sabolansky, Mateo Durante, Damián Rubio, Jeremías Pretto

Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI).
Facultad de Informática. Universidad Nacional de La Plata
50 y 120 La Plata

{jdiaz, pvenosa, nmacia, einar, asabolansky, mdurante, drubio, jpretto}@linti.unlp.edu.ar

RESUMEN

El grupo de investigación en ciberseguridad de la UNLP forma parte del LINTI[1] y participa e impulsa en forma ininterrumpida desde el año 2000 distintos proyectos nacionales e internacionales en relación a la temática. Entre los autores de este artículo se encuentra el equipo de CERTUNLP[2], primer CSIRT académico de la Argentina creado en 2008.

En el presente artículo se describe el trabajo que el equipo lidera en relación a detección de vulnerabilidades, detección de ataques, mitigación de ataques y gestión de incidentes. Un pilar fundamental para el crecimiento del equipo es trabajar en proyectos vinculados a la formación de recursos humanos en estas temáticas, es por ello que también se incluye la experiencia de este último año en relación a la organización de CTFs los cuales propician un ámbito de formación para la comunidad, de una forma innovadora y motivadora.

Palabras clave: ciberseguridad, monitoreo inteligente, threat intelligence, gestión de incidentes, CTF.

CONTEXTO

La línea de investigación "Ciberseguridad" presentada en este trabajo, se desarrolla en el marco del proyecto de investigación "De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región"[3] del Programa Nacional de

Incentivos. Este proyecto está acreditado por la UNLP y financiado por partidas del presupuesto nacional. De dicha línea participan docentes investigadores del LINTI de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP).

1. INTRODUCCIÓN

La detección de malware y ataques mediante el análisis de tráfico de red continúa siendo un desafío para los responsables del monitoreo de seguridad de una red de datos y de la gestión de los incidentes de seguridad[4]. Aunque existen varios mecanismos de detección bien conocidos para diferenciar con precisión los comportamientos maliciosos de los normales, todavía es extremadamente difícil contar con sistemas de detección eficientes.

Desde hace algunos años, los sistemas de detección de intrusiones han incorporado paradigmas inteligentes como las técnicas de aprendizaje automático. Hoy en día existen también algunas propuestas para implementar algoritmos de Ensemble Learning[5][6], a fin de combinar múltiples clasificadores para lograr una mejor precisión en la detección.

Además de los mecanismos y herramientas, para implementar una defensa eficaz, las organizaciones necesitan contar con información sobre los posibles atacantes, como sus técnicas, tácticas y procedimientos. Esta metodología, denominada inteligencia de amenazas, ayuda a las organizaciones a

comprender mejor su perfil de amenazas. Las fuentes de inteligencia permiten obtener indicadores que luego podrían ser utilizados por dispositivos como firewalls o sistemas de detección de intrusos para disponer de una reacción oportuna a las amenazas emergentes. Si se logra combinar la información de inteligencia con los mecanismos de detección de malware y detección de tráfico anómalo, permitiría gestionar los incidentes de seguridad de manera integral mejorando la mitigación de los ataques a los que estamos expuestos.

La gestión de incidentes de seguridad no debe ser un proceso separado de los procesos de monitoreo de seguridad y detección de incidentes a partir del uso de fuentes de información o feeds. Es por esto, que los sistemas de detección de incidentes utilizados, deben poder interactuar de manera fluida con otros procesos para poder automatizar dicha gestión, lo máximo posible.

En búsqueda de una evolución constante, y para no depender exclusivamente de fuentes externas de información para la gestión de incidentes, se requiere implementar herramientas de relevamiento continuo de información, que permitan identificar las fortalezas y debilidades de los servicios prestados a Internet por nuestra comunidad objetivo, manteniendo un historial que permita realizar consultas históricas y poder implementar análisis evolutivos.

Para poder llevar a cabo todas estas actividades, es necesaria la formación continua de recursos humanos tanto para el grupo de investigación como para la sociedad en general. En esta línea, es necesario generar marcos de capacitación formal a través del dictado de materias y cursos en el ámbito académico además de fomentar la formación a partir de un contexto lúdico a través del desarrollo de eventos de tipo CTF.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En la actualidad, las principales líneas de trabajo en las que el grupo de investigación en ciberseguridad desarrolla sus actividades y consolida su formación son las siguientes:

- Detección y análisis de vulnerabilidades en distintos tipos de dispositivos, protocolos y tecnologías.
- Desarrollo de herramientas propias para la automatización de los procesos de gestión de incidentes de seguridad y escaneo de vulnerabilidades. Integración con otras existentes.
- Pentesting de redes, aplicaciones y servicios.
- Monitoreo de seguridad de red. Uso de herramientas de Machine Learning para un monitoreo inteligente.
- OSINT. Uso de software libre para el uso, el procesamiento y la correlación de distintas fuentes de información de amenazas.
- Formación y actualización a través del desarrollo de competencias mediante la continua participación en concursos de tipo Capture The Flag.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Como principales objetivos se plantean:

- Desarrollar e implementar un método de detección de hosts infectados, basado en ensembling, que tenga en cuenta los resultados de detección de distintos clasificadores, que usen técnicas de aprendizaje automático y datos de Threat Intelligence y pueda funcionar con ventanas de tiempo y detección a lo largo del tiempo.

- Proponer una integración del mecanismo de ensembling para detectar hosts infectados, al servicio de monitoreo proactivo de seguridad de CERTUNLP [2], CSIRT Académico de la Universidad Nacional de La Plata.
- Desarrollar y mantener un sistema programable y configurable capaz de brindar soporte a la gestión de incidentes de seguridad en el ámbito de trabajo de un Computer Security Incident Response Team (CSIRT). Es deseable que dicho sistema sea integrable con otros componentes de software que son utilizados en la comunidad de CSIRTs
- Hacer posible la integración de los feeds al sistema de gestión de incidentes para así incorporar los enriquecedores de información que sirven para mejorar la base de conocimiento de un incidente y facilitar el intercambio de información con otros grupos.
- Capacitar al grupo de investigación y a terceros en distintas temáticas de ciberseguridad, mediante la implementación/organización y participación en CTFs con el objetivo de visibilizar, fomentar el interés e incorporar nuevas problemáticas y metodologías de resolución que puedan ser incorporadas y aplicadas en distintos ámbitos [7][8].
- Analizar la problemática y las alternativas de mitigación frente a ataques de denegación de servicio distribuidos en Internet con el objetivo de alcanzar una solución integral y abierta para que administradores de organizaciones con bajos recursos puedan acceder a una solución para mitigar este tipo de ataques utilizando un scrubbing center comunitario que permita captar y filtrar parte del tráfico relacionado con el ataque de DDoS.
- Implementar servicios reactivos que permitan contar con mayor visibilidad sobre los distintos recursos de nuestra organización para así mejorar las capacidades en la detección, prevención, gestión y análisis de incidentes de seguridad.
- Promover buenas prácticas en relación a la ciberseguridad que aplican en todas las etapas del ciclo de vida del desarrollo, de los servicios y de la gestión de las organizaciones.

Entre los resultados que se han obtenido en este último período se destacan:

- El diseño de una metodología para detectar hosts infectados en la red aplicando Ensemble Learning y la creación de un procedimiento asociado para testear la misma través de experimentos usando datasets reales y sus resultados.
- El diseño y propuesta de implementación del módulo de Ensembling integrado a Slips [9][10], en el marco de un trabajo de colaboración con el Laboratorio Stratosphere de la Universidad Técnica de República Checa de Praga.
- La actualización del sistema de incidentes NGEN integrando el mismo a IntelMQ con el fin de centralizar y normalizar la información obtenida de distintas fuentes de información las cuales se reciben por distintos canales de comunicación y en diferentes formatos.
- La implementación de un scanner de servicios activos de la UNLP. Este desarrollo hizo posible contar con una herramienta de relevamiento continuo que permite de manera centralizada, mantener actualizada en tiempo real y consultar información sobre los servicios brindados, los recursos de DNS usados, los productos de software, las

plataformas y las tecnologías utilizadas por la comunidad de la Universidad.

- La organización de diversas competencias de tipo CTF. En el último tiempo, el grupo de investigación organizó, desplegó y diseñó competencias de tipo CTF, desde la instalación de la infraestructura requerida hasta la creación de los retos de la competencia. Entre los eventos destacados se encuentran:
 - Etapa Argentina (segunda etapa) del CTF Internacional MetaRed 2020. Esta etapa contó con la participación de 342 equipos de 37 países distintos. [11]
 - CTF OWASP LATAM@Home 2020 organizado en el marco del OWASP LATAM TOUR 2020. Este CTF contó con 525 usuarios registrados de la comunidad internacional, representando a 181 equipos participantes [12].
 - 2 CTFs con alumnos de escuelas secundarias en el marco del Proyecto de Extensión “Vínculos con Escuelas Secundarias” de la Facultad de Informática de la UNLP. Uno de ellos en el marco del evento Chicas en TICs y el otro auspiciado por la embajada de Estados Unidos. Para ello se adaptó la plataforma creada en la Tesina de grado “Capture the flag aplicada a la enseñanza de ciberseguridad en escuelas secundarias” realizada por los alumnos Patricio Bolino y Gabriela Suárez [8] creando nuevos desafíos para esta instancia.

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). Parte de este grupo también forma parte de CERTUNLP, el CSIRT Académico de la Universidad Nacional de La Plata [2], ámbito en el cual aplican directamente las temáticas propuestas. En el marco de estas actividades, se ha finalizado la tesis para obtener la Maestría en Redes de Datos: “Detección de ataques de seguridad en redes usando técnicas de ensembling” de la Lic. Paula Venosa. Esta tesis realizó aportes al proyecto SLIPS del Stratosphere Laboratory en República Checa, que funciona en el ámbito de la CVUT (Czech Technical University in Prague)[13]. También se encuentra en su etapa final de desarrollo, la tesina de grado de Damián Rubio “Evolución del sistema de gestión de incidentes de seguridad orientado a CSIRT de la UNLP - Ngen”. Esta tesis, está íntimamente ligada al desarrollo y evolución de NGEN, el sistema de gestión de incidentes actualmente usado en CERTUNLP y liberado como software libre para su uso y contribución por parte de la comunidad. Además, se ha presentado la propuesta de la tesina de grado de Mateo Durante y Cristian Barbaro relacionada a la implementación de una herramienta para mitigar ataques de DDoS. Y como ocurre hace unos años, continuamos con la consolidación del equipo de CTF, denominado SYPER[14], formado por alumnos y docentes. Teniendo participación en distintos eventos, tanto nacionales como internacionales durante el 2020.

5. REFERENCIAS

[1] LINTI: Laboratorio de Investigación en

Nuevas Tecnologías Informáticas
www.linti.unlp.edu.ar

[2] CERTUNLP. Sitio institucional
<https://www.cert.unlp.edu.ar> (accedido en febrero de 2021).

[3] Proyectos I+D - 11/F028 (2020/2023)-
"De la Sociedad del Conocimiento a la
Sociedad 5.0: un abordaje tecnológico y ético
en nuestra región
<http://secyt.presi.unlp.edu.ar/Wordpress/wp-content/uploads/2020/03/DISPOSICION-130.pdf>

[4] Emmanouil Vasilomanolakis et al.
Taxonomy and survey of collaborative intrusion
detection. *ACM Computing Surveys*
(CSUR), 47:1–33, 2015.

[5] Emna Bahri et al. Approach based
ensemble methods for better and faster
intrusion detection. *Computational
Intelligence in Security for Information
Systems*, pág. 17–24, 2011.

[6] Emna Bahri et al. A survey of intrusion
detection systems based on ensemble and
hybrid classifiers. *Computers Security*,
65:135–152, 2017.

[7] Francisco Javier Díaz et al. (2018). WICC
2018 (Workshop de Investigadores en
Ciencias de la Computación). UNNE,
Corrientes, Argentina. Abril de 2018. Libro
de Actas XX Workshop de Investigadores en
Ciencias de la Computación. pp1056-1060.
ISBN 978-987-3619-27-4.

[8] Francisco Javier Díaz et al. Participación y
despliegue de CTFs como herramienta para
fortalecer la formación en ciberseguridad
WICC 2020 (Workshop de Investigadores en
Ciencias de la Computación). UNPA, Santa
Cruz, Argentina. Abril de 2020. Libro de
Actas XX Workshop de Investigadores en
Ciencias de la Computación. pp1056-1060.
ISBN 978-987-3714-82-5.

[9] Stratosphere Lab. Stratosphere IPS [https://
www.stratosphereips.org/stratosphere-ips-
suite.2](https://www.stratosphereips.org/stratosphere-ips-suite.2)

[10] Repositorio donde se encuentra la versión
de SLIPS con el módulo Ensembling
[https://github.com/pvenosa/StratosphereLinux
IPS](https://github.com/pvenosa/StratosphereLinuxIPS)

[11] Sitio del evento de Metared
[https://eventos.metared.org/55909/detail/ctf-
internacional-metared-2020.html](https://eventos.metared.org/55909/detail/ctf-internacional-metared-2020.html)

[12] Sitio del evento OWASP LATAM [https://
owasp.org/www-event-2020-latam-at-home/](https://owasp.org/www-event-2020-latam-at-home/)

[13] Stratosphere Thesis Projects homepage
[https://www.stratosphereips.org/thesis-
projects](https://www.stratosphereips.org/thesis-projects).

[14] Perfil del equipo SYPER
<https://ctftime.org/team/2003>

Sistema Inteligente de Detección de Anomalías para IoT

Bolatti Diego, Karanik Marcelo, Todt Carolina, Scappini Reinaldo, Gramajo Sergio

Universidad Tecnológica Nacional, Facultad Regional Resistencia
Departamento de Ingeniería en Sistemas de Información
Centro de Investigación Aplicada en Tecnologías de la Información
y la Comunicación (CInApTIC)
French 414 – Resistencia (3500) Chaco - Argentina
{diegobolatti, mkaranik, carolinatodt, rscappini, sergiogramajo}@gmail.com

RESUMEN

En los últimos años, con el avance de Internet de las Cosas (IoT), ha aumentado la cantidad de dispositivos conectados a la red y, consecuentemente, el incremento de los riesgos de violaciones de seguridad y ataques maliciosos. Estadísticamente la mayoría de estos ataques se producen en los dispositivos finales de IoT y existen múltiples alternativas detectarlos. En ese contexto, este proyecto tiene como objetivo el diseño de un Sistema Inteligente de Detección de Anomalías para IoT que utilice técnicas de Machine Learning (ML). Específicamente, el proyecto abarca el diseño y desarrollo de un sistema capaz de detectar ataques de seguridad en base a anomalías en los dispositivos finales de IoT, aplicando técnicas de aprendizaje automático que provean el mecanismo adecuado para dicha detección.

Palabras Clave: Internet de las cosas, Detección de anomalías, Machine Learning, Seguridad.

CONTEXTO

Este trabajo de investigación se desarrolla en el marco del proyecto “Análisis y Aplicaciones de Internet de las Cosas y Ciudades Inteligentes basadas en Telecomunicaciones y Seguridad” (Código del Proyecto: CCUTIRE0005353TC) del Centro de Investigación Aplicada en TICS (CInApTIC) de la Universidad Tecnológica Nacional, Facultad Regional Resistencia.

1. INTRODUCCIÓN

Actualmente el Internet de las cosas (IoT) vive una gran expansión, y se ha convertido en una tendencia irreversible, que se refleja en la conexión diaria a internet de miles de dispositivos y sensores los cuales obtienen y distribuyen información a través de la Web. Los ámbitos en los que se utiliza IoT son variados y van desde la agricultura [1], la salud [2], la hotelería [3], el monitoreo de tráfico [4] y la gestión de flotas [5] entre otros.

Debido a la gran diversidad de dispositivos, tecnologías y protocolos de comunicación (Lora, Zigbee, Wifi, etc.) es un gran desafío gestionar la seguridad de un ecosistema IoT. Sumado a esto, la mayoría de los dispositivos de IoT no se diseñan pensando en la seguridad, y muchos de ellos no poseen capacidades esenciales de encriptación y autenticación. Lo que ha llevado a una categoría completamente nueva de ataques dirigidos explícitamente a los dispositivos finales.

En definitiva, sin un buen nivel de protección, los usuarios no pueden adoptar muchas aplicaciones de IoT que son útiles para el desarrollo de sus actividades.

Debido a estos motivos, se propone el desarrollo de un sistema inteligente de detección de anomalías que permita detectar de forma automática actividades anormales que comprometan la integridad, la confidencialidad y la disponibilidad de un entorno de IoT.

Una anomalía es un comportamiento inusual, irregular y no habitual en el siste-

ma que puede indicar un ataque de seguridad o una falsa alarma.

Entre las anomalías que el sistema inteligente de detección de anomalías podrá detectar se encuentran:

- **Denegación de servicio (DoS):** En este tipo de ataque se envía una gran cantidad de paquetes para inundar un objetivo y hacer que sus servicios no estén disponibles para otros servicios [6].
- **Tipo de dato incorrecto:** En este tipo de anomalía, un dispositivo de IoT malicioso escribe un tipo de dato diferente del tipo de dato previsto [6].
- **Agotamiento de batería:** Ataque que intenta agotar las baterías de los dispositivos de IoT para dejarlos fuera de servicio [7].
- **Jamming:** Este tipo de ataque es una variante de los ataques DoS y consiste en desactivar o saturar los recursos del sistema consumiendo toda la memoria o enviando una gran cantidad de tráfico a la red para que nadie más pueda utilizarla [8].
- **Retraso de Paquetes:** Las transmisiones de datos válidos se retrasan maliciosamente, pero, a diferencia de los ataques de reenvío selectivo, no se eliminan. Por lo tanto, el ataque provoca un retraso en la entrega de datos y, en consecuencia, una degradación del rendimiento de la red.
- **Alteración de paquetes:** Estos ataques intentan alterar el contenido de un paquete enviado por el dispositivo para inyectar datos maliciosos en los nodos de la red.
- **Man In The Middle (Ataque de Intermediario):** Estos ataques se realizan para espiar la comunicación de red de los dispositivos y modificar el tráfico de la red para realizar ataques de inyección y reproducción [9].

Para detectar de forma automática e inteligente las anomalías, se propone la utilización de técnicas de Machine Learning

(ML). ML es una rama de la Inteligencia Artificial (IA), que a través de algoritmos proporciona a los sistemas la habilidad de identificar patrones entre los datos para hacer predicciones [10].

Para utilizar el aprendizaje automático o ML se necesita una gran cantidad de muestras de datos de entrada correctamente etiquetadas para entrenar al módulo de ML.

Sin embargo, incluso cuando un algoritmo de ML ha recibido una gran cantidad de datos, aún no hay garantía de que pueda identificar correctamente las nuevas anomalías. Por lo tanto, se requiere constantemente la experiencia y el control del ser humano.

El mismo problema surge si el algoritmo solo usa sus propios datos de salida como entradas para un mayor aprendizaje. Los errores se refuerzan y multiplican, ya que los mismos resultados incorrectos vuelven a ingresar a la solución en un bucle y crean más falsos positivos ("FP": categorizar incorrectamente las muestras limpias como maliciosas) y falsos negativos (marcar las muestras maliciosas como benignas).

Por esa razón se propone diseñar un detector de anomalías híbrido que combine las técnicas de detección:

- **Basadas en reglas:** Las anomalías se detectan en base a reglas definidas por expertos en seguridad y son ideales para detectar anomalías conocidas.
- **Basadas en ML:** Las anomalías se detectan con la ayuda de técnicas de aprendizaje automático.

La propuesta es que el detector de anomalías esté basado en redes definidas por software (SDN) y se ubique en la capa de dispositivo de la arquitectura de IoT, tal como se muestra en la Figura 1.



Figura 1: Arquitectura Propuesta.

En la Figura 1 se puede apreciar que el sistema está compuesto por:

- **Dispositivos finales de IoT:** de poca capacidad de procesamiento, como sensores de temperatura, luces inteligentes, entre otros.
- **Gateway SDN:** permiten que los dispositivos de IoT se conecten a la red. Para la conectividad entre los dispositivos finales y los Gateway se utilizan redes de baja potencia y área amplia (LPWAN) como LoRa, Sigfox, entre otros [11]. Además, el gateway cuenta con un interruptor de flujo abierto para monitorear el tráfico proveniente de los dispositivos finales.
- **Controlador SDN:** este componente administra y configura los recursos de la red.
- **Módulo de detección de anomalías:** este módulo recopila los datos de las puertas de enlace para así buscar anomalías. El rol de este módulo es agregar inteligencia al controlador SDN para reajustar la red y mantener las políticas de seguridad definidas por los

administradores al detectar anomalías.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación que se abordan en el proyecto están vinculadas con:

- **Arquitectura de redes de información para IoT:** Para el diseño del detector de anomalías se estudiarán las arquitecturas de redes de IoT.
- **Redes definidas por software:** Se investigará esta tecnología emergente con el objetivo de aplicar las ventajas de la misma, en el desarrollo del sistema.
- **Virtualización de redes:** Se utilizará herramientas de simulación y virtualización de dispositivos, para generar escenarios de pruebas y probar el sistema.
- **Inteligencia Artificial:** Se realizará un análisis de las técnicas de ML, con el objetivo de seleccionar la mejor opción para la implementación en el módulo de detección de anomalías.
- **Tecnologías LPWAN:** Se estudiarán la tecnología LPWAN, con el objetivo de identificar las características de los paquetes de datos que ayuden a detectar anomalías en los dispositivos finales de IoT.
- **Seguridad de IoT:** Se analizarán los ataques de seguridad y anomalías que se presentan en un entorno de IoT.

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto, se centra en el diseño de un sistema de detección de anomalías, capaz de detectar ataques de seguridad en dispositivos finales de IoT.

Específicamente se pretende:

1. Estudiar las diferentes opciones de seguridad para un entorno de IoT.

2. Evaluar los últimos desarrollos e investigaciones de detección de anomalías realizados para un entorno de IoT.
3. Diseñar un prototipo de un detector de anomalías para IoT.
4. Documentar las características, arquitectura, ventajas y desventajas del sistema diseñado.
5. Evaluar el alcance de la solución y su viabilidad.

4. FORMACIÓN DE RECURSOS HUMANOS

Con el proyecto se pretende contribuir a la formación de recursos humanos desde diversas áreas:

- **Formación de becarios:** El proyecto cuenta con la participación de alumnos becarios del último año de la carrera de Ingeniería en Sistemas de Información que están realizando su práctica supervisada.
 - **Alumnos de la carrera de Ingeniería en Sistemas de Información:** Se prevé realizar actividades de actualización y talleres con alumnos de las cátedras del área de redes de información, comunicaciones y seguridad informática. Además, por las propias características de los temas que involucra el proyecto se pueden realizar actividades en cátedras como inteligencia artificial.
 - **Formación de jóvenes profesionales:** Se prevé la incorporación de jóvenes profesionales de Ingeniería en Sistemas de Información con la intención de seguir con una carrera en investigación universitaria. Los cuales pueden incorporarse en carácter ad-honorem al proyecto o a través de becas de iniciación en la investigación.
 - **Formación de postgrado:** A partir de las líneas de investigación desarrolladas en el proyecto se prevé que el Ing. Diego Bolatti finalice su doctorado mediante una tesis vinculada a este proyecto.
- **Equipo de trabajo:**
 - **Director:**
 - Gramajo, Sergio.
 - **Investigadores de apoyo:**
 - Bolatti, Diego
 - Scappini, Reinaldo
 - Karanik, Diego
 - **Becario alumno:**
 - Todt, Carolina
 - Federico Aguirre

5. BIBLIOGRAFÍA

- [1]. Agriculturers.com Red de Especialistas en Agricultura. (2019). *Aplicaciones de IoT en agricultura*. El texto puede consultarse en la siguiente URL:
<http://agriculturers.com/aplicaciones-de-iot-en-agricultura/>
- [2]. Dostie, R. (2019). *El Internet de las Cosas (IoT) en el área de la salud en 2019*. Everything Rad. El texto puede consultarse en la siguiente URL:
<https://www.carestream.com/blog/2019/01/01/el-internet-de-las-cosas-iot-en-el-area-de-la-salud-en-2019/>.
- [3]. Sanz Baños, B. (2019). *IoT en los hoteles - Think Big Empresas*. Think Big. El texto puede consultarse en la siguiente URL:
<https://empresas.blogthinkbig.com/iot-en-los-hoteles/>.
- [4]. Fractal.com. (2019). *Las 9 aplicaciones más importantes del Internet de las Cosas (IoT)*. El texto puede consultarse en la siguiente URL:
<https://www.fractal.com/blog/2018/10/10/9-aplicaciones-importantes-iot>.
- [5]. Wireless, K. (2019). *Soluciones de IoT para la gestión de flotas*. Mx.korewireless.com. El texto puede consultarse en la siguiente URL:
<https://mx.korewireless.com/industries/iot-fleet-management-solutions>.
- [6]. Hasan, M., Islam, M. M., Zarif, M. I. I., & Hashem, M. M. A. (2019). *Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches*. Internet of Things, 7, 100059.
- [7]. Smith, Ryan et al. *Battery Draining Attacks Against Edge Computing Nodes in*

IoT Networks. Cyber-Physical Systems 6.2 (2020): 96–116. Crossref. Web.

[8]. Font, V. G. (2017). *Anomaly detection in smart city wireless sensor networks* (Doctoral dissertation, Universitat Oberta de Catalunya).

[9]. Anthi, E., Williams, L., Słowińska, M., Theodorakopoulos, G., & Burnap, P. (2019). *A supervised intrusion detection system for smart home IoT devices*. *IEEE Internet of Things Journal*, 6(5), 9042-9053.

[10]. Sen, P. C., Hajra, M., & Ghosh, M. (2020). Supervised classification algorithms in machine learning: A survey and review. In *Emerging technology in modelling and graphics* (pp. 99-111). Springer, Singapore.

[11]. Y.4101/Y.2067 (10/17). *Common requirements and capabilities of a gateway for Internet of things applications*. El texto puede consultarse en la siguiente URL: <https://www.itu.int/rec/T-REC-Y.4101-201710-I/en>.

Un Sistema de Voto Electrónico para la FCEyN (UNLPam.)

Pablo García¹ Germán Montejano² Andrés Farías³
Claudio Ponzio¹ Martín Lobos¹ Adrian García¹

¹Departamento 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
pablogarcia@exactas.unlpam.edu.ar

²Departamento 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@unsl.edu.ar – web: <http://www.unsl.edu.ar>

³Facultad de Ciencias Físicas Exactas y Naturales
Universidad Nacional de La Rioja
Av. De La Fuente S/N - La Rioja
afarias@yahoo.com.ar

RESUMEN

La utilización del voto electrónico sigue siendo un tema que genera fuertes controversias. En los ámbitos políticos, se utiliza la dicotomía planteada contra el voto manual como un elemento de permanentes disputas.

Desde hace varios años este equipo de trabajo propone un análisis imparcial de los costos y beneficios de implementar este tipo de sistemas, proponiendo métodos y técnicas que permitan que un sistema de voto electrónico responda a exigencias del más alto nivel y publicando periódicamente sus avances (por ejemplo, [1], [2], [3] y [4]).

Se afirma que un sistema de E-Voting no solamente debe ser absolutamente seguro, sino que además, tal característica debe ser plenamente comprobable. Pero no sólo para

los expertos en la materia; también para todos los votantes que participen de un proceso electoral.

La confiabilidad del sistema no solamente debe apuntar a la integridad de los resultados obtenidos, sino que aparecen otros aspectos que deben observarse, como por ejemplo la confidencialidad del elector (que debe protegerse indefinidamente) y la velocidad con la que se obtienen los resultados finales.

En consecuencia, se propone implementar un sistema de voto electrónico que pudiera aplicarse en la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa, a través de un nuevo proyecto de investigación que será presentado durante 2021 y que tendrá una duración de cinco años. Se busca implementar todos los avances realizados en publicaciones previas y agregar

elementos novedosos en algunos puntos, tal como lo describe el presente documento

Palabras clave: *Sistemas de Voto Electrónico, Anonimato, Transparencia, Criptografía Homomórfica, Verificabilidad E2E, Prueba Física.*

CONTEXTO

El presente trabajo tiene por objeto presentar un Proyecto de Investigación que será presentado durante 2021 en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa. El mismo se titulará: “Un Sistema de Voto Electrónico Basado en Criptografía Homomórfica para la FCEyN (UNLPam.)”. El mismo es una derivación de un proyecto anterior (“Aspectos de Seguridad en Proyectos de Software”, Resolución N° 488/14 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales) y que fue dirigido por el Doctor Germán Antonio Montejano (Universidad Nacional de San Luis) y codirigido por el Magister Pablo Marcelo García (FCEyN - UNLPam) e incluyó a la Magister Silvia Gabriela Bast, al Magister Daniel Vidoret, a la Profesora Estela Marisa Fritz, al Analista Programador Adrián García y al Programador Superior Claudio Ponzio como investigadores y a la estudiante Silvia Nicosia como asistente de investigación.

El nuevo proyecto incluirá inicialmente, a los autores del presente trabajo, aunque se buscará elevar el número de investigadores a los efectos de enfrentar los fuertes desafíos que se presentan desde el punto de vista teórico pero también para lograr una implementación eficaz de la aplicación final.

Queda claro que se trata de la continuidad de proyectos anteriores. El origen de esta línea de investigación se ubica en [5], que a su vez se enmarca en el Proyecto “Ingeniería de Software: Aspectos de Alta Sensibilidad en el ejercicio de la Profesión de Ingeniero de Software” de la Facultad de Ciencias Físico - Matemáticas y Naturales de la Universidad

Nacional de San Luis (UNSL), y que se desarrolla en cooperación con la Universidad Federal de Minas Gerais (UFMG, Brasil).

1. INTRODUCCIÓN

Una de las ventajas principales de un sistema de voto electrónico tiene que ver con la obtención de resultados en tiempos significativamente inferiores a los que existen si el proceso es manual. Si bien esa velocidad no es la cuestión más importante dentro del proceso, tampoco es un detalle que deba ignorarse. Cabe recordar que en las elecciones legislativas PASO de 2017, la provincia de Buenos Aires mostraba, ya muy tarde en la noche, resultados opuestos a los que finalmente se obtuvieron al final del recuento, 16 días después (<https://www.perfil.com/noticias/elecciones2017/paso-cristina-le-gano-a-bullrich-por-021-de-los-votos.phtml>). También se puede mencionar que Al Gore en 2000 tardó un mes en reconocer su ajustada derrota frente al candidato George W. Bush (<https://aristeguinoticias.com/0811/mundo/bush-vs-gore-la-eleccion-presidencial-que-casi-colapsa-a-eu/>).

Los efectos producidos por esa demora en la difusión de los resultados de los comicios resultan perjudiciales. Se produce una incertidumbre que sólo se despeja cuando la incógnita es develada. Tal situación es indeseable para cualquier sociedad.

Es precisamente en ese aspecto, que la criptografía homomórfica presenta una característica muy conveniente a los efectos de implementar sistemas de voto electrónico, dado que permite realizar operaciones directamente sobre los datos cifrados. Ese atributo le otorga un gran atractivo a su aplicación práctica, aunque, por supuesto, es necesario arbitrar los medios para garantizar la veracidad de los resultados obtenidos y, simultáneamente, proteger la privacidad del votante. Adicionalmente, todo deberá ser absolutamente demostrable.

La criptografía homomórfica aparece en 1978 [6]. Las primeras técnicas presentaban homomorfismo parcial, dado que sólo era posible aplicar un tipo de operación sobre los

datos cifrados (suma o producto). Son ejemplos de este tipo El Gamal [7], Benaloh [8] y Paillier [9].

El primer esquema totalmente homomórfico es de 2009 y se presenta en [10]. Esto resulta un avance de gran importancia, porque permite operaciones de adición y producto sobre datos cifrados.

El modelo que se va a implementar, utilizará alguna técnica de criptografía homomórfica, que será definida como parte del proceso de investigación. Se piensa en variantes de métodos existentes o combinación de más de uno, considerando que es posible que aparezca un esquema novedoso que aporte ventajas concretas.

Por otra parte, el sistema a implementar deberá cumplir con una serie de requisitos que se exigen actualmente a los sistemas de votación electrónica:

- Evidencia física que garantice la transparencia del proceso [11].
- Aplicación del concepto de independencia del software [12].
- Definición de un modelo concreto para la aplicación de verificabilidad “End to End” (E2E) [13]. Esto incluye la aplicación de estrategias que eviten la aparición de maniobras fraudulentas
- Selección de una interface apropiada, con un fuerte análisis de alternativas y una fundamentación sobre la elección.
- Implementación de esquemas de seguridad que permitan asegurar la total integridad en la base de datos asociada.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El grupo de investigación trabajará en forma paralela sobre cada uno de los aspectos que se mencionaron en el punto anterior:

- **Evidencia física:** se trabajará en la búsqueda de alternativas respecto de la impresión concreta del sufragio en papel como elemento de respaldo. En

cualquier caso, se deberá probar físicamente la integridad del proceso.

- **Independencia del software:** Deberá definirse un esquema que asegure que el modelo no podrá ser vulnerado por intentos fraudulentos (conocidos o novedosos).
- **Verificabilidad E2E:** El proceso podrá verificarse íntegramente. La verificabilidad “End to End” es uno de los puntos de mayor valor a los efectos de agregar transparencia al proceso de votación electrónica. Se define mediante las tres condiciones siguientes:
 - Verificabilidad individual: cualquier votante puede verificar que su sufragio fue incluido en el recuento.
 - Verificabilidad universal: cualquier persona puede determinar que el recuento total de los votos es correcto.
 - Secreto del voto: ningún votante podrá demostrar cuál fue la opción que eligió, a los efectos de evitar maniobras relacionadas con el “clientelismo político”.
- **Interface:** ya se están realizando relevamientos que permitan deducir qué aspecto debe presentar dicha interface y de qué manera los usuarios se relacionarán con el sistema.
- **Seguridad de las comunicaciones:** dos de los autores de este trabajo ya están trabajando en el diseño de un esquema basado en redes virtuales (VLANs).

3. RESULTADOS Y OBJETIVOS

Los avances del grupo de trabajo que han surgido durante 2020 fueron:

- Análisis crítico de los resultados obtenidos en una encuesta online

destinada a personas de todo nivel (desde expertos informáticos hasta votantes comunes) para obtener opiniones sobre la forma concreta que debería tener la interface de un sistema de voto electrónico.

- Desarrollo de una serie de entrevistas a expertos informáticos para complementar los resultados obtenidos en la encuesta online. Esta etapa aún se está realizando en la actualidad.
- Incorporación al proyecto de dos especialistas específicos en comunicaciones de datos para proporcionar metodologías de transmisión de datos que garanticen los niveles de seguridad exigibles. Los mismos se encuentran realizando el análisis de alternativas para implementar un esquema de basado en VLANs. El objetivo será proporcionar un modelo de comunicación que cumpla con los requisitos de máxima seguridad que se exigen en un sistema de este tipo.
- Se continuó trabajando en el análisis de métodos homomórficos existentes. Se pudieron obtener conclusiones parciales en busca de la selección final del esquema definitivo.

A futuro, se pretende llevar a cabo las siguientes acciones:

- Elegir la interface exacta del modelo en base a los datos obtenidos en encuestas y entrevistas.
- Aplicar la interface seleccionada en la aplicación a desarrollar y publicar los fundamentos de la opción elegida.
- Definir un modelo de transmisión de datos e implementarlo.
- Realizar un relevamiento de aplicaciones orientadas al voto electrónico, que permita detectar

falencias y proponer mejoras en el nuevo modelo.

- Seleccionar, finalmente, un método criptográfico con características homomórficas para aplicar en la implementación.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco del presente proyecto se presentan los siguientes puntos relacionados con la formación de recursos humanos:

- Pablo García realizó tres estadias de investigación en el ámbito del Departamento de Ciência da Computação (DCC) perteneciente al Instituto de Ciências Exatas (ICEEx) de la Universidade Federal de Minas Gerais (UFMG) en Belo Horizonte, Brasil. La primera fue en 2012 y duró diez meses. Las dos restantes tuvieron una duración de 30 días cada una y se llevaron a cabo en 2017 y 2018.
- Pablo García defendió con éxito en 2013 su tesis correspondiente a la Maestría en Ingeniería de Software (FCFMyN, UNSL). Obtuvo una calificación de sobresaliente y fue orientada por los Jeroen van de Graaf, PhD. (UFMG) y el Dr. Germán Montejano (UNSL).
- Pablo García completó el cursado de la totalidad de los créditos exigidos en el Doctorado en Ingeniería Informática en la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL) y presentó su Plan de Tesis Doctoral, en el marco del Doctorado en Ingeniería Informática en la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL). El mismo se encuentra en proceso de evaluación.
- Andrés Farías defendió con éxito en 2019 su tesis correspondiente a la

Maestría en Ingeniería de Software (FCFMyN, UNSL). Obtuvo una calificación de sobresaliente y fue orientada por los directores del futuro proyecto que se presenta en este documento.

- Claudio Ponzio y Martín Lobos desean presentar una Tesis de Especialización o maestría relacionada con los objetivos de este proyecto.

Se espera que otros integrantes del grupo de trabajo a conformar, realicen tesis de posgrado en el ámbito del proyecto a presentar.

5. BIBLIOGRAFÍA

[1] **van de Graaf J., Montejano G., García P.:** “Manejo de Colisiones en un Protocolo Non Interactive Dining Cryptographers”. Anales de las 42° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). 2013.

[2] **García P., van de Graaf J., Hevia A., Viola A.:** “Beating the Birthday Paradox in Dining Cryptographer Networks”. The third International Conference on Cryptology and Information Security in Latin America, Latin-crypt 2014. Florianopolis, Brasil. Lecture Notes in Computer Science, Springer (2014).

[3] **García P., van de Graaf, J., Montejano G., Riesco D., Debnath N., Bast S.:** “Storage Optimization for Non-Interactive Dining Cryptographers (NIDC)”. The International Conference on Information Technology: New Generations. 2015. Las Vegas, Nevada, USA.

[4] **García P., Bast S., Fritz E., Montejano G., Riesco D., Debnath N.,** “A Systematic Method for Choosing Optimal Parameters for Storage in Parallel Channels of Slots”. IEEE International Conference on Industrial Technology (ICIT 2016). / Taiwan, Taipei. <http://ieeexplore.ieee.org/document/7475019/>.

[5] **Uzal R., van de Graaf J., Montejano G., Riesco D., García P.:** “Inicio de la Línea de Investigación: Ingeniería de Software y Defensa Cibernética”. Memorias del XV WICC. Ps 769-773. ISBN: 9789872817961. 2013. <http://sedici.unlp.edu.ar/handle/10915/27537>.

[6] **Rivest, R. Adleman L., Dertouzos, M.:** “On Data Banks and Privacy Homomorphisms. Foundations of Secure Computation, Academia Press (1978)

[7] **El Gamal T.** “A public key cryptosystem and a signature scheme based on discrete logarithms”. In Proceedings of CRYPTO 84 on Advances in cryptology, pages 10–18. Springer-Verlag New York, Inc. 1985.

[8] **Benaloh, J.:** “Dense Probabilistic Encryption”. Workshop on Selected Areas of Cryptography. pp. 120–128. 1994.

[9] **O’Keeffe M.:** “The Paillier Cryptosystem: A Look Into The Cryptosystem And Its Potential Application”. The College of New Jersey Mathematics Department. 2008.

[10] **Gentry G.:** “Fully Homomorphic Encryption Using Ideal Lattices”. In the 41st ACM Symposium on Theory of Computing (STOC), 2009.

[11] **Hao, F, Ryan P.:** “Real -World Electronic Voting. Design, Analysis And Deployment”. Cr Press. ISBN-13: 978-1498714693. ISBN-10: 1498714692. 2017.

[12] **Rivest R.:** “On the notion of ‘software independence’ in voting systems”. Philosophical Transactions of The Royal Society A, 366(1881):3759–3767. 2008.

[13] **Ryan P., Schneider S., Teague V.:** “End-to-End Verifiability in Voting Systems, from Theory to Practice”. Voting Systems, from Theory to Practice. IEEE Security & Privacy, 13(3):59–62, 2015.

Tecnología Informática aplicada en Educación

Agregación de la Tecnología Móvil en los Aprendizajes

Mg. Roberto Bertone¹, Mg. José Luis Filippi², Lic. Guillermo Lafuente³, Mg. Carlos Ballesteros⁴,
Lic. Gustavo Lafuente⁵, I.S. Daniel Perez⁶, I.S. Sofia Aguirre⁷, A.S. Alejandra Mansilla⁸
LIAU⁹ - Facultad de Ingeniería – UNLPam.

pbertone@ada.info.unlp.edu.ar¹

{filippi², lafuente³, balleste⁴, gustavo⁵, perezd⁶, aguirres⁷, mansilla⁸}@ing.unlpam.edu.ar

⁹Laboratorio de Investigación de Ambientes Ubicuos

Resumen

La integración de las TICs en las instituciones educativas es un proceso que se desarrolla desde hace varias décadas en la República Argentina. Si bien existe un fuerte consenso en la universalización del acceso a las diferentes tecnologías de la información y las comunicaciones, la discusión vigente reside en la incorporación de los dispositivos móviles (teléfonos de última generación) como instrumento mediador del Proceso de Enseñanza-Aprendizaje.

En la actualidad la educación no se encuentra limitada a los entornos formales conformado por las instituciones educativas en sus diferentes niveles, se integra con procesos formativos que tienen lugar a través de los nuevos dispositivos tecnológicos de última generación, caracterizados por su miniaturización, movilidad y conectividad permanente; haciendo posible el acceso a múltiples contextos de aprendizaje virtuales. Resulta entonces imprescindible estudiar las tendencias del aprendizaje móvil en el mundo desarrollado, realizar las adaptaciones necesarias e implementarlas en el ámbito educativo.

El propósito del proyecto, de naturaleza teórico-práctico, es el de indagar las posibilidades que ofrecen los dispositivos móviles (teléfonos inteligentes, tabletas digitales) como instrumentos aplicados a diferentes situaciones de enseñanza y aprendizaje.

Bajo esa premisa se intenta llevar a cabo un trabajo experimental en la Facultad de

Ingeniería - UNLPam, desarrollando un escenario de aplicación real con la utilización de éste tipo de dispositivos.

Palabras claves: Aprendizaje Móvil. Dispositivo 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.

Introducción

Con más de 5900 millones de usuarios de telefonía móvil en el mundo, los dispositivos de última generación han transformado nuestra manera de vivir. Aunque en todos los sectores de la sociedad se hace uso intensivo de esta tecnología, los educadores y los responsables de formular las políticas educativas no han aprovechado su potencial para mejorar el proceso educativo. [1]

Múltiples necesidades se presentan en el ámbito educativo en sus diferentes niveles entre los cuales se pueden mencionar: desarrollo de contenidos bibliográfico digital de producción local y regional, incentivar el uso de aplicaciones móviles que posibiliten el intercambio de información entre los actores

que conforman la comunidad educativa (videoconferencias entre docentes, estudiantes, directivos, y comunidad en general), incorporar el uso de plataformas de formación virtual abiertas, que posibiliten incorporar recursos y/o materiales didácticos en diferentes formatos, mantener un canal de comunicación fluido entre la institución educativa y el contexto local, regional, y nacional, e implantar estrategias metodológicas para alcanzar un aprendizaje cooperativo y colaborativo.

Mark Weiser [2] menciona que: *“vamos camino a ambientes ubicuos, ambientes poblados de numerosos sensores que gracias a la miniaturización de los dispositivos son invisibles al usuario y están en permanente rastreo de la actividad humana”*. Aquí radica el objetivo primario del proyecto: indagar diferentes aplicaciones tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las necesidades de los usuarios que transitan en el ámbito educativo, profundizando en el aprendizaje móvil.

Situación Actual del Problema

La aparición de dispositivos tecnológicos caracterizados por su miniaturización, gran capacidad de almacenamiento y gran velocidad de procesamiento, dieron origen a nuevos entornos en la comunicación de las personas, con una hipercomunicación caracterizada por la multimedialidad. El envío de mensajes textuales seguidos de imágenes, sonidos y videos es parte de la actividad cotidiana.

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.

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 preciso 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” [3].

Mucho se habla de cuáles serían las competencias digitales para un docente, un acercamiento a ello lo propone Núñez-Torrón Stock [4], 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. [5] 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. [6] 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. [7]

b. Continuidad del aprendizaje.

El aprendizaje móvil permitirá a los estudiantes dar continuidad a su formación a

lo largo de toda su vida, a partir de plataformas que posibilitan el aprendizaje móvil caracterizado por:

- Poseer gran cantidad de información.
- Ser omnipresentes.
- Disponer de gran diversidad de recursos materiales.
- Fomentar la participación colaborativa.

Actualmente los estudiantes y futuros profesionales acceden a la información a través del aprendizaje informal, han dejado de ser únicamente consumidores para ser productores de información.

En el caso particular de 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. Sin embargo, en muchos países no se están desarrollando estas habilidades. [8]

c. Grandes volúmenes de datos. Big data.

Disponer y aprovechar ésta tecnología es una obligación para todas aquellas instituciones que manipulan grandes volúmenes de datos, en nuestro caso particular la universidad.

La gran cantidad de información accesible a partir de variadas herramientas conlleva una serie de consideraciones éticas relacionadas a la propiedad de los datos y la privacidad.

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. [9]

En éste contexto el punto de partida debe ser siempre 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.

La herramienta en sí misma no ofrece por sí sola resultados, pero un buen docente sabrá aprovechar las ventajas que brinda para un aprendizaje móvil de calidad.

Por ello el aprendizaje móvil (mLearning) puede incorporarse como refuerzo del aprendizaje formal, ampliando la oferta educativa y la modalidad. Lo importante es innovar en el proceso educativo en su conjunto y no solo con el dispositivo tecnológico.

El entorno educativo cambia [10], la educación se presenta como la formación de los educandos en competencias, destrezas, habilidades para desempeñarse en un nuevo espacio social, el digital [11]. Nuevos escenarios educativos, los mismos actores con un nuevo rol, y la implantación de las tecnologías móviles es el desafío actual [12].

Línea de Investigación 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 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.
- 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 transcurriendo, 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.

Resultados Obtenidos/Esperados

El proyecto da inicio a su actividad de investigación y desarrollo durante el año 2019, por lo que sus integrantes se encuentran efectuando las tareas enmarcadas en el tercer año.

A partir de la tarea realizada por los integrantes del proyecto de I+D durante el segundo año se concretó:

- a. El desarrollo de nuevos objetos de aprendizaje para las materias Introducción a la Informática, Programación Procedural, Programación Web y Automatas y Lenguajes. El producto alcanzado se incorporó a la plataforma de formación virtual como material educativo a utilizar durante el ciclo lectivo 2020, y se publicó en la revista TE&ET con el nombre Experiencia de Virtualización en la UNLPam, volumen nro. 26, septiembre 2020.
- b. Se configuró un enfoque metodológico contemplando la experiencia del usuario en el desarrollo de objetos de aprendizaje. El desarrollo se presentó en el congreso virtual CONAIISI 2020.
- c. Se construyó una aplicación que permite automatizar las noticias que la UNLPam envía a los docentes, estudiantes, no docentes y directivos. Se presentó la app en el congreso virtual CONAIISI 2020.
- d. Se desarrolló una herramienta de capacitación a distancia y recursos educativos abiertos bajo un enfoque de diseño universal orientados a personas con discapacidad visual. Se presentó en la VIII Jornada Institucional de Extensión Universitaria en la UNLPam.

Al finalizar el proyecto se espera contribuir en la incorporación de aplicaciones tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las necesidades de los usuarios que transitan en el

ámbito educativo de la UNLPam y la UNLP, profundizando en el aprendizaje móvil.

Formación de Recursos Humanos

Director de Proyecto
Co-Director de Proyecto
6 Investigadores

Referencias

[1] Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2013). El futuro del aprendizaje móvil. Implicaciones para la Planificación y la Formulación de Políticas. <http://unesdoc.unesco.org/images/0021/002176/217638s.pdf>

[2] Weiser M. (1991), *The Computer for the Twenty-First Century*.

[3] 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].

[4] 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].

[5] 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_fi nal_CVSL_SF.pdf

[6] 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.

[7] 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.

[8] (World Economic Forum, 2015).

[9] 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.

[10] Weiser, M. (1998) *The future of Ubiquitous Computing on Campus*. Communications of ACM, 41-1, January 1998, 41-42.

[11] Malani R., Griswold W, Simon B, (2009) *Public Digital Note-Taking in Lectures*. Ubicomp 2009.

[12] Richards M, Woodthorpe J, (2009), *Introducing TUI00 "My Digital Life": Ubiquitous computing in a distance learning environment*. Ubicomp 2009.

Análisis de los recursos de visualización de datos usados en un marco de trabajo de inteligencia de negocios aplicados en la educación superior.

Curso Cynthia, Constable Leticia, Colacioppo Nicolás.

Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información
Departamento Ingeniería en Sistemas de Información
Facultad Regional Córdoba. Universidad Tecnológica Nacional
Maestro M. López esq. Cruz Roja-Ciudad Universitaria-Córdoba
cynthia@bbs.frc.utn.edu.ar/leticiaconstable@gmail.com/nicolas.colacioppo@hotmail.com

RESUMEN

El objetivo de este trabajo es realizar un análisis de los recursos de visualización de datos, que son elegidos en las instituciones de educación superior, cuando implementan un proceso de inteligencia de negocios con la finalidad de fortalecer el proceso de toma de decisiones. Para ello se analizaron trabajos publicados en esta temática que corresponden a los tres últimos años (2018-2020). Se seleccionaron diversas dimensiones de análisis como son: la elección de la herramienta de inteligencia de negocios, recursos de visualización usados y grado de participación de los usuarios finales en la selección de los mismos. Del análisis realizado se desprende como aporte una guía que facilita la identificación de dimensiones a tener en cuenta en la elección de los recursos de visualización de datos en un proyecto de inteligencia de negocios aplicado en el contexto de la educación superior.

Palabras claves: *inteligencia de negocios, recursos de visualización, educación superior.*

CONTEXTO

Este trabajo hace referencia al proyecto “Integración de recursos del Paradigma Analítico y de la Inteligencia de Negocios como estrategia para el fortalecimiento en el proceso de toma de decisiones” PID-SIUTNCO0005101, que ha sido homologado 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 U.T.N Facultad Regional Córdoba.

1. INTRODUCCIÓN

Uno de los grandes desafíos que enfrentan las instituciones de educación superior es encontrar la manera de integrar la cantidad significativa de datos que se generan de sus procesos diarios para la toma de decisiones. Algunos de los sistemas transaccionales que se usan en este contexto tienen una serie de limitaciones; algunas de ellas la imposibilidad de brindar síntesis, análisis, consolidación de los datos y proyecciones. Esta situación puede llevar a las instituciones de educación superior a tomar decisiones inoportunas [1].

Frente a esta problemática surge la necesidad de contar con una herramienta tecnológica que facilite la obtención de indicadores y proporcione una serie de recursos visuales como reportes, dashboard que se estén diseñados y estructurados de una manera rápida, simple y eficaz. Justamente la Inteligencia de Negocios “se entiende por un conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de la información que permite tomar mejores decisiones a los usuarios de una organización” [2].

Un proceso de Inteligencia de Negocios lo componen cuatro fases, la última de estas fases está relacionada con la explotación del repositorio de datos, este proceso implica el análisis y distintas maneras de visualizar y resumir dicha información a los usuarios finales [1]. Esta última fase puede ser abordada con diferentes recursos y formas de visualización de los datos [3].

De hecho en [4] señalan que la elección del gráfico o la manera de estructurar la información entre otros son aspectos críticos para que los usuarios puedan interpretar la información de manera simple e intuitiva. Se puede pensar que si los profesores tuvieran acceso a visualizaciones efectivas de sus datos educativos, podrían usarlos para proporcionar retroalimentación formativa a sus estudiantes o para mejorar los materiales didácticos que emplean. Del mismo modo, si los estudiantes pudieran tener acceso a este tipo de visualizaciones, se podría favorecer el desarrollo de sus habilidades de autorregulación del aprendizaje y ayudarles a conseguir sus objetivos.

En base a lo antes mencionado surgen los siguientes interrogantes: ¿existen mejores

recursos para representar y visualizar los datos para fortalecer el proceso de toma de decisiones en el contexto de la educación superior? ¿Existe alguna guía que permita definir un diseño apropiado de estos recursos de visualización en el ámbito bajo estudio?

Este trabajo tiene como objetivo realizar un análisis de la última fase del proceso de inteligencia de negocios que utilizan las instituciones de educación superior, que refiere precisamente a los recursos de visualización de datos seleccionados para generar conocimiento en los usuarios finales en el contexto bajo estudio.

Para tal fin se realizó un relevamiento sistemático de trabajos limitado a los tres últimos años (2018-2020). Para la búsqueda se seleccionó el buscador Google Scholar. Cabe aclarar que los resultados de las búsquedas están limitados a los filtros usados, pudiendo quedar fuera otros trabajos pertenecientes a esta temática. Se realizó una búsqueda y se establecieron los siguientes filtros: el trabajo debe contener las siguientes frases “Inteligencia de Negocios” *and* “Instituciones de Educación Superior”, estar entre las fechas 2018-2020 y el idioma en español. Se obtuvieron dieciocho trabajos de 16.700 trabajos analizados, este resultado fue depurado aplicando los criterios de inclusión/exclusión relacionados con:

- Formato: el idioma de publicación es español, debe tener una extensión mayor a dos páginas, deben estar completos y el periodo de publicación debe estar comprendido desde el año 2018 al 2020 incluidos.
- El contenido: responde de manera directa a las preguntas de investigación, dominio referido a instituciones de educación superior, se enfoca en trabajos

que han aplicado de manera exitosa procesos de inteligencia de negocios en el dominio bajo estudio.

De esta manera se logró un reducto de diez trabajos sobre los cuales se comenzaron a realizar el análisis y extracción de datos. Para la extracción de datos se confeccionó un formulario que contenía una tabla, en la que se especificaron dimensiones específicas y de interés para este estudio. Los datos de clasificación específicos apuntaban a obtener información puntual de cada trabajo. Algunas de las preguntas de investigación consideradas son:

P1: ¿Cuál es la herramienta de inteligencia de negocios seleccionada?

P2: ¿Los trabajos plantean un marco formal para la elección de la herramienta de inteligencia de negocios?

P3: ¿Cuáles son los recursos de visualización seleccionados?

P4: Para la elección de los recursos visualizados ¿Se utilizó alguna guía que facilitará la selección? ¿Consideran principios de diseño formales para la implementación de los mismos?

P5: ¿En el proceso de diseño se trabajó en forma colaborativa con los usuarios finales como un mecanismo de validación?

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

De este proyecto se desprenden al menos cuatro líneas de investigación y desarrollo.

- Plataformas educativas
- Inteligencia de negocios
- Analítica académica y de aprendizaje
- Analíticas visuales de aprendizaje

3. RESULTADOS OBTENIDOS/ESPERADOS

A continuación se presenta el análisis realizado sobre los trabajos seleccionados de la búsqueda realizada. Es importante mencionar que las dimensiones analizadas están alineadas a las preguntas de investigación planteadas en la sección Introducción.

El primer análisis que se decidió realizar en relación a los trabajos obtenidos de la búsqueda fue determinar cuál fue la herramienta de inteligencia de negocios seleccionada. El resultado fue que las herramientas de inteligencia de negocios usadas en el ámbito de las instituciones de educación superior en el periodo considerado son: Power BI en [5] [6][7] [8], Pentaho en [9][10][11][12] y Spago BI en [13]. Solo en [14] no explicita cuál es la herramienta de inteligencia de negocios con la cual se trabajó. Se puede evidenciar una leve tendencia en la elección de herramientas de inteligencia de negocios libres como lo son Pentaho y Spago BI. Si bien cada institución de educación superior tiene sus necesidades particulares de información, se considera de suma importancia establecer criterios claros que fundamenten la elección de la herramienta de inteligencia de negocios. Solo en [5] explicita un análisis para la selección de la plataforma de inteligencia de negocios, especificando características generales de cada una de ellas y sobre todo enfocándose en los servicios ofrece. Una cuestión interesante de analizar de los trabajos seleccionados, es que no siempre se usó la herramienta de inteligencia de negocios para implementar todas las fases del proceso de inteligencia de negocios. En [1][12][13] realizan el proceso de ETL con la herramienta de inteligencia de negocios seleccionadas,

mientras que la fase de explotación de los datos la implementan con herramientas complementarias como OLAP Qlickview, Spago BI y Saiki Analytcs (visor OLAP). Quizás la necesidad de integrar con otras herramientas, en estos casos para la visualización de los datos, puede que se deba al escaso análisis que se manifiesta en los trabajos en general a la hora de seleccionar la herramienta de inteligencia de negocios a usar.

También se analizaron los recursos visuales que se usan para abordar la fase de explotación del repositorio de datos. Los resultados del análisis arrojaron: dashboards en [1][5][6][10][11][12], como los más usados. En segundo lugar los reportes parametrizados en [5] [6][8] [10][11][12][13]. Y en menor medida usan gráficos de barras, circulares y cubos OLAP [1][8][11][12][14]. De este resultado se puede evidenciar que en la gran mayoría de los trabajos combinan más de un recurso visual. Sin embargo solo en [13] utilizan únicamente como recurso de visualización los reportes.

Del análisis realizado surgió como aporte una guía con algunas recomendaciones a tener en cuenta a la hora de encarar el proceso de diseño para visualización de datos en un proyecto de inteligencia de negocios. Algunas de ellas son:

- Considerar en el proceso de diseño la utilización de heurísticas relacionadas con la usabilidad y accesibilidad, por ejemplo las propuestas por Jacob Nielsen.
- En este análisis se evidencia que los dashboards son los recursos visuales más frecuentemente elegidos para visualizar los indicadores de interés en el contexto de la educación superior. En este punto es importante tener en cuenta y reflexionar acerca de los errores más comunes a la hora de su diseño. Stephen Few ,referente

en esta temática, puede ser tomado en cuenta para profundizar sobre este aspecto.

- Tener en cuenta la participación de los usuarios finales durante el proceso de diseño de los recursos visuales seleccionados. Es posible la utilización de técnicas que favorezcan su participación como “focus group” que es usada en uno de los trabajos relevados [6] para la validación de aspectos relacionados con la forma de visualización de los indicadores y datos.

Se espera que estas recomendaciones sean de utilidad para apoyar el proceso de visualización de datos en un proyecto de inteligencia de negocios, fortaleciendo el proceso de toma de decisiones en diversos ámbitos, no solo en el contexto de la educación superior.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto está conformado por docentes-investigadores pertenecientes a la carrera de grado de Ingeniería en Sistemas de Información. Todos los integrantes docentes del PID han participado del proceso de categorizaciones en investigación dentro del Programa de Incentivos del MECyT; así como en la categorización interna que posee la U.T.N. Además se prevé la participación de alumnos avanzados en la carrera que realizan su práctica supervisada como requisito para el otorgamiento del título de grado de Ingeniero. En este proyecto participa un becario alumno con el objetivo de complementar su formación académica con un acercamiento al ámbito de la investigación científica.

5. REFERENCIAS

- [1] Medina, F., Fariña, F., & Castillo-Rojas, W. (2018). Data Mart para obtención de indicadores de productividad académica en una universidad. *Ingeniare. Revista chilena de ingeniería*, 26, 88-101.
- [2] Curto, J. (2012). *Introducción al Business Intelligence*. UOA 2da edición.
- [3] Mamani, Y. (2018). Business Intelligence: herramientas para la toma de decisiones en procesos de negocio. *Universidad Nacional Micaela Bastidas de Apurímac*.
- [4] Haro Valle, V. A. (2018). Diseño e implementación de un dashboard de soporte académico basado en datos de entornos virtuales de aprendizaje.
- [5] Ortiz Maldonado, O. P. (2018). *Desarrollo de un sistema de registro de datos, análisis y toma de decisiones para el proceso de evaluación y acreditación de carreras de la Universidad Nacional de Chimborazo* (Master's thesis, Pontificia Universidad Católica del Ecuador).
- [6] Chiluisa, Q., & Rolando, O. (2019). *Gestión de información socio económica con bussines intelligence para los estudiantes de la UPEC* (Master's thesis).
- [7] Vega, G. A. H. (2019). Sistema de Alerta Temprana basado en Inteligencia de Negocios para detectar riesgo académico en estudiantes de la Universidad de La Serena.
- [8] Falcón de la Cruz, A. J. (2020). Implementación de un datamart para la gestión de indicadores de deserción universitaria relacionados a la calidad docente administrativa de la modalidad CPE de la Universidad Científica del Sur.
- [9] Medina, F., Fariña, F., & Castillo-Rojas, W. (2018). Data Mart para obtención de indicadores de productividad académica en una universidad. *Ingeniare. Revista chilena de ingeniería*, 26, 88-101.
- [10] Reyes-Mena, F. X., Fuertes-Díaz, W. M., Guzmán-Jaramillo, C. E., Pérez-Estévez, E., Bernal-Barzallo, P. F., & Villacís-Silva, C. J. (2018). Aplicación de Inteligencia de Negocios para el análisis de vulnerabilidades en pro de incrementar el nivel de seguridad en un CSIRT académico. *Revista Facultad de Ingeniería*, 27(47), 21-29.
- [11] Farfán Gavancho, D. V. (2019). Incorporación de Inteligencia de negocios en el proceso de admisión y matrícula en la Universidad Andina Néstor Cáceres Velásquez.
- [12] Ortiz Yumisaca, L. A. (2018). *Desarrollo de un data mart para el monitoreo de indicadores de acreditación universitaria* (Bachelor's thesis, Quito, 2018.).
- [13] Valderrama Triviño, F. A., & Garces Bohada, A. S. (2018). Diseño e Implementación de un Datamart para las Notas Históricas de los Estudiantes en la Universidad Distrital Francisco José de Caldas.
- [14] Riaño Bejar, E. L. (2019). Inteligencia de negocios aplicada a la deserción de los programas posgraduales de universidades de Pamplona, Norte de Santander.
- [15] Alcaraz-Martínez, R., Ribera-Turró, M., Granollers-Saltiveri, T., & Pascual, A. (2020). Accesibilidad para personas con baja visión de los gráficos estadísticos en la prensa digital: una propuesta metodológica basada en indicadores heurísticos. *El Profesional de la Información*, 29(5).

DESARROLLO DE INSTRUMENTACIÓN VIRTUAL PARA USO EN CARRERAS DE INGENIERÍA

Paola Beltramini^{1,2}, Marcos A. Aranda^{1,3}, Eduardo Cano^{1,2}, Marcelo D'amore^{1,2}, Mario Alaniz^{1,3}, Aureliano Herrera^{1,4}, Pedro Foresi^{1,2}

pbeltramini@teco.unca.edu.ar, markosdarioaranda@gmail.com

(1) Grupo de Investigación en Internet de las Cosas (GIIoT), Facultad de Tecnología y Ciencias Aplicadas, UNCa.

(2) Departamento Electrónica, Facultad de Tecnología y Ciencias Aplicadas, UNCa.

(3) Departamento de Informática, Facultad de Tecnología y Ciencias Aplicadas, UNCa.

(4) Alumno Becario, Facultad de Tecnología y Ciencias Aplicadas, UNCa.

AREA TEMATICA: Tecnología Informática Aplicada en Educación

RESUMEN

Las prácticas de laboratorio son una actividad indispensable en la formación integral de un estudiante de carreras tecnológicas. Bajo esta concepción, el equipo de trabajo que integra el proyecto, viene trabajando en el desarrollo tecnológico y pedagógico de instrumental de laboratorio que permita al estudiante “ver” lo que estudia mediante la construcción de los ensayos y el uso de una interfaz gráfica, como apoyo al proceso de visualizar señales y parámetros no visibles al ojo.

Si bien originalmente se propuso la construcción de un “Laboratorio Portátil de Escritorio”, la pandemia del COVID-19, que impidió la realización de actividades educativas presenciales durante todo el año 2020 y posiblemente afecte a gran parte del año 2021, puso de manifiesto la importancia de disponer de instrumentación virtual, que permitan a nuestros estudiantes ser artífices de su propio aprendizaje, y realizar prácticas de laboratorio aun en sus hogares.

Ante esta realidad, se consideró oportuno adaptar el proyecto original y, a partir del uso de software libre, encarar el desarrollo de instrumentos virtuales para que alumnos que cursan asignaturas de los primeros años de carreras de ingeniería o similares conozcan y aprendan a utilizar instrumental que luego usarán en asignaturas de años posteriores.

Palabras Claves: TIC, instrumentación virtual, Enseñanza,

CONTEXTO

Nuestra propuesta surge como continuación del trabajo que viene realizando el grupo de investigadores desde el año 2016, a través de proyectos de investigación presentados en diferentes convocatorias de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca (UNCA) y ante convocatorias de Organismos provinciales. El proyecto inicial (2016) se enfocó en el diseño y construcción de un prototipo de un “Laboratorio de escritorio” y en el diseño de las actividades de enseñanza. Posteriormente el trabajo se enfocó en salvar las dificultades y falencias detectadas y optimizar el hardware y software realizados, manteniendo la estructura base del prototipo, realizando algunas modificaciones y mejoras en cada una de las etapas en función de las nuevas tecnologías accesibles.

Como principal organismo beneficiario resulta la Universidad Nacional de Catamarca, siendo las carreras de grado y pregrado de la Facultad de Tecnología y Ciencias Aplicadas los espacios donde se realizan las pruebas y se implementa los prototipos. Una vez concluidos los ensayos y pruebas de funcionamiento, el instrumental desarrollado podrá ser puesto a disposición de los colegios técnicos de la provincia que se encuentren interesados en esta nueva modalidad de enseñanza.

El ámbito de trabajo es la Facultad de Tecnología y Ciencias Aplicadas, principalmente el Laboratorio que posee el grupo de investigación en Internet de las Cosas (GIIoT) y el Laboratorio de Electrónica, los

cuales aportan los materiales, equipamiento e instrumental necesarios.

1. INTRODUCCIÓN

En las carreras de ingeniería se está prestando especial atención a las estrategias y formas de aprendizaje activo, investigaciones han demostrado que los estudiantes pueden mejorar su aprendizaje si están activamente comprometidos con el material que están estudiando. Integrar laboratorios especiales ya sean remotos, virtuales o de escritorio, en el diseño de actividades de enseñanza, permite que los estudiantes conduzcan mejor su aprendizaje para lograr los resultados deseados. [1][2]

Para el caso particular de la Facultad de Tecnología y Cs. Aplicadas, las prácticas de Laboratorio se realizan de manera presencial en un entorno de “laboratorio real”, y en general, es difícil que cada asignatura/estudiante disponga de sus propios instrumentos. Una solución a la falta de disponibilidad de instrumental suficiente podría ser utilizar los recursos propios de los estudiantes, es decir, sus computadoras y teléfonos personales, reemplazando el instrumental de laboratorio por instrumentos virtuales. Esta idea hoy toma mayor relevancia, ante la pandemia que está atravesando la sociedad mundial, y en particular la Argentina, donde, en el 2020 debido a las medidas de aislamiento adoptadas por los gobiernos nacional y provincial no se pudieron realizar actividades prácticas presenciales en forma tradicional, siendo que aún se desconoce si se podrá retornar a las mismas durante el transcurso del 2021.

La utilización de TICs en la educación, como ser programas de simulación, de software y hardware, permiten generar ambientes de aprendizaje basados en experimentos de laboratorio y resolución de problemas de ingeniería reales, fortaleciendo el desarrollo de competencias y desempeños necesarios para la práctica profesional del futuro ingeniero.[2] Así se abren camino dos nuevos conceptos muy importantes: la instrumentación virtual y los sistemas de adquisición o toma de datos. [3]

1.1 Instrumentación Virtual

La instrumentación virtual es un concepto introducido por la compañía National Instruments, definiendo a un Instrumento Virtual como "un instrumento que no es real, que se ejecuta en una computadora y tiene sus funciones definidas por software" (National Instruments, 2001), y a la “Instrumentación Virtual” como un sistema de medición, análisis y control de señales físicas con un PC por medio de instrumentos virtuales. [4]

La instrumentación virtual permite implementar “laboratorios virtuales” para realizar experiencias similares utilizando aplicaciones informáticas que incluyen el armado esquemático del circuito, manipulación de instrumentos, visualización de animaciones, obtención numérica y gráfica de resultados, incluyendo en algunos casos el almacenamiento digital para procesamiento posterior con programas de graficación más complejos y bases de datos. [5]

Algunas ventajas de los instrumentos virtuales frente a los instrumentos tradicionales son: [4] [5]

- Flexibilidad: un instrumento virtual puede ser diseñado por el usuario de acuerdo con sus necesidades, sus funciones pueden modificarse y adaptarse a las necesidades particulares del usuario, modificando el programa, sin necesidad de reemplazar todo el instrumento.
- Portabilidad: los instrumentos virtuales pueden correr en las computadoras portátiles.
- Menor costo
- Ocupa menor espacio
- En general no requieren fuente de alimentación externa.
- Se pueden adaptar fácilmente a los cambios tecnológicos

Estos beneficios convierten a los Instrumentos Virtuales en una herramienta didáctica muy importante que permite mejorar el proceso de enseñanza y aprendizaje en las carreras de ingeniería, y además los experimentos diseñados bajo este esquema pueden estar disponibles no sólo localmente sino a distancia a través de Internet.[4]

1.2 Osciloscopio

El osciloscopio es el instrumento fundamental de un laboratorio, básicamente es un

dispositivo de visualización gráfica que muestra señales eléctricas analógicas variables en el tiempo

Los osciloscopios son fundamentalmente de dos tipos Analógicos y Digitales, los primeros, trabajan directamente con la señal aplicada, se amplifica y posteriormente, se ingresa por medio de la creación de campos eléctricos y magnéticos, que desvía un haz de electrones en sentido vertical y de forma proporcional a dicho valor de entrada. En contraste los osciloscopios digitales utilizan un conversor analógico-digital (ADC) para adquirir y almacenar digitalmente la señal de entrada, reconstruyendo posteriormente esta información en la pantalla. [7][8]

Un osciloscopio virtual es una especie de osciloscopio digital, que, en lugar de mostrar la señal en una pantalla, la grafica mediante un software que corre en una computadora y que ingresa por lo general mediante conexión USB. En si son tarjetas de adquisición de datos de alta velocidad que funcionan igual que un osciloscopio. [8]

La etapa de adquisición de datos es vital para que el osciloscopio reconstruya con precisión la forma de onda de la señal a mostrar. La adquisición de datos puede realizarse con cualquier placa de desarrollo o sistema embebido, como por ejemplo las plataformas Arduino [10], EDU-CIAA [11] y STM32F4 Discovery [12]. En general, las entradas analógicas de las placas de desarrollo admiten valores de tensión positivos que oscilan de 3V a 5V, por lo que previo a la digitalización en el ADC, se debe realizar una etapa de acondicionamiento de la señal a medir, ajustando sus niveles a los admisibles por cada placa.

El software es el componente más importante en un instrumento virtual, ya que a partir de él se crean las aplicaciones, diseñando e integrando las rutinas que requiere un proceso en particular. [5]. En función del objetivo del proyecto se trabaja con “Processing” que es un lenguaje de programación y entorno de desarrollo integrado de código abierto basado en Java, de fácil utilización, y que sirve como medio para la enseñanza y producción de proyectos multimedia e interactivos de diseño

digital. Fue iniciado por Ben Fry y Casey Reas, ambos miembros de Aesthetics and Computation Group del MIT Media Lab dirigido por John Maeda. [6]

Se eligió este entorno por ser de código abierto, de fácil programación y porque además permite generar un archivo (o aplicación) portable que no requiere la instalación del programa para ejecutar el instrumento en cualquier computadora con Sistemas Operativos Windows y Linux.

La comunicación entre la placa utilizada para la adquisición y digitalización de las señales se realiza mediante el puerto serie USB (Universal Serial Bus), que además sirve para alimentación eléctrica de la placa. Nuestro Osciloscopio, para poder graficar, recibe por el puerto serie tres paquetes.

1.3 Metodología de uso

Finalizado el desarrollo, para que el alumno pueda utilizar el instrumento en su hogar sólo deberá disponer de:

- Una computadora a la cual le instalará un archivo ejecutable,
- Una placa de desarrollo que posea un converso A/D de 8, 10 o 12 bits (por ejemplo, Arduino o similar) En caso de no disponer de una placa podrá adquirir alguna de bajo costo o solicitarla en el laboratorio de GIIoT de la facultad. En esta placa el alumno deberá grabar un código que le entregará el docente (atento que el desarrollo está pensado para estudiantes de ingeniería esta tarea no debería generar dificultades extras),
- El circuito que utilizará para desarrollar el trabajo práctico, previamente indicado por el docente, y cables de conexión.

Toda esta información será entregada al docente de la asignatura, junto con un instructivo o manual de uso, que contendrá además algunos ejemplos de trabajos prácticos que puede desarrollar.

De esta manera el docente podrá publicar la documentación necesaria en el aula virtual de la asignatura para que el alumno pueda acceder al mismo con la debida antelación, para leerlo y adquirir el material necesario para la realización de la práctica.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Este proyecto tiene como principal objetivo el desarrollo y construcción de un recurso educativo de formación práctica relacionado con los circuitos eléctricos y electrónicos para contribuir a la enseñanza centrada en el estudiante y basada en el logro de competencias. El punto de partida es el diseño de un laboratorio de escritorio portátil que resuma la instrumentación básica y necesaria para el desarrollo de prácticas experimentales guiada, su optimización y adaptación a la instrumentación virtual.

Teniendo en cuenta la motivación adicional que las nuevas tecnologías de la información y la comunicación (TICS) producen en forma natural en las actuales generaciones de estudiantes, se busca organizar nuevas estructuras de laboratorio que resulten estimulantes para la enseñanza aprendizaje del nuevo siglo.

El desarrollo del mismo requiere de tres ingenierías muy relacionadas, la electrónica, la informática y las comunicaciones, de esta manera el proyecto se divide en tres fases de desarrollo:

- Desarrollo electrónico, que provee los equipos o dispositivos con los cuales se ingresa, extrae, presenta o procesan datos,.
- Desarrollo de software, uno para la interfaz que permita leer datos directamente de la placa electrónica y el software para la computadora o PC que permitirá interactuar docentes con alumnos en forma amigable, simple y eficaz.
- Desarrollo de las comunicaciones, se deberán manejar los sistemas de comunicación entre la interfaz y la computadora del alumno y además todo el grupo de interfaces o laboratorios con el profesor.

Una segunda línea de investigación es el correcto uso de estas nueva TIC en la educación. La misma se está abordando desde un inicio por especialistas docentes de las cátedras a beneficiarse. Consiste en ir investigando y diseñando las clases que se propondrán. Para cada clase o tema se deberán

desarrollar: aprendizaje teórico, prácticas a realizar por el alumno y mecanismos de evaluación.

3. RESULTADOS OBTENIDOS / ESPERADOS

El diseño propuesto busca asemejarse lo más posible a un osciloscopio digital convencional, replicando las funciones más comunes de dicho instrumento. De esta manera se busca que los alumnos comiencen a familiarizarse con el uso de dicho instrumento para facilitar su posterior manejo, y realicen prácticos donde puedan visualizar, medir, analizar y guardar señales variables en el tiempo.

Respecto a los resultados alcanzados, podemos mencionar los trabajos y pruebas realizadas en años anteriores, con el laboratorio de escritorio, lo que marcó un inicio en el desarrollo del instrumento virtual.

En cuanto al instrumento virtual, en la figura 2 se muestra el desarrollo en el que el grupo está trabajando, el cual simula un osciloscopio de dos canales y permite utilizar diferentes sistemas embebido para la adquisición y conversión de la señal a medir.

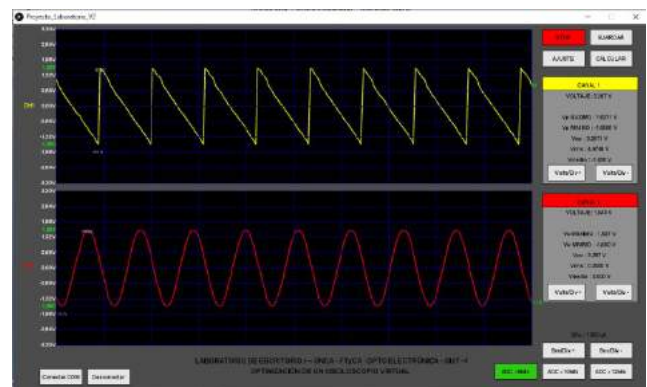


Figura 2. Primer prototipo de Osciloscopio virtual

Se realizaron diversas pruebas de funcionamiento utilizando para la adquisición de las señales placas electrónicas disponibles en el laboratorio, se capturaron diferentes ondas de un generador de señales comercial, señales autogeneradas y salidas de circuitos eléctricos sencillos, como por ejemplo un circuito RC. Se pretende que a partir de esta herramienta los alumnos puedan determinar por lectura directa el periodo y el valor del voltaje, y de manera indirecta la frecuencia de una

señal. Observar desfasaje entre dos señales y valores de DC y AC.

Este diseño, que constituye un primer prototipo de osciloscopio, funciona correctamente en la banda de frecuencia de señales de audio, e incluso hasta 30 kHz y que se adapta a diferentes placas de desarrollo, incluyendo la plataforma Arduino Uno, la cual es de bajo costo y fácil programación para los alumnos de los primeros años de las carreras.

La herramienta de programación utilizada permite generar un archivo ejecutable portable que corre en cualquier computadora con Sistemas Operativos Windows y Linux, sin necesidad de la instalación de ningún software. El desarrollo resulta de gran importancia dado que, con el objeto de plantear a los estudiantes situaciones similares a las de una práctica convencional en el Laboratorio, permite que docentes de diferentes asignaturas tengan la posibilidad de generar actividades prácticas de aprendizaje no presenciales que puedan realizar los alumnos desde sus hogares. Los resultados de estas experiencias pueden guardarse para un posterior análisis por parte de los alumnos, y el envío del mismo a los docentes para su evaluación.

Además se espera generar una ayuda en línea, que guíe a los alumnos y que posteriormente les ayude a interpretar los resultados obtenidos y en caso de que no sean los que debieran esperar, a

De esta manera la propuesta de desarrollo representará un aporte novedoso y primordial ya que va a ir más allá del mero diseño tecnológico y abarcará también el fundamental aspecto pedagógico de implementación para el desarrollo de las clases asegurando de que se cumplan los objetivos de mejoras del proceso de enseñanza aprendizaje.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto principalmente por docentes de las carreras Ing. Electrónica e Ing. en Informática de la Facultad de Tecnología y Ciencias Aplicadas, pertenecientes al grupo de Investigación en internet de las cosas (IIoT). Todos los años se agregan alumnos, incentivados a través de un

programa de becas que ofrece la Facultad para estudiantes investigadores.

La diversidad del equipo de investigación es muy importante. Hay docentes de diferentes asignaturas de las mencionadas carreras, en las áreas de física y análisis matemático, electrónica analógica, sistemas digitales, arquitectura de computadoras, instrumentación, control y redes, por lo que se cuenta con un conocimiento detallado de la problemática de la educación técnica en el área.

En cuanto a la formación de los integrantes, hay docentes formados en el área de educación, uno de los cuales posee Título de Especialista en Docencia Universitaria de Disciplinas Tecnológicas. Por otro lado, aquellos que poseen una vasta experiencia de divulgación y de formación en temas de desarrollo tecnológico con microprocesadores de última generación.

En función de ello, para el desarrollo del proyecto, el equipo se organiza en grupos de trabajo: a) del hardware y b) del software. Ambos grupos coinciden en las tareas de montaje y testeado del equipo y en el desarrollo de las herramientas educativas.

5. BIBLIOGRAFIA

- [1] Beltramini, P; Poliche, M; Aranda, M; D'Amore M et. all. (2018). "Aportes al desarrollo de competencias en Ingeniería utilizando un laboratorio portátil de escritorio". IV Congreso Argentino de Ingeniería – X Congreso Argentino de Enseñanza de la Ingeniería. Córdoba,
- [2] Cano, J; Poliche, M; Beltramini, P; Aranda, M et. all. (2017). "Laboratorio portátil de escritorio para la enseñanza de la electrónica". Revista Argentina de Ingeniería. Buenos Aires, v.9, n.5, p. 86-93.
- [3] (Chicala, 2015, "Adquisición de datos: Medir para conocer y controlar. Handbook de adquisición de datos" Primera Edición. Cengage Learning Editores, 2015.
- [4] Chacón Rugeles R.. "La Instrumentación Virtual en la Enseñanza de la Ingeniería Electrónica". Acción Pedagógica, Vol. 11, No. 1 / 2002, pp 74-84
- [5] National Instrumentos Corp. (2020) "La Instrumentación Virtual" National Instrumentos Corp., 2003.

- [6] Processing [online] Wikipedia. Available: <https://es.wikipedia.org/wiki/Processing>
- [7] Cortés Osorio J., Chaves Osorio J., Medina A.. “Diseño y Construcción de un Osciloscopio Digital implementado en Matlab” *Scientia et Technica* Año XIII, No 34, mayo de 2006. Universidad Tecnológica de Pereira. ISSN 0122-1701
- [8] Logicbus [online] <https://www.logicbus.com.mx/osciloscopio.php>
- [10] <https://www.arduino.cc/>
- [11] Proyecto CIAA: <http://www.proyecto-ciaa.com.ar/devwiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp>
- [12] <https://stm32f4-discovery.net/>
- [13] Manuel A., Biel D., y otros. “Instrumentación virtual: Adquisición, procesado y análisis de señales”. EDICIONS UPC. 2001.
- Aranguren G. *Nuevos métodos de enseñanza: una experiencia en diseño electrónico*. IEEE-RITA Vol. 3, Núm. 1 Páginas 39 a 46, Mayo 2008.
- Sánchez Ruvirosa R., *Enseñar y aprender con nuevos métodos. La revolución cultural de la era electrónica*, UAM-X – MÉXICO, pp. 321-326. 2002.
- Monge Nájera, J.; Méndez Estrada, V. H. *Ventajas y desventajas de usar laboratorios virtuales en educación a distancia: la opinión del estudiantado* Revista Educación, ISSN: 0379-7082, vol. 31, núm. 1, pp. 91-108 Universidad de Costa Rica San Pedro, Montes de Oca, Costa Rica. 2007.
- Ferreiro R., *Nuevos ambientes de aprendizaje*. INPUT. No. 21, Barcelona. 2006.
- La metodología del Aprendizaje Basado en Problemas, http://www.ub.edu/dikasteia/LIBRO_MURCIA.pdf, mayo, 2016

DESARROLLO Y EVALUACIÓN DE COMPETENCIAS EN LA INGENIERÍA DE SOFTWARE EN UN ENTORNO VIRTUAL DE APRENDIZAJE COLABORATIVO

Nicolás Battaglia, Carlos Neil, Marcelo De Vincenzi

Universidad Abierta Interamericana. Facultad de Tecnología Informática.
Centro de Altos Estudios en Tecnología Informática. Buenos Aires. Argentina
{nicolas.battaglia, carlos.neil, medevincenzi}@uai.edu.ar

RESUMEN

Los cambios constantes de las tecnologías emergentes nos hacen reflexionar sobre cuán efectivo es el método tradicional de enseñanza. Los graduados deben ser competentes; no solo deben conocer los saberes de un área de conocimiento, sino también aplicarlos en futuros contextos profesionales. Para impulsar el aprendizaje basado en competencias, en particular en la ingeniería de software, es importante utilizar herramientas y aplicaciones específicas propias de la disciplina que permitan el aprendizaje de manera contextualizada y ubicua. Por otro lado, es necesario adaptar los criterios y estrategias de evaluación motivados por esta disrupción que permitan evaluar competencias y no saberes aislados. Un doble desafío: diseñar una infraestructura para poder brindar servicios de educación mediada por la tecnología y coadyuvar para que los estudiantes puedan aprender en forma colaborativa optimizando al máximo el proceso formativo. En este proyecto se propone la integración de plataformas tecnológicas para el aprendizaje ubicuo colaborativo, basado en competencias, en cursos vinculados con la ingeniería de Software, en particular, durante el proceso de modelado de software.

Palabras clave: Aprendizaje Colaborativo, Aprendizaje Basado en Competencias, Ingeniería de Software, Trabajo Colaborativo, UML.

CONTEXTO

El proyecto de investigación se implementa en la facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI). Alumnos de la universidad desde 2° año hasta 5° año de la carrera de Ingeniería en Sistemas Informáticos, trabajan de manera articulada en las siguientes asignaturas relacionadas: 1) Metodologías de Desarrollo de Sistemas I y II (2° año). 2) Bases de Datos, Trabajo de

Diploma y Trabajo de Campo I (3° año). 3) Seminario de Aplicación Profesional y Trabajo Final de Ingeniería (5° año).

Materializando el proyecto se desarrolló una herramienta CASE que permite a estudiantes aprender e interactuar entre ellos con temas relacionados con la Ingeniería de Software (IS) y el modelado UML en un proyecto de software. Esto lo realizan de manera colaborativa y en forma evolutiva con un nivel de complejidad que se va incrementando en los sucesivos años conforme incrementan su nivel de competencia.

El proyecto, por otro lado, está en línea con el documento recientemente publicado por el Consejo Federal de Decanos de Ingeniería (CONFEDI) llamado “Propuesta de Estándares de Segunda Generación para la Acreditación de Carreras de Ingeniería en la República Argentina” [1], donde se incorporan las condiciones curriculares comunes sobre las Competencias de Egreso que un alumno debe acreditar para acceder a su graduación.

1. INTRODUCCIÓN

El 2020 será recordado como el “año de la pandemia”. En este contexto, la sociedad tuvo que reinventarse de manera obligada en casi todos los aspectos de la vida cotidiana.

Como se pudo observar durante el 2020, uno de los aspectos que más sufrió el impacto del COVID-19, sin duda, fue la educación. Las instituciones educativas tuvieron que reestructurarse en muy poco tiempo para poder continuar operando de manera virtual y así lograr, parcialmente, sus objetivos. Por medio de la tecnología y las comunicaciones, gran parte de los alumnos continuaron estudiando, pero de manera virtual. Esto presentó nuevos desafíos: formar a los docentes en esta nueva modalidad y, además, implantar y adaptar sistemas pedagógicos e informáticos que permitieron lograr estos objetivos, en general, de manera parcial. Si bien esta realidad aceleró el proceso de cambio, también es cierto que durante los últimos años la educación ha ido cambiando sus métodos de aprendizaje

centrado en el docente a métodos de aprendizaje centrados en el alumno [2].

Desde el comienzo del proyecto de investigación [3]–[5] afirmamos que el cambio y la evolución de la tecnología generaron la necesidad de repensar el sistema educativo (cuyas raíces se originaron hace cientos de años). Esta afirmación y el COVID-19 como catalizador, nos obliga a ahondar, aún más, en las implicancias de la tecnología en la educación.

Uno de los cambios más relevantes de los últimos tiempos es la capacidad actual de poder acceder a información de manera ubicua. Esto está vinculado, sin dudas, con la evolución en las Tecnologías de Información y Comunicación (TIC) [6], [7]. De esta manera, gracias a Internet y a los dispositivos móviles, el estudiante puede aprender de manera virtual, utilizar herramientas sincrónicas (videoconferencias, chat, etc.) y asincrónicas (foros y redes que promueven la interacción con otras personas), facilitándoles el aprendizaje colaborativo.

Por este motivo, el trabajo colaborativo, las TIC y la enseñanza se pueden integrar para crear entornos de aprendizaje colaborativo asistido por computadora. Esta integración se basa en el impacto social del trabajo colaborativo y las TIC en la educación que transforma la manera tradicional de enseñanza/aprendizaje. Esto conlleva varios cambios, entre ellos el Aprendizaje Basado en Competencias (ABC) y centrada en el alumno y el impacto de las TIC en los procesos de enseñanza y aprendizaje, para lo cual es necesario también modificar y adaptar los criterios y estrategias de evaluación motivados por esta disrupción [8].

Por un lado, el trabajo colaborativo apoyados en el concepto de Zona de Desarrollo Próximo (ZDP) propuesto por Vygotski a principios del siglo XX [9] y, por el otro, las TIC en el marco de las propuestas de *e-learning* o *blended learning*. Esto sumado al desarrollo creciente de las TIC junto con el concepto de trabajo colaborativo, conforman los entornos de Trabajo Colaborativo Asistido por Computadora (CSCW). Este concepto, integrado en entornos de enseñanza y aprendizaje colaborativo, dio origen a los entornos de Aprendizaje Colaborativos Asistidos por Computadora (CSCL) [10].

En el ámbito académico, los nuevos estándares para la acreditación de carreras de ingeniería en la República Argentina (Libro Rojo de CONFEDI) [1] que proponen el ABC y están centrado en el alumno, dan cuenta de

la necesidad de la integración de ambos modelos, además de considerar todas las herramientas de seguimiento y evaluación que sean requeridas por este paradigma (figura 1).



Figura 1. Conformación del modelo CSCL

Sin dudas, el estudio teórico de la Ingeniería de Software (IS) no es suficiente para comprender y resolver los problemas de cooperación y colaboración que surgen durante el desarrollo de un proyecto informático [11].

Los estudiantes suelen centrar su esfuerzo en aspectos técnicos y asumen que los inconvenientes vinculados al trabajo en equipo no impactarán en el proyecto. En virtud de resolver el problema planteado previamente, Daniele et al. [12] proponen, para mejorar los aspectos comunicacionales, la integración de las plataformas CSCL con entornos especializados en la resolución de problemas prácticos y técnicos en un proyecto de desarrollo de software; además, esta propuesta potencia el aprendizaje de la IS y permite el desarrollo de competencias profesionales y sistemáticas en el desarrollo de software [13].

De la misma manera, tal como sucede en los procesos tradicionales de enseñanza, en los entornos CSCL, la evaluación cumple un rol muy importante. Según el estudio realizado en [14], en el modelo de enseñanza y aprendizaje de la IS mediado por tecnología, los estudiantes cumplen un papel de mayor protagonismo y responsabilidad y, en consecuencia, los docentes asumen un nuevo rol cuando es el alumno quien adquiere mayor autonomía en su proceso de aprendizaje a través de entornos virtuales. Esta autonomía es una característica distintiva de la enseñanza centrada en el estudiante.

En este contexto, la evaluación es entendida como un proceso que promueve el aprendizaje con una finalidad formativa, más que como un proceso de control de resultados. Por

consecuente, un proceso específico de enseñanza y aprendizaje en un entorno virtual colaborativo requiere también de un proceso de evaluación, autoevaluación y coevaluación y seguimiento acorde al modelo de aprendizaje colaborativo planteado por el concepto CSCL [8]. El énfasis en la educación centrada en el estudiante permite retomar el modelo de ABC, que también promueve reformular y adaptar los criterios y estrategias de evaluación motivados por la disrupción tecnológica que da origen a los entornos CSCL.

En resumen, en este proyecto nos enfocamos en la enseñanza y el aprendizaje del modelado de software y utilizamos el estándar UML como lenguaje para especificar, construir, visualizar y documentar sistemas informáticos. Para esto proponemos el rediseño de un entorno CSCL específico para dicha área de conocimiento, con un enfoque centrado en el ABC.

2. uCASE-CL

Los modelos de aprendizaje colaborativo asistido por computadora están inmersos en un universo tecnológico ubicuo, basado en las TIC. Este tipo de modelo fue denominado uCSCL por Coto, et al [15], entre otros autores. El modelo CSCL está compuesto por tres dimensiones (Ciencias de la Computación, Psicología y Pedagogía), a éstas se agrega una cuarta dimensión que representa la tecnología subyacente que materializa el concepto de ubicuidad.

La propuesta de uCASE-CL [6], [8], [16] se basa en la idea de ampliar el modelo uCSCL mencionado para contar con un bloque funcional de evaluación integrando las herramientas de enseñanza y aprendizaje necesarias para áreas de conocimiento específico, como sucede con la IS y el modelado UML.

El modelo uCASE-CL propuesto inicialmente no consideraba las herramientas necesarias para poder adoptar el sistema de ABC. Si bien estaba diseñado para centrar la educación en el alumno, no permitía definir competencias, resultados de aprendizaje y tampoco brindaba herramientas para obtener indicadores particulares sobre la evolución del aprendizaje y las competencias vinculadas.

Por lo tanto, este modelo amplía el objetivo principal incorporando diferentes técnicas que permitan integrar la educación basada en competencias dentro de entornos virtuales colaborativos que, además, brinden información en tiempo real sobre la evolución

de los resultados de aprendizaje vinculados con las competencias genéricas y específicas.

3. APRENDIZAJE BASADO EN COMPETENCIAS

La universidad debe preparar al estudiante en un entorno de competencias genéricas y específicas que les permitan afrontar los problemas que plantea esta sociedad cambiante.

McClelland [17], en la búsqueda de una alternativa a las pruebas de aptitud e inteligencia tradicionales, desarrolló el concepto de “competencia” definido como una característica subyacente de una persona que le permite demostrar un desempeño superior en un determinado rol o situación, haciendo la diferencia entre personas con desempeño excelente versus personas con desempeño promedio.

Por otro lado, el concepto de competencia engloba todo un conjunto de conocimientos, procedimientos y actitudes que se combinan, coordinan e integran, para que el individuo logre «saber conocer», «saber hacer», «saber ser» [18]. El dominio de estos saberes le permiten ser «capaz de» actuar con eficacia y eficiencia en situaciones que se le presenten [19]. En este sentido, la competencia no reside en los recursos (capacidades) sino en la movilización misma de los recursos. Para ser competente es necesario poner en juego el repertorio de recursos. Además, el aprendizaje por competencias ha promovido toda una renovación de las teorías psicopedagógicas de los aprendizajes, dando lugar a la evolución de los esquemas de referencia de la formación de profesionales [19].

En esta línea, Tobón [18] propone concebir las competencias como: “Procesos complejos de desempeño con idoneidad en determinados contextos, integrando diferentes saberes (saber ser, saber hacer, saber conocer y saber convivir), para realizar actividades y/o resolver problemas con sentido de reto, motivación, flexibilidad, creatividad, comprensión y emprendimiento, dentro de una perspectiva de procesamiento metacognitivo, mejoramiento continuo y compromiso ético, con la meta de contribuir al desarrollo personal, la construcción y afianzamiento del tejido social, la búsqueda continua del desarrollo económico-empresarial sostenible, y el cuidado y protección del ambiente y de las especies vivas”.

4. EVALUACIÓN Y COMPETENCIAS

Tal como planteó Perrenoud [20], las competencias son invisibles y sólo son abordables a través de los desempeños observables, por lo tanto, en línea con Zapata [21] éstas deben ser observables y evaluables y esto sucede, básicamente, por medio de las conductas que los sujetos generan ante diversos problemas en contextos cambiantes. Esto plantea la necesidad de la sustitución o coexistencia de procedimientos, herramientas y formas de evaluar los saberes por medio de procesos descriptivos o conceptuales (exámenes, pruebas de destreza, problemas teóricos, etc.) por otros métodos donde se tenga en cuenta la simulación real de casos prácticos, con metodologías de indagación, de investigación formativa, con elaboración de proyectos, etc. Todo esto bajo la observación del docente sobre aquellos aspectos de desenvolvimiento personal del alumno en la ejecución de tareas y resolución de problemas.

Como consecuencia, el ABC requiere un sistema de evaluación variado, debido a que cada competencia tiene componentes muy distintos que necesitan procedimientos diversos para ser evaluados correctamente [22]. La evaluación por competencias se basa en el acceso a fuentes múltiples y variadas de información con el fin de determinar si los estudiantes han alcanzado el nivel esperado de desarrollo de competencias, así como un grado suficiente de dominio de los recursos vinculados a cada competencia [21].

Si bien existen diversas metodologías de evaluación de competencias [23], en la actualidad uno de los mejores instrumentos, son las rúbricas, debido a su facilidad de uso y a la oportunidad de obtener aspectos complejos, imprecisos y subjetivos [2], [14].

Por otro lado, los entornos CSCL amplían las posibilidades de trabajo grupal sincrónico y asincrónico, con la consecuente necesidad de realizar la evaluación bajo el mismo esquema.

Según Zapata [24], cuando se evalúa en un entorno virtual de aprendizaje, es necesario también la comunicación de los resultados que ayudan al estudiante a comprender el nivel de adquisición de una competencia determinada. En estos entornos, es fundamental el concepto tradicional de retroalimentación o *feedback* propuesto por los entornos virtuales. Sin embargo, en los entornos virtuales colaborativos, es necesario medir también el nivel de participación de un estudiante dentro de un equipo de trabajo, por ejemplo, por

medio de la cantidad de mensajes enviados, tomando en cuenta aspectos semánticos y de relevancia en cuanto a cómo son estos mensajes, su influencia en otros estudiantes, si son respondidos, etc.

Por otro lado, un aspecto a tener en cuenta en la evaluación en los entornos virtuales está vinculado con la dificultad, por un lado, de garantizar la acreditación de la autoría del alumno como, así también, identificar si las competencias manifestadas corresponden con el individuo que se evalúa [25]. Esto implica un desafío en el diseño y utilización de métodos e instrumentos que contemplen los aspectos anteriormente planteados en la evaluación competencias en entornos CSCL.

En esta línea, Heberling y Flint [26] afirman que el mantenimiento de la integridad académica es un reto en ambas modalidades: la tradicional y en línea. Además, amplían el concepto explicando que, si bien es imposible erradicar el engaño por completo, se puede minimizarlo concientizando al comienzo de cada curso las implicancias del fraude en cuanto a sanciones y resultados académicos.

5. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente trabajo está identificado como proyecto de investigación y desarrollo en el Centro de Altos Estudios en Tecnología Informática (CAETI), dentro de la línea de investigación Sociedad del Conocimiento y Tecnologías aplicadas a la educación.

Además, el proyecto se ramifica en diferentes áreas temáticas específicas, a saber: 1) Desarrollo y evaluación de competencias en la ingeniería de software en entornos virtuales de aprendizaje colaborativo, 2) Aprendizaje por competencias y computación cognitiva, 3) Analítica de aprendizaje y análisis de datos. 4) Enseñanza de IS a personas con capacidades limitadas. 5) Estudios empíricos sobre modelos de enseñanza de aprendizaje. 6) Estilos de aprendizaje y personalización de la enseñanza y, por último, 6) Desarrollo de herramientas CASE integrada con los conceptos de aprendizaje basado en competencias.

6. RESULTADOS OBTENIDOS/ESPERADOS

i. Resultados obtenidos:

1) Desarrollo, diseño e implementación del prototipo de una herramienta CASE

colaborativa ubicua multiplataforma, denominada UAI Case. 2) Diseño de una plataforma académica colaborativa multiplataforma para evaluación, seguimiento, interacción y coordinación de proyectos informáticos. 3) Definición y especificación de los bloques funcionales necesarios para determinar un proceso específico en la enseñanza y aprendizaje de modelados en la IS.

ii. Objetivos Futuros / Resultados esperados:

1) Diseñar herramientas que permitan integrar el concepto de ABC junto con herramientas que faciliten evaluar y realizar seguimientos en entornos colaborativos (UAI Case). 2) Generar indicadores en tiempo real que permitan medir la evolución de los saberes y competencias en el ABC. 3) Obtener un conjunto de indicadores que permitan medir el rendimiento de un equipo virtual de trabajo durante el proceso de enseñanza y aprendizaje de la IS. 4) Realizar la evaluación empírica del prototipo de la herramienta UAI Case por medio de la definición de un método de evaluación basado en métricas. 5) Desarrollar sistemas de tutorías inteligentes mediante la implementación de tutores cognitivos que imiten el rol del profesor, guíen el desarrollo del aprendizaje y ofrezcan pistas a los estudiantes cuando están atascados en un problema. 6) Utilizar analíticas del aprendizaje que permitan, mediante la medición, recopilación y análisis de datos de los estudiantes durante el proceso de enseñanza aprendizaje, detectar los errores más comunes y proporcionen una respuesta en tiempo real al estudiante.

7. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado, además de los integrantes que encabezan este trabajo, por 9 estudiantes de posgrado (doctorado en ciencias informáticas, maestría en tecnología informática y especialización en ingeniería de software) quienes están desarrollando sus respectivas tesis y trabajos finales en temas vinculados al proyecto: Fernando Parra, Jorge Lomoro, Silvia Poncio Gabriela Iannantuoni, Jorge Zarate, Marcelo Monferrato, Juan Facundo Oliva, Charles Maldonado, Nelson Garrido.

8. BIBLIOGRAFÍA

- [1] CONFEDI, *Propuesta de estándares de segunda generación para la acreditación de carreras de ingeniería en la República Argentina "Libro Rojo de CONFEDI"*. 2018.
- [2] A. Schiter, N. Battaglia, and C. Neil, "El Porqué de las Rúbricas para la Evaluación y la Autoevaluación," *Congr. Nac. Ing. Informática - Sist. Inf. CONAIISI*, 2018.
- [3] C. Neil, M. De Vincenzi, N. Battaglia, and R. Martínez, "Herramientas Colaborativas Multiplataforma en la Enseñanza de la Ingeniería de Software," *XVIII Work. Investig. en Ciencias la Comput. (WICC 2016, Entre Ríos, Argentina)*. 2016.
- [4] M. De Vincenzi, C. Neil, N. Battaglia, and R. Martínez, "Uaicase: Enseñanza De Ingeniería De Software En Entornos Virtuales Colaborativos," *XX Work. Investig. en Ciencias la Comput. (WICC 2018, Univ. Nac. del Nord.*, pp. 1124–1129, 2018.
- [5] N. Battaglia, C. Neil, M. De Vincenzi, R. Martínez, and J. P. Beltramino, "UAI case: desarrollo y evaluación de competencias en la ingeniería de software en un entorno virtual de aprendizaje colaborativo," *XXI Work. Investig. en Ciencias la Comput. (WICC 2019, Univ. Nac. San Juan)*., 2019.
- [6] N. Battaglia, C. Neil, R. Martínez, and M. De Vincenzi, "UAI case: An uCASE-CL model implementation," *J. Technol. Sci. Educ. (i-CITE). Senai, Johor, Malaysia.*, vol. 9, no. 1, p. 96, 2019, doi: 10.3926/jotse.543.
- [7] G. Lavigne, M. Ovando, J. Sandoval, and L. M. Salas, "Exploración preliminar del aprendizaje colaborativo dentro de un entorno virtual," *Actual. Investig. en Educ.*, 2012.
- [8] N. Battaglia, "Integración de una Herramienta CASE en un Entorno Académico Colaborativo para la Enseñanza de Ingeniería de Software," Universidad Abierta Interamericana, 2017.
- [9] L. Vygotsky, "The collected works of LS Vygotsky: Problems of the theory and history of psychology," 1997, Accessed: Dec. 12, 2020.
- [10] G. Stahl, T. Koschmann, and D. Suthers, "Computer-supported collaborative learning: An historical

- perspective,” 2006. Accessed: Dec. 21, 2020.
- [11] P. Bouillon, J. Krinke, and S. Lukosch, “Software engineering projects in distant teaching,” 2005, doi: 10.1109/CSEET.2005.31.
- [12] M. Daniele, M. Uva, P. Martelloto, and G. Picco, “Aplicación de herramientas CASE a la enseñanza de Ingeniería de Software: Gestión de la Configuración de Software y Testing Funcional,” 2010.
- [13] N. Battaglia, C. Neil, M. De Vincenzi, and J. P. Beltramino, “Competency-based learning in collaborative virtual platforms,” *Int. Inst. Innov. Technol. 10th IITEC Int. Symp. Innov. Technol. ISIT2019. Cuzco, Perú.*, 2019.
- [14] N. Battaglia, R. Martínez, M. Otero, C. Neil, and M. De Vincenzi, “Autoevaluación Colaborativa por medio de Rubricas en Entornos Virtuales de Enseñanza y Aprendizaje,” 2016.
- [15] M. Coto, C. A. Collazos, and S. Mora-Rivera, “Modelo Colaborativo y Ubicuo para apoyar los procesos de enseñanza-aprendizaje a nivel Iberoamericano,” *Rev. Educ. a Distancia*, 2016, doi: 10.6018/red/48/10.
- [16] N. Battaglia, C. Neil, M. De Vincenzi, R. Martínez, and D. González, “UCASE - CL: aprendizaje colaborativo de la ingeniería de software en entornos virtuales ubicuos,” in *XII Congreso de Tecnología en Educación y Educación en Tecnología*, 2017, pp. 439–452.
- [17] D. C. McClelland, “Testing for competence rather than for ‘intelligence’,” *Am. Psychol.*, 1973, doi: 10.1037/h0034092.
- [18] S. Tobón, “La Formación Basada en Competencias en la Educación Superior: El enfoque complejo,” 2009.
- [19] J. T. Fernández and C. R. Bueno, “Evaluación de competencias profesionales en educación superior: Retos e implicaciones,” *Educ. XXI*, 2016, doi: 10.5944/educXX1.12175.
- [20] P. Perrenaud, *Construir competencias desde la escuela*. JC Sáez, 2008.
- [21] A. Fernández March, “La evaluación orientada al aprendizaje en un modelo de formación por competencias en la educación universitaria,” *REDU. Rev. Docencia Univ.*, 2011, doi: 10.4995/redu.2010.6216.
- [22] A. Villa Sánchez and M. Poblete Ruiz, *Aprendizaje Basado En Competencias Una Propuesta Para La Evaluación*. 2007.
- [23] J. Cubero-Ibáñez, M. S. Ibarra-Sáiz, and G. Rodríguez-Gómez, “Propuesta metodológica de evaluación para evaluar competencias a través de tareas complejas en entornos virtuales de aprendizaje,” *Rev. Investig. Educ.*, vol. 36, no. 1, pp. 159–184, 2018, doi: 10.6018/rie.36.1.278301.
- [24] M. Zapata, “Evaluación de competencias en entornos virtuales de aprendizaje y docencia universitaria,” *Revista de Educación a Distancia. Sección de Docencia Universitaria en la Sociedad del Conocimiento.*, vol. 1, no. 1DU, 2010.
- [25] J. Bailie, J. L. Bailie, and M. A. Jortberg, “Online Learner Authentication: Verifying the Identity of Online Users,” *MERLOT J. Online Learn. Teach.*, vol. 5, no. 2, 2009, Accessed: Dec. 26, 2020.
- [26] M. Heberling and M. Flint, “Maintaining academic integrity in online education,” *Online J. Distance Learn. Adm.*, vol. V, no. 1, pp. 1–6, 2002.

Diseño y construcción de objetos de aprendizaje basada en criterios de calidad

Fernanda CARMONA¹, Antonio CASTRO LECHTALER^{1,2}, José TEXIER, Marisa GAGLIARDI¹, Emmanuel PORTUGAL¹, Patricia MANRIQUES¹, José SORIA¹

¹Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito 9 de Julio 22, Chilecito, La Rioja, Argentina; ²CISTIC/FCE - Universidad de Buenos Aires, Córdoba 2122, CABA

{fbcarmona, jtexier, mgagliardi, eportugal}@undec.edu.ar, antonio.castro@fce.uba.ar, patriciamanriques311995@gmail.com, soriaajose@gmail.com

RESUMEN

La incorporación de las Tecnologías de la Información y Comunicación (TIC) en la Educación ha permitido extender los ambientes de enseñanza y aprendizaje, para así poder desarrollar ambientes virtuales colaborativos e interactivos, con el uso de diferentes materiales educativos digitales que permitan el logro de los objetivos.

Adicionalmente, En los últimos años, el desarrollo de Repositorios Institucionales (RI) de acceso abierto ha sido un tema prioritario en las políticas de educación, ciencia y técnica de muchos países, y en particular, en las universidades públicas de Argentina, aprobándose la Ley 26899 de Repositorios Institucionales (RI), y en los otros niveles educativos las Bibliotecas Escolares con la aprobación de la Ley 26.917 sobre el Sistema Nacional de Bibliotecas Escolares y Unidades de Información Educativa.

Esta línea de I+D+i corresponde al diseño, desarrollo e implementación de proyectos que fortalezcan la investigación y las diferentes labores relacionadas con la gestión del conocimiento en cuanto a la visibilidad web y la preservación de la producción educativa y académica de la UNdeC y de las instituciones intervinientes.

Dentro de esta línea de trabajo se encuentran el desarrollo de Repositorios Institucionales (RI), Objetos de Aprendizaje (OA) y el movimiento de Acceso Abierto (AA); tanto en lo que respecta a los aspectos técnicos vinculados a ellos, como a su

utilización y uso como productos útiles en el proceso de enseñanza – aprendizaje en los diferentes niveles educativos.

Palabras clave: repositorios institucionales, objetos de aprendizaje, material educativo digital, acceso abierto.

CONTEXTO

Actualmente la UNdeC se encuentra trabajando sobre una política de fortalecimiento de las carreras en su modalidad presencial con el fin de poder llegar a lugares más distantes de la región, comenzando paso a paso a transitar el camino de la virtualidad con el desarrollo de nuevas carreras. A partir del año 2013 se implementó el campus virtual institucional en <http://campus.undec.edu.ar>, sobre la plataforma E-ducative, con aulas virtuales como soporte a la enseñanza presencial. Se crea el área de Tecnologías de la Información y la Comunicación en la Educación, que tiene a cargo la implementación de los servicios de educación a distancia, campus virtual UNdeC, sistemas de videoconferencias y capacitación y soporte para la inclusión de TIC en propuestas pedagógicas. Modalidad que se fue fortaleciendo durante todo el año 2020 y que posibilitó, durante del aislamiento social y preventivo impuesta por la pandemia por COVID-19, que los estudiantes pudieran mantener el cursado con regularidad, inclusive posibilitando que muchos de

ellos pudieran egresar con su título universitario.

En el año 2018 se inician las gestiones y el trabajo para la implementación del Repositorio Institucional adquiriendo durante el primer semestre 2019, un servidor de gran porte para tal fin.

En este marco esta línea de I+D+i corresponde al diseño, desarrollo e implementación de proyectos que fortalezcan la investigación y las diferentes labores relacionadas con la gestión del conocimiento en cuanto a la visibilidad web y la preservación de la producción educativa y académica de la UNdeC y de las instituciones intervinientes.

La Resolución Rectoral UNdeC N° 155/20 aprueba el financiamiento y ejecución del proyecto “Metodología para el diseño y construcción de objetos de aprendizaje basada en criterios de calidad” presentado en la convocatoria FICyT UNdeC (Financiamiento para estímulo y desarrollo de la Investigación Científica y Tecnológica), período de ejecución 2020-2022. El proyecto propone una metodología para el diseño y construcción de OA, que incluya técnicas de evaluación de calidad orientado a realizar el proceso de evaluación desde diferentes enfoques, definiendo un conjunto de características que debe cumplir el objeto, asociadas a diferentes visiones y actores que intervienen en el proceso de análisis, construcción y evaluación. Esta temática forma parte del proyecto de tesis de maestría de uno de los integrantes del equipo de trabajo que vienen trabajando en la UNdeC en colaboración con otras instituciones del país y del extranjero a través de los proyectos “Red para la Integración de Universidades en el uso de TIC para la Inclusión en la Educación Superior” aprobado en la 7 Convocatoria a Redes Internacionales, año 2013, el proyecto “Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales” aprobado en la 9 Convocatoria a

Redes Internacionales, 2016-2017, y el Proyecto “Repositorios Digitales con Contenidos Orientados a las Necesidades de Escuelas Rurales (ER)” por Resolución CE N° 1055/15 convocatoria de Proyectos de Desarrollo Tecnológico y Social (PDTs), 2016-2018, presentado en el marco del Plan de Fortalecimiento de la Investigación Científica, el Desarrollo Tecnológico y la Innovación en las Universidades Nacionales (Ac. Pl. N° 676/08 y N° 687/09) y del Reglamento PDTs-CIN (Ac. Pl. N° 901/14), del Consejo Interuniversitario Nacional y el CONICET.

En el marco de esta línea de investigación en diciembre de 2020 se defendió y aprobó un Trabajo Final para la obtención del título de grado Licenciado/a en Sistemas denominado “Sistema de evaluación de métricas de calidad de Objetos de Aprendizaje”.

INTRODUCCIÓN

La incorporación de las Tecnologías de la Información y Comunicación (TIC) en la Educación ha permitido extender los ambientes de enseñanza y aprendizaje, para así poder desarrollar ambientes virtuales colaborativos e interactivos, con el uso de diferentes materiales educativos digitales que permitan el logro de los objetivos. Es por ello, que en el ámbito educativo se ha impuesto un nuevo concepto que busca la reutilización, permanencia, interoperabilidad, accesibilidad y compatibilidad de recursos digitales para el desarrollo de cursos y programas de formación en línea a través de la Web, los llamados Objetos de Aprendizaje (OA). Lo que distingue a un OA de un material educativo digital es la introducción de información autodescriptiva, expresada a través de los metadatos, conjunto de atributos o elementos que permiten describir al objeto, es decir son entidades generalmente entregadas a través de Internet y diseñadas con el fin de que sean utilizadas y reutilizadas en múltiples contextos educativos [1] [2].

En los últimos años, el desarrollo de Repositorios Institucionales (RI) de acceso abierto ha sido un tema prioritario en las políticas de educación, ciencia y técnica de muchos países, y en particular, en las universidades públicas de Argentina, aprobándose la Ley 26899 de Repositorios Institucionales (RI), y en los otros niveles educativos las Bibliotecas Escolares con la aprobación de la Ley 26.917 sobre el Sistema Nacional de Bibliotecas Escolares y Unidades de Información Educativa.

Los RI tienen como propósito recopilar, catalogar, gestionar, acceder, difundir y preservar información, permitiendo entre otras cosas el acceso libre y gratuito a todos los recursos que los conforman [3] [4].

El desarrollo de Repositorios Institucionales (RI) de acceso abierto ha sido un tema prioritario en las políticas de educación, ciencia y técnica de muchos países [5]. Un RI puede integrar a los OA en los denominados Repositorios de Objetos de Aprendizaje (ROA), bibliotecas digitales especializadas, orientados a facilitar la búsqueda y recuperación de los OA de manera que puedan ser utilizados en diversos ambientes educativos [6].

Sin embargo, no es suficiente con que estos materiales educativos estén disponibles, se requiere que cumplan con un alto nivel de calidad de modo que permita mayores posibilidades de lograr los objetivos de aprendizaje esperados en los estudiantes [7]. El factor más importante para alcanzar el éxito en la educación es el grado con el cual los docentes son capaces de producir actividades estructuradas, con una buena aplicación de la tecnología. Aquí es donde entran a jugar un papel importante los contenidos y su diseño en forma de OA.

Para optimizar la creación de los OA se han diseñado varias metodologías que apoyan su proceso de desarrollo y construcción [8]. Si bien se han propuesto enfoques que permiten evaluarlos teniendo en cuenta diferentes criterios y técnicas, no se cuenta con

modelos de evaluación generalizados que permitan identificar el nivel de calidad de estos recursos integrando diferentes dimensiones [7] [9].

La mayoría de las estrategias de evaluación propuestas se enfocan en pocos criterios y están orientadas a un conjunto específico de OA, son pocos los modelos de evaluación que integren evaluaciones automáticas con respuestas de usuarios y expertos, que permitan evaluar aspectos visuales, de contenido y pedagógicos, entre otros [9]. Estos involucran diferentes usuarios relacionados con los OA almacenados en repositorios y los resultados sirven para ordenar los resultados de búsqueda, es decir que se perfilan como herramientas de gestión de la calidad de OA publicados en repositorios.

En esta línea de I+D+i se propone una metodología para la construcción de OA, que está orientada a realizar el proceso de evaluación de calidad desde diferentes enfoques y durante el desarrollo de cada una de sus fases. Es decir, una metodología que se convierta en referente a la hora de planear y desarrollar OA aportando al proceso una organización interna, la identificación y selección de las competencias a desarrollar, así como el tipo de actividades cognitivas y de evaluación, garantizando la calidad de los OA construidos mediante la implementación de técnicas de evaluación de calidad tanto en el diseño instruccional y como en el diseño tecnológico de los OA en las diferentes fases de su desarrollo.

LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Objetos de aprendizaje.
- Repositorios institucionales y bibliotecas digitales.
- Gestión de la información y el conocimiento.
- Sistemas de información web y bases de datos.
- Índices bibliométricos.
- Interoperabilidad.

- Preservación digital.
- Recuperación de la información.

OBJETIVOS

El desarrollo de esta línea de investigación está permitiendo cumplir con los siguientes objetivos:

- Comprender los problemas y las tendencias asociadas a la producción abierta del conocimiento, el uso y la difusión.
- Fortalecer la visibilidad web de la UNdeC y de las instituciones intervinientes.
- Fomentar la formación de habilidades y conocimientos relacionados con los OA y los RI.
- Caracterizar las diferentes estrategias y técnicas de evaluación de calidad de los OA.
- Definir las fases, criterios de calidad y técnicas de evaluación de calidad de para la construcción de OA.
- Desarrollar una metodología para el diseño y construcción de OA basada en criterios de calidad.
- Entender y valorar el movimiento mundial de Acceso Abierto y señalar las vías de su materialización.
- Implementar un Repositorio Institucional para visualizar los diferentes recursos educativos y académicos que se produzcan.
- Depositar y preservar los producción educativa y académica de la UNdeC y de las instituciones intervinientes.
- Definir e implementar políticas de recopilación, distribución y mantenimiento para el funcionamiento del RI.

RESULTADOS

OBTENIDOS/ESPERADOS

- Curso a nivel de postgrado llamado “Gestión de la Información Académica y Científica”.
- Desarrollo de un proyecto de asignaturas sobre las Bibliotecas Digitales para

presentarse en dos universidades [10] y otras que lo puedan requerir.

- Desarrollo de talleres para la formación de habilidades y conocimientos relacionados con los RI, los OA y su construcción.
- Caracterización de las diferentes estrategias y técnicas de evaluación de calidad de los OA.
- Desarrollo de una metodología de construcción de OA que incluya técnicas de evaluación de calidad.
- Implementación de un repositorio de para la gestión los recursos educativos y académicos de la UNdeC y de las instituciones intervinientes.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por docentes investigadores categorizados y otros en formación especializados bibliotecas digitales, repositorios institucionales y en enseñanza en entorno virtuales de aprendizaje y estudiantes de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas.

Tres de los miembros del equipo poseen formación de postgrado: un Doctor en Ciencias Informáticas, un Magister en Informática y un Especialista en Entornos Virtuales de aprendizaje. Dos integrantes se encuentran en la etapa de desarrollo de su trabajo de tesis para obtener el título de Magister en Informática uno y otro de Magister en Enseñanza en Entornos Digitales. Tres miembros están cursando la Especialización en inteligencia de datos orientada a Big Data.

En diciembre de 2020 una de las integrantes del equipo defiende y aprueba su Trabajo Final para la obtención del título de grado Licenciado/a en Sistemas, denominado “Sistema de evaluación de métricas de calidad de Objetos de Aprendizaje”, con la dirección y asesoramientos de los integrantes de este equipo de trabajo. Otro de

los integrantes es estudiante avanzado de la carrera Licenciatura en Sistemas.

BIBLIOGRAFÍA

- [1]. D. Wiley, “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos, SPEDECE, 2001.
- [2]. E. Morales, F. García, A. Barron, A. Berlanga, and C. López, “Propuesta de Evaluación de Objetos de Aprendizaje,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE), Barcelona, 2005.
- [3]. J. Texier, “Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior,” presented at the 11th Latin American and Caribbean Conference for Engineering and Technology - 2013, Cancun, México, 2013, p. 9.
- [4]. M. De Giusti, N. Oviedo, A. Lira, A. Sobrado, J. Martínez, and A. Pinto, “SEDICI – Desafíos y experiencias en la vida de un repositorio digital,” *RENATA*, vol. 1, no. 2, pp. 16–33, Aug. 2011.
- [5]. Texier, J. (2016). Los repositorios institucionales y su importancia en las Universidades Venezolanas. *Revista Scitus*. Venezuela.
- [6]. R. McGreal, “A Typology of Learning Object Repositories”, In *Handbook on Information Technologies for Education and Training* (pp.5–28).doi:10.1007/978-3-540-74155-8_1. 2008
- [7]. V. Morales Tabares, “Modelo por Capas para Evaluación de la Calidad de Objetos de Aprendizaje en Repositorios de Objetos de Aprendizaje”, Instituto de Sistemas y Ciencias de la Decisión. 2013
- [8]. C. Sanz, L. Moralejo, & F. Barranquero, Curso de Doctorado “Metodología CROA”, Universidad Nacional de la Plata, Argentina, 2014
- [9]. V. Morales Tabares, N. Duque Méndez & D. Ovalle Carranza, “Modelo por capas para evaluación de la calidad de Objetos de Aprendizaje en repositorios”. *Revista electrónica de investigación educativa*, 19(3), 33-48. 2017.
- [10]. J. Texier, J. Zambrano, & F. Carmona, “Las Bibliotecas Digitales en el Currículum de las carreras de Ciencias de la Computación: una propuesta posible para Argentina y Venezuela”, XII Congreso Argentino de Ciencias de la Computación CACIC 2016.

Diseño y desarrollo de herramientas y entornos digitales para escenarios educativos híbridos

Sanz Cecilia^{1,3}, Madoz Cristina¹, Gorga Gladys¹, Gonzalez Alejandro¹, Zangara Alejandra¹, Iglesias Luciano¹, Ibáñez Eduardo¹, Violini Lucía^{1,2}, Fachal Adriana¹, Archuby Federico^{1,2}, Abásolo María José¹, Manresa-Yee Cristina⁴, Pesado Patricia¹

¹Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC.

Facultad de Informática – Universidad Nacional de La Plata

²Becario/a UNLP

³Investigador Asociado de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

⁴Departamento de Ciencias Matemáticas e Informática. Universidad de las Islas Baleares, España

{csanz, cmadoz, ggorga, agonzalez, li, eibanez, lviolini, farchuby, mjabasolo, ppesado}@lidi.info.unlp.edu.ar, alejandra.zangara@gmail.com, afachal@hotmail.com, cristina.manresa@uib.es

RESUMEN

Este trabajo presenta y describe algunas de las líneas de investigación y desarrollo del subproyecto “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos”, correspondiente al Instituto de Investigación en Informática LIDI. Se detallan más específicamente los avances y resultados del año 2020, e inicios del 2021.

Este subproyecto aborda la investigación, desarrollo e innovación en el área de tecnologías digitales para escenarios educativos. Se estudian metodologías y herramientas para el diseño y desarrollo de diferentes sistemas de interés para escenarios educativos, en los que se entran diferentes dispositivos, contextos, formas de acceso y dinámicas de trabajo. Además, se participa fuertemente en la formación de recursos humanos en el área y en la cooperación con otras universidades del país y del exterior.

Palabras clave: materiales educativos digitales, entornos digitales para el aprendizaje, juegos serios, herramientas colaborativas

CONTEXTO

Este subproyecto llamado “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos” forma parte de un proyecto más general titulado: “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” (período 2018-2021), perteneciente al Instituto de Investigación en Informática LIDI, de la Facultad de Informática de la Universidad Nacional de La Plata y acreditado por el Ministerio de Educación de la Nación. En el marco de este proyecto se trabaja interdisciplinariamente y se llevan adelante diferentes cooperaciones.

1. INTRODUCCION

El subproyecto se estructura en diferentes ejes temáticos en el que se tienen en común la investigación en metodologías y herramientas para:

- A. El diseño y desarrollo de entornos para el escenario educativo, y materiales hipermediales para el estudio, con diferente granularidad y paradigmas.
- B. La mediación y el seguimiento de actividades colaborativas en escenarios educativos, con indicadores y visualización del proceso subyacente.
- C. La creación de juegos serios educativos con diferentes paradigmas de interacción que combinan el mundo físico y digital.

Estos ejes son de actualidad en la agenda de investigación nacional e internacional. Como es sabido el aprendizaje no solo se enmarca en las interacciones que ocurren cuando los estudiantes y docentes se encuentran en el marco de una clase, en espacios y tiempos específicos (Carvalho and Yeoman, 2021). En los procesos educativos cada vez toman más relevancia la variedad de actividades que se llevan adelante, que implican el uso de diferentes ecosistemas de herramientas y experiencias con tecnologías digitales, en diferentes entornos y entramados sociales (Dabbagh & Fake, 2017; Osorio, 2010). En el marco de estas actividades se combinan instancias de reflexión personal con tareas de colaboración.

Es por ello que bajo estas consideraciones, se investigan en este proyecto diferentes diseños de materiales educativos digitales, que favorecen la reflexión, y la apropiación por parte del estudiante de los contenidos y habilidades que se le proponen. Se está trabajando en el tema de diseño de objetos de aprendizaje, en frameworks para su diseño y desarrollo y metodologías específicas para su creación (Violini,

Sanz & Pesado, 2020). Se aborda también el concepto de granularidad de estos materiales, y su estandarización (Allen & Mugisa, 2010) para la integración y recuperación desde distintos entornos tecnológicos, y su ensamblaje para la conformación de itinerarios. Para su recuperación y para la creación de itinerarios se estudian sistemas recomendadores y ensambladores, para ellos se trabaja con repositorios estructurados, entornos virtuales de enseñanza y aprendizaje y plataformas de MOOC (Gasparetti, De Medio, Limongelli, Sciarrone, Temperini, 2018; Astudillo, Sanz & Santacruz, 2017). Aquí un tema relevante es la metaanotación de los materiales educativos (Mouriño-García, Pérez-Rodríguez, Anido-Rifón, Fernández-Iglesias, & Darriba-Bilbao, 2018), que permiten la aplicación de diferentes tipos de sistemas recomendadores

Se está realizando, además, una tesis de maestría que propone un prototipo de simulador de entrevistas laborales como herramienta de formación inclusiva para adultos postlocutivos que utilizan ayudas auditivas conforme a los principios de accesibilidad.

También se trabaja en el diseño de entornos orientados a la enseñanza, el aprendizaje y la colaboración. Se abordan el diseño de entornos 3D con avatares, e investigaciones sobre indicadores de mediación de la colaboración que pueden aportar a la autorregulación del estudiante y del grupo (Dieser, Sanz, Zangara, 2019). Se aplican técnicas y herramientas de *mirroring* que permiten reflejar el proceso de trabajo de un grupo, su interacción y producción (van Leeuwen & Rummel, 2020; Zangara & Sanz, 2020, 2019)

Finalmente, se viene trabajando fuertemente en el eje de juegos serios digitales, en metodologías para su diseño, desarrollo y evaluación. Los juegos serios son aquellos que tienen un objetivo caracterizante además del entretenimiento (Dörner, Effelsberg, Göbel & Wiemeyer, 2016). En el proyecto se crean juegos serios educativos que se aplican diferentes contextos y que integran diferentes paradigmas de interacción persona – ordenador, buscando aprovechar las posibilidades del contexto físico y virtual.

En todos los ejes se abordan enfoques participativos y de codiseño para involucrar a docentes y a estudiantes en el diseño de la tecnología apropiada para generar soluciones centradas en las necesidades de las personas que se verán afectadas al utilizarlas (Gros & Durall, 2020).

Cabe señalar que las líneas de I+D+I que se abordan se vinculan con las temáticas de la Maestría y Especialización en Tecnología Informática Aplicada en Educación de la misma Facultad, por lo que esta vinculación favorece el desarrollo de tesis y trabajos finales en estos temas.

2. LINEAS DE INVESTIGACION / DESARROLLO/ INNOVACIÓN

Se presentan aquí las principales líneas de investigación, desarrollo e innovación abordadas en el marco del proyecto:

- Entornos digitales para la mediación de procesos educativos. Integración de nuevas funcionalidades y formas de acceso a estos espacios, trazabilidad de las actividades, estándares. Estrategias para el diseño e implementación de estos tipos de entornos.
- Materiales educativos digitales. Metodologías para su diseño y producción. Objetos de aprendizaje. Multimedia e hipermedia en escenarios educativos. Nuevos entramados de medios, soportes y lenguajes.
- Integración de TIC en procesos educativos. Hibridación de las modalidades educativas. Diseño de MOOC.
- Juegos Serios con diferentes paradigmas de interacción. Metodologías para su creación.
- Trabajo colaborativo mediado por TICs. Autorregulación y capacidades metacognitivas como factores claves para su desarrollo. Conceptualización, análisis y desarrollo de software y metodologías. Actividades colaborativas aprovechando dispositivos móviles.

En el proyecto participan investigadores formados, en formación, becarios, tesisistas y alumnos de grado.

3. RESULTADOS OBTENIDOS

Se presentan aquí los principales resultados que se han alcanzado en el período 2020 - inicios de 2021. Para ello se organizan de acuerdo a los principales ejes mencionados del proyecto.

- A. El diseño y desarrollo de entornos para el escenario educativo, y materiales hipermediales para el estudio, con diferente granularidad y paradigmas

En esta línea se avanzó en las mejoras del entorno IDEAS¹ (entorno virtual de enseñanza y aprendizaje - EVEA). Se está realizando una actualización de las tecnologías subyacentes (migración a versiones actuales de *symphony*), y el rediseño de su *look & feel* para atender a las demandas de su constante utilización. Durante el período de la pandemia se ha profundizado el aprovechamiento de este desarrollo que se vincula con este proyecto de investigación. Se ha puesto el foco en el uso de sus herramientas de autoevaluación, de tareas, y de estadísticas, encontrando necesidades de mejora. Se han definido estrategias de mediación de los contenidos de las carreras de la Facultad de Informática utilizando este

¹ Proyecto IDEAS:
<https://proyectoideas.info.unlp.edu.ar/>

EVEA. Se avanzó también en el uso con fines de autoevaluación del juego Desafiarte integrado con la herramienta de evaluación de IDEAS (Archuby, Sanz & Manresa, 2020).

Se finalizó un trabajo de especialización sobre calidad de entornos virtuales de enseñanza y aprendizaje, dirigido por un miembro del proyecto (Ochoa Roblez, Astudillo & Esponda, 2020), y se dará continuidad al tema con una tesis de maestría (Ochoa Roblez, Esponda, Gorga & Astudillo, 2020).

Al mismo tiempo, durante 2020 se avanzó en la tesis doctoral referida a entornos 3D para el aprendizaje de personas con discapacidad auditiva (Fachal, Abásolo, Sanz, 2019). Se creó un servidor propio de OpenSim² con diferentes escenarios preparados para el trabajo de campo. Se estudiaron diferentes alternativas para la comunicación de personas sordas en este espacio y se está creando un tablero para expresar emociones mediante movimientos del avatar.

Se llevó a cabo un avance en el estudio de aplicaciones móviles para personas con discapacidad auditiva. Esta línea se realizó en el marco de un trabajo de cooperación con la UIB y la UNSE.

En relación con los temas vinculados al diseño y desarrollo de materiales educativos hipermediales, se realizó la producción de nuevos materiales educativos, en particular se crearon objetos de aprendizaje orientados a diferentes áreas, y se integraron al repositorio creado en el marco del proyecto para su difusión: <http://roa.info.unlp.edu.ar>. Además, se finalizó una tesis de maestría relacionada con el diseño y producción de un material hipermedial para el área de Química, dirigida por dos miembros del proyecto (García, Bertone, & Gorga, 2020). En el tema de objetos de aprendizaje se avanzó en el framework para la creación de OA (Violini, Sanz, & Pesado, 2020) para el que se está llevando adelante un proceso de prototipos evolutivos con diseño participativo.

Al mismo tiempo, como parte de este eje se trabajó en la creación de una metodología para guiar el diseño de MOOC, que se plasmó en un documento y se utilizó en el marco del proyecto Cap4city (Erasmus). Su aplicación se realizó tanto para la creación de cursos en el marco de este proyecto como en instancias de capacitación para formación de formadores (Zangara & Sanz, 2020).

B. La mediación y el seguimiento de actividades colaborativas en escenarios educativos, con indicadores y visualización del proceso subyacente

Como resultado de este eje se está dirigiendo una tesis doctoral en la que se diseñan herramientas para el seguimiento y visualización de procesos colaborativos en escenarios educativos. La tesis se desprende de trabajos previos del proyecto, relacionándose en forma

directa con la metodología MetSCIN que propone, entre otras cosas, un conjunto de indicadores para analizar procesos colaborativos y reflejarlos a partir del uso de estrategias de *mirroring* (Zangara & Sanz, 2019). MetSCIN se aplica todos los años en el marco de una actividad colaborativa en un curso de postgrado de la Facultad de Informática. Al mismo tiempo, se finalizó una tesis de maestría que se vincula con estrategias de colaboración entre docentes y alumnos para potenciar visitas guiadas a centros interactivos de ciencia y tecnología (Dávila, Fernández & Gorga, 2020). En 2020, también se abordó un tema transversal a los distintos ejes del proyecto, que se enfoca en analizar la autorregulación de los estudiantes en procesos educativos con mediación tecnológicas, y su vinculación con el rendimiento académico. Se difundieron los resultados de la investigación realizada en el marco del trabajo de Especialización de Dieser, Sanz & Zangara (2019), se escribió un capítulo de libro (Dieser, Sanz & Zangara, 2020) y se encaminó su profundización a partir de su continuación en su tesis de maestría.

C. La creación de juegos serios educativos con diferentes paradigmas de interacción que combinan el mundo físico y digital.

Se continúa en este eje con la investigación sobre el diseño, la creación, la evaluación y la puesta en práctica de juegos serios educativos. Se finalizó una tesis de maestría relacionada en la que se investigaron metodologías para el diseño de juegos serios educativos, relevando un total de 38 trabajos sobre esta temática. Se propuso además una nueva metodología, llamada DIJS, con el fin de abordar aspectos no considerados en los estudios previos. Se llevó a cabo un juicio de expertos sobre esta propuesta y su validación a través de un estudio de caso (Archuby, Sanz & Pesado, 2020; Archuby, Sanz, Manresa-Yee, 2020). También se aborda el tema de juegos serios en cooperación con la Universidad Nacional de Río Negro, donde se dirige un proyecto vinculado a estas temáticas (Lovos et al., 2020).

Se ha finalizado una tesis de maestría, dirigida por uno de los miembros del proyecto, que se orientó a gamificar procesos educativos, si bien no se relaciona en forma directa con este eje, algunas de las estrategias usadas en la gamificación provienen del mundo de los juegos serios (Hunicken & Gonzalez, 2020).

En el último año, se han creado diferentes juegos serios como parte de la investigación, entre ellos, se mencionan: Albores que combina la interacción tangible con las posibilidades de los entornos inmersivos, y el juego Innovática basado en realidad virtual, entre otros. Parte de estos desarrollos se aplicaron en talleres y charlas con escuelas que se abordan a partir del Proyecto Nexos y de las acciones

² http://opensimulator.org/wiki/Main_Page

del Centro de Innovación y Transferencia de la Facultad de Informática.

En cuanto a los proyectos vinculados con la temática y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se cuenta con un acuerdo de colaboración en estos temas con la Universidad de Zaragoza y se trabaja en forma conjunta.
- Se tienen un acuerdo de cooperación con la Universidad de Islas Baleares, en particular se trabaja con el Departamento de Ciencias Matemáticas e Informática. En este contexto se dirigen tesis y se participa de proyectos conjuntos.
- Se finalizó la participación en el proyecto REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente. Proyecto que se desarrolla en el marco de los Proyectos de Innovación y Transferencia en Áreas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA) con presentación de informe.
- Se participa del Proyecto Nexos de articulación entre la escuela y la universidad, a partir del cual se llevan a cabo acciones con diferentes escuelas de la región de La Plata y se participa de ferias y exposiciones de ciencia, tecnología y educación.
- Se coopera con la Universidad Nacional de Santiago del Estero y con la Universidad Nacional de Río Negro con asesoría en el primer caso y dirección de un proyecto en el segundo, en vinculación a los ejes presentados aquí.
- Se participa en el proyecto PERGAMEX y en particular en el subproyecto RTI2018-096986-B-C31: “Design of pervasive gaming experiences for intergenerational social and emotional well-being (PERGAMEX-INTERGEM)”.

4. FORMACION DE RECURSOS HUMANOS

Se realiza formación de recursos humanos en el área a través de la dirección de becas, tesis de doctorado, maestría y trabajos finales de especialización y tesinas de grado.

En 2020, se cuenta como resultado con 4 trabajos de Maestría aprobados en relación a los temas aquí presentados (dirigidas por miembro del proyecto, y 1 de ellas correspondiente a un integrante), 1 de Especialización en el área de Tecnología y Educación, también dirigido por miembros de este proyecto y la dirección de una tesina de grado, además se continúa con otras que están en desarrollo. Entre las tesis finalizadas se destaca la tesis de Maestría en Tecnología Informática Aplicada en Educación, titulada “Metodologías de diseño y desarrollo para la creación de juegos serios digitales”, finalizada en septiembre de 2020 (Archuby et al., 2020). Debajo se presenta un resumen.

5. BIBLIOGRAFIA

- Allen, C & Mugisa, E. (2010). Improving Learning Object Reuse Through OOD: A Theory of Learning Objects. In *Journal of Object Technology*, vol. 9, no. 6, 2010, pages 51–75, doi:10.5381/jot.2010.9.6.a3
- Archuby, F., Sanz, C. & Manresa-Yee, C. (2020). Metodologías de diseño y desarrollo para la creación de juegos serios digitales. Tesis de maestría finalizada. Sep. 2020. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/111123>
- Archuby, F., Sanz, C. & Pesado, P. (2020). Experience Analysis for the Use of Desafiate Serious Game for the Self-assessment of Students. In: Pesado P., Arroyo M. (eds) *Computer Science – CACIC 2019*. CACIC 2019. Communications in Computer and Information Science, vol 1184. Springer, Cham. https://doi.org/10.1007/978-3-030-48325-8_8
- Astudillo, G., Sanz, C. & Santacruz-Valencia, L. (2017). Proceedings of the Twelfth Latin American Conference on Learning Technologies (LACLO), 2017, ISBN: 978-1-5386-2376-3, págs. 1-4, doi. 10.1109/LACLO.2017.8120939.
- Carvalho, L.; Yeoman, P. (2021) Performativity of Materials in Learning: The Learning-Whole in Action. *Journal of New Approaches in Educational Research*, [S.l.], v. 10, n. 1, p. 28-42, jan. 2021. ISSN 2254-7339. Available at: <https://naerjournal.ua.es/article/view/627>
- Dabbagh, N. & Fake, H. (2017) Percepciones de los estudiantes universitarios sobre los Entornos Personales de Aprendizaje a través de un prisma de herramientas digitales, procesos y espacios. *Journal NAER*. Vol. 6. No. 1. Enero 2017. pp. 30–38 ISSN: 2254-7339 DOI: 10.7821/naer.2017.1.215
- Dávila, A.; Fernández, A.;Gorga, G. (2019) La colaboración mediada por tecnología como estrategia para potenciar visitas guiadas. El caso del Centro Interactivo de Ciencia y Tecnología Abremate, desde la mirada de sus actores. Tesis de Maestría en TIAE de la Facultad de Informática de la UNLP
- Dieser, P; Sanz, C. & Zangara, A. (2019) Estrategias de autorregulación del aprendizaje en escenarios educativos mediados por tecnologías de la información y la comunicación. Una revisión y análisis en la Educación Superior Iberoamericana. Trabajo final de Especialización en TIAE. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/85104>
- Dieser, P., Sanz, C., Zangara, A. Capítulo: Autorregulación del aprendizaje en contextos educativos mediados por tecnologías digitales. Teoría, investigación y aplicaciones en la

- educación superior iberoamericana. Libro: Aprendizaje y tecnologías: habilidades del presente, proyecciones de futuro. Editorial Noveduc. Julio 2020. 978-987-538-764-5
- Dörner, R., Effelsberg, W., Göbel, S., Wiemeyer, J. (2016). Serious Games. Foundations, concepts and practice. Springer. Alemania
 - Fachal, A. S., Abásolo, M. J., & Sanz, C. V. (2019). Experiences in the Use of ICT and Digital Ramps for Students in Tertiary Education with Visual or Hearing Impairment. In Argentine Congress of Computer Science (pp. 369-388). Springer, Cham.
 - Filsecker, M., & Hickey, D. T. (2014). A multilevel analysis of the effects of external rewards on elementary students' motivation, engagement and learning in an educational game. *Computers & Education*, 75, 136-148. doi:10.1016/j.compedu.2014.02.008
 - García, R.; Bertone, R. & Gorga, G. (2020) Producción de material multimedia interactivo con contenido de Química General. Tesis de Magister TIAE. Facultad de Informática. UNLP.
 - Gasparetti, F., De Medio, C., Limongelli, C., Sciarrone, F., Temperini, M. (2018) Prerequisites between learning objects: Automatic extraction based on a machine learning approach, *Telematics and Informatics*, Volume 35, I3, 595-610, ISSN 0736-5853, <https://doi.org/10.1016/j.tele.2017.05.007>.
 - González, A. H., Quintana, N., Vallejo, A. E., & Pereyra, J. M. (2020). Capacitación de adultos hipoacúsicos mediante un entorno multimedia basado en la simulación de una entrevista laboral. In XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020). <http://sedici.unlp.edu.ar/handle/10915/103672>
 - Gros, B., & Durall, E. (2020). Retos y oportunidades del diseño participativo en tecnología educativa. *EduTec. Revista Electrónica De Tecnología Educativa*, (74), 12-24. <https://doi.org/10.21556/edutec.2020.74.1761>
 - Hünicken, L. A., Gonzalez A. (2020). Gamificación y aprendizaje adaptativo para el desarrollo de competencias: el caso de la asignatura Algoritmos y Estructuras de Datos (Doctoral dissertation, Universidad Nacional de La Plata). Disponible en: <http://sedici.unlp.edu.ar/handle/10915/112436>
 - Kiili, K., & Ketamo, H. (2018). Evaluating Cognitive and Affective Outcomes of a Digital Game-Based Math Test. *IEEE Transactions on Learning Technologies*, 11 (2), 255-263. doi: 10.1109/TLT.2017.2687458
 - Lovos, E., Sanz, C., Goin, M., Ricca, M., Molina, C., Gil, E., Basciano, I., Gastaminza, M. (2020). Juegos Serios Móviles. Diseño, Desarrollo e Integración En Escenarios Educativos. Workshop de Investigadores en Ciencias de la Computación. WICC2020
 - Mouriño-García, M., Pérez-Rodríguez, R., Anido-Rifón, L., Fernández-Iglesias, M., Darriba-Bilbao, V. (2018). Cross-repository aggregation of educational resources, *Computers & Education*, Volume 117, 2018, Pages 31-49, ISSN 0360-1315, <https://doi.org/10.1016/j.compedu.2017.09.014>
 - Ochoa Roblez, J., Esponda, S. & Astudillo, G. (2020). Análisis del estado del arte de los modelos de calidad de Entornos Virtuales de Enseñanza y Aprendizaje. Trabajo final de Especialización en TIAE. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/89019>
 - Ochoa Roblez, J., Esponda, S., Gorga, G. & Astudillo, G. Diseño de un modelo de evaluación de calidad de entornos virtuales de enseñanza aprendizaje basado en la familia de calidad ISO/IEC 25000. Propuesta de tesis de Maestría en TIAE.
 - van Leeuwen, A. & Rummel, N. (2020) Comparing teachers' use of Mirroring and advising dashboards. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt Germany, pp. 26-34, March 2020. <https://doi.org/10.1145/3375462.3375471>
 - Violini, L.; Sanz, C. & Pesado, P. (2020) metodologías y framework para el diseño y la creación de objetos de aprendizaje. *Revista Investigación joven* (ISSN 2314-3991), vol. 6, num. Especial, págs. 176-177, 2019.
 - Zangara, A. & Sanz, C. (2020) Trabajo colaborativo mediado por tecnología informática en espacios educativos. Metodología de seguimiento y su validación. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, no. 25, pp. 8-20, 2020. doi: 10.24215/18509959.25.e1
 - Zangara, A. & Sanz, C. (2019). Del trabajo grupal al colaborativo. antecedentes, conceptualización y propuesta de abordaje didáctico. *Signos Universitarios*, (54). Referato Nacional.

Metodologías de diseño y desarrollo para la creación de juegos serios digitales

**Tesis de Maestría en Tecnología Informática Aplicada en Educación
Facultad de Informática, Universidad Nacional de La Plata**

Tesista: Federico Archuby (III LIDI, Facultad de Informática, UNLP)
farchuby@lidi.info.unlp.edu.ar

Directoras: Cecilia Sanz (III LIDI, Facultad de Informática, UNLP, Argentina) – Cristina Manresa-Yee (Universidad de Islas Baleares, España)

Fecha de exposición: septiembre de 2020

MOTIVACIÓN

Esta tesis se centra en una investigación sobre metodologías que guían el diseño y desarrollo de juegos serios educativos digitales. La investigación en juegos serios empezó a proliferar gracias al aumento en popularidad de los videojuegos. Un juego serio busca un equilibrio para alcanzar su objetivo caracterizante y el de entretenimiento por lo que su diseño debe atender a estos aspectos desde su origen. Así es necesario abordar metodologías específicas que posibiliten considerar los componentes propios de los juegos con la inclusión de estrategias vinculadas al diseño instruccional. Esto ha motivado la investigación de esta tesis y sus aportes.

APORTES

En esta tesis se indagan inicialmente conceptos de base como el de juegos serios, sus clasificaciones según distintos criterios, los beneficios encontrados en la literatura sobre su integración en diferentes áreas, y se revisan cuáles son los principales componentes de estos juegos. Como aportes a partir de la revisión realizada se dan definiciones y clasificaciones propias que ayudan a echar luz sobre la temática.

Luego, se analiza la importancia de guiar el diseño de los juegos serios educativos digitales, con una mirada que combine tanto aspectos tecnológicos como pedagógicos, tamizando los saberes del diseño instruccional, del diseño de juegos y de la Ingeniería de Software. A partir de una revisión sistemática de literatura se estudian más de 35 metodologías, que se presentan en forma completa, o parcial y en algunos casos se han aplicado. Este estudio resulta un aporte, ya que muestra categorizaciones propias y presenta un estado del arte en esta línea de investigación. Además, se encontró, que las metodologías no cubrían completamente el proceso de diseño y desarrollo. Por ejemplo, algunas se enfocan en contextos muy específicos, otras presentan miradas parciales, atendiendo solo a algunos aspectos del proceso de diseño. Esto motivó la creación de la metodología DIJS, que se consolida en el aporte central de esta tesis.

La metodología DIJS propuesta en este trabajo, se basa en el estudio previo, e integra algunas de las ideas encontradas en la literatura. Propone atender tanto al diseño, como al desarrollo y la evaluación del juego serio que se desea crear. Incluye una metáfora encontrada en una de las metodologías estudiadas, en donde se plantea que este proceso es como un menú que cuenta con platos (etapas), y para cada plato se utilizan ingredientes y utensilios para realizarlo. Así las etapas de DIJS recomiendan ingredientes y utensilios en cada una, para guiar y viabilizar su realización. Las etapas involucran desde la definición de los objetivos pedagógicos del juego, la definición del perfil del jugador (atendiendo a cuestiones de interés para su aprendizaje), hasta la evaluación del juego, con estudiantes y docentes, y con foco en la jugabilidad.

La metodología propuesta se evaluó de dos formas diferentes: 1. A través de aplicarla para el análisis y extensión del juego serio Desafiate, previamente creado por el autor de esta tesis; 2. Con un juicio de expertos con participación de 10 profesionales con formación específica mayormente en el área de Tecnología Informática aplicada en Educación.

Los resultados dan cuenta de que la metodología ha guiado adecuadamente tanto el análisis de Desafiate como la integración de una nueva funcionalidad. Mediante la etapa de evaluación con sus ingredientes y utensilios se pudieron hallar nuevas oportunidades de mejora para el juego. Además, se encontraron algunos ingredientes que para el análisis podrían resultar redundantes. En el juicio de expertos, hay un consenso en cuanto a la utilidad y el valor otorgado a la metodología DIJS y a las etapas propuestas. Sin embargo, se abre el camino para profundizar respecto de la flexibilidad para cambiar los ingredientes y utensilios.

LÍNEAS FUTURAS DE TRABAJO

Se concluye que la DIJS es un aporte metodológico e instrumental que no ha sido aplicado y ha dado buenos resultados. Asimismo, abre nuevas oportunidades para profundizar en esta investigación que continuará a través de una tesis doctoral. Se aplicará DIJS para trabajar en el diseño de juegos serios con docentes y/o estudiantes, para avanzar en las mejoras de esta metodología a partir de su aplicación. Además, se abordará el desarrollo de herramientas que puedan acompañar la propuesta de esta metodología, sistematizando el proceso de creación de los juegos serios, esto formará parte de las tareas del plan de trabajo de doctorado del mismo tesista.

Agradecimientos: esta tesis ha sido realizada con una beca de la Universidad Nacional de La Plata bajo la dirección de Cecilia Sanz y Patricia Pesado. La tesis se encuentra disponible en el repositorio de la UNLP (SEDICI)

EduScrum un Marco de Trabajo que puede Propiciar Aprendizaje Significativo

APU Macarena Quiroga ¹ APU Pablo Dibez ², Lic. Angela Belcastro ³, Mg. Rodolfo Bertone ⁴

¹Auxiliar de segunda. Departamento de Matemáticas.FI. UNPSJB. Comodoro Rivadavia.

^{1,2} Estudiante de grado. Departamento de Informática.FI. UNPSJB. Comodoro Rivadavia.

³Departamento de Informática. Facultad de Ingeniería (FI). UNPSJB. Comodoro Rivadavia.

⁴ III LIDI – Facultad de Informática – UNLP. (1900) La Plata, Buenos Aires, Argentina

¹msinf.quiroga@gmail.com, ²pdibez@gmail.com, ³ angelab@ing.unp.edu.ar,

⁴ pbertone@lidi.info.unlp.edu.ar

RESUMEN

Este artículo presenta aspectos de la Teoría Constructiva del Aprendizaje, algunas estrategias para generar aprendizaje significativo, para promover el fortalecimiento del pensamiento computacional y generar motivación por aprender con videojuegos. Se examina el uso que en el ámbito educativo se ha dado al marco de trabajo EduScrum, que apoya el trabajo colaborativo, metodología con la que se desarrollará una propuesta de uso en una materia de una carrera de Informática. Destaca algunas acciones realizadas en 2020 con elementos de propuestas educativas ideadas dentro del proyecto de investigación acreditado (PIA), que es continuación de tres proyectos previos del área de Tecnología Informática Aplicada en Educación (TIAE), una de ellas para el dictado de un taller, y otra para una actividad de una asignatura de grado de Informática, con la preparación y uso del videojuego: “El juego de la OKA CYCC”.

Palabras claves:

Evaluación Formativa. Videojuegos. EduScrum.

CONTEXTO

Esta línea de investigación pertenece al PIA: “TIAE, y Aprendizaje Significativo”, UNLP-UNPSJB, reconocido por el programa de Incentivos, en el Sistema de Ciencia y Técnica de la UNPSJB, se realiza en el ámbito de la Facultad de Ingeniería, Departamento de Informática de la UNPSJB, donde se dictan las carreras: “Analista Programador Universitario” y “Licenciatura en Informática”. Este proyecto contribuirá a mejorar los recursos humanos con la

formación en los temas: constructivismo, aprendizaje significativo, la tecnología móvil en la educación, el pensamiento computacional, juegos serios y videojuegos.

1. INTRODUCCION

La *Teoría de Aprendizaje Constructivista* está basada en el aprendiz, que tiene un rol activo, basado en la reflexión y en la toma de conciencia. El aprendizaje es intencional, es una acción dirigida hacia fines. Es importante lograr motivación por aprender en el alumno, promover el trabajo colaborativo generando intercambios, y vinculaciones con el material, los docentes, los alumnos y el medio, planteando una actividad representativa para el aprendiz, que ayude a generar conocimiento y a fortalecer competencias esperadas. Dentro de las estrategias para generar *aprendizaje significativo* encontramos distintos tipos de evaluaciones: diagnósticas, formativas y sumativas.



Figura 1. La EF. Las rúbricas como herramientas de apoyo al proceso de enseñanza aprendizaje.

Vemos en la figura 1, que la *evaluación formativa (EF)* permite recoger información a medida que los procesos se están desarrollando. Camilloni alude a la contemporaneidad de la evaluación con los procesos de enseñanza y aprendizaje, e indica

que la información recogida, debe mejorar dichos procesos. Carless afirma que la evaluación debería contribuir eficazmente en la mejora del aprendizaje. Sus principios sugieren que: a) La evaluación de las tareas debe diseñarse para estimular prácticas correctas de aprendizaje entre los alumnos. La alineación entre objetivos, contenidos y tareas de evaluación facilita la experiencia de aprendizaje profundo hacia los logros deseados. b) La evaluación debe involucrar activamente a los alumnos, mediante criterios de calidad sobre el propio rendimiento y el de los pares. La participación de los estudiantes en la evaluación: autoevaluación, evaluación por pares, retroalimentación de los compañeros, promueve una mejor comprensión de los objetivos de aprendizaje. c) La retroalimentación de la evaluación debe ser oportuna, brindando apoyo en los aprendizajes actuales y futuros. [1][2][3][4]

La retroalimentación adecuada propicia el compromiso del alumno con su proceso de aprendizaje. [5]

La *figura 1*, presenta el doble objetivo de las *rúbricas*, para los alumnos, es un instrumento de autoevaluación en tiempo real, para verificar si el producto se ajusta o no a los requerimientos, a medida que lo van haciendo, también es útil para los docentes. Son guías precisas que valoran los aprendizajes y productos realizados. [6][7]

Con las nuevas tecnologías, para los jóvenes, los *videojuegos* tienen un rol protagónico.

El *juego* es una actividad universal, no hay ninguna persona que no haya practicado esta actividad en alguna circunstancia. Con él, se pueden crear situaciones de valor educativo y cognitivo que permitan experimentar, investigar, resolver problemas, descubrir y reflexionar. Las implicaciones de tipo emocional, el carácter lúdico, el desbloqueo emocional, la desinhibición, son fuentes de motivación que proporcionan una forma distinta a la tradicional de acercarse al aprendizaje. El juego como recurso educativo, ha sido usado por el hombre desde la antigüedad y está presente en el pensamiento de muchos filósofos, educadores y psicólogos, entre ellos, Piaget y Vygotsky. Piaget (1986)

hace mención del juego, como elemento predominante del desarrollo de los estadios cognitivos. [8][9] [10][11]

Salen y Zimmerman definen un juego como un sistema en el que los jugadores se enfrentan o participan en un conflicto artificial, definido por reglas, que se traducen en un resultado cuantificable. [12]

Un *videojuego* es todo juego electrónico que posee imágenes animadas interactivas, acompañadas de un ambiente sonoro y un controlador, él que permite la interacción entre el usuario y el medio de reproducción. (Gonzales Tardón, 2014)

En los videojuegos se ponen en juego componentes tecnológicos, con los que el jugador interactúa, como dispositivos electrónicos (computador, tablet, Smartphone) y, además, medios audiovisuales (imágenes, audios y videos). También disponen de una dinámica propia del juego, con reglas definidas.

Freitas (2006) define al **juego educativo** como: "Las aplicaciones que usan ciertas características del hardware de dispositivos ya sean móviles, computadoras de escritorio o juegos de ordenador, para crear experiencias de aprendizaje interesantes, motivadoras, lúdicas y envolventes para la entrega de objetivos de aprendizaje específicos, resultados y experiencias". Los videojuegos educativos, ayudan a desarrollar habilidades cognitivas, y motoras, contribuyendo así a la enseñanza de hechos y la resolución de problemas complejos, aumentando la creatividad (Figueredo, 2015).

Aprender a programar es una llave que puede abrir la puerta para el desarrollo de estas habilidades y competencias. Los "juegos orientados al aprendizaje de la programación", se refiere al uso de entornos como Scratch, Alice, Pilas Engine por numerosos docentes y alumnos de diferentes países y niveles educativos para acercarse a la programación, y al pensamiento computacional. Entre los componentes del pensamiento computacional, encontramos: análisis descendente, análisis ascendente, heurística, resolución de problemas, creatividad, recursividad, iteración y metacognición. [14][15][16][17]

2. LINEA DE INVESTIGACION Y DESARROLLO

El equipo es dirigido por el Mg. Rodolfo Bertone de la UNLP, integrado por docentes investigadores y alumnos de carreras de grado de Informática. La codirectora avanza en carrera de postgrado en TIAE. Entre las principales líneas de investigación, encontramos: el modelo constructivista, aprendizaje significativo, metacognición, tecnología móvil en la educación. El objetivo es el de contribuir a la formación de personas competentes, fomentar el desarrollo de la sociedad y aumentar el conocimiento, analizando y confeccionando propuestas educativas que propicien el aprendizaje significativo, diseñando y desarrollando sistemas interactivos de apoyo al aprendizaje con tecnología móvil, con el objeto de incentivar a los alumnos a investigar, y a participar en PI, y mejorar las competencias de nuevos profesionales de Informática.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los antecedentes encontramos publicaciones en WICC 2015, WICC 2018, WICC 2020, TE&ET 2015, y CONAISI (de 2016 a 2020). En 2020 se realizaron talleres presenciales, uno de ellos, el “evento de transferencia DeCoDev: desarrollo colaborativo de videojuego”, con aval (Res. FI-SJB-1097/2019), que se diseñó para promover capacidades del pensamiento computacional en los alumnos con el desarrollo de videojuegos en pilas engine, con actividades virtuales, aplicando trello y Scrum, el 50% de los participantes, obtuvieron certificado de aprobación. Y el “evento de Transferencia JEIYAP, Juegos Educativos para ingresantes y alumnos de primer año de Informática”, con aval (Res. RDFI 73/2020). Se desarrollaron propuestas educativas basadas en la Teoría Constructivista en asignaturas de Informática, en una de ellas, “Fundamentos Teóricos de Informática”, se creó y utilizó en 2020, un videojuego: el Juego de la OKA CYCC, de Computabilidad y Complejidad Computacional, creado con (<https://mobbyt.com/>) para generar motivación

en el entrenamiento en el tema: “Computabilidad y Complejidad Computacional”. Se realizó revisión bibliográfica, y se desarrollará una propuesta de uso de EduScrum en una asignatura de una carrera de grado de Informática de la UNPSJB. En las metodologías ágiles, se construye el producto mientras se modifican y aparecen nuevos requisitos. El manifiesto ágil valora “a los individuos y su interacción, por encima de los procesos y las herramientas”, “El software que funciona, por encima de la documentación exhaustiva” y “La colaboración con el cliente, por encima de la negociación con contrato previo detallado”. *El objetivo de un proyecto ágil no es controlar la ejecución conforme a procesos y cumplimiento de planes, sino proporcionar el mayor valor posible al producto.* [18][19][20][21]

Scrum es un marco de trabajo que ayuda a las personas, equipos y organizaciones a generar valor a través de soluciones adaptativas para problemas complejos. *Es una de las metodologías ágiles más populares, empleadas para el desarrollo de Software, es muy adaptable, por lo cual se puede utilizar en diferentes ámbitos y contextos.* La teoría de Scrum se basa en el empirismo, que asegura que el conocimiento proviene de la experiencia y que la toma de decisiones se realiza en base a lo observado, y también en el pensamiento de Lean, lo que elimina lo innecesario y se enfoca en lo importante. El enfoque que emplea es iterativo e incremental. *Scrum facilita el trabajo en equipo, fomenta el ejercicio del liderazgo, propicia un mayor rendimiento de los miembros del equipo, se caracteriza por una evaluación constante y un aumento en la interacción entre los miembros del equipo.* Entre las características claves de los miembros del equipo encontramos el trabajo colaborativo, las habilidades de toma de decisiones, la capacidad para resolver problemas difusos, la confianza y respeto mutuo, y la organización propia. [18][19]

La figura 2 muestra aspectos claves de Scrum. Los roles que intervienen son: (1- el propietario del producto), que fija las prioridades; (2- el equipo de desarrollo), que construye el producto; y (3- Scrum Master),

que gestiona y facilita la ejecución de las reglas de scrum; y (4- interesados), que son los restantes implicados, que asesoran y observan. La pila del producto es un artefacto definido por el propietario del producto, que prioriza, no es muy detallada la relación de requisitos del producto, se trata de una lista en evolución y abierta a todos los roles. La pila de sprint incluye la lista de requisitos comprometidos por el equipo, para el sprint con nivel de detalle suficiente para su ejecución. El incremento es la parte del producto desarrollada en un sprint, en condiciones de ser usada (pruebas, codificación limpia y documentada). Entre los eventos encontramos las reuniones diarias de no más de 15 minutos, analizando el estado actual del sprint, empleando un tablero para tareas y/o funcionalidades. Al finalizar el sprint se realiza una reunión para la revisión del mismo, se verifica el cumplimiento de metas y objetivos, garantizando la entrega del producto, luego se realiza una nueva reunión retrospectiva en la cual se analizan resultados del Sprint anterior, detectando falencias para aplicar mejoras al avanzar al siguiente Sprint. Automáticamente se inicia un nuevo sprint empleando otra funcionalidad de la pila del producto, e inicia nuevamente el proceso hasta tener un nuevo producto funcional. Al finalizar cada sprint se entrega al cliente un sistema con el que puede interactuar y ver el avance del proyecto, y al finalizar todos los sprint, se obtiene el producto terminado. [18][20]



Figura 2. Características de Scrum

EduScrum es una metodología adaptada de scrum que se ha usado en el ámbito educativo, para el desarrollo de actividades con trabajo en equipo, tanto de desarrollo de software como para otras actividades educativas con la construcción de un trabajo final en equipo.

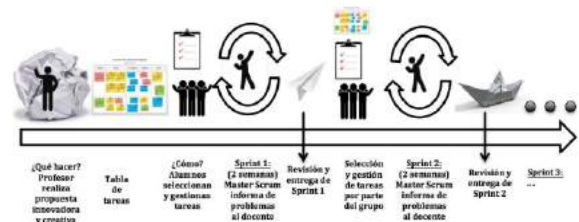


Figura 3. Metodología EduScrum por fases. [27]

La figura 3 esquematiza la metodología scrum aplicada en el contexto educativo. En la que se establecieron tres roles, cada uno con diferentes responsabilidades. El profesor, que gestiona, asesora y evalúa el proyecto. El Scrum Master (o facilitador), que asumirá un alumno del equipo para ayudar a gestionar el proceso de trabajo, y que hará de enlace con el profesor cuando haya que entregar los informes bisemanales, o bien para consultar dudas. El rol de Scrum Master no es el de liderar y organizar a su equipo, es un facilitador, y para ello debe asegurarse de que todos los miembros del equipo realicen lo prometido en la última reunión, tratar de gestionar la resolución de problemas desde el equipo, y si no es posible, con la ayuda del docente, garantizar que las reuniones al inicio de clase no duren más de 15 minutos, o bien que cada alumno del equipo intervenga como máximo, de entre 60 o 90 segundos, y reportar al docente periódicamente información destacada sobre el progreso de las diferentes tareas propuestas.

Los alumnos planifican y definen sus actividades y llevan un seguimiento del progreso. El profesor determina las tareas y brinda apoyo. Se busca generar creatividad, propiciar la buena comunicación, colaboración y reforzar el pensamiento crítico. [24][25][26] EduScrum genera información periódicamente que apoya el desarrollo de evaluación formativa, y la mejora de la calidad de los trabajos, propiciando la reflexión y el ejercicio de metacognición.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo incluye estudiantes avanzados de carrera de grado de Informática que, desde estos proyectos continuados, apoyados por docentes investigadores

categorizados, uno de los cuales desarrolla carrera de postgrado del área TIAE, han participado en congresos, como disertantes en talleres, como desarrolladores de juegos interactivos, en publicaciones con referato, y como auxiliares de segunda. Se han desarrollado transferencias a la comunidad educativa en la UNPSJB, se espera formar recursos humanos en los temas tratados.

5. BIBLIOGRAFIA

- [1] Sylvia Libow Martínez. Gary Stager. Inventar para aprender. Guía práctica para instalar la cultura maker en el aula. Siglo Veintiuno editores. Colección educación que aprende. 2019.
- [2] Sanz Cecilia, y otros. (2016). Escenarios educativos mediados por Tecnologías de la Información y la Comunicación. Actas WICC 2016, ISBN: 978-950-698-377-2, págs. 983-988, 2016.
- [3] Calero Pérez, M. Constructivismo pedagógico. Teorías y aplicaciones básicas. Alfaomega. 2008.
- [4] Díaz, F. Y Barriga, A. (2002) Estrategias Docentes para un Aprendizaje Significativo: una interpretación constructivista. México: McGraw Hill
- [5] Vanina Bottini. Guillermo Kalocai. Juan María Palmieri. (2019). Desarrollo de Competencias Sociales en Estudiantes de Carreras de Ingeniería Mediante Estrategias de Aprendizaje Centrado en el Estudiante. Disponible en <https://conaiisi2019.unlam.edu.ar/pdf/2019-CONAIISI-ACTAS-7MA-EDICION.zip> Último acceso: 19-10-20.
- [6] Hugo M. Castellano. (2011). Integración de la tecnología educativa en el aula. Enseñando con TIC. Cengage Learning.
- [7] Gatica-Lara, F. y otros. ¿Cómo elaborar una rúbrica? Elsevier. (2013). Disponible en: <https://www.elsevier.es/es-revista-investigacion-educacion-medica-343-articulo-como-elaborar-una-rubrica-S200750571372684X> Acceso 22-11-19.
- [8] Archuby Federico Héctor. Sanz Cecilia. Pesado Patricia. 2018. Desafiate: juego serio para la autoevaluación. ISBN: 978-950-766-124-2.
- [9] Cristina Rojas. Javier Raúl Alejandro Cantó. Aprendizaje Centrado en el Estudiante aplicado en la asignatura Investigación Operativa. 2018.
- [10] SÁNCHEZ ARBRIZ, Mercedes Leticia (2013). Profesores frente a los videojuegos como recurso didáctico, En Revista Didáctica, Innovación y Multimedia, núm. 25 <http://www.pangea.org/dim/revista25> Revista DIM / Año 9 - Nº 25 - Abril 2013 - ISSN: 1699-3748. Mercedes Leticia Sánchez Ambriz
- [11] Antonia Bernat Cuello. La construcción de conocimientos y la adquisición de competencias mediante el uso de los videojuegos. 2008. Acceso (febrero 2020) <https://dialnet.unirioja.es/servlet/articulo?codigo=2561257>
- [12] Rocío Serna Rodrigo. El papel del videojuego en la formación de universos transmedia. Universidad de Alicante. Escenarios de aprendizaje basados en TIC. (ed.). Tecnología, innovación e investigación en los procesos de enseñanza-aprendizaje. Barcelona: Octaedro, 2016. ISBN 978-84-9921-848-9, 3173 p.
- [13] Moreno Arias Juan Carlos. Programación del videojuego “Ilumpak” para dispositivo móvil de sistema operativo Android. Trabajo de graduación previo a la obtención del título de ingeniero en computación gráfica. Universidad Central del Ecuador. Facultad de Ingeniería, Ciencias Físicas y Matemática. Carrera de Ingeniería en Computación Gráfica. Tutor: Fis. Gonzalo Bayardo Campuzano Nieto. Quito – 28 de junio de 2016.
- [14] Juan Carlos Matos Franco. Utilización de Videojuegos como Apoyo en el Aprendizaje de la Programación (ID2012/042). Acceso (febrero 2020). Disponible en: https://gredos.usal.es/bitstream/handle/10366/122556/MI_D_12_042.pdf?jsessionid=F2CE69CD57ECA63F76ACA91385DB3297?sequence=1
- [15] Cecilia Sanz. Aprender a programar en tiempos digitales. Acceso (marzo 2020). http://sedici.unlp.edu.ar/bitstream/handle/10915/57357/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y
- [16] Miguel Zapata-Ros. Vista de Pensamiento computacional Una nueva alfabetización digital. RED-Revista de Educación a Distancia, 46(4). 15-Sept.-2015 DOI: 10.6018/red/46/4 (Acceso junio 2020). <http://www.um.es/ead/red/46/zapata.pdf>
- [17] Ahumada, Hernán y otros. (2019). Programación por Bloques en cátedras de Ingeniería en Informática. Actas WICC 2019, ISBN: 978-987-3619-27-4. Págs. 636-640. 2019.
- [18] Juan Palacios. Scrum Manager I. Las reglas de scrum. Scrum Manager. 2015.
- [19] Roger S. Pressman. Ingeniería de software. Un enfoque práctico. Séptima edición. Mc Graw Hill. 2010.
- [20] Guía oficial de scrum: Acceso febrero 2021. Disponible en: <https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-Spanish-Latin-South-American.pdf>
- [21] Nicolás Tymkiw y otros. SCRUM como metodología de enseñanza y aprendizaje de la Programación. Acceso febrero 2021. Disponible en: http://sedici.unlp.edu.ar/bitstream/handle/10915/104104/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y
- [22] Nazareno, Roberto y otros. (2015). Trazabilidad de Procesos Scrum. Actas del Simposio Argentino de Ingeniería de Software (ASSE 2015). ISSN: 2451-7593. Págs. 284-298. 2015.
- [23] Gálvez Díaz, María del Pilar. Tolaba, Carolina. (2019). Como Alcanzar Competencias Usando Scrum. Actas SAEI 2019. ISSN: 2683-8958. Págs: 62-70.
- [24] Jair Steven Calderón Huertas, Manuel Ernesto Bolaños González, Nelson Antonio Jaramillo Enríquez. FORMACIÓN CON SCRUM, METODOLOGÍAS ÁGILES DE DESARROLLO DE SOFTWARE Y SU APLICACIÓN EN LA ENSEÑANZA – APRENDIZAJE. 2016.
- [25] David Hernando de la Plaza. Scrum como marco de trabajo para el aprendizaje de competencias en los ámbitos universitarios y empresariales. 2020.
- [26] Juan Lucas Onieva López. SCRUM COMO ESTRATEGIA PARA EL APRENDIZAJE COLABORATIVO A TRAVÉS DE PROYECTOS. PROPUESTA DIDÁCTICA PARA SU IMPLEMENTACIÓN EN EL AULA UNIVERSITARIA. 2018.
- [27] Rueda, Sonia y otros. (2014). Herramientas para apoyar el descubrimiento de vocaciones en Ciencias de la Computación. CACIC 2014. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/42122> Acceso: 21-11-19

ESTUDIO DE HERRAMIENTAS PARA LA IMPLEMENTACIÓN DE LABORATORIOS REMOTOS

Godoy, Pablo Daniel^{1,2,3}; Marianetti, Osvaldo¹; Mansilla Roberto Alejandro¹; García Garino, Carlos^{1,3}

¹Universidad Nacional de Cuyo, Facultad de Ingeniería

²Universidad Nacional de Cuyo, Facultad de Ciencias Exactas y Naturales

³Universidad Nacional de Cuyo, ITIC

pablo.godoy@ingenieria.uncuyo.edu.ar

RESUMEN

Este artículo presenta líneas de trabajo actuales y futuras sobre laboratorios remotos destinados a tareas de docencia e investigación. Para ello, primero se resume el trabajo realizado durante 10 años, que consiste en el diseño y despliegue de laboratorios remotos de redes de sensores inalámbricos y de computación, destinados a investigación científica y docencia. Se presentan trabajos de investigación científica y experiencias de aplicación en tareas de enseñanza en distintas asignaturas de carreras de computación. Se describen diferentes tecnologías empleadas, entre estas, cloud computing, VPNs y escritorios remotos, con sus ventajas y desventajas. Se concluye con una enumeración de problemas encontrados y posibles soluciones, propuestas de mejoras y de integración de las tecnologías utilizadas. Estas posibles soluciones a problemas encontrados y propuestas de mejoras constituyen las líneas de investigación futuras. **Palabras clave:** laboratorios remotos, acceso remoto, laboratorios virtuales, experimentos remotos.

CONTEXTO

El trabajo que se presenta en este artículo ha sido financiado por dos proyectos denominados: “Implementación de laboratorios remotos basados en cloud computing” desarrollado entre 2016 y 2018 y “Laboratorio remoto y nómada de arquitectura de computadoras destinado a enseñanza e investigación” durante 2019 y 2021. Ambos proyectos financiados por la Universidad Nacional de Cuyo. El primer proyecto se centro en estudiar distintas alternativas para integrar laboratorios remotos

con plataformas de cloud computing, como también ventajas y desventajas de esta integración. El segundo proyecto se orientó a estudiar y aplicar de estos laboratorios remotos a tareas de enseñanza en distintas asignaturas de carreras de grado y posgrado relacionadas con computación de la Universidad Nacional de Cuyo y la Universidad de Mendoza.

Además se colaboró con otros proyectos denominados “Dispositivos de lógica programables como alternativa de solución en problemas de seguridad en Internet de las Cosas” desarrollado en el periodo 2019-2021 y “Análisis de topologías de comunicación en Internet de las Cosas” durante 2016-2018.

Los investigadores que forman parte de estos proyectos son profesores de asignaturas como Arquitectura de Computadoras, Redes de Computadoras, Arquitecturas Distribuidas y Sistemas Operativos en la Universidad Nacional de Cuyo y la Universidad de Mendoza, en carreras de pregrado, de grado y de posgrado.

1. INTRODUCCIÓN

Los laboratorios remotos son plataformas que permiten el acceso de manera remota a equipamientos o sistemas de distinto tipo, con fines de investigación científica o enseñanza. Constan fundamentalmente de 3 partes:

- El sistema bajo prueba, compuesto por el equipamiento o sistema sobre el cual se necesita realizar experimentos remotamente.
- Un sistema de gestión, que permite el acceso de manera ordenada al laboratorio remoto. Realiza tareas como autenticación de usuarios, gestión de permisos y turnos,

almacenamiento de información relacionada con el uso del laboratorio remoto, etc.

- La interfaz de usuario remoto, que es la parte del laboratorio remoto con la cual el usuario remoto está en contacto. Puede ser una interfaz web, línea de comandos, puede mostrar imágenes en directo del sistema bajo prueba, mostrar valores de variables medidas, permitir al usuario realizar acciones sobre el equipo bajo prueba, etc.

Cualquiera de estas tres partes pueden estar implementadas en un solo sitio o de forma distribuida en varias computadoras [1].

Se implementaron varias arquitecturas de laboratorios remotos, las cuales se resumen a continuación:

1.1 Laboratorio remoto integrado con el Cloud a través de RPC

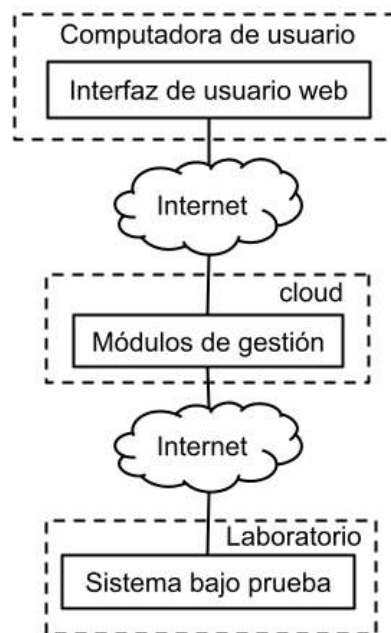
Esta arquitectura se utilizó para implementar un laboratorio remoto de redes de sensores inalámbricos destinado a educación e investigación científica [2]. Se empleó una red de sensores inalámbricos Zigbee, con 10 nodos como sistema bajo prueba. Se utilizó una interfaz web con el usuario, la cual muestra los valores de los parámetros medidos y de configuración de los dispositivos, y a través de la cual los usuarios pueden modificar los parámetros de configuración. La comunicación entre las diferentes partes del laboratorio remoto se realizan a través de RPC (llamadas a procedimientos remotos).

Se implementaron los siguientes módulos de gestión en el cloud:

- Control de acceso para autenticación de usuarios y gestión de permisos.
- Sistema de reserva de turnos.
- Registro de actividad de los usuarios.
- Base de datos.
- Material de ayuda.
- Interacción en tiempo real.

La figura 1 muestra un diagrama en bloques de esta arquitectura.

Figura 1: Laboratorio remoto integrado al cloud a través de RPC.



Esta arquitectura se utilizó para realizar experimentos científicos, consistentes en medir diferentes parámetros temporales y de performance de nodos sensores Zigbee, destinados a evaluar su aplicabilidad a sistemas de sensores de variables agrícolas y ambientales [3][4].

El proveedor de servicios de cloud computing fue Google Cloud Platform.

1.2 Laboratorio remoto y nómada

Este laboratorio remoto se implementó para ser utilizado en tareas de docencia en asignaturas relacionadas con arquitecturas y redes de computadoras. Sus dos objetivos principales son:

- Que pueda ser trasladado al aula para que los estudiantes puedan realizar sus trabajos prácticos sobre el equipo bajo prueba de manera presencial.
- Que los estudiantes puedan acceder de forma remota a través de Internet.

Para lograr el primer objetivo, se implementaron tres formas de acceso al equipo bajo prueba:

- Red Ethernet disponible en el aula.
- Red Wifi disponible en el aula.
- Red Wifi propia.

Para lograr el segundo objetivo, los estudiantes acceden a través de Internet empleando los servicios de plataformas de

control remoto de dispositivos de IoT, como Remote IoT (<https://remoteiot.com>).

El sistema bajo prueba consiste en computadoras Raspberries Pi 3 y Pi 4 y placas Arduino Uno conectadas a las computadoras Raspberries.

Este laboratorio remoto está destinado a estudiantes con conocimientos medios o avanzados sobre computación. Por tal motivo, es deseable que los estudiantes tengan control total sobre el sistema bajo prueba, pudiendo configurar cualquier parámetro, instalar o desinstalar aplicaciones, etc. Para lograr este propósito, la interfaz con el usuario se implementa mediante SSH o escritorios remotos, según prefiera el usuario.

El sistema de turnos se implementa solo para el acceso a través de Internet. El mismo se implementa en el cloud. El estudiante debe solicitar un turno a través de una página web brindando un nombre de usuario. Una de las computadoras Raspberry del sistema bajo prueba solicita periódicamente el archivo de turnos almacenado en el cloud. Con esa información, permite el acceso solo los usuarios autorizados a la hora de sus turnos. La figura 2 muestra un esquema de esta arquitectura [5].

1.3 Laboratorios remotos basados en VPN y escritorios remotos

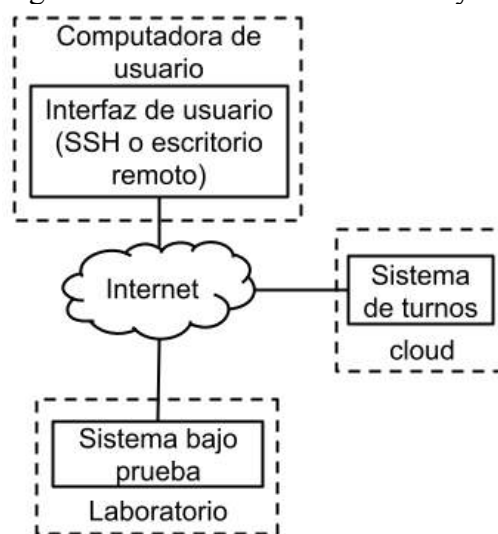
Durante el año 2020, debido a que el dictado de clases fue no presencial, se requirió un uso intensivo de los laboratorios remotos construidos.

Esta experiencia mostró que las plataformas para control remoto de dispositivos IoT no permiten una experiencia en tiempo real para laboratorios remotos de computación, y el número de usuarios es limitado.

Por lo tanto, se optó por dos métodos alternativos de accesos:

- Túneles VPN (redes privadas virtuales).
- Escritorios remotos.

Figura 2: Laboratorio remoto y nómada



Para la primera opción, se utilizaron los servicios del proveedor de servicios VPN Hamachi LogMe Inc (www.vpn.net), y para la segunda opción, se emplearon los servicios de RealVNC y Google Remote Desktop. Se lograron experiencias de uso satisfactorias tanto para acceso a través de túneles VPN y para escritorios remotos.

En cuanto a plataformas VPN, existen en la actualidad dos tipos de servicios ofrecidos:

- Túneles que permiten acceso a Internet a través de un punto de acceso diferente al real.
- Conectar computadoras ubicadas en puntos geográficos diferentes como si estuvieran conectadas a través de la misma red LAN.

Solo proveedores de servicios VPN que provean la segunda opción son útiles para los laboratorios remotos presentados en este trabajo [1].

El diagrama en bloques de la figura 2 es válido también para este laboratorio remoto. La diferencia está en que el acceso para los laboratorios remotos y nómades (sección 1.2) es a través de un proveedor de plataformas para control remoto de dispositivos IoT, mientras que para los laboratorios remotos basados en VPN y escritorios remotos, el acceso remoto es a través de proveedores de escritorios remotos a través de Internet o por túneles VPN.

2. LINEAS DE INVESTIGACION y DESARROLLO

2.1 Laboratorios remotos

Los laboratorios remotos han sido utilizados desde hace aproximadamente 20 años para tareas de investigación y docencia. Constituyen una alternativa intermedia a los simuladores y los experimentos presenciales en laboratorios. Como ventaja frente a simuladores se destaca que los experimentos son realizados sobre equipamientos o sistemas reales, y no sobre simuladores, susceptibles a errores por simplificaciones en los modelos sobre los cuales se construyen. Con respecto a la experimentación en laboratorios presenciales, tienen la ventaja de que el usuario no necesita adquirir equipamiento, ni trasladarse físicamente al laboratorio, y puede acceder en cualquier momento.

Como desventajas se menciona la pérdida de contacto con el sistema bajo prueba, y que los experimentos están limitados a las acciones que permite el sistema de acceso.

2.2 Integración de laboratorios remotos al Cloud

Cloud computing posee ventajas que son de interés para la implementación de laboratorios remotos. Estas ventajas son:

- Alta disponibilidad.
- Confiabilidad.
- Seguridad.

La alta disponibilidad permite asegurar a los usuarios que los componentes desplegados en el cloud estarán siempre disponibles. Por lo tanto, ante una falla en la conectividad o en equipamientos del sistema bajo prueba, el usuario podrá acceder a sus datos y resultados de experimentos. Además, Pueden desplegarse componentes en el cloud que frecuentemente verifiquen el buen funcionamiento del sistema bajo prueba y ante una falla, ejecuten rutinas de recuperación ante fallas o notifiquen lo sucedido.

Otras de las características deseables de cloud computing es la confiabilidad. Los componentes de software desplegados en el cloud no están expuestos a fallas en el hardware.

En cuanto a seguridad, los proveedores de servicios de cloud computing implementan mecanismos de seguridad para evitar accesos no autorizados, robo de información, etc.

La integración con el cloud supone un elemento adicional entre el equipo bajo prueba y la interfaz con el usuario. Por lo tanto, la integración debe implementarse cuidadosamente para evitar aumentos de latencia o pérdida de ancho de banda.

2.3 Nuevas herramientas para la implementación de laboratorios remotos.

Las plataformas de escritorios remotos, VPNs y videoconferencias se han utilizado desde hace varios años. Sin embargo, debido a las condiciones de aislamiento motivadas por la pandemia de Covid-19, han sido muy utilizadas durante el año 2020, y varias empresas han implementado mejoras en sus servicios. Por otro lado, el trabajo realizado durante el último año ha mostrado que estas plataformas son adecuadas para el despliegue de laboratorios remotos.

Existen varios problemas que resolver, como la forma de integrar con servidores virtuales en el cloud o realizar experimentos de performance.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1 Resultados obtenidos

Se han obtenidos resultados experimentales sobre los sistemas bajo prueba y los laboratorios remotos, como también experiencia de uso en docencia. Todos estos resultados han sido documentados en diferentes publicaciones.

Se realizaron experimentos para analizar el comportamiento del protocolo Zigbee en situaciones de carga elevada. Se midieron la ocupación real del canal de comunicaciones y latencias. El laboratorio remoto permitió que los experimentos pudieran realizarse de forma continua durante un periodo de tiempo de varias semanas, siendo monitoreado tanto de forma presencial como remota [3][6].

Se realizaron experimentos de performance sobre los laboratorios remotos implementados a través del cloud, midiendo parámetros de

performance, latencias y contribución a la misma de cada uno de los componentes [4]. En cuanto a experiencias en docencia, los laboratorios remotos descritos en este artículo han sido utilizados en 6 asignaturas de 3 carreras diferentes [1].

3.2 Resultados esperados

El trabajo futuro estará centrado en tres ejes:

- Realizar experimentos de performance sobre los laboratorios remotos basados en VPN y escritorios remotos.
- Integrar estos laboratorios remotos con plataformas de cloud computing.
- Investigar nuevas herramientas o plataformas que puedan mejorar las prestaciones de los laboratorios remotos.
- Realizar experimentos de prueba de concepto.

En cuanto al último eje, debido a que se espera que durante el año 2021 el cursado a nivel universitario sea mixto, combinando cursado presencial y remoto, y que los investigadores que forman parte de los proyectos presentados en la sección “Contexto” dictan varias asignaturas relacionadas con computación en varias universidades, se requerirá un uso intensivo de las plataformas de laboratorios remotos implementadas. Estas prácticas permitirán recolectar experiencias de usuario y detectar fallas o posibles mejoras.

4. FORMACIÓN DE RECURSOS HUMANOS

Por un lado, se mencionan los estudiantes que fueron usuarios de los laboratorios remotos. Si bien estos estudiantes no fueron parte de los proyectos que financian el trabajo presentado en este artículo, se destaca que los laboratorios remotos presentados contribuyeron a su formación práctica, la cual, en situación de pandemia, hubiera sido muy difícil sin contar con estas herramientas.

Por otro lado, dos de los investigadores que forman parte del proyecto son estudiantes de carreras de maestría.

AGRADECIMIENTOS

Los autores agradecen el financiamiento recibido de la Universidad Nacional de Cuyo a través de los proyectos B071 “Laboratorio remoto y nómada de arquitectura de computadoras destinado a enseñanza e investigación”, y B076 “Dispositivos de lógica programables como alternativa de solución en problemas de seguridad en Internet de las Cosas”.

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; Communication channel occupation and congestion in wireless sensor networks. *Computers & Electrical Engineering*. 72, 846–858 (2018).
- [4] 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)
- [5] 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.
- [6] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; Zigbee WSN Round Trip Latency in Function of Channel Occupation and Nodes Configuration, In *IEEE ArgenCon* 2016.

EXPERIENCIA DE ACCESIBILIZAR UN MOOC CONSIDERANDO LA PLATAFORMA Y RECURSOS MULTIMEDIALES

Javier Díaz, Alejandra Schiavoni, Paola Amadeo, Ivana Harari,
Alejandra Osorio, Federico Carrilao Ávila
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, federicoca95@gmail.com

RESUMEN

El presente artículo describe la investigación que se viene llevando a cabo en el LINTI, Laboratorio de Investigación en Nuevas Tecnologías Informáticas de la Facultad de Informática de la UNLP, sobre Accesibilidad Web y su aplicación en diferentes ámbitos, considerando principalmente las plataformas de aprendizaje abiertas.

Actualmente se continúa trabajando sobre un curso masivo en línea, MOOC (MOOC – Massive Open Online Courses) sobre Accesibilidad Web. El surgimiento de los MOOCs ponen a disposición un conjunto de cursos en diferentes áreas, y flexibiliza el aprendizaje superando los obstáculos del tiempo y espacio. Sin embargo, al diseñar un MOOC deben atenderse las necesidades de las personas que sufren algún tipo de discapacidad. A partir del resultado de distintas etapas de análisis de los aspectos de usabilidad y accesibilidad de la plataforma edX utilizada y de la naturaleza de los contenidos y su organización dentro del curso, se van realizando modificaciones y adaptaciones que permiten alcanzar las normas y estándares establecidos. Además, se analizan las pautas y procedimientos para la construcción de recursos y materiales textuales y multimediales accesibles.

Palabras clave: MOOC, Recursos educativos abiertos, accesibilidad web.

CONTEXTO

En la Facultad de Informática se utilizan plataformas de código abierto para las gestiones académicas desde hace más de quince años, incluyendo sistemas de gestión de aprendizaje como Moodle, repositorios abiertos y sistemas de gestión administrativa, como SIU Guaraní. Además, se vienen realizando mecanismos de integración entre diferentes plataformas, lo que permite una mayor flexibilidad y aprovechamiento en el uso de las mismas. A partir del surgimiento en el mundo de los cursos abiertos masivos en línea, MOOCs, se comenzó a utilizar la plataforma edX para implementar el curso sobre Accesibilidad Web, que ya lleva cuatro ediciones.

El proyecto descripto en este artículo se desarrolla en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas, LINTI de la Facultad de Informática de la UNLP y está enmarcado en el proyecto I+D 11/F028 “De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región”, aprobado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz. Este proyecto hace hincapié entre otros puntos en la formación de competencias y habilidades digitales para todos los ciudadanos, en una sociedad que se plantea como digital.

Dada su relevancia, el tema de accesibilidad web se viene trabajando en la Facultad desde el año 2002, y se incorporó esta temática en el plan de estudios de las carreras que se dictan en la institución, a través de la asignatura Diseño Centrado en el Usuario. El libro “Guía de recomendaciones para diseño de software centrado en el usuario” registra aproximadamente 10.000 descargas desde el año 2013 [1]. También se institucionalizó su abordaje mediante la creación de una Dirección de Accesibilidad desde el año 2010, se desarrollaron tesis, 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” (2017), “Por una Web Inclusiva” (2018), “Por una Web inclusiva: abordaje en escuelas secundarias” (2019) [2]. También, se aprobó el Proyecto de Desarrollo e Innovación sobre “Rampas Digitales Innovativas para Personas con Discapacidad”, en la convocatoria Agregando Valor 2018-2019, del Ministerio de Ciencia, Tecnología e Innovación de la Nación Argentina.

1. INTRODUCCIÓN

Actualmente se está dando un proceso de cambio en la educación basado en los nuevos modelos de e-Learning y el desarrollo de la tecnología. En esta etapa el surgimiento de los MOOCs representa una manifestación muy evidente y de gran recorrido, convirtiéndose en protagonistas importantes en la educación en el año 2012 [3], aunque existieron algunas experiencias anteriores. Los MOOCs surgieron como consecuencia de las nuevas tendencias internacionales en el aprendizaje basado en la tecnología, como parte del movimiento educativo abierto [4]. La flexibilidad del aprendizaje a través de los cursos MOOC permite superar los obstáculos de tiempo, espacio y ritmo de estudio al igual que en

cualquier contexto e-learning, así como la promoción de la comunicación continua y la interacción entre todos los participantes. Pero son precisamente los MOOCs aquellos que pueden llegar a audiencias globales, y en este contexto, es esencial que se tenga en cuenta que se trata de los grupos de usuarios más vulnerables aquellos que pueden quedar excluidos de la sociedad del conocimiento [5]. Por otro lado, los MOOCs imponen cambios metodológicos, diseños colaborativos e interactivos, materiales ubicuos y atractivos que faciliten y promuevan la navegación y el descubrimiento, en entornos diseñados para tal fin. Gran parte del contenido suele ser multimedia basado en video y en muchos casos incluyen juegos serios.

No debemos olvidar que el acceso a los MOOCs supone una dificultad añadida a las personas con diversidad funcional debido a la necesidad de desarrollar habilidades específicas y cambiantes por los contenidos audiovisuales y elementos interactivos disponibles en estos cursos añadiendo una nueva dificultad a los requisitos de accesibilidad [6].

Respecto a la Accesibilidad, es una temática compleja que reúne características de índole social, moral, legal, educativa como técnica. Va desde los derechos de las personas con discapacidad al acceso a la información, normas y recomendaciones internacionales sobre accesibilidad web, como cuestiones de implementación en el código. Esto también ha permitido a la sociedad ser más inclusiva con las personas con discapacidad. Ahora podemos identificar la posibilidad de permanecer conectados y tecnológicamente informados mediante la instalación de apps y las plataformas de redes sociales. Además, las redes sociales han influido enormemente en la comunidad de personas con discapacidad en los enfoques académico, social, de entretenimiento y de negocios. Las personas con discapacidad quieren ser partícipes igualmente de este fenómeno sociológico, pero encuentran mayores dificultades, y pueden presentar una

amplia gama de habilidades y necesidades muy heterogéneas, dependiendo de la naturaleza de la discapacidad y del grado de afectación.

En el ámbito educativo, hay un consenso general que es necesario analizar y plantear políticas de accesibilidad en plataformas que albergan Recursos Educativos Abiertos (Open Educational Resources-OER), lo que significa que los repositorios abiertos deben ser diseñados teniendo en mente el concepto de accesibilidad [7]. Los MOOCs están estrechamente relacionados con los OER en tanto representan contenido abierto al que se le agrega la característica de masividad.

Para promover la accesibilidad, el Consorcio W3C creó la Web Accessibility Initiative (WAI) para crear los lineamientos para el contenido web, las herramientas de autoría y los navegadores y otras plataformas [8]. Estos lineamientos son un buen punto de partida para entender las necesidades de accesibilidad de los usuarios. Si bien existen herramientas específicas para evaluar la accesibilidad, los comentarios y opiniones de los usuarios también tienen un valor muy importante para el establecimiento de pautas de diseño de MOOCs accesibles.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Dentro de la línea de trabajo que venimos describiendo, las tareas de investigación, desarrollo e innovación que se llevan a cabo en este proyecto incluyen los temas relacionados a e-learning y accesibilidad web. Esto último, en relación también a los aspectos de usabilidad. En la línea de e-learning se está trabajando con plataformas de código abierto, tanto sistemas de gestión de aprendizaje (LMS) como herramientas para el soporte de MOOCs. También se trabaja lo relacionado a la integración de distintas plataformas que se utilizan en la gestión académica. El tema de accesibilidad web se trata en forma transversal,

evaluando el cumplimiento de las normas respectivas en cada uno de los casos.

En esta etapa se está trabajando con la plataforma OpenEdX [9] para el soporte de MOOCs. Se eligió dicha plataforma por las facilidades de instalación y configuración, por los módulos disponibles y la amplia comunidad de usuarios que contribuyen al proyecto. Provee toda la funcionalidad básica de las herramientas de MOOCs, en lo relativo a la creación de contenido, gestión de usuarios, confección de evaluaciones y utilización de foros y wikis. La comunidad de edX lleva adelante importantes políticas para garantizar la accesibilidad web a través de las cuales se compromete a priorizar los errores o barreras de accesibilidad que pueda presentar la herramienta y ofrece también una guía de buenas prácticas para los desarrolladores de los contenidos de los cursos [10].

Usando la plataforma edX se creó un MOOC sobre Accesibilidad Web, que se puso en producción en el año 2017 y se dicta todos los años desde esa fecha. El curso se encuentra disponible en: <https://actividades.linti.unlp.edu.ar>. La plataforma de base está instalada y es administrada por becarios del proyecto, lo que permite sucesivas actualizaciones y adaptaciones según las necesidades que se presentan. Se realizan modificaciones para que cumpla con las normas de accesibilidad vigentes. Además, se evalúan cuestiones de usabilidad y facilidad de uso en cuanto a la localización y acceso a los contenidos; aspectos de instalación y configuración y cuestiones de accesibilidad.

A partir de las sucesivas ediciones, se vienen realizando evaluaciones con el objetivo de analizar el acceso e impacto del contenido, las actividades y ejercitaciones planteadas a través de las experiencias y devoluciones de los participantes, en algunos casos fueron personas con alguna discapacidad. Se realizan encuestas al comenzar y al finalizar el curso con el objetivo de recabar información del perfil de los estudiantes y de las características de

interacción con la plataforma y los resultados obtenidos [11]. Es importante destacar en este punto que en las últimas dos ediciones del curso se inscribieron en total 13 estudiantes con algún tipo de discapacidad visual. En una encuesta realizada con los alumnos que cursaron en 2019, 8 con discapacidad visual contestaron las preguntas, 4 de ellos ciegos y el resto con poco residuo visual, que interactuaron con el curso a través de lectores de pantalla y uso exclusivo de teclado.

Al momento de testear la accesibilidad se aplican criterios de normativa tanto a la naturaleza y presentación del contenido, como a la estructuración y organización del curso respecto a los temas abordados y a las estrategias de evaluación. Se trata de proveer múltiples formas de representar el contenido, ya que varía la forma en que cada estudiante percibe y comprende el material propuesto. Además, se hace hincapié en la accesibilización de los recursos multimediales incluidos, considerando imágenes, videos y documentos textuales estandarizados. Se realiza una reestructuración de la organización de las unidades, planteando diferentes secuencias de acuerdo al perfil del alumno. Esto permite personalizar el trayecto de aprendizaje según sus necesidades.

El uso de espacios de comunicación entre los alumnos representa una de las actividades más destacadas y es importante para fomentar la interacción y la generación de conocimiento compartido, logrando una forma de aprendizaje colaborativo.

En la organización de un curso de esta naturaleza se requiere aplicar diferentes estrategias pedagógicas para poder, en el proceso de enseñanza a distancia, concientizar a los estudiantes, proponer el ponerse en el lugar del otro, incentivar la investigación de la problemática de los usuarios con discapacidad. Asimismo, se trabaja en pos de garantizar una experiencia del usuario inclusiva, promover la colaboración para la búsqueda conjunta de soluciones accesibles y formas de aplicarlas en

las producciones informáticas que se desarrollen.

3. RESULTADOS Y OBJETIVOS

Según las líneas de trabajo descritas, se plantean los siguientes objetivos:

- Analizar y aplicar las nuevas guías de accesibilidad propuestas por la comunidad de edX, con el fin de mantener actualizada la plataforma según estos criterios.
- Investigar nuevas metodologías de diseño de MOOCs accesibles aplicadas en universidades del mundo, analizando sus ventajas y posibles adecuaciones al ámbito local.
- Realizar entrevistas individuales con diferentes personas con discapacidad que interactúen con la herramienta.
- Como base de la planificación estipulada, analizar los resultados obtenidos en cada etapa, corrigiendo posibles falencias en un ciclo de mejora continua.
- Redefinir secuencias de aprendizaje y evaluaciones personalizadas para los diferentes perfiles de alumnos.
- Promover la producción de material accesible, según las inquietudes, expectativas y necesidades de cada estudiante.
- Llevar a cabo estudios comparativos entre la experiencia del MOOC y los cursos en un LMS tradicional, que permitan aportar nuevos conceptos a los entornos de aprendizaje y enseñanza que se generan con estas herramientas.
- Utilizar evaluaciones heurísticas con herramientas específicas para testear la accesibilidad, teniendo en cuenta el rol del usuario en su interacción con la plataforma. Evaluar la posibilidad de intervención y participación de una

persona con discapacidad en los diferentes roles.

- Medir el grado de aceptación del curso por parte de los participantes a través de encuestas y encuentros y realizar análisis de datos considerando distintas variables como formación, edad, género, procedencia, participación y rendimiento, entre otros.
- Definir técnicas de diseño inclusivo, a partir de las evaluaciones realizadas y de la información recolectada, y utilizar estas técnicas en futuros cursos, partiendo del caso de uso desarrollado.
- Replantear componentes del proceso de aprendizaje, relacionadas a la naturaleza del contenido y al diseño temporal, a fin de adaptar el curso según las evaluaciones realizadas.
- Definir y establecer pautas de diseño y construcción de MOOCs que se utilicen como punto de partida para la creación de cursos masivos sobre temas que se investigan y sobre los cuales se viene trabajando en el LINTI.

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, dos alumnos becarios de la carrera de Licenciatura en Informática 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.

Las Jornadas de Accesibilidad que se realizan todos los años permiten intercambiar experiencias entre los participantes del curso y exponer los trabajos realizados. El video de la Jornada del año 2020 se encuentra disponible

en: <https://www.youtube.com/watch?v=0JWOP6D1GPo>.

5. REFERENCIAS

[1] Díaz, J., Amadeo, P., Harari, I. “Guía de recomendaciones para diseño de software centrado en el usuario”. Editorial de la Universidad Nacional de La Plata (EDULP). ISBN: 978-950-34-1030-1. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/32172>

[2] Dictamen Proyectos de Extensión Acreditados en la UNLP. [https://unlp.edu.ar/proyctosex/proyectos_de_extension_acreditados_y_subsidios-4708](https://unlp.edu.ar/proyctosex/ proyectos_de_extension_acreditados_y_subsidios-4708)

[3]. Daniel, J. “Making Sense of MOOCs: Musings in a Maze of Myth”, Paradox and Possibility. (2012). Disponible en: <http://blog4222.blogspot.com.ar/2012/09/making-sense-of-moocs-musings-in-maze.html>.

[4] Ramírez, M., Burgos, J. “Movimiento Educativo Abierto. Acceso, colaboración y movilización de recursos educativos abiertos.” México: CIITE (2012).

[5] De Waard I., Gallagher M.S., Zelezny-Green R., Czerniewicz L., Downes S., Kukulska-Hulme A., Willems J. “Challenges for conceptualising MOOC for vulnerable learner groups.” eMOOC2014 MOOC stakeholder summit, pp. 33-41. Lausanne, Switzerland. P.A.U. Education, S.L. U. Cress, C. Delgado-Kloos (2014).

[6] Iniesto, F., Rodrigo, C. “Pautas para la evaluación de la accesibilidad en las plataformas MOOC.” Actas del VI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (Bengochea, Luis; Gutiérrez, José María; García, Antonio and García, Eva eds.), 29 - 31 october 2014, Universidad de Alcalá,

Alcalá de Henares (Spain), Universidad de Alcalá, pp. 57–64 (2014)

[7] Law, P., Perryman, L.-A., and Law, A. “Open educational resources for all? Comparing user motivations and characteristics across The Open University’s iTunes U channel and OpenLearn platform.” 2013.

[8] Web Content Accessibility Guidelines – WCAG, 2018. <https://www.w3.org/WAI/intro/wcag>

[9] Open edX Portal | Open Source MOOC Platform <https://open.edx.org>

[10] “Accessibility Best Practices Guidance for Content Providers”. Disponible en: <https://edx.readthedocs.io/projects/edx-partner-course-staff/en/latest/accessibility/index.html>

[11] Díaz, J., Schiavoni, A., Amadeo, P., Harari, I. “Evaluación de un MOOC sobre Accesibilidad Web desde la experiencia del estudiante. Un caso de estudio”. JAIIO 2020, Jornadas Argentinas de Informática, CABA, Argentina, 19 - 30 Octubre, 2020.

IDENTIFICACION DE REGULARIDADES EN MATEMATICA, MEDIANTE DETERMINACION DE PATRONES Y CON USO DE SOFTWARE LIBRE

Elisa Silvia Oliva¹, Vanesa Gallardo², María Inés Ciancio², Andrea Salas³

¹Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales

²Departamento de Geología, Facultad de Ciencias Exactas, Físicas y Naturales

³Departamento de Geofísica y Astronomía, Fac.de Ciencias Exactas, Físicas y Naturales

Universidad Nacional de San Juan

elisaoliva65@gmail.com, vanesagallardol@gmail.com, miciancio@hotmail.com, andreamarielsalas@gmail.com

RESUMEN

Realizar propuestas didácticas que permitan llevar al aula experiencias de aprendizaje que involucren procesos de descubrimiento, en la búsqueda de la formalización de patrones en el proceso de generalización en Educación matemática es una tarea que impacta : tanto en la calidad de la educación, como en la curiosidad y motivación del alumno, al incluir uso de Tecnologías de la Información y la Comunicación (TIC) en los procesos de aprendizaje: anima a los estudiantes a realizar trabajos exploratorios, a potenciar la indagación, investigación y creatividad; y así poder transitar del proceso fenomenológico a la acción epistemológica.

Este proyecto aborda una investigación aplicada en el área de Educación en Matemática, se orienta a analizar la presencia de núcleos generadores en contenidos matemáticos, como: cálculo de límites, integrales; también a investigar la existencia de patrones en Transformaciones lineales, o regularidades en los autovalores y/o autovectores de algunas matrices ó en procesos del Análisis Numérico, de Matemática Discreta, entre otras áreas de Matemática.

La propuesta se sustenta en adicionar al trabajo operacional sobre desarrollos teórico-prácticos realizados en forma manual, herramientas tecnológicas, que ayuden a anticipar

regularidades sobre nuevas funciones o transformaciones lineales, por ejemplo. Es de destacar que durante la experiencia, son importantes aspectos a potenciar tales como la observación, intuición y la predicción; pues pronosticar comportamientos y descubrir relaciones al comparar, son los pasos iniciales para la inducción y generalización de patrones.

Palabras clave: Regularidad-inducción; Experiencias de aprendizaje; TIC

CONTEXTO

La línea de investigación descrita en este artículo se desarrolla en el proyecto de investigación “Generalización de Patrones en Matemática, con ayuda de Nuevas Tecnologías y Método Inductivo “, en el área de Educación matemática. Se implementa en grupos de estudiantes mediante experiencias de aprendizaje que involucren procesos de descubrimiento, en la búsqueda de la formalización de patrones en procesos inductivos que permitan alcanzar una fase de generalización. Potenciando la indagación y permitiendo iniciar al estudiante en la investigación, al incluir procesos que pueden ser enriquecidos con trabajo con TIC.

El proyecto mencionado es parte de la convocatoria del Consejo de Investigaciones Científicas y Técnicas y de Creación Artística (Cicitca)- Universidad Nacional de San Juan (UNSJ), para el período 2020-2022, y tiene su base de trabajo en cátedras de matemática de

las carreras de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFN).

1. INTRODUCCIÓN

La generalización es uno de los sistemas para producir conocimiento, se logra de la interacción de tres pilares: el primero es un problema fenomenológico basado en la selección de las determinaciones sensibles, en el que participan: la intuición, la atención, la intención, la sensibilidad, etc. El segundo es un problema epistemológico, que consiste en la extrapolación o generalización propiamente dicha y a través de la cual se produce el nuevo objeto. El tercero es un problema semiótico, que resulta de los medios a través de los cuales se denota el objeto generalizado.

La idea básica implicada en la noción de generalización, es que toda situación repetida con regularidad da lugar a un patrón, que es la traducción elegida para el término inglés pattern. Los patrones suelen formarse a partir de un núcleo generador; en algunos casos el núcleo se repite, en otros el núcleo crece de forma regular.

Hoy se pueden investigar ciertas situaciones matemáticas con uso de tecnologías que no sólo permiten proceder de maneras que no se podría hacer sin ella, por la velocidad y variedad de exploraciones [1] que lleva coleccionar ideas, acumular conocimiento de tal manera que cuando finalmente se obtenga una fórmula se pueda identificar en ella (leyendo los símbolos entre líneas) lo que ya se sabe, y constatar que las distintas representaciones (una tabla, cambios geométricos, la gráfica de la función y/o su expresión simbólica) ofrecen la misma información, pero además detalles complementarios.

Estas tecnologías innovadoras permiten modelar situaciones no sólo algebraicamente sino también mediante herramientas más cercanas a los significados que queremos representar y estudiar. Posponer el uso de símbolos enriquece nuestro entendimiento de la situación como asimismo nos permite escudriñar ciertos aspectos que son opacos en la fórmula. Es decir son los recursos que

permiten transitar del proceso fenomenológico a la acción epistemológica, de la búsqueda de la formalización de patrones en el proceso de generalización.

En los estándares de competencias sobre el manejo de TICs, propuestos por UNESCO, se afirma que: "Para vivir, aprender y trabajar con éxito en una sociedad cada vez más compleja, rica en información y basada en el conocimiento, los estudiantes y los docentes deben utilizar la tecnología digital con eficacia". La habilidad de "aprender a aprender" [2], [3], socialmente requiere la formación de personas capaces de realizar un manejo autónomo de herramientas cognitivas. Estos conceptos son propios de una educación centrada en quien aprende, que es transversal para actuar en cualquier área del mundo laboral o en la continuidad de estudios de nivel superior.

Este escenario actual impone acciones que deberán, a través de la integración de las TIC, cubrir las competencias digitales que los alumnos deben incorporar en su bagaje cultural y de conocimientos para desarrollar su autonomía en el camino de aprender a aprender.

Meirieu [4] sugiere entusiasmar a los alumnos: para lo cual deben emplearse estrategias "accesibles y difíciles al mismo tiempo". Estas estrategias son tales cuando el alumno siente que es capaz de lograr el objetivo propuesto y percibe la existencia de una hipótesis que todavía no le es propia. Por su parte, Maggio [5] distingue dos modalidades al incorporar TIC a las prácticas de enseñanza, inclusiones efectivas e inclusiones genuinas. La inclusión genuina, se refiere a un proceso de integración de TIC de orden epistemológico que "reconoce el complejo entramado de la tecnología en la construcción del conocimiento en modos específicos por campo disciplinar y emula ese entramado en el plano de la práctica de la enseñanza". De esta forma, en lugar de ocupar el lugar subsidiario de la superficie y el agregado, los desarrollos tecnológicos pasan a formar parte del cuerpo mismo del área en la que han sido incluidos.

En el proyecto se hace uso del software libre Geogebra, con el objeto de que el alumno pueda hacer un aprendizaje de la matemática más dinámico; dado que permite proponer ac-

tividades de investigación y experimentación en diferentes niveles educativos, y desarrollar actividades que involucren reflexión, descubrimiento, formalización; contando con el adecuado manejo de los contenidos que se van abordar.

Una de las utilidades de Geogebra [6] es la construcción de conocimiento, pues es una herramienta que permite visualizar simbólica y gráficamente información matemática. Así, Buendía Avalos [7], indica que las gráficas son excelentes formas de comunicación para plantear y resolver problemas, ejemplificar, ilustrar y comunicar resultados. Los encontramos en medios tan diversos como los periódicos, los libros de texto o los artículos de investigación. Por eso la educación tiene entre sus objetivos enseñarnos a hacer e interpretar gráficas.

Nuestra práctica docente permite señalar que siempre al utilizar las herramientas tecnológicas en los procesos de enseñanza – aprendizaje, como el uso del software Geogebra (aunque es de sencillo manejo), no resolverá todos los ajustes que son necesarios en el contexto educativo, pero se debe tener en cuenta que es un medio para presentar información el estudiante: que capta su atención, y además le permite ejercitar habilidades de pensamiento y estrategias de resolución de problemas, iniciarse e involucrarse en contextos de investigación con los que puede conjeturar, deducir, generalizar, reflexionar sobre contenidos algebraicos y geométricos. También el uso del software permite generar ambientes abiertos e inclusivos de aprendizaje de la matemática en sus diferentes contextos.

El uso de TIC, en este caso el software Geogebra, en el ámbito de enseñanza universitaria aporta al estudiante la variedad y velocidad en la repetición de hechos, para generar espacios de apropiación de contenidos encontrando los rasgos comunes en un grupo definido, para llegar a conclusiones de los aspectos que lo caracterizan. Las generalizaciones a que se arriban tienen una base empírica, partiendo de observaciones que se hacen sobre fenómenos particulares de una clase y luego a partir de ellos se obtiene inferencias de la clase entera. Su colaboración en la ejecución

de generalización permite obtener nuevas metodológicas de trabajo [8] que no limiten las prácticas sólo al uso de recursos tradicionales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación en que se enmarca este proyecto es de investigación aplicada en Educación. Resulta importante la determinación de núcleos generadores en contenidos de las áreas del cálculo, del álgebra lineal, etc; en matemática, pues en algunos casos el núcleo se repite, en otros el núcleo crece de forma regular, identificando estrategias y formas de razonamiento, distinguiendo la relación parte-todo presente en los procesos de generalización que validan un proceso o fórmula.

Entre los objetivos se plantea:

- Reconocer regularidades en matemática como fuente de aprendizaje, despertando la capacidad de inspeccionar y cuestionar los patrones de generalización
- Animar a los estudiantes a realizar trabajos exploratorios, con aplicación del software Geogebra, en actividades de Álgebra y Análisis, para potenciar la indagación, exploración y creatividad
- Aprovechar el software como entorno de experimentación y reflexión, por sus potencialidades de visualización gráfica y simbólica, para propiciar una participación activa del alumno, nativo digital, en la construcción de procesos de generalización

Para lo cual se han planificado las siguientes etapas:

- a) Determinar el problema, las condiciones y necesidades (diagnóstico)
- b) Planteamiento de los interrogantes
- c) Consideración de dificultades
- d) Elaboración de materiales didácticos

- e) Intervenciones didácticas
- f) Recolección de información
- g) Procesamiento de información
- h) Reajuste de estrategias
- i) Difusión y publicación de resultados

Los integrantes de este proyecto participan, de variadas cátedras de Matemática, en las distintas carreras donde forman parte de equipos de cátedra. Por lo que, las áreas donde se puede implementar son muy diversas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto se encuentra iniciando el segundo año de desarrollo y debido a la pandemia de covid 19, al no disponer de una situación de cursado con presencialidad, en los estudiantes, se observa:

- 1) Los tiempos de contacto con estudiante-docente son muy limitados, a su posibilidad de conectividad de ambos; priorizando siempre la comunicación on-line sólo estrictamente a los contenidos priorizados de las asignaturas de cursado.
- 2) La comunicación sólo se da con una parte de los estudiantes, los que sí tienen conectividad.
- 3) En las comunicaciones por video conferencia, no quedan registros que podrían servir de partida a estudios de análisis de los procesos de inducción de los estudiantes.
- 4) Imposibilidad de trabajar variados planteos, en un mismo grupo de estudiantes, para estudiar distintos abordajes de problemas para ver como plantean la determinación de patrones y regularidades.

Se han realizado los siguientes trabajos [9],[10]:

- Con alumnos de primer año de Lic. en Geología en Matemática I, sobre la identificación de regularidades en el cálculo del error para funciones aproximadas por polinomios de Taylor

- Con alumnos de primer año de Lic. en Geofísica y Lic. en Astronomía, en Álgebra, sobre el reconocimiento de regularidades en Transformaciones lineales.

Se espera poder implementar más experiencias de aprendizaje y avanzar en otras etapas del proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado por cuatro docentes- investigadores, participan del proyecto dos alumnos de facultad, que son ayudantes y adscriptos a cátedras, respectivamente; con lo cual su intervención les ayuda a aprender a realizar actividades de investigación, y cómo integrarse en un equipo de investigación existente.

Además, la producción científica realizada se puede volcar a temas de actualización de las asignaturas en las que forman parte de los equipos de cátedra, con lo cual los alumnos también se retroalimentan de las investigaciones.

5. REFERENCIAS

- [1] A. Arcavi, N. Hadas (2000). *Computer mediated Learning: An example of an approach*. International Journal of Computers for Mathematical Learning, 2000.
- [2] A. Pelliccia, M. Vasta, *Aprender a Aprender, Ciencia e investigación*, Tomo 66 N° 5, 2016
- [3] L. Lluch Molins, Ma. Portillo Vidiella, *La competencia de aprender a aprender en el marco de la educación superior*, Revista Iberoamericana de Vol. 78, No. 2, pp. 59-76, ISSN: 1022-6508 / ISSN: 1681-5653 Organización de Estados Iberoamericanos (OEI/CAEU), 2018
- [4] Ph. Meirieu, *Aprender, sí. Pero ¿Cómo?*, Octaedro, Madrid, 1992.
- [5] M. Maggio, *Enriquecer la enseñanza. Los ambientes con alta disposición tecnológica*

- ca como oportunidad*. Editorial Paidós, Buenos Aires, 2012.
- [6] A. Carrillo de Albornoz Torres, *Geogebra*. Editorial RA-MA S.A, Madrid.2009.
- [7] G. Buendía Avalos, *¿Como ves?* Revista de divulgación de la Ciencia de la UNAM. México. Año 17. No. 203. Octubre 2015.
- [8] A. Zangara, *Uso de nuevas tecnologías en la educación: una oportunidad para fortalecer la práctica docente*. Revista Puertas Abiertas - ISSN 1853-614X.
- [9] V. Gallardo, M. Ciancio, E. Oliva, *Identificando regularidades en el cálculo del error para funciones aproximadas por polinomios de Taylor*. Noticiero de la UMA, Vol.55. Argentina. 2021.
- [10] E. Oliva, J. Avila, V. Gallardo, M. Ciancio, *Reconocimiento de regularidades en Transformaciones lineales*. Noticiero de la UMA, Vol.55. Argentina. 2021.

**IMPACTO DE LAS CONDICIONES SOCIOECONÓMICAS
EN EL RENDIMIENTO Y LA DESERCIÓN ESTUDIANTIL**
ISTVAN, Romina¹; BACIGALUPE, Maria de los Ángeles²; LASAGNA, Valeria¹

¹Universidad Tecnológica Nacional, Facultad Regional La Plata
Grupo de I&D Aplicado a Sistemas Informáticos y Computacionales, GIDAS
Av. 60 s/n° esquina 124, CP 1900, La Plata, Buenos Aires, Argentina

²CONICET, División Etnografía, Museo de La Plata, UNLP
ristvan@frlp.utn.edu.ar; mariabacigalupe@conicet.gov.ar; valerial@frlp.utn.edu.ar

RESUMEN

Se define el nivel socioeconómico (NSE) como el conjunto de bienes materiales y características no económicas, entre las que se incluye el prestigio social y la educación, las cuales se asocian al nivel de estrés y calidad de vida, así como a aspectos generales de salud y de habilidad cognitiva [1]. Una amplia cantidad de investigaciones dentro de las Neurociencias Cognitivas, ha demostrado que las diferencias en las experiencias de estimulación asociadas a distintos Niveles Socioeconómicos (NSE) afectan el desarrollo cerebral, desde el nivel bioquímico hasta la organización de las funciones psicológicas más complejas [2]; evidenciándose una correlación entre el nivel socioeconómico de una familia y el desarrollo del cerebro de sus hijos.

Sobre esta línea de investigación se desarrolla el presente estudio, el cual tiene como objetivo realizar una revisión de los factores entendidos como mediadores del NSE y propuestos como condicionantes del desarrollo cognitivo; con el fin de contrastar los resultados con una muestra de datos de estudiantes de las carreras de Ingeniería brindadas en la Universidad Tecnológica Nacional - Facultad Regional La Plata (UTN-FRLP).

Se plantea como método de análisis de datos la Minería de Datos mediante la cual se analiza la correlación entre: NSE; rendimiento académico y deserción estudiantil universitaria.

Palabras clave: Estatus Socioeconómico, Nivel Socioeconómico, Neurodesarrollo, Deserción Estudiantil, Rendimiento Académico.

CONTEXTO

La línea de investigación presente se encuentra inserta en el Proyecto de Investigación y Desarrollo: «Determinación de perfiles de riesgo de deserción estudiantil en UTN-FRLP utilizando técnicas de minería de datos».

El proyecto se encuentra homologado por Rectorado con código TEUTNLP0007653 y financiado por la Universidad Tecnológica Nacional, UTN.

1. INTRODUCCIÓN

Según Arán Filippetti [3] las investigaciones que analizan la relación entre el NSE y el desempeño cognitivo se encuentran clasificadas en dos tipos :

1) Aquellas que analizan qué indicadores socioeconómicos se relacionan con el desempeño cognitivo. Dentro de ellas, Duncan & Magnuson [4] proponen tres factores de estudio para esta perspectiva: (a) nivel de educación, (b) ingreso familiar y (c) nivel ocupacional .

En particular, Noble y Farah en 2005 [5] destacan que existe una relación entre la cantidad de años de formación educativa de los padres y el tamaño de regiones cerebrales relacionadas con el lenguaje, la lectura y las funciones ejecutivas de sus hijos. Estas

diferencias llegan hasta un 3% si se comparan los hijos de padres con secundaria completa (por lo menos 12 años de educación continua) con los hijos de padres universitarios (con más de 15 años de formación).

En el año 2015 el mismo equipo de investigadores publicó los resultados obtenidos sobre ingresos económicos y pobreza y los fundamentos biológicos de estos efectos. El estudio consistió básicamente en la realización de resonancias magnéticas a los cerebros de 1099 niños residentes en Estados Unidos con el fin de comparar las diferencias de su estructura cerebral. Los resultados arrojaron que aquellos niños pertenecientes al grupo de ingresos más bajos tenían hasta un 6% menos de superficie que los niños de las familias de altos ingresos, presentando diferencias en la estructura del cerebro, particularmente en áreas asociadas con el lenguaje y las habilidades de toma de decisiones. Los puntajes de los niños en las pruebas que miden las habilidades cognitivas, como la capacidad de lectura y memoria, también disminuyeron con el ingreso de los padres [6].

II) En segunda instancia se encuentran aquellas investigaciones que estudian cuáles son las variables mediadoras y cómo influyen en el desarrollo. Surgen para analizar los diferentes mecanismos mediadores de los indicadores y demostrar que éstos no actúan de la misma manera en el desempeño cognitivo.

Si bien no se encuentran definidos de manera estricta los mediadores involucrados, los enfoques que surgen de esta línea comparten mecanismos en común. Como ser, Brooks-Gunn y Duncan [7] proponen como mecanismos mediadores: (a) la salud y la nutrición, (b) el ambiente en el hogar, (c) la interacción padres-hijos, (d) la salud mental de los padres y (e) las condiciones del barrio. Y Guo y Harris [8] analizan: (a) el ambiente físico del hogar, (b) el estilo parental, (c) la estimulación cognitiva en el hogar, (d) la salud

del niño al momento del nacimiento y (e) la salud del niño durante la infancia.

Carrillo López [9] presenta una revisión bibliográfica acerca de cómo afectan las diferencias socioeconómicas en el desarrollo neuropsicológico infantil. En ella pueden observarse mediadores, que si bien pueden considerarse mecanismos diferentes, se encuentran próximamente vinculados y difícilmente delimitables en algunos casos.

Factores prenatales. Existe relación directa entre el NSE Bajo (NSE-B) de la madre en el embarazo y el desarrollo fetal. El mecanismo biológico responsable es la liberación de corticotropina y glucocorticoides en la madre y por consiguiente, en el feto. Esto ocurre debido a altos niveles de estrés, a una mala calidad en la nutrición y al número de infecciones ocurridas durante el embarazo. Las consecuencias de la secreción de glucocorticoides en madres gestantes se asocia con un aumento en conductas de distracción y falta de atención, así como un menor coeficiente intelectual en los niños [10].

Estilo Parental. Calidad del ambiente doméstico. En ambientes con NSE-B las relaciones padres-hijos se asocian con mayor irritabilidad, ansiedad y estados de ánimo depresivos que inciden negativamente en la interacción y formas de vinculación [11]. En estos ambientes se hace presente una menor sensibilidad hacia a las necesidades del niño, que se manifiestan luego en problemas socioemocionales y de comportamiento.

Interacción Padres-Hijos. Estimulación cognitiva. Existe evidencia de que la interacción padres-hijos es una fuente de estimulación cognitiva para el niño. En ambientes con interacción reducida se produce un efecto negativo con consecuencias en el desarrollo, especialmente en el control de los impulsos y competencias de autorregulación.

Dentro de este mediador se observa que madres con NSE-Alto (NSE-A), en comparación con

las madres de NSE-Medio (NSE-M) y NSE-B (NSE-B), crean ambientes lingüísticamente más enriquecidos para sus hijos. En ambientes creados por aquellas madres con NSE-B se destaca la ausencia de regularidad en las interacciones y las dificultades maternas para intervenir adecuada y oportunamente propiciando el desarrollo cognitivo de los niños. Esta situación junto al acceso a libros, la dedicación a la lectura, la participación de las familias en actividades de la escuela y la exposición a materiales estimulantes crean un ambiente estimulante para el desarrollo cognitivo.

Resulta necesario destacar cómo la brecha de aprendizaje entre los niños que cuentan con entornos de aprendizaje favorables y aquellos con entornos desfavorables aumenta con la edad. La relación es relevante que niños de 5 años que cuentan con ambientes favorables demuestran casi el doble de habilidades (lectura y números) que los niños que cuentan con entornos de aprendizaje inferiores. Esta brecha continúa creciendo a medida que los niños entran a la escuela primaria, dejando a los niños de entornos familiares débiles en mayor riesgo de deserción y fracaso escolar [12].

Nutrición. El desarrollo cognitivo también resulta sensible al tipo de alimentación, la cual se encuentra directamente vinculada con el indicador Ingreso Familiar. Bajo esta dimensión, es posible destacar cómo los alimentos que poseen bajo contenido glucémico mejoran la atención y la memoria; en contrapartida con los que poseen alto índice de azúcares simples que se asocian con dificultad de concentración y atención. En el mismo sentido, la ingesta elevada de grasa saturada se relaciona con un deterioro cognitivo, mientras que el consumo de ácidos grasos poliinsaturados tiene efectos beneficiosos en su prevención. Las vitaminas B1, B6, B12, B9 (ácido fólico) y D, hierro y yodo ejercen efectos neuroprotectores y mejoran el rendimiento intelectual [13].

Condiciones del barrio. El barrio condiciona el tipo de escuela a la que concurren los niños y con ella es posible identificar los recursos a los que tienen acceso durante su desarrollo cognitivo temprano.

Se ha comprobado que el compartir aula en el colegio desde la infancia temprana con compañeros con mayores niveles socioeconómicos afecta a un mejor desarrollo de dichas funciones [14].

El aprendizaje dentro y fuera del hogar explica gran parte de la relación entre el NSE y las habilidades cognitivas, incluidos el lenguaje, el razonamiento no verbal y las funciones ejecutivas [15].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Dentro de las Neurociencias Cognitivas una amplia cantidad de investigaciones demostró que las diferencias en las experiencias de estimulación asociadas a distintos NSE afectan el desarrollo cerebral en varios niveles, desde el bioquímico hasta la organización de las funciones psicológicas más complejas; evidenciándose una correlación entre el nivel socioeconómico de una familia y el desarrollo del cerebro de sus hijos.

Sobre esta línea de investigación se desarrolla el presente estudio, el cual tiene como objetivo aportar evidencia empírica sobre el rendimiento académico y la deserción estudiantil universitaria sobre el caso de estudio conformado por los ingresantes de Ingeniería de la UTN FRLP.

Con el fin de obtener una correlación entre los indicadores del NSE y las variables rendimiento académico y deserción estudiantil universitaria, se utiliza la Minería de Datos o Data Mining. Esta técnica, se define formalmente como “un conjunto de herramientas aplicadas al proceso no trivial de extraer y presentar conocimiento implícito, previamente desconocido, potencialmente útil y humanamente

comprensible, a partir de grandes conjuntos de datos, con el objeto de predecir de forma automatizada, tendencias o comportamientos y descubrir modelos previamente desconocidos” [16].

3. RESULTADOS OBTENIDOS / ESPERADOS

La influencia de la pobreza sobre el desarrollo cognitivo es identificada y evaluada en estudios neurocientíficos encontrándose una vinculación directa entre NSE y diferencias en el tamaño, forma y funcionamiento del cerebro.

Los resultados generales de las investigaciones en las que se analiza este fenómeno, evidencian que los niños con nivel socioeconómico bajo tienden a obtener peores resultados en un gran número de pruebas, respecto a niños de nivel socioeconómico medio o alto; estadísticamente, son más propensos a no terminar sus estudios secundarios, a no ingresar en la universidad y a obtener empleos precarios.

Dentro de estas investigaciones diferentes indicadores son definidos como básicos para el análisis de los niveles socioeconómicos: Ingreso Familiar, Ocupación y Nivel de Instrucción de los Padres. Y es a partir de ellas, que surgen más específicamente los estudios de diversos mecanismos mediadores para demostrar que los indicadores no actúan de la misma manera en el desarrollo cognitivo. Si bien no se encuentran definidos de manera estricta, los enfoques que surgen de esta línea comparten variables en común: los factores prenatales, la nutrición, el ambiente en el hogar, la interacción padres-hijos, estimulación cognitiva en el hogar, la salud mental de los padres y las condiciones del barrio.

Con el fin de contrastar los resultados precedentes, se utiliza la técnica de Minería de Datos y se aplica a un conjunto de datos de ingresantes de Ingeniería de la UTN FRLP; a fin de evaluar el desempeño académico de cada estudiante durante su tramo de su formación académica considerando su NSE.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto está formado por docentes investigadores del Grupo de I&D Aplicado a Sistemas Informáticos y Computacionales (GIDAS) y del Departamento de Ingeniería en Sistemas de Información (DISI) de la Universidad Tecnológica Nacional Facultad Regional La Plata, un investigador de apoyo, un tesista de postgrado y dos alumnos becarios de investigación.

El proyecto anualmente capacita y forma a alumnos becarios que participan y aprenden desarrollando diversas tareas de investigación. Constituyéndose, de esta manera, en una etapa de entrenamiento para la futura actividad profesional y de inicio a la investigación.

El proyecto brinda un marco propicio para el desarrollo de las Prácticas Supervisadas (PS) de los estudiantes, necesarias para la obtención del título de grado de la carrera de Ingeniería en Sistemas de Información. A la fecha presenta dos PS finalizadas.

5. BIBLIOGRAFÍA

- [1] Braveman, P. A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K. S., Metzler, M. & Posner, S. (2005). Socioeconomic status in health research: one size does not fit all. *Journal of American Medical Association*, 294, 2879-2888.
- [2] McEwen, B. S. (2001). From molecules to mind: stress, individual differences, and the social environment. *Annals of the New York Academy of Sciences*, 935, 42-49.
- [3] Arán Filippetti, Vanessa (2012). Estrato Socioeconómico y Habilidades Cognitivas en Niños Escolarizados: Variables Predictoras y Mediadoras. *Psykhe (Santiago)*, 21(1), 3-20. <https://dx.doi.org/10.4067/S0718-22282012000100001>
- [4] Duncan, G. J., & Magnuson, K. A. (2003).

Off with Hollingshead: Socioeconomic resources, parenting, and child development. In M. H. Bornstein & R. H. Bradley (Eds.), *Monographs in parenting series. Socioeconomic status, parenting, and child development* (pp. 83-106). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

[5] Noble, K. G., Norman, M. F. & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, 8, 74-87.

[6] Noble, K.; Houston, S.; Brito, N.; Bartsh, H.; Kan, E.; Kuperman, J.; ... Sowell, E. (2015); "Family income, parental education and brain structure in children and adolescents", *Nature Neuroscience*, 30 de marzo de 2015, <http://dx.doi.org/10.1038/nn.3983>

[7] Brooks-Gunn, J. & Duncan, G. J. (1997). The effects of poverty on children. *The Future of Children*, 7(2), 55-71. doi:10.2307/1602387

[8] Guo, G. & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography*, 37, 431-447. doi:10.1353/dem.2000.0005

[9] Carrillo López, Ángela; (2014); "Ejecución neuropsicológica de niños y niñas de diferentes estatus socioeconómico", Universidad de Granada, Facultad de Ciencias de la Educación

[10] Hackman DA, Farah MJ, Meaney MJ (2010); "Socioeconomic status and the brain: mechanistic insights from human and animal research."; *Nat Rev Neurosci*. 2010 Sep;11(9):651-9. Accedido en abril de 2019: doi: 10.1038/nrn2897.

[11] Belsky, J., & Jaffee, S. R. (2006). The multiple determinants of parenting. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Risk, disorder, and adaptation* (pp. 38-85). Hoboken, NJ, US: John Wiley & Sons Inc.

[12] Pisani, Lauren; Dyenka, Karma; Sharma, Parvati; Chhetri, Nar; Dang, Sara; Gayleg,

Karma; Wangdi, Chencho (2017); "Bhutan's National ECCD Impact Evaluation: Local, National, and Global Perspectives"; *Early Child Development and Care*, v187 n10 p1511-1527








[13] Martínez García; Rosa María; Jiménez Ortega, Ana Isabel; López Sobaler, Ana M. ; Ortega, Rosa M. (2018); "Estrategias nutricionales que mejoran la función cognitiva"; *Nutrición hospitalaria: Organo oficial de la Sociedad española de nutrición parenteral y enteral*, ISSN 0212-1611, Vol. 35, Nº. Extra 6, págs. 16-19

[14] Weiland C, Yoshikawa H (2014) Does higher peer socio-economic status predict children's language and executive function skills gains in prekindergarten? *Journal of Applied Developmental Psychology* 2014; 35(5): 422-432. doi: doi.org/10.1016/j.appdev.2014.07.001.

[15] McCoy DC, Zuilkowski SS, Fink G. Poverty, physical stature, and cognitive skills: Mechanisms underlying children's school enrollment in Zambia. *Dev Psychol*. 2015; 51(5):600-14. doi: 10.1037/a0038924.

[16] Frawley, W.J.; Piatetski-Shapiro, G.; Matheus, C.J. "Knowledge Discovery in Databases: an Overview". *AI Magazine*, 1992.

Interacción natural, entornos inmersivos y otras tecnologías emergentes aplicadas a contextos educativos

Sanz Cecilia^{1,4} , Gorga Gladys¹ , Artola Verónica^{1,3} , Salazar Mesía Natalí¹ , Iglesias Luciano¹ , Archuby Federico^{1,2} , Nordio Mauricio^{1,4} , Buffarini Abril¹ , Astudillo Gustavo⁵ 
Baldassarri Sandra⁶ 

¹Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC.
Facultad de Informática – Universidad Nacional de La Plata

²Becaria TIPO A UNLP

³Becaria Doctoral CONICET

⁴ Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

⁵ Facultad de Ciencias Exactas y Naturales - Universidad Nacional de La Pampa

⁶GIGA AffectiveLab, Universidad of Zaragoza, España

{csanz, vartola, nsalazar, farchuby, mnordio}@lidi.info.unlp.edu.ar,

li@info.unlp.edu.ar, abrilbuffarini@gmail.com, astudillo.gustavo@gmail.com, sandra@unizar.es

RESUMEN

Se presenta aquí una línea de investigación, desarrollo e innovación que se enmarca en un subproyecto del Instituto de Investigación en Informática LIDI. A partir de esta línea se investigan diferentes modelos de interacción persona ordenador y su interrelación con procesos educativos. En los últimos años, se ha iniciado el estudio de la combinación de diferentes paradigmas de interacción, dando lugar a entornos inmersivos que posibilitan aprovechar las posibilidades del mundo físico y el virtual. Como parte de este proyecto se diseñan, desarrollan, evalúan y transfieren aplicaciones y juegos serios basados en realidad aumentada (RA), interacción tangible con objetos pasivos y activos, y realidad virtual (RV). Se combinan sus posibilidades con el fin de ofrecer diferentes experiencias educativas. También, se abordan algunas investigaciones vinculadas a la computación afectiva y al desarrollo de sistemas recomendadores orientados a estudiantes y docentes, que aportan a la interacción y al aprendizaje con tecnologías. Como parte del proyecto, se participa en la formación de recursos humanos en el área, a través de tesis de postgrado, trabajos finales de grado, becarios de investigación y proyectos de innovación con alumnos. En el trabajo se detallan los principales resultados alcanzados durante 2020 e inicios de 2021.

Palabras clave: interacción persona-ordenador, interacción tangible, realidad aumentada, realidad virtual, entornos inmersivos, computación afectiva, escenarios educativos

CONTEXTO

Esta línea de investigación se integra al subproyecto llamado “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos”, y forma parte de un proyecto más

general titulado: “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” (período 2018-2021), en el que se estudian y diseñan metodologías y herramientas de la Ingeniería de Software para escenarios híbridos que combinan diferentes entornos, dispositivos, formas de acceso y de interacción. Se trata de un proyecto del Instituto de Investigación en Informática LIDI, de la Facultad de Informática de la Universidad Nacional de La Plata, acreditado por el Ministerio de Educación de la Nación.

1. INTRODUCCION

Desde los inicios de la computación, la búsqueda por mejorar las formas de interacción entre las personas y las computadoras ha sido un tema relevante de investigación. Con el avance tecnológico, el uso de sensores, la variedad de dispositivos móviles, y el auge de la computación ubicua se han abierto nuevas posibilidades en este campo, que ofrecen oportunidades para generar experiencias educativas innovadoras. En el marco de este proyecto se vienen estudiando modelos de interacción basados en realidad aumentada, en realidad virtual, e interacción tangible. Diferentes marcos teóricos permiten echar luz sobre los aportes de cada uno de estos modelos. Hornecker (2002) discute las propiedades que hacen que las interfaces tangibles sean una herramienta valiosa para la cooperación y la colaboración. Su trabajo aborda la relación entre las interfaces tangibles y los efectos sociales que traen aparejados. Esta autora indica que la visibilidad constante en este tipo de interfaces también puede apoyar la cognición individual, ya que los objetos físicos sirven como ayuda para la memoria externa. Así estos objetos son referencia de comunicación o medio de demostración. Zuckerman y cols. (2005) presentan un marco para pensar a las interfaces tangibles en el ámbito

educativo, con un enfoque específico en los dominios de problemas abstractos. El marco teórico de estos autores se centra en el uso de materiales manipulativos (tangibles), que son aquellos objetos físicos específicamente diseñados para fomentar el aprendizaje. Marshall (2007) presenta un marco analítico derivado de un análisis de la literatura en Ciencias Cognitivas, Computación Tangible y Educación, creado para apoyar el diseño de TUI en procesos de aprendizaje. El marco propone 2 tipos de actividades en las que un estudiante puede participar cuando usa un elemento tangible: A) una actividad exploratoria o B) una actividad expresiva. En el caso de las actividades exploratorias los participantes pueden conocer modelos, relaciones entre sus componentes y formas de comportamiento preestablecidas, que son exploradas a través de la interfaz. Los sistemas tangibles expresivos, son aquellos que permiten a la persona realizar actividades y crear sus propias representaciones. Esto posibilita exhibir las inconsistencias, los conflictos y las asunciones incorrectas. En el marco de una tesis doctoral de un miembro de este proyecto se han revisado más de 60 trabajos relacionados con el área de interacción tangible que permiten fundamentar su importancia en el campo de la educación y en diferentes disciplinas (Artola, Sanz, Pesado, 2020). Otro trabajo que se ha desarrollado desde 2018 en el proyecto, se vincula con los beneficios de los objetos activos combinados con las interfaces tangibles, los tipos de *feedback* y su consideración durante la etapa de diseño de estos sistemas tangibles (Alvarado Cruz, Sanz & Baldassarri, 2020). También se estudian los fundamentos de la RA en el campo educativo, metodologías y herramientas para el diseño de aplicaciones de RA, y en el último año se ha trabajado en el área diseño y desarrollo de aplicaciones educativas de realidad virtual. Específicamente, se realizan juegos serios con estas modalidades de interacción y sus combinaciones. Las investigaciones hasta el momento, han demostrado que las aplicaciones de RV son efectivas en múltiples niveles de educación y formación, con un alto grado de aceptación por parte de los estudiantes (Kavanagh, Luxton-Reilly & Wuensche, 2017). Una característica importante en las aplicaciones de RV para el escenario educativo es la interacción por parte de usuario con el mundo virtual. Esta interacción fomenta el compromiso activo, que es deseable para el aprendizaje (Pantelidis, 2009). Diversas teorías del aprendizaje permiten dar sustento al uso de la RV en procesos educativos: el aprendizaje experimental, el constructivismo, la teoría de la autoeficacia, entre otras, han sido mencionadas por autores de referencia en el área de RV para analizar experiencias educativas con esta tecnología (Kavanagh, Luxton-Reilly & Wuensche, 2017; Anopas & Wongsawat, 2014). Además, en este proyecto se han estudiado diversos marcos para diseñar juegos de realidad virtual orientados al escenario educativo, uno de ellos considera los componentes de los juegos y los niveles de procesamiento de la información presentados por (Norman, 2004). Los juegos serios de RV deben considerar las reglas

establecidas por la jugabilidad (González Sánchez & Gutiérrez Vela, 2010), que ponen un límite a las acciones del jugador, y que a su vez lo guían en una dirección, buscando llevarlo a cumplir un objetivo. Los objetivos son llevados a cabo en un escenario dentro del cual se desarrollan las acciones. La historia del juego guía la experiencia y se vincula con las mecánicas que se adaptan a un entorno de cierta verosimilitud. En cuanto a la realidad aumentada, también se han indagado trabajos teóricos, herramientas y librerías (Salazar, Sanz & Gorga, 2019), que permiten el diseño de actividades educativas de realidad aumentada y/o su integración en juegos serios (Lovos et al., 2020). En el 2020, se ha avanzado en el diseño de dos plantillas que forman parte de una herramienta de autor AuthorAR, para facilitar la creación de actividades educativas basadas en posicionamiento. También, se continuó la difusión de la app Ruta Darwin, desarrollada en 2019, y disponible actualmente en el Centro de Innovación y Transferencia Tecnológica de la Facultad de Informática (Lizarralde et al., 2019).

Finalmente, se abordan trabajos en el área de Computación Afectiva. Según Baldassarri (2016): “la Computación Afectiva es un área de investigación emergente cuyo objetivo es el desarrollo de dispositivos y sistemas capaces de reconocer, interpretar, procesar y/o simular las emociones humanas para mejorar la interacción entre el usuario y la computadora.” (p. 14). En el proyecto se aplican estos temas para mejorar la adaptabilidad, el *feedback* y las situaciones de interacción con sistemas orientados al ámbito de la educación. Se está trabajando específicamente en la meta- anotación de recursos educativos considerando emociones para poder diseñar sistemas recomendadores de estos recursos, que consideren la dimensión emocional (Astudillo, Sanz & Baldassarri, 2020). También se ha avanzado en el estudio de antecedentes de una tesis doctoral, en la que se analiza el modelado del estudiante y la metaanotación social de recursos educativos de matemática, para mejorar su recomendación (Del Río, Álvarez, Sanz & Baldassarri, 2020).

Cabe señalar que el subproyecto se vincula 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. Además, se llevan adelante diferentes cooperaciones con universidades nacionales y de otros países en estas temáticas.

2. LINEAS DE INVESTIGACION / DESARROLLO

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Juegos educativos basados en Realidad Aumentada, Interacción Tangible y Realidad Virtual. Juegos pervasivos.

- Metodologías y herramientas para el diseño de aplicaciones educativas basadas en estos paradigmas.
- Interacción Tangible. Marcos que fundamentan este paradigma, relación entre objetos físicos y digitales, aplicaciones en educación.
- Objetos Pasivos y Activos para IT. Tipos de *feedback* en objetos activos.
- Realidad Aumentada para juegos educativos.
- Entornos inmersivos e interactivos basados en realidad virtual.
- Sistemas educativos adaptativos y sistemas recomendadores para recursos educativos.
- Computación afectiva en entornos digitales para el escenario educativo.

Cabe señalar que los distintos miembros del proyecto participan en vinculación a las diferentes líneas de investigación que aquí se mencionan. Se desarrollan acciones con becarios, tesis, pasantes y alumnos.

3. RESULTADOS OBTENIDOS

Se presentan aquí los principales resultados que se han alcanzado en el subproyecto, y corresponden al período 2020 e inicios de 2021.

En 2020 se aprobó la tesis doctoral de uno de los miembros del proyecto, vinculada a los temas de interacción tangible. Se avanzó en la difusión de los resultados de esta tesis, y se abrieron oportunidades para cooperar en estas temáticas (Artola, Pesado & Sanz, 2020). En relación también a sistemas con interfaces tangibles, se ha desarrollado a partir de un proyecto de innovación con alumnos, el juego Albores, que se considera un juego serio y que integra interfaces tangibles con objetos activos, y realidad aumentada, creado un ambiente combinado físico-virtual. Albores se propone mostrar figuras innovadoras de la historia de la Informática, con desafíos, mini-juegos que se desbloquean según el avance de los usuarios. El juego tiene relatos de interés educativo vinculados a los innovadores y sus aportes a la Informática (se puede ver un tráiler en <http://163.10.22.174/innova/albores-un-juego-interactivo-e-inmersivo/>). Este juego se presentó en las Jornadas de Ciencia y Tecnología 2020 de la Facultad de Informática. Se continuó además con la difusión del juego Murales basado en interacción tangible (Sanz et al., 2020; Iglesias & Sanz, 2020).

También se ha avanzado en la investigación relacionada con la tesis de maestría sobre juegos serios y su vinculación al desarrollo de competencias digitales en docentes (Teo, 2009; Sandí Delgado, Sanz & Lovos, 2020).

Se estudiaron metodologías de diseño de juegos serios (Archuby, Sanz & Manresa-Yee, 2020), y en particular de juegos serios con realidad virtual. En 2020, se finalizó una tesis de maestría relacionada con la realidad virtual en educación (Chirinos, Sanz & Dapoto, 2020). En el marco de esta tesis se abordó el desarrollo de HUVI, una aplicación e realidad virtual orientada a dar a conocer y acercar el patrimonio argentino a niños con algún tipo de vulnerabilidad social. La aplicación está

disponible en Google Play (<https://play.google.com/store/apps/details?id=com.Company.HP.VR>) (Chirinos, Sanz, Rucci, Comparato, Gonzalez, Dapoto, 2020). HuVi se inserta en un proyecto de extensión de la Facultad de Ciencias Económicas de la UNLP en el que se participa. En 2020, se ha aprobado una propuesta de tesis doctoral vinculada a metodologías de diseño de juegos serios de RV, que profundiza estas líneas de investigación.

También en el marco de una tesina de grado se desarrolló el juego serio de realidad virtual, llamado Innovática (Mazza, Sanz & Artola, 2020). El juego se implementa como una aplicación móvil y solo requiere de gafas económicas de realidad virtual. Tiene como objetivo educativo dar a conocer a figuras destacadas de la Informática, que se entremezclan en una historia, en la que el protagonista deberá restaurar una anomalía de espacio-tiempo, ocurrida en un museo. Esto da la oportunidad de viajar a través de portales y conocer aspectos personales de las figuras históricas de interés. Como parte de la investigación se estudiaron heurísticas de jugabilidad que guiaron el diseño a Innovática. Al mismo tiempo, se atendieron a aspectos de la estética, las mecánicas, y de los niveles de procesamiento de la información, según el marco dado en Norman (2004). Innovática fue evaluado con estudiantes y docentes, mediante el cuestionario *SUS* para analizar su usabilidad (Brooke, 2013), con muy buenos resultados. La aplicación está disponible aquí: <https://play.google.com/store/apps/details?id=com.DefaultCompany.Innovatica>

Se presentó esta aplicación en charlas, por videoconferencia, con estudiantes de escuela secundaria y de primeros años de la Facultad, vinculando la propuesta de Innovática con relatos, y recuperando la importancia del pensamiento computacional en los aportes de las figuras que se destacan dentro del juego.

En relación a los temas de realidad aumentada, se está trabajando en una tesis de maestría donde se estudia la incidencia de la integración de actividades con realidad aumentada en el rendimiento académico (Romano, Gorga & Sanz, 2020). También se está desarrollando una tesis en la que se diseñan y desarrollan plantillas para que los docentes puedan crear actividades de realidad aumentada. Se hizo en este contexto un estudio de librerías de RA (Salazar, Gorga & Sanz, 2019, a y b), que fue profundizado durante el año 2020.

En cuanto a los temas de computación afectiva, se continúa con dos tesis doctorales en las que se trabaja en la recomendación de recursos (música, y videos educativos) considerando las emociones (Ospitia-Medina, Baldassarri, Sanz, Beltrán & Olivas, 2020; Astudillo, Sanz & Baldassarri, 2020). También se continuó profundizando en el estado del arte de otra tesis doctoral que tienen en cuenta la metaanotación social y el modelado de usuario para la recomendación de recursos matemáticos (Del Río, Alvarez, Sanz & Baldassarri, 2020).

Finalmente, se está participando en estas temáticas a través de otros proyectos con universidades del país

(Lovos et al., 2020) y del exterior (Baldassarri, Sanz, Coma, Aguelo-Arguis, & Alvarez, 2019).

En cuanto a los proyectos vinculados con este proyecto y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se cuenta con un acuerdo de colaboración en estos temas con la Universidad de Zaragoza y la Universidad de Islas Baleares para cooperar en estas temáticas.
- Se ha obtenido financiamiento en el marco de una convocatoria K107 de Erasmus para estancias de viaje entre la Universidad de Zaragoza y la Universidad Nacional de La Plata.
- Se participa en el proyecto “*Pervasive Gaming Experiences For @ll* (Pergamex)”, en particular en el subproyecto (RTI2018-096986-B-C31) de la UZ.
- Además, la Dra. Sanz es miembro colaborador del grupo de Investigación en Interfaces Avanzadas (AffectiveLab).
- Se participa en la Red constituida por universidades de Iberoamérica en el marco del programa "Pablo Neruda" dentro del Espacio Iberoamericano del Conocimiento (EIC) y de la Organización de los Estados Iberoamericanos (OEI) orientada a la movilidad de estudiantes y docentes de doctorado.
- Se participa en la RedAUTI: Red temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva. En este ámbito se analizan materiales educativos para la TVDI.
- Se participa en un proyecto con la Universidad Nacional de Río Negro en relación a estas temáticas y como asesor en un proyecto de la Universidad Nacional de Santiago del Estero.

4. FORMACION DE RECURSOS HUMANOS

En este proyecto se participa en la formación de recursos humanos a través de la dirección de tesis de doctorado, maestría y trabajos finales de especialización y tesinas de grado, con becas de diferentes organismos de ciencia y técnica. En 2020, se ha aprobado 1 tesis doctoral de uno de los miembros del proyecto, 2 tesis de Maestría y 1 trabajo de Especialización en el área de Tecnología y Educación, dirigidos por miembros de este proyecto. Se desataca la tesis de maestría de Yesica Chirinos, titulada: “La realidad virtual como mediadora de aprendizajes. Desarrollo de una aplicación móvil de realidad virtual orientada a niños”, dirigida por la Dra. Cecilia Sanz, correspondiente al Magister en Tecnología Informática Aplicada en Educación de la Facultad de Informática de la Universidad Nacional de La Plata y se expuso en diciembre de 2020 (Chirinos et al., 2020). Además se ha finalizado 1 tesina de grado dirigida por miembros del proyecto y se ha realizado un proyecto de innovación con alumnos.

5. BIBLIOGRAFIA

- Alvarado Cruz, M. C., Sanz, C., & Baldassarri, S. (2020). Análisis de experiencias con objetos activos en actividades educativas basadas en interacción

tangible. *Revista de la Asociación Interacción Persona Ordenador (AIPO)*, (1):43 -58.

- Anopas, D., & Wongsawat, Y. (2014). Virtual reality game for memory skills enhancement based on QEEG. *Proceedings of the 7th 2014 Biomedical Engineering International Conference*, pp. 1–5
- Archuby, F., Sanz, C. & Manresa-Yee, C. (2020). Metodologías de diseño y desarrollo para la creación de juegos serios digitales. Tesis de maestría finalizada. Sep. 2020. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/111123>
- Artola, V.; Pesado, P. & Sanz, C (2019). EDIT: una herramienta de autor para la creación de actividades educativas basadas en interacción tangible. *Actas del XIV Congreso sobre Tecnología en Educación & Educación en Tecnología (TE&ET 2019)*, ISBN: 978-987-733-196-7, págs. 115-124. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/90735>
- Artola, V., Sanz, C. & Pesado, P. (2020). Tesis Doctoral. Interacción tangible en escenarios educativos. Diseño de una herramienta de autor para la creación de Actividades educativas basadas en interacción tangible. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/98135>
- Astudillo, G.; Sanz, C. & Baldassarri, S. (2020). Definición de un modelo de relaciones entre tipos de videos educativos, perfiles de usuarios y emociones. Propuesta de tesis de doctorado aprobada en 2020 e Informe técnico de avance.
- Baldassarri, S. (2016). Computación afectiva: Tecnología y emociones para mejorar la experiencia del usuario. *Bit & Byte*, 2.
- Baldassarri, S., Sanz, C. Coma, T., Aguelo, A. & Alvarez, P. (2019) Involving students in the generation of automatic assessment tests. 12th annual International Conference of Education, Research and Innovation. ICERI2019, isbn 978-84-09-14755-7, issn 2340-1095, doi: 10.21125/iceri.2019.1687, España, 7105-7113
- Brooke, J. (2013). SUS: A Retrospective. *Journal of Usability Studies*. 8. 29-40
- Chirinos, Y., Sanz, C. & Dapoto, S. (2020) La realidad virtual como mediadora de aprendizajes. Desarrollo de una aplicación móvil de realidad virtual orientada a niños. Tesis de Maestría en Tecnología Informática Aplicada en Educación finalizada en Diciembre de 2020. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/111879>
- Chirinos Delfino, Y., Sanz, C. V., Rucci, A. C., Comparato, G. J., Gonzalez, G., & Dapoto, S. H. (2020). HUVI: una aplicación de realidad virtual para acercar el patrimonio argentino. In XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020).
- Del Río, L.; Alvarez, X.; Sanz, C. & Baldassarri, S. (2020). Etiquetado libre y análisis del comportamiento de los usuarios para el diseño de un sistema recomendador de recursos educativos matemáticos. Tesis doctoral en desarrollo.

- De Russis, L. (2015). *Interacting with Smart Environments: Users, Interfaces, and Devices*. IOS Press. Vol7, Nro.1, ISSN: 1876-1364. *J. Ambient Intell. Smart Environ*, pp 115-116.
- Galeano Echeverri, O. J. (2014). Consideraciones en el Desarrollo de Interfaces Naturales Gestuales. *Revista CINTEX*, 19, 183-193. <https://revistas.pascualbravo.edu.co/index.php/cintex/article/view/46>
- González Sánchez, J. y Gutierrez Vela, F. (2010) Jugabilidad: caracterización de la experiencia del jugador en videojuegos. Tesis doctoral de la Univ. Granada.
- Hernández Sampieri, R.; Fernández Collado, C.; & Baptista Lucio, M. del P. (2010). *Metodología de la investigación*. Mc Graw Hill (5th ed.). México D.
- Herrera, S.; Sanz, C. & Manresa-Yee, C. (2019). *Mobile Computing for Hearing-Impaired Children*. Artículo enviado a revista para su evaluación. Desarrollado como parte de la cooperación con la Universidad de Islas Baleares.
- Hornecker, E. (2002). Understanding the benefits of graspable interfaces for cooperative use, pp. 71–87.
- Iglesias, L. & Sanz, C. (2020). Aplicación de inmersión en Murales con Kinect. Informe de Tareas.
- Kavanagh, S., Luxton-Reilly, A., Wuensche, B. (2017). A Systematic Review of Virtual Reality in Education - The Open University. *Themes in Science & Technology Education*. Vol. 2. 85-119.
- Liu, D.; Dede, C.; Huang, R. & Richards, J. (editors) (2017). *Virtual, Augmented, and Mixed Realities in Education*. Springer.
- Lizarralde, A.; Sanz, C.; Gorga, G.; Buffarini, A.; Beltrán, E. & Kraselsky, R. (2019). *Ruta Darwin: un juego con realidad aumentada para conocer las experiencias de Charles Darwin en su travesía a bordo del Beagle*. TE&ET 2019. ISBN: 978-987-733-196-7. Pp. 262-264.
- Lovos, E., Sanz, C., Goin, M., Ricca, M., Molina, C., Gil, E., Basciano, I., Gastaminza, M. (2020). *Juegos Serios Móviles. Diseño, Desarrollo e Integración En Escenarios Educativos*. WICC2020.
- Marshall, P. (2007). Do tangible interfaces enhance learning? En *Proceedings of the 1st international conference on tangible and embedded interaction* (pp. 163–170). New York, USA: ACM. <http://doi.acm.org/10.1145/1226969.1227004> doi: 10.1145/1226969.1227004
- Mazza, M., Sanz, C., & Artola, V. (2020). *Juego serio con realidad virtual para jóvenes orientado a conocer hitos de la historia de la Informática Tesina de grado finalizada*.
- Naiouf, T.; Loza Bonora, L.; Salazar Mesia, N.; Archuby, F. & Sanz, C. (2019). *RD3D – Ruta Darwin 3D una aplicación educativa con Realidad Aumentada*. Informe del proyecto con alumnos.
- Nordio, M. & Sanz, C. (2019). *Impresión de objetos 3D*. Informe de tareas de personal de apoyo CIC.
- Norman, D. *El Diseno Emocional: Por Que Nos Gustan (O No) los Objetos Cotidianos*. Ed. Paidós. (2004)
- Ospitia-Medina, Y.; Baldassarri, S., Sanz, C. Beltrán, J. R.; Olivás, J. (2020). *Fuzzy Approach for Emotion Recognition in Music*. In *IEEE Congreso Bienal de Argentina (ARGENCON)*.
- Osorio G. Luz Adriana (2010). *Ambientes híbridos de aprendizaje: elementos para su diseño e implementación*. SISTEMAS: Ambientes Educativos Modernos Basados en Tecnología (117) (2010), pp. 70-79. http://www.acis.org.co/fileadmin/Revista_117/Un_o.pdf
- Pantelidis, V. (2009). *Reasons to use Virtual Reality in education and training courses and a model to determine when to use Virtual Reality*. *Themes in Science and Technology Education*, 2(1-2), 59-70
- Romano, L., Gorga, G., Sanz, C. (2020). *Realidad Aumentada en contextos educativos y su relación con el rendimiento académico universitario*. Propuesta de Tesis de Maestría en Tecnología Informática Aplicada en Educación. En desarrollo.
- Salazar, N.; Sanz, C. & Gorga, G. (2019 a). *Análisis comparativo de librerías de realidad aumentada. Sus posibilidades para la creación de actividades educativas*. Especialización en TIAE.
- Salazar Mesia, N.; Sanz, C. & Gorga, G. (2019 b). *Diseño de plantillas para la creación de actividades educativas con Realidad Aumentada en AuthorAR*. XXV Congreso Argentino de Ciencias de la Computación. Octubre, Córdoba. ISBN: 978-987-688-377-1.
- Salazar Mesia, N.; Sanz, C. & Gorga, G. (2019 c). *Posibilidades de las librerías de Realidad Aumentada en el desarrollo de actividades educativas*. XIV Congreso Tecnología en Educación y Educación en Tecnología (TEYET 2019). San Luis.
- Sandí Delgado, J. C., & Sanz, C. V. (2020). *Juegos serios para potenciar la adquisición de competencias digitales en la formación del profesorado*. *Revista Educación*, 44(1), 471-489.
- Sanz, C. V., Artola, V., Nordio, M., Pirondo, F., Ibáñez, B. R., & Corro, B. (2020). *Murales: creando puentes entre el mundo digital y el físico*. In *XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020)*.
- Teo, T. (2009). *Modelling technology acceptance in education: A study of pre-service teachers*. *Computers & Education*, 52(2), 302–312.
- Zuckerman, O., Arida, S., y Resnick, M. (2005). *Extending tangible interfaces for education: Digital montessori-inspired manipulatives*. En *Proceedings of the sigchi conference on human factors in computing systems*, pp. 859–868).ACM. <http://doi.acm.org/10.1145/1054972.1055093>

La realidad virtual como mediadora de aprendizajes. Desarrollo de una aplicación móvil de realidad virtual orientada a niños

**Tesis de Maestría en Tecnología Informática Aplicada en Educación, Facultad de Informática,
Universidad Nacional de La Plata**

Tesista: Yesica Chirinos
yesich86@gmail.com

Directora: Cecilia Sanz (III LIDI, Facultad de Informática, UNLP) – Asesor: Sebastián Dapoto (III LIDI, Facultad de Informática, UNLP)

Fecha de exposición: diciembre de 2020

MOTIVACIÓN

Las investigaciones hasta el momento, han demostrado que las aplicaciones de RV son efectivas en múltiples niveles de educación y formación, con un alto grado de aceptación por parte de los estudiantes. Una característica importante en las aplicaciones de RV, es la interacción por parte de usuario con el mundo virtual. Esta interacción fomenta el compromiso activo, que es deseable para el aprendizaje. Las propiedades intrínsecas y el mecanismo cognitivo de la RV, permiten a los estudiantes concentrarse conscientemente en lo que están experimentando y participar en un aprendizaje más significativo. Las aplicaciones de RV pueden permitir a los estudiantes explorar y aprender a su propio ritmo y repetir las experiencias tantas veces como lo deseen. Esto resulta atractivo tanto para los estudiantes como para los educadores, a quienes les permite enseñar diferentes habilidades.

De las tecnologías educativas que se utilizan actualmente, la RV es considerada prometedora debido a su capacidad de “sumergir” a los alumnos en el entorno que están estudiando. El valor agregado de la RV es involucrar al estudiante en contextos de aprendizaje que son difíciles de comprender debido a su abstracción, dificultad o incluso debido a que son peligrosos.

De acuerdo con lo anteriormente expuesto, y del interés por buscar mecanismos alternativos en los procesos de enseñanza y aprendizaje, surge la motivación de esta tesis, de investigar sobre estas temáticas y del desarrollo de una aplicación móvil educativa de RV, basada en diversas experiencias obtenidas de la investigación previa y un análisis exhaustivo de distintas aplicaciones de RV. A partir de estas motivaciones se plantean, los objetivos de esta tesis.

El objetivo general de este trabajo es investigar y analizar sobre la Realidad Virtual (RV) y su aplicación en entornos educativos, como una “tecnología” mediadora en los procesos de enseñanza y aprendizaje.

Como objetivos específicos se plantean:

- Estudiar el concepto de RV y sus fundamentaciones.
- Analizar ventajas y desventajas de su uso.
- Estudiar las teorías de aprendizaje sobre las cuales se forjan las bases de su implementación en el contexto educativo y de formación.
- Identificar prácticas pedagógicas, particularmente relacionadas con la experiencia, que justifican el uso de la RV en procesos de enseñanza y aprendizaje.
- Estudiar y analizar distintas aplicaciones de RV y su aplicación en el ámbito educativo.
- Desarrollar una aplicación educativa de RV, a partir de los lineamientos del análisis teórico realizado, y haciendo uso de herramientas de distribución libre, que permitan el desarrollo de aplicaciones de RV, orientadas a temáticas educativas.

APORTES

Esta tesis se centra en la investigación sobre la realidad virtual (RV) y su aplicación en contextos educativos y de formación. Presenta un estudio de corte teórico y experimental sobre la temática, que inicia por la conceptualización de RV bajo la mirada de diferentes autores, analiza sus principales características, su aplicación en el ámbito educativo y otros contextos de formación, y las fundamentaciones de cómo diferentes teorías de aprendizaje se relacionan con la integración de la RV en procesos educativos. Al inicio del recorrido, se realizó una revisión sistemática de antecedentes de experiencias educativas y de entrenamiento de habilidades, que integran la RV y presentan resultados desde diferentes ópticas sobre las posibilidades y limitaciones de esta tecnología. Se realizó un estudio, como parte de esta revisión, de 28 experiencias. De éstas, 10 presentaban una descripción completa, por lo que a partir de un conjunto de criterios definidos en la tesis, agrupados en 3 categorías, se llevó a cabo un análisis a nivel de cada experiencia y posteriormente, uno

transversal incluyendo las 18 experiencias restantes. Esto dio la oportunidad de echar luz sobre algunos resultados importantes que pueden aportar al diseño de aplicaciones y juegos educativos de RV, sobre las posibilidades y limitaciones que ofrece esta tecnología en diferentes niveles educativos, y también sobre formas de evaluar su integración. A partir de este estudio teórico, y en base a las motivaciones que dieron origen a la tesis, se desarrolló una aplicación móvil de RV llamada: HuVi (Huellas Virtuales), con formato de juego, cuyo objetivo es dar a conocer las manifestaciones patrimoniales de Argentina, a niños/as en situaciones de vulnerabilidad social.

El desarrollo de HuVi estuvo a cargo de un equipo interdisciplinario, en el marco del proyecto de extensión: “Huellas Patrimoniales”, de la Facultad de Ciencias Económicas, más específicamente vinculado con la carrera de Licenciatura en Turismo. Para dar respuesta a las preguntas de investigación que orientan este trabajo, se llevó a cabo un estudio de caso, que involucró una serie de sesiones en el marco de talleres con niños/as, realizados como parte del proyecto de extensión antes mencionado. Este estudio de caso, tuvo como protagonistas a 21 niños/as que utilizaron una primera versión de HuVi, la cual previamente fue sometida, en forma previa, a diversas pruebas de laboratorio y experiencias piloto. Durante el estudio de caso, se indagó sobre la motivación intrínseca, la usabilidad, el cumplimiento de los objetivos de aprendizaje, y la experiencia lúdico - educativa en general. Se analizaron preferencias de los/as niños/as, la atención a diferentes estímulos visuales y auditivos, y a su interacción con estos.

Los resultados alcanzados dan cuenta de que HuVi permitió acercar el Parque Nacional Iguazú (patrimonio involucrado en esta primera versión de la aplicación), despertando una elevada motivación en los/as niños/as que jugaron reiteradas veces en forma voluntaria, logró la recuperación de los conocimientos que se buscó trabajar, y convocó al juego y la diversión. Durante las sesiones también se pudieron observar algunos aspectos que se deben atender para mejorar experiencias de este tipo. Los/as niños/as experimentaron inicialmente nervios por enfrentarse a algo nuevo, haciendo que al comenzar con el uso de HuVi atendieran menos a la información presentada y más a la forma de interacción, a moverse, y mirar el contexto 3D. Una vez pasada esta etapa de adaptación se concentraron en los desafíos e interactuaron más con los objetos 3D utilizando la información que se presentaba en su camino.

Tanto los estímulos visuales como auditivos fueron aspectos destacados por los/as niños/as. Las conclusiones presentadas abren el camino para profundizar la investigación en estas temáticas y dejan una serie de lineamientos que pueden ser tomados para el diseño de este tipo de aplicaciones.

LÍNEAS DE TRABAJO FUTURO

Como líneas de trabajo futuro, se proponen tres ejes:

1. Profundizar el estudio de caso ampliando la muestra y la población con la que se ha trabajado. Es importante destacar que durante 2020 no se han podido desarrollar los talleres previstos en escuelas, debido a la situación de público conocimiento sobre el Covid-19. La continuidad en esta línea posibilitará profundizar en las conclusiones halladas al momento.
2. Continuar en el desarrollo de la aplicación, a partir de la integración de los ejes faltantes y de nuevas manifestaciones patrimoniales. Los hallazgos y estudios realizados en esta tesis, permitirán tomar decisiones de diseño más justificadas para quienes se involucren en su desarrollo.
3. Generar metodologías y estrategias para diseñar aplicaciones y juegos de RV educativos que se basen en las teorías de aprendizaje que fundamentan la utilidad de esta tecnología y en resultados encontrados en el análisis de antecedentes. Este eje será abordado por la tesista a lo largo de su propuesta de doctorado, la cual ya ha sido presentada y aprobada.

INTRANETS EDUCATIVAS PARA ESCUELAS RURALES AISLADAS DE SALTA

Sergio Rocabado¹, Emilse Tacacho², Carina Reyes³, María Laura Masse³, Carlos Cadena⁴

(1) Consejo de Investigación Universidad Nacional de Salta (CIUNSa)
Universidad Nacional de Salta
srocabado@di.unsa.edu.ar

(2) Centro de Investigaciones Sociales y Educativas del Norte Argentino (CISEN)
Facultad de Humanidades, Universidad Nacional de Salta
emi.tacacho@gmail.com

(3) Departamento de Informática
Facultad de Ciencias Exactas, Universidad Nacional de Salta
reyescarina@gmail.com, mlmassep@gmail.com

(4) Instituto de Investigaciones en Energía No Convencional (INENCO),
Universidad Nacional de Salta y Consejo Nacional de Investigaciones Científicas y Técnicas
cadenacinenco@gmail.com

Resumen

La Provincia de Salta posee numerosas regiones aisladas o remotas a las cuales no llegan las redes convencionales de distribución de energía eléctrica y de datos móviles.

En algunos casos, estas regiones tienen sistemas alternativos (Pe. zonas alimentadas por energía solar) que permiten que los pobladores tengan un servicio limitado de energía y de redes celulares. Sin embargo, tanto la condición de ausencia como la de limitación a estos recursos, impiden el desarrollo normal de procesos básicos del ser humano, como el proceso de aprendizaje. Esta dificultad se vio acentuada en el período de pandemia en dichas regiones.

Con el propósito de contribuir a una educación inclusiva, que no deje de lado a las personas que habitan en zonas remotas, en esta investigación se plantea el diseño e implementación de un modelo de red que permita el acceso a la sociedad de conocimiento en estas zonas.

El modelo de red se sustenta en tecnologías de bajo consumo energético que posibilitan el aprovechamiento de energías renovables. Los contenidos educativos se instalan en los servidores de placa simple (Raspberry) y son accedidos desde dispositivos móviles utilizando redes inalámbricas Bluetooth o WiFi. En función de los requerimientos energéticos de los equipos, se dimensionan pico sistemas fotovoltaicos para abastecer de energía a los dispositivos que forman parte de la red.

La solución estudiada puede ser utilizada en otras zonas rurales del país que presenten características similares de aislamiento.

Palabras clave: Intranet, dispositivos móviles, consumo de energía, escuelas rurales aisladas, energía solar fotovoltaica.

1 Contexto

El presente trabajo se lleva a cabo en el marco del proyecto de investigación N° 2607 “Intranets educativas para escuelas rurales aisladas de Salta” en colaboración con el proyecto de investigación N° 2319 “Energía solar fotovoltaica y eólica: desarrollo y transferencia de equipos a pobladores de zonas rurales de la Provincia de Salta, y su impacto en la calidad de vida”. Los dos proyectos están financiados por el Consejo de Investigación de la Universidad Nacional de Salta, el primero durante el período enero 2019 a diciembre 2020 y el segundo de enero de 2017 a diciembre de 2020. Para ambos se solicitó prórroga hasta diciembre de 2021.

El equipo de investigación se encuentra conformado por investigadores del Consejo de Investigación de la UNSa (CIUNSa), del Centro de Investigaciones Sociales y Educativas del Norte Argentino (CISEN) y del Instituto de Investigaciones en Energía no Convencional (INENCO - CONICET).

Una de las investigadoras del proyecto, la profesora Emilse Tacacho, es coordinadora de Educación Rural de la Provincia de Salta. La profesora Tacacho colabora en la selección de las escuelas rurales aisladas, vinculación con directores y maestros de las escuelas seleccionadas y en la definición de aplicaciones y contenidos educativos que son instalados en los servidores de las Intranets.

Además, existen vinculaciones con las siguientes instituciones:

- Ministerio de Educación, Ciencia y Tecnología de la Provincia de Salta. Se encuentra en desarrollo un protocolo de colaboración recíproca entre la coordinación de Educación Rural del Ministerio de

Educación de la Provincia de Salta y el Departamento de Informática de la facultad de Ciencias Exactas de la Universidad Nacional de Salta.

- UNICEF. Por intermedio de los referentes pedagógicos para escuelas rurales mediadas por TIC en Salta. Quienes brindan asesoramiento sobre el uso de tecnologías de información en escuelas rurales aisladas. [1]

2 Introducción

En Argentina la ruralidad comprende realidades muy diferentes. Por un lado, las poblaciones rurales agrupadas que habitan en pequeñas localidades de menos de dos mil habitantes, y por otro las poblaciones rurales dispersas que habitan en zonas a territorio abierto sin constituir centros poblados. Estas zonas presentan, entre otras, las siguientes problemáticas:

- Aislamiento geográfico. Grandes distancias entre las viviendas de los pobladores y de estas a los centros urbanos.
- Dificultades de acceso. No se dispone de caminos ni medios de transporte, los habitantes se movilizan a pie o usando caballos y mulas.
- Baja densidad demográfica. Número reducido de pobladores habitando un territorio extenso.
- Carencia de servicio de distribución de energía eléctrica. Los habitantes de estas zonas utilizan energías alternativas, como paneles solares y grupos electrógenos, para cubrir necesidades energéticas elementales.
- Dificultades de comunicación. Disponen de servicios de telefonía celular con cobertura limitada, razón por la cual los pobladores se tienen que ubicar en posiciones estratégicas (elevadas y con visión directa) para mejorar la ganancia de señal.



Figura 1. Escuela Albergue “4546” – Paraje El Rosal

Estas problemáticas motivaron la creación de albergues dentro de las escuelas como una forma de garantizar que los chicos asistan cotidianamente a clases. En algunos casos, los alumnos viven en los albergues durante toda la semana (de lunes a viernes) y en otros, permanecen durante todo el ciclo lectivo o en períodos específicos. Para ilustrar las condiciones de aislamiento de estas

escuelas, en la Figura 1 se muestra una imagen de una escuela rural aislada de la Provincia de Salta.

En estos establecimientos educativos la enseñanza mediada por tecnologías digitales es escasa debido a las limitaciones energéticas de la zona y al elevado consumo de los equipos computacionales [2]. Sin embargo, los dispositivos móviles (Celulares y Tablets) constituyen una alternativa viable para este tipo de escuelas, por su bajo consumo energético respecto de computadoras convencionales [3], haciendo posible que los alumnos accedan a contenidos educativos digitales y se nutran de estrategias de enseñanza/aprendizaje basadas en TIC [4].

El uso de celulares, para acceder a contenidos digitales educativos, en escuelas rurales aisladas introduce las siguientes cuestiones:

1. ¿Dónde almacenar y cómo acceder a los contenidos educativos?

En los parajes dispersos en los que se localizan las escuelas rurales aisladas, la provisión del servicio de Internet resulta demasiado compleja, debido al tipo de territorio, la ausencia de infraestructura necesaria en la zona y las dificultades de acceso a servicios básicos. Para estos casos, se propone el uso de Intranets de bajo consumo energético como solución tecnológica que posibilita el acceso a contenidos educativos digitales a los integrantes de la comunidad educativa. En esta solución, los contenidos educativos se almacenan en servidores locales de bajo consumo (tipo Raspberry PI) y son accedidos desde celulares a través de una red Bluetooth LE [5] (Figura 2).

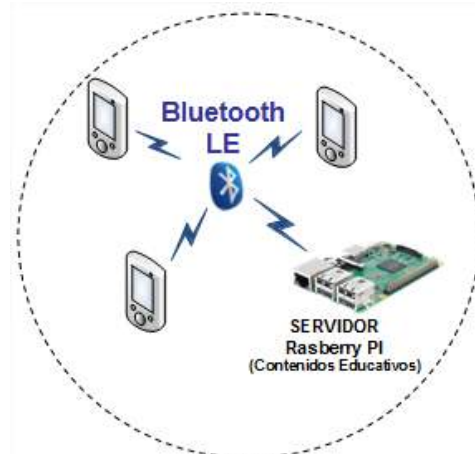


Figura 2. Intranet educativa de bajo consumo energético

2. ¿Cómo administrar la energía disponible?

Las zonas rurales aisladas presentan limitaciones energéticas que dificultan la recarga de baterías de los dispositivos que conforman la Intranet.

Para subsanar este inconveniente, se realiza un estudio sobre el consumo de energía en la Intranet, utilizando herramientas de software y/o hardware para medir/estimar la energía que consume cada dispositivo. Los resultados del estudio permitirán

establecer diferentes mecanismos para reducir el consumo energético de los dispositivos y optimizar el uso de la energía disponible para la Intranet.

3. ¿Cómo proporcionar energía a los dispositivos de la intranet?

El recurso energético natural de las regiones aisladas de Salta es la energía solar. La aridez del clima y la latitud tropical hacen que se cuente con una alta radiación solar la mayoría de los días del año. Este recurso puede ser aprovechado para proporcionar energía a los dispositivos de la Intranet, mediante el uso de Sistemas Fotovoltaicos que transformen la energía solar en energía eléctrica [6].

La investigación realizada busca solucionar las siguientes carencias de las poblaciones rurales aisladas con recursos energéticos limitados:

- Imposibilidad de acceso a información digital utilizando dispositivos de bajo consumo energético.
- Acceso a Internet inexistente. Las empresas de telefonía celular no brindan cobertura en estas zonas y la opción de Internet Satelital tiene un elevado costo.
- Falta de equipamiento basado en energías renovables que garantice el uso continuo de dispositivos móviles.

3 Líneas de Investigación y Desarrollo

Las principales líneas de investigación de esta propuesta son:

- Tecnologías de la información y la comunicación para escuelas rurales.
- Computadoras de Placa Base de bajo consumo energético.
- Computación móvil. Dispositivos móviles.
- Despliegue de redes inalámbricas en zonas aisladas.
- Reducción del consumo energético en dispositivos móviles. Técnicas para administrar y conservar energía.
- Cargadores fotovoltaicos portátiles para dispositivos móviles.

4 Objetivos y Resultados

4.1 Objetivos

Objetivo general:

- Diseñar estrategias para el despliegue de Intranets educativas en regiones rurales aisladas, haciendo uso de tecnologías de bajo consumo energético que faciliten el aprovechamiento de la energía solar disponible.

Objetivos específicos:

- Analizar y seleccionar tecnologías de bajo consumo energético que permitan implementar Intranets en regiones aisladas con recursos energéticos limitados.
- Determinar los requerimientos energéticos de los equipos que forman la Intranet y especificar el equipamiento basado en energía solar que entregue la potencia demandada por los dispositivos.
- Establecer procedimientos para reducir el consumo energético de dispositivos móviles, sin afectar su rendimiento y confiabilidad.
- Seleccionar las aplicaciones y contenidos educativos (en colaboración con los maestros) que serán instalados en los servidores de la Intranet.

4.2 Resultados

Entre los resultados más relevantes se encuentran:

- Diferentes mecanismos para la optimización del consumo energético en dispositivos móviles para su uso en zonas rurales aisladas abastecidas con energía solar fotovoltaica, presentados en [7].
- Realización de experiencias educativas mediadas por TIC en comunidades educativas aisladas, publicadas en congresos educativos de la temática ([8], [9] y [10]).
- Diseño de un Pico Sistema Fotovoltaico de tamaño y peso reducido, optimizado para entregar la energía eléctrica requerida por los dispositivos que forman la Intranet, presentado en [11] y [12].

5 Impacto

La realización del proyecto permitió recoger experiencia en regiones con poblaciones de culturas muy diferentes, situadas en zonas geográficas muy distintas por su clima y orografía. Esta experiencia permitió establecer una metodología de trabajo multidisciplinaria para la selección de aplicaciones y contenidos educativos a utilizar en las escuelas rurales, su instalación en la Intranet y su posterior operación y mantenimiento.

La implementación de Intranets de bajo consumo energético abastecidas con energía solar fotovoltaica, permitió a los maestros rurales el empleo de TIC en sus clases y acercó las tecnologías de la información y comunicación a los alumnos.

La aplicación de esta investigación generó un significativo impacto social en las comunidades educativas rurales aisladas donde se desarrolló el trabajo experimental. A continuación se presentan los aportes más relevantes:

- Disminución de la brecha digital existente entre establecimientos educativos rurales y urbanos, posibilitando que alumnos de escuelas rurales aisladas (que no cuentan con acceso a Internet)

accedan a contenidos digitales educativos alojados en los servidores de la Intranet.

- Mejoras en el proceso de enseñanza y aprendizaje en comunidades escolares rurales aisladas, mediante la implementación de estrategias educativas de m-learning sobre la intranet.
- Aprovechamiento de la energía solar fotovoltaica, para posibilitar el uso de dispositivos móviles en regiones aisladas.
- Mejoras en el uso de la tecnología disponible en escuelas aisladas, teniendo en cuenta que algunos alumnos son propietarios de equipos celulares que utilizan como reproductores de música o cámaras fotográficas y no como dispositivos de comunicación.



Figura 3. Comunidad educativa de El Rosal con integrantes del proyecto de investigación N° 2607

La cuarentena generada por el Covid-19 ha puesto aún más en evidencia la brecha digital existente entre zonas urbanas y rurales aisladas. Debido a las medidas de aislamiento social las escuelas permanecieron cerradas. Mientras que en los centros educativos urbanos se realizan actividades virtuales y se continúa con el dictado en colaboración con padres de familia, la educación en las escuelas rurales aisladas fue completamente interrumpida, ya que los alumnos no cuentan con acceso a Internet, o el mismo es muy limitado. Ante esto, las Intranets de bajo consumo se constituyen en una alternativa viable para facilitar a los alumnos el acceso a contenidos educativos sin necesidad de trasladarse a lugares con cobertura de red celular o acceso a Internet.

6 Formación de recursos humanos

La propuesta involucra la integración de los conocimientos en esta área por parte de investigadores del Departamento de Informática de la Universidad Nacional de Salta (UNSa) y del Instituto de Investigaciones en Energía no Convencional (INENCO).

Uno de los investigadores finalizó el doctorado en Ciencias Informáticas de la UNLP, el trabajo de investigación de la tesis [13] se encuentra directamente relacionado con este proyecto.

El proyecto cuenta con la participación de estudiantes avanzados de carreras de grado en Informática,

pertenecientes a la UNSa, los cuales reciben formación en las áreas de computación móvil, energía solar y eficiencia energética; además, experiencia en el desarrollo de investigaciones. Dos alumnos están finalizando, con temáticas relacionadas al proyecto, la tesis de grado de la Licenciatura en Análisis de Sistemas de la Facultad de Ciencias Exactas de la UNSa.

Bibliografía

1. UNICEF. (2019). *Enseñar y aprender en las Secundarias Rurales mediadas por Tecnologías*. UNICEF. Buenos Aires.
2. Magadán, C. (2008). *Las TIC en Escuelas rurales de La Argentina*. Las TIC: del aula a la agenda política, Ponencias del Seminario internacional Cómo las TIC transforman las escuelas. IPE-UNESCO, Sede Regional Buenos Aires, UNICEF.
3. Rocabado, Sergio; Cadena, Carlos. (2015). *Uso de dispositivos móviles inteligentes en zonas rurales aisladas abastecidas con energía solar fotovoltaica*. Paper presented at the CACIC 2015.
4. Rocabado, Sergio; Cadena, Carlos; Ottavianelli, Emilce. (2017). *Propuesta para el uso de TDICs en comunidades educativas rurales. Una experiencia en escuelas albergue*. Paper presented at the Congresso Brasileiro de Informática na Educação, Recife - Brasil.
5. SPECIAL INTEREST GROUP Bluetooth. (2010). Bluetooth v4.0 low energy (LE). from <http://www.bluetooth.com/Pages/Low-Energy.aspx>
6. Attia, Yousif I. Al-Mashhadany; Hussain A. (2014, November 12). *High performance for real portable charger through lowpower PV system*. Paper presented at the International Journal of Sustainable and Green Energy.
7. Rocabado, Sergio. (2020). *Optimización del consumo energético en dispositivos móviles para su uso en zonas rurales aisladas abastecidas con energía solar fotovoltaica*. (Tesis Doctoral), Universidad Nacional de La Plata, La Plata - Argentina.
8. Rocabado, Sergio; Cadena, Carlos; Ottavianelli, Emilce. (2018). *Propuesta para el uso de las TDIC en comunidades educativas rurales. Una experiencia en escuelas albergue*. Paper presented at the CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO. CBIE 2017, Recife - Brasil. ISSN 2316-6533.
9. Rocabado, Sergio; Cadena, Carlos; Ottavianelli, Emilce. (2018). *Propuesta tecnológica para introducir el aprendizaje mediado por tics en escuelas rurales aisladas del NOA*. Paper presented at the EDUTEC 2018, Lerida - España.
10. Rocabado, Sergio; Cadena, Carlos. (2020). *Uso de TIC en comunidades educativas aisladas: Una experiencia pedagógica mediada por dispositivos móviles abastecidos con energía solar fotovoltaica*. In UMA (Ed.), *Tecnologías educativas y estrategias didácticas* (pp. 552-562). Málaga - España.

11. Rocabado, Sergio, & Cadena, Carlos. (2015). *Cargadores solares portátiles para el uso de dispositivos móviles en zonas rurales aisladas del NOA*. Paper presented at the ASADES 2015.
12. Rocabado, Sergio, & Cadena, Carlos. (2016). *Mini sistemas fotovoltaicos para el uso de dispositivos móviles en zonas rurales: Optimización de la potencia entregada y consumida*. Paper presented at the Congreso Brasileiro de Energía Solar - CBENS 2016.
13. Rocabado, Sergio. (2015 - 2017). *Optimización del consumo energético en dispositivos móviles para su uso en zonas rurales aisladas abastecidas con energía solar fotovoltaica*. Doctorado en Ciencias Informáticas. UNLP.
14. Tarkoma, Sasu; Siekkinen, Matti; Lagerspetz, Emil; Xiao, Yu (2014). *Smartphone Energy Consumption: Modeling and Optimization* (pp. 234).
15. Attia, Y. I. Al-Mashhadany; H. A. (2014). Novel Design and Implementation of Portable Charger through Low-Power PV Energy System. *Advanced Materials Research*, 925, pp. 495-499 (ISSN 1662-8985).
16. Zhang, Lide. (2013). *Power, Performance Modeling and Optimization for Mobile System and Applications*. (Thesis - Doctor of Computer Science and Engineering), University of Michigan.
17. Schuss, C; Rahkonen, T. (2013, 11-15 Nov.). *Solar Energy Harvesting Strategies for Portable Devices such as Mobile Phones*. Paper presented at the 2013 14th Conference of Open Innovations Association (FRUCT) - IEEE, Espoo.
18. Távora, Filipe; Maia, Andrea Sarmiento. (2012). *Solar Battery charger for portable devices application*. SiliconReef Consultoria, Pesquisa e Projetos em Tecnologia da Informação. Retrieved from www.siliconreef.com.br
19. Carroll, Aaron; Heiser, Gernot (2010). *An Analysis of Power Consumption in a Smartphone*. Paper presented at the USENIX Annual Technical Conference, Boston, MA, USA.
20. Bekaroo, G., & Santokhee, A. (2016, 3-6 Aug. 2016). *Power consumption of the Raspberry Pi: A comparative analysis*. Paper presented at the 2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech).
21. Wai Mar Myint Aung, Yadanar Win, Nay Win Zaw. (2018). *Implementation of Solar Photovoltaic Data Monitoring System*. International Journal of Science, Engineering and Technology Research. Vol. 7. Issue 8.
22. Ministerio de Educación, Cultura, Ciencia y Tecnología. (2015 - 2019). Segundo Proyecto de Mejoramiento de Educación Rural. <https://www.argentina.gob.ar/promer-ii>.
23. Secretaria de Energia - Ministerio de Planificación. (2015). PERMER, Proyecto de Energías Renovables en Mercados Rurales. from <https://www.se.gob.ar/permer/>

Jabutí EDU, plataforma de Robótica Educativa IOT open Hardware

Lucas Kucuk¹, Leonardo Añais²

¹ Departamento de Informática, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones, kucuk@hotmail.es,

² Universidad Tecnológica Nacional, Facultad Regional Buenos Aires, infonegandresito@gmail.com

RESUMEN

En el presente trabajo se despliega el desarrollo de una plataforma virtual denominada Jabutí EDU Nube como una herramienta para la enseñanza de robótica educativa a distancia, la cual se basa en el concepto de IoT (internet of Things) y permite a los alumnos controlar dispositivos físicos, mediante un portal web en el cual cuentan con un entorno de programación en línea, basado en el lenguaje LOGO. La implementación de dicha plataforma se realizó en el marco de una competencia internacional de robótica educativa, obteniendo la aprobación satisfactoria por parte de los participantes.

CONTEXTO

El trabajo se desarrolló en el marco de un trabajo de investigación de docentes y alumnos de la Universidad Nacional de Misiones, Universidad Tecnológica Nacional y la empresa EJR Robótica Educativa.

1. INTRODUCCIÓN

La robótica educativa es un recurso pedagógico actualmente utilizado con mucha frecuencia en diversos niveles educativos con el cual pueden trabajarse en diversas áreas curriculares, por lo tanto, existen múltiples experiencias educativas que utilizan la programación de robots para mejorar procesos de enseñanza-aprendizaje y sobre todo para desarrollar en los estudiantes el pensamiento computacional [1].

El principal potencial de la robótica educativa como herramienta es la posibilidad de

comprender conceptos abstractos de la de la computación como así también favorecer en gran medida la construcción de conocimiento en el área de las ciencias exactas [2].

Con la pandemia del coronavirus, todas las escuelas han adoptado la educación remota y la tecnología se ha convertido en un gran aliado. Sin embargo, incluso con la reanudación de las clases presenciales, hay varias medidas de protección a seguir contra la enfermedad. De esta forma, no será posible una devolución del 100% presencial y se mantendrá la posibilidad de mantener parte de la formación de forma remota [3].

La forma en que enseñamos y aprendemos está en constante cambio y el aprendizaje a distancia ya es un hecho, así como el Machine Learning y el Internet de las cosas (IoT). Estas tendencias se adaptarán para uso pedagógico y las EdTech que sirvieron al mercado educativo en 2020 se han revolucionado con un enfoque en la conectividad, la versatilidad y el aprendizaje centrado en el estudiante.

La construcción de objetos de aprendizajes mediados por tecnologías de la Información, requieren de un diseño y ejecución dinámicos, existen actualmente marcos de trabajos que fijan las pautas al momento del diseño, ejecución y análisis de usabilidad de estos productos educativos tecnológicos [4][5].

Partiendo del contexto educativo, se presenta la plataforma jabutí EDU como una alternativa para la enseñanza de la robótica de forma remota, la cual fue diseñada como un entorno principalmente enfocado a la enseñanza en los primeros niveles educativos,

tanto para inicial y primaria, lo cual no quita que pueda utilizarse en otros niveles. Los integrantes de la empresa EJR Robótica Educativa y autores de este trabajo, desarrollaron dicha plataforma. Dicho ambiente se compone de 6 módulos, los cuales tienen diferentes niveles de complejidad en los cuales se desarrolla la práctica educativa.

La plataforma Jabuti EDU nube es un entorno Open Hardware, lo cual permite incorporar múltiples proyectos de robótica, basados en Arduino, ESP32 y ESP 8266, por lo tanto, cada institución educativa puede trabajar con sus propios desarrollos [6].

El concepto de IOT (internet of things) está basado en la conectividad de los objetos cotidianos al internet, actualmente podemos ver desde electrodomésticos hasta coches que se conectan a internet para que la experiencia de interacción sea enriquecida de gran forma [7][8]. Del mismo modo, la plataforma permite al alumno la posibilidad de interactuar con un dispositivo físico de forma remota programando sus acciones a través de una web.

La plataforma soporta múltiples dispositivos, modos de programación y usuarios; lo que quiere decir que más de un alumno puede programar el mismo robot al mismo tiempo utilizando una herramienta de alternancia, en la cual el docente se encarga de ejecutar el código programado por cada alumno, una vez este es validado por el docente. Esta posibilidad brinda al educador la democratización del recurso educativo, ya que lo que suele suceder en la mayoría de los casos es que al trabajar en grupos solamente un alumno puede programar o ejecutar su código.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco de ASPO (Aislamiento social, preventivo y obligatorio), la presencialidad se

tornó inviable puesto a que los estudiantes no podían asistir de forma presencial a las clases de la escuela de Robótica de Misiones, por lo cual se tuvo que optar por una alternativa que permitirá el trabajo desde casa, lo viable fue la utilización de la plataforma de robótica educativa Jabuti EDU, la cual dada sus características de funcionamiento, que permite que múltiples usuarios puedan interactuar con un mismo hardware al mismo tiempo pero con turnos definidos pueda ser utilizada como herramienta educativa. Para la implementación de la plataforma se desarrolló una competencia internacional de robótica a distancia denominada CIRDI en la que participaron alumnos de Argentina, Brasil y Uruguay.

Para la realización de la competencia se utilizaron 6 sedes denominadas Arenas, las cuales se dispersaron a lo largo de 6000 km, y se Ubicaron en las Ciudades de Comandante Andresito Misiones Argentina, Cascavel Estado de Paraná Brasil, Novo Hamburgo Estado de Santa Catarina Brasil, Fortaleza-Ceará Brasil, Santarém Pará, Recife, Pernambuco Brasil.

La dinámica de trabajo fue dividir a los 23 equipos participantes en grupos de 3 los cuales fueron rotando hasta participar de cada una de las arenas localizadas en las diferentes ciudades. Los integrantes de cada equipo trabajaban en grupo para controlar un Robot ubicado en un lugar geográfico distinto y enviar comandos de forma remota para que este complete un desafío específico para cada arena.

Los desafíos propuestos para dicha competencia fueron un laberinto, una competencia de obstáculos en la cual los robots tenían que esquivar obstáculos, y explotar globos y una competencia en la cual los robots deberían “patear penales”, el cual consiste en ubicar 3 bloques de 3x3cm impresos en 3d, en un área de varios

semicírculos concéntricos los cuales cada uno de ellos tenía un puntaje específico (Figura 1).



Figura 1 - Desafío globos

La primera instancia de la competencia fue clasificatoria y se realizó en noviembre de 2020 y la instancia final se dio en diciembre 2020 en el marco de la Feria Internacional de Ciencia y tecnología MOSTRATEC, organizada por la Fundación Liberato en Novo Hamburgo Rio Grande do Sul, Brasil.

La competencia se pensó para que cada equipo pueda competir por lo menos una vez en cada una de las ciudades y experimentar con los robots específicos de cada arena, los cuales son a su vez proyectos de open hardware desarrollados por cada uno de los anfitriones de las sedes.

3. RESULTADOS OBTENIDOS

Con respecto a los resultados de esta primera experiencia de trabajo con la herramienta, podemos mencionar que a lo largo del año 2020 se pudieron capacitar utilizando la plataforma 2563 estudiantes, con un total de 182 escuelas registradas en la misma y además de 230 docentes registrados en tres países del Mercosur. La competencia realizada permitió obtener información respecto a la usabilidad de la plataforma, el cual demostró satisfactorios resultados. En mismos términos, demostró ser viable utilizar la tecnología IoT para la enseñanza y ejecución de proyectos de robótica

Estas experiencias en términos técnicos representan un gran desafío ya que el

volumen de usuarios simultáneos requiere por un lado de capacidad instalada en cuanto a ancho de banda, Así como también plantea un reto en cuanto a diseño y usabilidad, con lo que podemos concluir que el proceso de validación de la plataforma como herramienta para el trabajo de la robótica está totalmente aprobado y fue adoptada por la comunidad educativa.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo participan los docentes, alumnos e investigadores de la Universidad Nacional de Misiones, Facultad de Ciencias Exactas, Químicas y Naturales. y de la Universidad Tecnológica Nacional, Facultad Regional Buenos Aires con antecedentes de investigación en los temas que se abordan.

5. BIBLIOGRAFÍA

- [1] González, S. M. (2011). Estudio sobre la utilidad de la robótica educativa desde la perspectiva del docente. *Revista de Pedagogía*, 32(90), 81-117.
- [2] Salamanca, M. L. P., Lombana, N. B., & Holguín, W. J. P. (2010). Uso de la robótica educativa como herramienta en los procesos de enseñanza. *Ingeniería Investigación y Desarrollo: I2+ D*, 10(1), 15-23.
- [3] Cortés Díaz, G., Henao Godoy, N., & Osorio Linero, V. Trabajo remoto en tiempos de covid-19 y su impacto en el trabajador, Pontificia Universidad Javeriana (2020).
- [4] L. Kucuk, J. S. Ierache y G. Dapozo. “Marco de trabajo para la usabilidad en objetos de aprendizajes basados en realidad aumentada”, In XX Workshop de Investigadores en Ciencias de la Computación, ISBN: 978-987-3619-27-4, Abril 2018.
- [5] Kucuk, L., & Ierache, J. S. (2020). Aplicación de rúbrica COdA para

evaluación de calidad objetos de aprendizajes basados en realidad aumentada. In XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz).

- [6] Ángel-Díaz, C. M., Segredo, E., Arnay, R., & León, C. (2020). Simulador de robótica educativa para la promoción del pensamiento computacional. *Revista de Educación a Distancia (RED)*, 20(63).
- [7] Alvear-Puertas, V., Rosero-Montalvo, P., Peluffo-Ordóñez, D., & Pijal-Rojas, J. (2017). Internet de las cosas y visión artificial, funcionamiento y aplicaciones: Revisión de literatura. *Enfoque UTE*, 8, 244-256.
- [8] Mota-Macías, J. M., Baena Pérez, R., Person Montero, T., & Ruiz-Rube, I. (2020). Visual creation of educational mobile applications for robotics and the internet of things.

JUEGOS SERIOS Y REALIDAD AUMENTADA. OPORTUNIDADES Y RETOS PARA SU INCLUSIÓN EN PROPUESTAS EDUCATIVAS

Cecilia Sanz²; Edith Lovos¹; Martin Goin³, Mónica Ricca¹, Claudio Molina³, Evangelina Gil¹, Ivan Basciano¹, Ismael Ponce Cevoli¹, Manuel Gastaminza¹

¹ Universidad Nacional de Río Negro, Sede Atlántica, Centro Interdisciplinario en Derechos, Inclusión y Sociedad (CIEDIS),

² Investigador asociado de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Instituto de Investigación en Informática LIDI (III-LIDI), Facultad de Informática, Universidad Nacional de la Plata

³ Universidad Nacional de Río Negro, Sede Andina, Centro Interdisciplinario en Derechos, Inclusión y Sociedad (CIEDIS),

csanz@lidi.info.unlp.edu.ar - {elovos, mgoin,mricca, cmolina}@unrn.edu.ar - {EvaEG19,ivaanbas19,ismaelponcecevoli77,Manuelgastaminza}@gmail.com

RESUMEN

Cabero y Fernández (2018) sostienen que las tecnologías disruptivas en los escenarios educativos son aquellas que permiten transformar y mejorar los mismos. Así, a partir del acceso a dispositivos como celulares inteligentes por parte de los adolescentes y jóvenes que transitan los espacios de formación, es posible avanzar en el diseño de propuestas educativas más cercanas al paradigma del aprendizaje móvil, incluyendo tecnologías como la realidad aumentada (RA) y los juegos serios. En esta línea, desde el proyecto de investigación (PI-40C-750) acreditado por la Universidad Nacional de Río Negro (UNRN), se propone generar conocimiento respecto al diseño, desarrollo y aplicación de juegos educativos móviles (JEM), que incluyan interacciones usando RA en particular en aquellos espacios de formación de nivel medio y superior que se insertan en el contexto de la UNRN, y a través de una metodología de investigación-acción colaborativa.

Desde sus inicios el proyecto, se centró en la exploración de tecnologías como la realidad aumentada y los juegos educativos en particular aquellos considerados móviles, analizando las posibilidades y barreras que presenta su inclusión en experiencias educativas concretas. En este sentido, se revisaron aspectos tales como metodologías para el diseño de juegos serios, herramientas para la producción de los mismos, así como también la evaluación de juego serios móviles en propuestas pedagógicas afines a las temáticas del grupo de trabajo. En este trabajo se presentan detalles de la investigación realizada hasta el momento.

Palabras clave: Aprendizaje Móvil, Juegos Serios, Realidad Aumentada, Diseño.

CONTEXTO

Esta investigación, se lleva adelante en el marco del proyecto de investigación bianual denominado “Juegos educativos móviles con realidad aumentada. Aspectos de diseño, desarrollo e integración en escenarios

educativos”, acreditado y financiado por la UNRN y bajo dirección externa. El proyecto inició en 2019 y está prevista su finalización en Mayo de 2021.

El mismo se incluye entre las actividades de la unidad ejecutora: Centro Interdisciplinario de Estudios sobre Derechos Inclusión y Sociedad (CIEDIS) perteneciente a la UNRN, siguiendo la línea de investigación sobre epistemología, investigación e innovación docente

1. INTRODUCCIÓN

El contexto de pandemia en el que se llevaron adelante las actividades académicas durante 2020, evidenció por una parte aspectos vinculados a las desigualdades de acceso y a las habilidades y competencias necesarias para el uso de las TIC por parte de los diferentes actores, y por otra parte, se presentó una como una oportunidad /desafío para los docentes, qué nos vimos empujados a salirnos del escenario convencional no solo físico sino de estructuración de la clase, y avanzar hacia otro totalmente virtual, donde en algunos casos el dispositivo móvil se convierte en el recurso tecnológico de mayor acceso entre los estudiantes, y que les permite continuar con sus estudios (Cabero y Llorente, 2020). En este sentido, y desde la perspectiva de la enseñanza, se renueva la necesidad de trabajar en el diseño de propuestas educativas innovadoras que tengan en cuenta estos aspectos. Así, en este proyecto se pone el foco en el diseño de aquellas propuestas que permitan enriquecer escenarios reales con información virtual, a través de juegos educativos que aporten al factor motivacional tan necesario en los procesos de enseñanza y aprendizaje.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto tiene como objetivo principal desarrollar conocimiento sobre el diseño, desarrollo y aplicación de juegos educativos móviles en particular aquellos que incluyan realidad aumentada, y que puedan integrarse a

espacios de enseñanza y aprendizaje de nivel medio y superior en el contexto de la UNRN.

Los objetivos específicos son:

- Revisar marcos teóricos y analizar antecedentes sobre interacción persona ordenador (HCI) y el uso de juegos educativos móviles con RA en contextos de enseñanza y aprendizaje.
- Indagar metodologías de diseño de juegos educativos usando RA y dispositivos móviles.
- Contribuir al desarrollo de metodologías para el diseño, desarrollo e integración de juegos educativos en situaciones específicas, en particular haciendo uso de dispositivos móviles, y RA entre otras.
- Diseñar, desarrollar y evaluar experiencias de inclusión de juegos educativos móviles con RA en situaciones educativas concretas.

A partir de los objetivos trazados, se abordan las siguientes líneas de investigación:

- Marcos teóricos sobre el diseño de juegos serios (JS), en particular aquellos que incluyen RA.
- Exploración de herramientas para el desarrollo de juegos móviles (JM) incluidos los juegos serios móviles (JSM). Profundizar en aquellas que permitan incluir interacciones usando RA y/o puedan ser entendidas como herramientas de autor.
- Diseño de propuestas didácticas que incorporen JS y/o JSM.
- Puesta en práctica y evaluación de experiencias educativas que incluyan JSM.
- Instrumentos para recolección de datos y evaluación: sobre el diseño de JSM y de experiencias de aprendizaje usando JSM.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto está en su segundo año de desarrollo, y hasta aquí se ha logrado avanzar acorde a lo planificado y a los objetivos establecidos. Se ha optado por la metodología de investigación-acción participativa (IAP),

buscando como sostienen Ruiz et al. (2002) articular teoría y práctica entre quienes llevan adelante la investigación y quienes buscan poner en práctica los aportes de ésta, a través de un proceso cíclico de diagnóstico, planificación, acción y reflexión, teniendo en cuenta las disciplinas específicas con las que está vinculado el equipo de investigación en el contexto de la UNRN.

A continuación se presentan los resultados obtenidos hasta el momento:

- Revisión bibliográfica que incluyó temas como: metodologías de diseño de JSM, JSM en contextos de enseñanza y aprendizaje, en particular en el nivel medio y superior, herramientas para la producción de JMs que permitan incluir interacciones usando RA. Específicamente se indagaron aquellas herramientas que presentan posibilidades para su uso por perfiles no técnicos. Algunos de estos resultados se presentaron en el Congreso CACIC 2020 (Lovos et al, 2020).
- Diseño, implementación y evaluación de una propuesta didáctica de articulación entre el nivel medio y superior, que incluye el uso del juego denominado MAHI (Padilla et al, 2017) para repaso de álgebra. La propuesta apuntó específicamente a recuperar saberes sobre el proceso de resolución de ecuaciones y estuvo destinada a estudiantes ingresantes de las Sede Andina y Atlántica, que participaron del Ciclo Introductorio a Carrera (CIC) dispuesto por la UNRN durante el año académico 2020. A partir de este trabajo se generó un acuerdo de colaboración con el autor de MAHI para participar en la evaluación de la nueva versión del juego. Asimismo los docentes del área de matemáticas del proyecto realizaron aportes a los contenidos de esta nueva versión (Lovos, Goin, Molina y Sanz, 2020).
- Diseño e implementación de un estudio comparativo entre

herramientas que permite enriquecer materiales de lectura (Lovos y Sanz, 2019), así como también la producción de juegos para dispositivos móviles, a perfiles que no cuentan con conocimientos de programación. En ambos casos se focalizó en la posibilidad de incluir interacciones usando realidad aumentada, a través del uso de códigos QR, la ubicación del jugador y/o imágenes.

- Diseño y evaluación de un prototipo de JSM usando una herramienta de autor, cuyo objetivo es el repaso de conceptos básicos de resolución de problemas usando algoritmos y el manejo de estructuras de control. Respecto a la evaluación, la misma incluyó a docentes y estudiantes de carreras vinculadas a la informática tanto de la UNRN como de otras instituciones de nivel superior.
- Diseño e implementación de dos experiencias formativas en modalidad virtual : una, denominada “Materiales Educativos Digitales. Posibilidades para incluir aspectos lúdicos y tecnologías emergentes en el espacio de enseñanza y aprendizajes”. Ésta fue destinada a los docentes de nivel medio que se encuentran cursando la asignatura "Práctica docente" de la carrera Profesorado de Educación Secundaria de la Modalidad Técnico Profesional en concurrencia con la formación de base, que se dicta en el Instituto de Formación Docente Continua en Educación Física de la ciudad de Viedma. La otra denominada: “Capacitación Virtual en Programación Scratch”, cuyos destinatarios son docentes de nivel primario en el alcance geográfico de la UNRN.

Ambas experiencias tuvieron la intención de presentar a los docentes participantes recursos y herramientas que les permitan avanzar en el diseño de propuestas pedagógicas que incluyan interacciones lúdicas.

Asimismo, han sido avaladas por la UNRN (dispo. ATL 218/20, AND 420/202) y se desarrollaron en el primer y segundo semestre del 2020.

Por último, cabe señalar que se han realizado publicaciones (Lovos y Aballay, 2020) y presentaciones de trabajos en eventos científicos afines a la temática (Sanz et al, 2020, Lovos et al, 2020, Ponce y Lovos, 2020). Recientemente ha sido aceptado para su presentación en I Congreso Internacional “Ingresos e Ingresantes a la Universidad” (UNRN) un trabajo sobre la inclusión de JSM en el área de matemáticas como estrategia de acompañamiento y retención a los estudiantes ingresantes.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores del proyecto se conformó inicialmente por dos docentes de la Sede Atlántica (Viedma) y dos docentes de la Sede Andina (San Carlos de Bariloche) de la UNRN, dos estudiantes que cursan las carreras Lic. en Sistemas y Lic. en Comunicación Social de la Sede Atlántica. Más adelante se incorporaron al proyecto, a través de becas de formación en investigación, un estudiante de la Lic. en Sistemas, y como tesista de grado, un estudiante de la Lic. en Kinesiología y fisiatría ambos con pertenencia institucional a la Sede Atlántica de la UNRN. Los docentes trabajan en asignaturas del área de matemáticas, literatura y programación de computadoras a través de las carreras vinculadas a ingenierías, sistemas y comunicación social. El proyecto está a cargo de una investigadora externa con una reconocida trayectoria en la temática, cuya pertenencia institucional corresponde a un laboratorio de investigación de la Facultad de Informática de la Universidad Nacional de La Plata.

Asimismo, una de las integrantes del proyecto, está realizando el trabajo final de la Especialización en Docencia Universitaria de la UNRN, y ha participado durante el proyecto de los cursos de postgrado del

Doctorado en Ciencias Informáticas (UNLP), mientras que algunos de los estudiantes han tomado cursos específicos sobre desarrollo de juegos móviles usando el motor Unity y testing de software.

Por último, se espera avanzar a partir del regreso a las aulas establecido por el Ministerio de Educación, con el desarrollo de un proyecto de extensión (Res UNRN. 589/19) que tiene por objetivo acompañar a los docentes de nivel primario en el uso de un juego de mesa ad-hoc para el desarrollo del pensamiento computacional en niños de los últimos años de la escuela primaria rionegrina. El juego ha sido diseñado en el marco del proyecto de investigación.

5. REFERENCIAS

- 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>
- Lovos, E., & Aballay, L. (2020). Deserción Académica y Emociones en Ambientes E-learning. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, (27), e10-e10.
- Lovos, E., & Sanz, C. V. (2019). Herramientas de autor para enriquecer materiales de lectura: análisis comparativo. In XXV Congreso Argentino de Ciencias de la Computación (CACIC)(Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019).
- Lovos, E., Goin, M., Molina, C., & Sanz, C. V. (2020). Evaluación de un juego serio móvil para repaso de álgebra. In XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020) (Neuquén, 6 y 7 de julio de 2020).
- Lovos, E., Basciano, I., Gil, E. & Sanz, C. V. (2020). La Producción de Juegos Serios Móviles. Posibilidades y Desafíos para el Docente de Nivel Superior. In Actas del Congreso Argentino de Ciencias de la Computación (pp. 144-153). Universidad Nacional de La Matanza
- Padilla, Y. A. M., Vázquez-Reyes, S., González, A. M., & Hernández, A. G. (2017,

October). Mahi: Support tool for practicing first-degree algebraic equations. In 2017 6th International Conference on Software Process Improvement (CIMPS) (pp. 1-7). IEEE.

Ponce Cevoli Ismael, Lovos, Edith .(2020). Tecnologías aplicadas a la Kinesiología. El uso de la Realidad Virtual en la Rehabilitación post ACV. En las V Jornadas de Investigadores Noveles Becaries y Tesistas . UNRN Sede Atlántica, Noviembre, 2020

Ruiz, F., Polo, M., Piattini, M., & Alarcos, G. (2002). Utilización de Investigación-Acción en la Definición de un Entorno para la Gestión del Proceso de Mantenimiento del Software. In 1er. Workshop en: Métodos de Investigación y Fundamentos Filosóficos en Ingeniería del Software y Sistemas de Información.(MIFISIS'2002). Madrid

Sanz, C. V., Lovos, E., Goin, M., Ricca, M., Molina, C., Gil, E., ... & Gastaminza, M. (2020). Juegos serios móviles. In XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz).

Congreso Internacional Ingreso e Ingresantes a la Universidad. Del 22 de febrero al 05 de marzo de 2021. Sede Atlántica UNRN. Disp. Nro 077/2020

LA GAMIFICACIÓN COMO ESTRATEGIA DIDÁCTICA EN GEOMETRÍA

Pauletti C., Astudillo G.

Facultad de Ciencias Exactas y Naturales. Universidad Nacional de la Pampa

Avda. Uruguay 151 - Santa Rosa - La Pampa - 02954-425166

clarisapauletti@gmail.com astudillo@exactas.unlpam.edu.ar

RESUMEN

El Taller de Resolución de Problemas (TRP) es una asignatura del segundo año del Profesorado en Matemática, dictada en la Facultad de Ciencias Exactas y Naturales (UNLPam).

Desde el TRP se espera que los estudiantes tengan conocimiento sobre ciertos temas geométricos que se suponen abordados en el nivel secundario y en asignaturas previas al desarrollo del taller. Al observar una profundización en las dificultades de resolución, se realizaron encuestas sobre los alumnos matriculados en 2019. Las cuales evidenciaron que los estudiantes pertenecientes al nuevo plan de estudios indican no conocer, o no recordar, varios temas o conceptos geométricos. Luego, se identifica la amplia distancia conceptual entre los saberes necesarios para plantear y resolver problemas mediante métodos geométricos de manera autónoma y los conocimientos previos con los que cuentan los estudiantes como un problema que se debe resolver. Ante esto, es absolutamente necesario intervenir con una propuesta innovadora que permita lograr los objetivos de la cátedra sin descuidar otras unidades de la asignatura.

El objetivo de esta propuesta es fortalecer algunos conocimientos geométricos prescindidos, a partir de la gamificación como estrategia didáctica, mediante el diseño y la implementación de un trayecto pedagógico.

Palabras clave: gamificación, geometría, profesorado en matemática, trayecto pedagógico

CONTEXTO

Este trabajo de innovación es producto de un proyecto de Tesis correspondiente a la carrera *Maestría en Procesos Educativos mediados por Tecnologías* dictada por el Centro de Estudios Avanzados de la Universidad de Córdoba. El mismo obtuvo su aprobación en diciembre de 2020.

La implementación del trayecto pedagógico, así como la investigación, se desarrollarán sobre la asignatura TRP correspondiente al Profesorado en Matemática dictada en la Facultad de Ciencias Exactas y Naturales (UNLPam).

Dicha investigación se financia, en parte, mediante el programa *Becas de Investigación y Posgrado 2020*, subprograma: *Becas posgrado para iniciar doctorados y maestrías*. (Res.051/20 C.S.)

1. INTRODUCCIÓN

En el área de matemática, uno de los desafíos más difíciles que enfrenta un docente universitario es generar una propuesta áulica cuyo enfoque de enseñanza no sea el tradicional. Esto significa que las clases teóricas no deberían caracterizarse por ser expositivas, rutinarias y unidireccionales. A su vez, las clases prácticas deberían ofrecer situaciones que motiven al alumno, evitando caer en la mera repetición de

procedimientos tediosos y carentes de sentido. Charnay (1997) explica que “el alumno debe ser capaz no sólo de repetir o rehacer sino también de resignificar en situaciones nuevas, de adaptar, de transferir sus conocimientos para resolver nuevos problemas” (p. 53). A su vez, el autor utiliza el término problema para referir a aquellas situaciones en cuya resolución, el estudiante percibe una dificultad u obstáculo. Luego, lo que puede resultar un problema para algunos, tal vez no lo sean para otros.

Desde el TRP, se pretende que los estudiantes puedan tener una experiencia de aprendizaje diferente, por lo que se encuentra en la gamificación una estrategia innovadora y un recurso potente para atender dificultades emergentes. Para formular acciones que promuevan un cambio favorable es necesario conocer las experiencias de los actores involucrados, las potencialidades de los recursos disponibles y la viabilidad de la propuesta.

Marco teórico

Karl Kapp (2012) define la ludificación o gamificación como “el uso de las mecánicas del juego, su estética y el pensamiento de juego para involucrar a la gente, motivar la acción, promover el aprendizaje y resolver problemas”. En el ámbito educativo se utiliza la gamificación, “para presentar al alumno una serie de retos de aprendizaje, que cuanto el alumno lo haya cumplido, generará una recompensa a corto plazo dimensionada a la complejidad del reto. No tiene por qué haber un videojuego de por medio, o cualquier tipo de tecnología involucrada para que funcione correctamente.” (Parente, 2016, p.18). Oriol Ripoll (citado por Carreras, 2017) propone una definición alternativa “Hacer

vivir experiencias de aprendizaje gratificantes usando elementos de juego (p.109). La misma, a decir de Parente (2016) no se enfoca en los elementos y las mecánicas del juego, sino en el resultado y en estimular el interés por aprender, en lugar de orientar el comportamiento.

Varios autores (Torrente, et al., 2009; McGonigal, 2011; Kapp, 2012) coinciden en que los juegos digitales deben contar con las siguientes características: objetivos, reglas, desafíos o conflictos, competencia, colaboración y/o cooperación, retroalimentación y re-jugabilidad (*re-playability*). Las mismas deberían ser tenidas en cuenta al llevar adelante un proceso de gamificación.

Por otro lado, Carreras (2017) establece diferencias entre gamificación y juego serio. Para ello, la autora toma los aportes de Oriol Ripoll, para definir juego serio “como un desarrollo tecnológico de aquella preocupación pedagógica que comparten muchos docentes por la necesidad de incorporar a sus clases un componente motivador y estimulante que garantice la atención y el interés de sus estudiantes.” (p. 112). A partir de este posicionamiento la autora explica que un juego serio puede considerarse una actividad de gamificación en la medida que comparta los objetivos de ésta aunque no toda gamificación consiste en un juego serio.

La gamificación como estrategia didáctica, mediatizada por tecnologías, ofrece un conjunto de acciones posibles que potenciará la apropiación de ciertos saberes geométricos mediante una propuesta que genere entusiasmo e involucramiento..

Antecedentes

La búsqueda, selección y análisis de diferentes publicaciones ha permitido recolectar información acerca de la implementación de gamificación, preferentemente en el nivel universitario y en el área de matemática. Es de interés identificar los motivos por los cuales se elige esta estrategia como recurso en el proceso de enseñanza y aprendizaje, el objetivo para el cual se utiliza, los niveles educativos o áreas de conocimiento donde se desarrolla, los contenidos propios del área que se abordan, la experiencia del alumno, el rol docente, las problemáticas emergentes y los resultados obtenidos.

Los objetivos planteados en diferentes experiencias realizadas apuntan a mejorar el rendimiento de los estudiantes, motivarlos y fortalecer los aprendizajes. El uso de dinámicas de juegos como estrategia de enseñanza es considerado un gran agente motivador que, además, fomenta la adquisición de habilidades como la resolución de problemas, el trabajo en equipo y la colaboración.

Soberano Martín, Castillo Mendoza y Peña Martín (2016), ratifican la propuesta de Prensky (2017) quien sostiene que para “diseñar una experiencia de educación es fundamental partir desde un cambio de roles, en donde el papel del profesor es hacer preguntas y el del alumno contestarlas, en donde las actividades propuestas sean desafiantes para el estudiante y no para el profesor.” (p. 10).

En líneas generales, los resultados obtenidos en las diferentes experiencias realizadas, en diversos niveles educativos, han sido favorables. Los estudiantes se muestran entusiasmados ante la propuesta que los ubica en un rol protagónico y activo, esto implica mayor

involucramiento de éstos en la clase y en su vínculo con sus compañeros. Por lo tanto, de acuerdo a la bibliografía consultada, la gamificación contribuye a estimular la motivación y el compromiso de los estudiantes ante el aprendizaje de contenidos matemáticos.

En la exploración de información realizada hasta el momento, no se han encontrado publicaciones originarias de nuestro país que dieran cuenta de la implementación de la estrategia de gamificación en un contexto universitario y en el área de matemática.

Aspectos Metodológicos

En principio, y durante todo el transcurso de la innovación, se efectuará la exploración y análisis de la literatura científica pertinente. Luego, se realizará una transcripción del conjunto de datos relevados con la intención de interpretarlo, extraer conclusiones.

Luego, es de vital importancia rastrear los saberes geométricos que disponen los estudiantes al momento de iniciar el cursado del TRP. No sólo se espera conocer si recuerdan definiciones y propiedades específicas sino también el grado de vinculación y complejización que pueden lograr a partir de las mismas. Esta información será recolectada a partir de las respuestas dadas a un conjunto de consignas elaboradas cuidadosamente.

Por otro lado, mediante un cuestionario se indagará acerca de las experiencias con entornos virtuales y con juegos en un marco de aprendizaje de contenidos matemáticos.

La propuesta didáctica se elaborará respetando la planificación correspondiente, el enfoque de enseñanza. Las situaciones problemáticas del campo geométrico se focalizarán en la

interpretación de consignas, extracción de información de una figura o enunciado, comparación de diferentes magnitudes (amplitudes de ángulos, longitudes de segmentos curvos y lineales, etc.), establecimiento de relaciones y comunicación de resultados haciendo uso de la simbología propia del campo disciplinar. Una vez definida la propuesta didáctica se elegirán los recursos digitales a utilizar, considerando sus alcances y limitaciones, y se procederá a la construcción de los niveles que darán cuerpo a la propuesta de gamificación.

Los recursos tecnológicos son fundamentales para el desarrollo de este proyecto. Se debe explorar cuáles son los más apropiados y así seleccionar aquellos considerados más convenientes para la situación.

El plan de acción se ejecutará en su mayor parte de manera virtual y contará con algunos encuentros presenciales bajo la modalidad de taller, donde se compartirán las producciones realizadas. En estos encuentros también se desarrollarán aquellas consignas consideradas más adecuadas para la modalidad. El docente mantendrá canales virtuales de comunicación e interacciones formativas asincrónicas.

Se espera la participación comprometida de todos los estudiantes así como el tránsito autónomo por los diferentes niveles propuestos, lo cual implicaría un fortalecimiento de saberes geométricos considerados necesarios por la cátedra.

Para conocer el desempeño de los estudiantes durante la implementación de la propuesta de enseñanza, se realizarán observaciones cuantitativas y observaciones participantes. Este seguimiento, permitirá analizar y evaluar

la situación, para así realizar las modificaciones que se consideren oportunas. Los registros generados a partir de las observaciones realizadas, serán recuperados para su organización. De allí, se espera determinar unidades de análisis que serán posteriormente categorizadas.

Finalmente, se evaluará si efectivamente se logró un fortalecimiento de los saberes de los estudiantes, así como el impacto de la estrategia didáctica de gamificación en este proceso de adquisición

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación/innovación que se abordan en este proyecto son:

- El uso de gamificación en el diseño e implementación de una propuesta de enseñanza.
- Identificación y uso diferentes recursos tecnológicos para la implementación de una propuesta de gamificación.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos son:

- Se llevó adelante una actividad exploratoria en la que se pudo verificar que los estudiantes matriculados en el TRP indican no poseer conocimientos sobre ciertos temas de geometría.
- A partir del análisis bibliográfico no se halló aún evidencia de abordajes de contenidos geométricos, en nuestro país, en el nivel universitario.

Otros resultados esperados se vinculan con:

- Continuar con la exploración de propuestas gamificadas en matemática en

diferentes niveles educativos en nuestro país.

- Explorar y definir las herramientas tecnológicas que soporten la propuesta gamificada.
- Diseñar e implementar una propuesta gamificada que articule los conocimientos previos de los estudiantes con los esperados por la cátedra.
- Enriquecer la propuesta didáctica de TRP.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación presentada aporta a la tesis de maestría de la Prof. Clarisa Pauletti y está enmarcada en la cátedra del TRP como parte del desarrollo de actividades innovadoras.

5. BIBLIOGRAFIA

Carreras Planas, C. (2017). Del Homo Ludens a la gamificación. *Quaderns de Filosofia*, 4(1). <https://doi.org/10.7203/qfia.4.1.9461>

Charnay, R. (1997). Aprender (por medio de) la resolución de problemas. En C. Parra & I. Saiz (Eds.), *Didáctica de matemáticas. Aportes y Reflexiones* (5.a ed., p. 299). Paidós Educador.

Kapp, K. M. (2012). The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education. *International Journal of Gaming and Computer-Mediated Simulations*, 4(4), 81-83. <https://doi.org/10.4018/jgcms.2012100106>

McGonigal, J. (2011). *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. Penguin Press.

Parente, D. (2016). Gamificación en la educación. En R. S. Contreras & J. L. Eguía (Eds.), *Gamificación en las aulas*

universitarias. Institut de la Comunicació. Universitat Autònoma de Barcelona.

Soberano Martín, A., Castillo Mendoza, J. L., & Peña Martín, A. (2016). Aprendizaje matemático mediante aplicaciones tecnológicas en un enfoque de Gamificación. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 3(5), Article 5. <http://www.pag.org.mx/index.php/PAG/article/view/449>

Torrente, J., Marchiori, E. J., Frossard, F., Trifonova, A., & Barajas, M. (2009). Production of Creative Game-Based Learning Scenarios—A Handbook for Teachers. *ProActive: Fostering Teachers Creativity through Game-Based Learning*, 43.

Marco de referencia para la formulación, gestión e implementación de carreras a distancia según el SIED UNSJ. Desarrollo de Manual de procedimientos.

Pósito, Rosa María; Leiva, Alfredo Enrique; González, Liliana Mirna; Sanchez. Horacio; Dominguez, Ana Patricia; Calvo, Ines; Marquez, Victorina; Flaque, Valeria; Orellana Vassallo, Alejandra; Emilse Carmona

Gabinete de Tecnología e Innovación Educativa -GATIE- y Departamento Informática Facultad de Ciencias Exactas, Físicas y Naturales – UNSJ
Área de Educación a Distancia - SIED UNSJ

Av. Ignacio de la Roza 590 (O). Complejo Universitario "Islas Malvinas". Rivadavia, San Juan,
Teléfonos: 4234129, Fax 0264-4234980
{ rosapósito, Leiva.alfredo, lilianamirna, sanchezhd, anapato.domin }@gmail.com ; icalvo@unsj-
cuim.edu.ar; vmarquez3@hotmail.com , valeriaflaque@hotmail.com; {aoalitaorellana,
emilsecoaching}@gmail.com

Resumen

La UNSJ ha validado ante CONEAU, el Sistema Institucional de Educación a distancia de la Universidad (SIED UNSJ), tal como lo solicita la normativa del Ministerio de Educación N°2641/17ME. En torno a las normativas generales definidas en el SIED UNSJ, surgen los interrogantes para la debida implementación de las mismas a nivel de cada unidad académica: ¿las unidades académicas cuentan con las normativas, recursos y procesos de gestión adecuados para la formulación, gestión e implantación de carreras a distancias, según los requerimientos planteados? ¿las reglamentaciones existentes, están vinculadas al SIED UNSJ? ¿existen mecanismos operativos para la debida articulación entre la unidad académica y el área de EAD en relación a las dimensiones definidas en el SIED UNSJ? Arribar a las respuestas de los

interrogantes planteados permitirá generar un marco de referencia para la articulación de las unidades académicas con el Área EAD con dependencia de Rectorado. Los integrantes del equipo de investigación del presente proyecto, en su mayoría cuentan con una trayectoria de 20 años en proyectos de investigación afín a Educación a distancia e integraron la Comisión designada para la elaboración el SIED UNSJ, lo que representa una base de conocimiento base importante para resolver la problemática planteada. Los resultados del proyecto representarán un valioso aporte para la implementación del SIED UNSJ.

Palabras clave: Educación a distancia. Sistema Institucional de Educación a Distancia. SIED UNSJ. Manual de Procedimientos.

Contexto

Este trabajo se enmarca en el proyecto de investigación “Marco de referencia para la formulación, gestión e implementación de carreras a distancia según el SIED UNSJ.” Inserto como proyecto de investigación del GATIE - Gabinete de Tecnología e Innovación Educativa- y del Departamento de Informática, ambas unidades de la Facultad de Ciencias Exactas de la UNSJ. Los resultados tienen una transferencia inmediata en el Área de Educación a distancia del SIED UNSJ. La investigación está financiada por la Universidad Nacional de San Juan.

Introducción

La UNSJ ha validado ante CONEAU, el Sistema Institucional de Educación a distancia de la Universidad (SIED UNSJ) [1] tal como lo solicita la normativa del Ministerio de Educación N°2641/17ME. En torno al sistema desarrollado se han aprobado en el Consejo Superior dos ordenanzas: Ord N°011/19- [2] y la Ord. 012/19-CS [3]. Según el Reglamento, el SIED UNSJ es el conjunto de acciones, normas, procesos, equipamiento, recursos humanos didácticos que permiten el desarrollo de proyectos educativos a distancia. En el Art.3 [3], establece que las carreras que se desarrollen en el marco de la modalidad a distancia y aquellas carreras presenciales, cuya cantidad de horas no presenciales esté comprendida dentro de los porcentajes que establece la normativa nacional vigente, deben cumplir con los lineamientos y componentes del SIED UNSJ y contar con informe favorable del Área EAD. Así mismo, en el art. N°6 [3], se establece que las unidades académicas deberán arbitrar los medios para articular su vinculación

con el Área EAD de la UNSJ, para la implementación de carreras a distancia.

En este sentido, surgen los interrogantes ¿las unidades académicas cuentan con las normativas, recursos y procesos de gestión adecuados para la formulación, gestión e implantación de carreras a distancias? ¿las reglamentaciones existentes, están vinculadas al SIED UNSJ? ¿existen mecanismos operativos para la debida articulación entre la unidad académica y el área de EAD en relación a las dimensiones definidas en el SIED UNSJ?

Arribar a las respuestas de los interrogantes planteados permite generar un marco de referencia para la articulación de las unidades académicas con el Área EAD con dependencia de Rectorado.

Los integrantes del equipo de investigación, en su mayoría cuentan con una trayectoria de 20 años en proyectos de investigación a fin a Educación a Distancia (EaD) e integraron la comisión de elaboración del SIED UNSJ [4] y actualmente aportan a los equipos de trabajo del SIED UNSJ, lo cual representa una base de conocimiento muy importante para resolver la problemática planteada. Los resultados del proyecto representarán un valioso aporte para el SIED UNSJ.

Líneas de Investigación, Desarrollo e Innovación

Esta investigación se realiza en torno a la Línea de investigación, desarrollo e innovación definidas en el GATIE- Gabinete de Tecnología e Innovación Educativa: Educación a distancia/ Gestión y Administración [5].

Objetivos

El proyecto de investigación que se presenta tiene los siguientes objetivos:

- **Objetivo General:** Generar un marco de referencia en torno a la formulación, gestión e implementación de carreras de pregrado, grado y posgrados de Educación a distancia en el SIED UNSJ.

- **Objetivos específicos:**

- Describir el estado institucional de las unidades académicas para la implementación de carreras en el marco del SIED UNSJ.

- Identificar las necesidades de creación y/o modificación de normativas, recursos y procesos para la formulación, gestión e implantación de carreras a distancias.

- Generar mecanismos operativos que faciliten la articulación de las Unidades Académicas con el área EAD del SIED.

- Definir los procesos implicados en la formulación, gestión e implementación de una carrera en el marco del SIED UNSJ.

Metodología

Se lleva cabo una investigación cualitativa de tipo proyectiva, dado que se busca encontrar soluciones a los problemas en torno a la formulación, gestión e implementación de carreras en el marco del SIED UNSJ. Este tipo de investigación implica explorar, describir y proponer acciones. Se aplica un diseño de investigación de fuente mixta -de campo y documental-. La investigación documental se realizará a través de la consulta de la documentación institucional de la unidad académica que dé cuenta de las políticas y mecanismos operativos referidos a la temática objeto de estudio. La investigación de campo implica el relevamiento de los procesos de formulación, gestión e implementación de carreras que se realizan actualmente en

cada uno de los lugares -departamentos, sección alumnos, secretarías, etc.

Resultados

Se desarrollaron satisfactoriamente las siguientes actividades previstas con importantes acciones de transferencia, citadas a continuación:

1. Búsqueda y revisión documental relacionada con la formulación de carreras a distancia, en normativas de CONEAU, del Ministerio de educación y de otras Universidades.

2. Análisis de las normativas y ordenanzas definidas en el marco del SIED UNSJ

3. Relevamiento y sistematización del estado Institucional -actores, procesos, normativas, recursos- de la Unidad académica definida como caso base, en relación con las dimensiones definidas en SIED UNSJ: Pedagógico-didáctica, Tecnológico-comunicacional, Gestión administrativa, Seguimiento y evaluación y Formación y actualización.

4. Desarrollado un marco de referencia, en el marco de una tesis de Maestría, para la implementación de Learning Analytic en educación superior. Los resultados de esta investigación, aspiran a promover la cultura del análisis del aprendizaje en la institución educativa de la UNSJ de manera eminentemente práctica y concreta.

5. Identificación, análisis y creación/modificación de las normativas propias de la Unidad Académica que tienen incumbencia en la formulación, gestión e implementación de carreras a distancia

6. Identificación y modelado de los procesos implicados en la formulación, gestión e implementación de las carreras a distancia. Generación del Mapa de Procesos y Manual de Procedimientos.

7. Elaboración de mecanismos operativos que faciliten la articulación de

la Unidad Académica con el Área EAD de la Universidad

8. Acciones de transferencia de los resultados del proyecto a las autoridades del SIED.

Acciones de transferencia a la UNSJ:

-Participación en la Comisión de elaboración del SIED UNSJ [4]

-Presentación del Plan de emergencia sanitaria para abordar la virtualización de las carreras de la UNSJ. El equipo del proyecto, que se encontraba trabajando en la comisión designada para la elaboración del SIED, realizó la presentación de un plan de emergencia ante la situación de pandemia presentada en marzo 2020. Lo cual representó una acción de transferencia de gran importancia desde el proyecto de investigación que acompañó todas las actividades del plan ejecutado en el 2020. Se mencionan las actividades principales realizadas en el marco del Plan de emergencia propuesto por el equipo del proyecto

*Capacitación de docentes en el diseño y utilización de aulas virtuales como espacio virtual de intercambio y circulación de información entre docentes y estudiantes, dentro de la/s plataforma/s educativa/s de la UNSJ.

*Asignación de aulas virtuales para las distintas cátedras de las carreras de la Universidad en el nuevo Entorno Virtual de Aprendizaje (EVA) desarrollado en la versión actualizada de Moodle (Versión 3.6) www.campusvirtual.unsj.edu.ar . Este entorno fue configurado el año anterior, pero carecía de mecanismos operativos, para lo cual el equipo del proyecto realizó importantes aportes.

*Generación de mecanismos que posibiliten/ garanticen la habilitación de las aulas, y el alta de docentes y estudiantes, de las aulas virtuales creadas en el nuevo EVA.

*Creación de un portal informativo, para docentes y estudiantes www.sied.unsj.edu.ar

*Propuesta al Área de Educación a Distancia de mecanismos adecuados para acompañar a las y los docentes, en los procesos de diseño y desarrollo de las aulas virtuales, en el uso de las herramientas y recursos disponibles en las distintas plataformas, asesoramiento o en la elaboración de guiones de clase, y en la selección de bibliografía digital.

*Propuesta al Área de Educación a Distancia de mecanismos adecuados para acompañar los estudiantes, a fin de que sean capacitados en el uso del aula de apoyo a la presencialidad, de manera virtual, a través de los sistemas de tutorías y/o video tutoriales de las herramientas y recursos disponibles en entorno virtual.

*El equipo del proyecto participó en la elaboración del PlanVES [6]. La Resolución de participación de los integrantes del proyecto está en trámite. En el marco de este Plan, las actividades más relevantes propuestas por el equipo del proyecto, para desarrollar en el 2021 son:

1-Elaboración de los instrumentos para que las Unidades Académicas puedan, en cada una de las carreras que incluyen horas no presenciales, chequear cada uno de los componentes que solicita la normativa del Ministerio de Educación N° 2641/17 ME y se dé cumplimiento al Art 3°, de la respuesta de validación de CONEAU[1].

2-Elaboración del Mapa de Procesos y Manual de procedimiento para el SIED UNSJ . Toda institución de prestigio en la enseñanza universitaria, debe trabajar firmemente en la excelencia, la sociedad lo reclama. Como única alternativa para lograr ese cometido, más allá de los valores educativos que exhibe, la capacidad y el compromiso de los

recursos humanos utilizados en el proceso de enseñanza – aprendizaje, se hace necesario la adecuación de la gestión organizacional, su desarrollo y control como una manera de avanzar sistemáticamente hacia la calidad a través de un proceso permanente de mejora continua. En este sentido, desde el presente proyecto se dio inicio al proceso de desarrollo de un Manual de Procedimientos del SIED UNSJ, que en su primera etapa, plantea el objetivo del diseño, redacción, actualización e implementación de Procedimientos para todas las dimensiones del SIED UNSJ.

La tarea consiste en el diseño en algunos casos y el rediseño en otros, de los procedimientos que describen los innumerables procesos en la complejidad del servicio que el SIED UNSJ debe brindar a las carreras futuras de grado y posgrado de la UNSJ, como así también a la asistencia a la opción pedagógica de presencialidad, que en el marco de la presente coyuntura tiene un nivel de requerimiento nunca antes observado. El presente manual de procedimientos, se está desarrollando siguiendo lineamientos técnico-operativos generales del Estandart Internacional del grupo de Normas ISO 9000, para el desarrollo de un Sistema de Gestión de la Calidad, porque se pretende que este Manual pueda ser la base documental de un futuro proceso de certificación.

El Manual de Procedimientos que se ha comenzado tiene la siguiente estructura:

- Procedimientos SIED
- Registros SIED
- Instructivos SIED
- Formularios y plantillas
- Anexos

Para el modelado de procesos se ha elegido el Lenguaje estandarizado BPM (Business Process Management, utilizando el software Bizagi.

Formación de Recursos Humanos

El proyecto marco de este trabajo promueve la capacitación y actualización permanente de los miembros del equipo de investigación y a la consolidación del grupo de investigadores en la temática bajo estudio. Se destaca la sig. formación:

- Tesis de Maestría en Informática de Alejandra Orellana “*Marco de Referencia para la implementación de Learning Analytic en educación superior*”
- 2 Becas SIED de Estudiantes avanzados: Salas, Jimena; Dipane , Miguel Angel.
- 5 Pasantías en el equipo del SIED UNSJ: Molina, Andres; Peralta, Maria L; Moreno,Nahuel; Ruggieri, Pietro; Radicetti, Lucia
- Adscripta Egresada: Lic. Valeria Mallea.

Referencias:

- [1] Validación del Sistema Institucional de Educación a distancia de la Universidad (SIED UNSJ) RESOL-2021-9-APN-SECPU#ME. para reconocimiento oficial y la consecuente validez nacional de los títulos de carreras desarrolladas bajo la modalidad a distancia o bien presenciales en las que la cantidad de horas no presenciales se encontrara entre los porcentajes definidos Resolución Ministerial 2641.
- [2] Ord N° 011/19-CS de Creación del Área de Educación a Distancia UNSJ.
- [3] Ord. 012/19-CS de Creación del Reglamento General del SIED UNSJ
- [4] Resolución N° 3939-19-R- Designación de la comisión SIED.
- [5] Plan estratégico de desarrollo en GATIE .Res 1041-19-CD – FCEFYN
- [6] Plan VES -Plan de virtualización de la Educación Superior-. Mediante Nota N° NO-2020-85849813-APN-SSPU#ME emitida diciembre2020 por el Equipo de Evaluación Técnica, de la Dirección de Calidad Universitaria de la Secretaría de

Políticas Universitarias del Ministerio de Educación de la Nación

[7] “Artículo N°3 - Al momento de la evaluación de las carreras de pregrado, grado o posgrado para su acreditación y/o reconocimiento oficial del título, deberá asegurarse que los componentes del SIED estén debidamente implementados en el funcionamiento de dichas carreras.”

Título de la tesis: “Marco de Referencia para la implementación de Learning Analytic en educación superior”

Tesista: Alejandra Orellana Vassallo

Título alcanzado: Magister en Informática-Universidad Nacional de San Juan

Fecha de la defensa: 16-septiembre-2020

Director: Mag. Rosa María Pósito

Co-director: Mag. Liliana Mirna González

Gabinete de Tecnología e Innovación Educativa -GATIE- y Departamento Informática Facultad de Ciencias Exactas, Físicas y Naturales – UNSJ Área de Educación a Distancia - SIED UNSJ

Av. Ignacio de la Roza 590 (O). Complejo Universitario "Islas Malvinas". Rivadavia, San Juan, Teléfonos: 4234129, Fax 0264-4234980 aoalitaorellana@gmail.com

Resumen

El presente trabajo de tesis ha incursionado en el campo de la analítica del aprendizaje, en inglés, Learning Analytic (LA) un campo aun en desarrollo que se ocupa de la interpretación de un amplio rango de datos educativos producidos durante el proceso de enseñanza-aprendizaje por los estudiantes para orientar su progreso. Se abordó la problemática de satisfacer la necesidad del docente de medir, recolectar, visualizar e interpretar estos datos. Los resultados obtenidos son de carácter eminentemente práctico, pues se ha propuesto una taxonomía de métricas para medir ciertos aspectos del proceso enseñanza-aprendizaje, se ha identificado y generado buenas prácticas sobre el uso de herramientas analíticas asociadas a cada métrica, herramientas de limpieza y transformación de datos como así también herramientas de visualización de los mismos, obteniendo un prototipo: un “dashboard docente”. Estos resultados, en conjunto, conforman el marco de referencia que aspira contribuir a mejorar la calidad educativa.

Palabras clave: Analítica del aprendizaje, Learning Analytic, métricas de seguimiento, cultura del análisis del aprendizaje.

Contexto

La tesis se presenta como uno de los resultados de la línea de investigación dentro del proyecto de investigación “Marco de referencia para la formulación, gestión e implementación de carreras a distancia según el SIED UNSJ.” Inserto como proyecto de investigación del GATIE - Gabinete de Tecnología e Innovación Educativa- y del Departamento de Informática, ambas unidades de la Facultad de Ciencias Exactas de la UNSJ. Los resultados tienen una transferencia inmediata en el Área de Educación a distancia del SIED UNSJ.

Línea de Investigación

Este trabajo de tesis ha seguido la línea de investigación, de LA, un campo aun en desarrollo que se ocupa de la interpretación de un amplio rango de datos educativos producidos y recogidos para orientar el progreso de los estudiantes.

En educación, el término LA, se acuñó por primera vez en el año 2000 en un artículo sobre e-learning escrito por Mitchel J, Costello S.,[1] y se extendió a partir del año 2005 gracias a la

compañía Blackboard que se dedicó en especial a proporcionar a instituciones educativas de entornos de aprendizaje, los LMS (Learning Management System). De acuerdo a Santamaría G. F [2], además de estar enfocada en el aprendizaje, es una disciplina que interactúa con otras de gran relevancia como la minería de datos educativos, en inglés: Educational Data Mining, (EDM), la inteligencia empresarial (BI), el análisis de redes sociales, en inglés: Social Network Analysis (SNA) y lo relativo a Machine Learning (ML).

De acuerdo a Suthers y Rosen [3] el aprendizaje y la creación de conocimiento a menudo se distribuyen en múltiples medios y sitios en entornos en red, de modo que los rastros de dicha actividad pueden fragmentarse en múltiples registros y pueden no coincidir con las necesidades analíticas. A medida que las herramientas y recursos de aprendizaje se están moviendo cada vez más en la nube, el desafío es cómo integrar estos datos sin procesar, provenientes de múltiples fuentes, para crear un conjunto de datos educativos útiles que refleje las actividades distribuidas. Éste es un gran desafío que se le presenta al docente: la recolección e interpretación de los datos educativos fragmentados y dispersos. A su vez Dietz-Uhler y Hurn, [4] expresan: *“LA es una disciplina emergente relacionada con el desarrollo de métodos para explorar una serie de datos procedentes de ecosistemas educativos, con el uso posterior de los resultados del análisis para entender mejor al alumnado, sus comportamientos y así mejorar el diseño de los entornos en los que aprenden. Estos postulados conforman el desafío de aplicar métodos, técnicas y buenas prácticas en el aula para alcanzar la mejor innovación educativa”*.

Bibliografía:

- [1] Mitchell, J., & Costello, S. (2000). International e-VET market research report: A report on international market research for Australian VET online products and services. Sydney, Australia: John Mitchell & Associates and Education Image.
- [2] Santamaría Glez, F. (2013). Análisis del aprendizaje: práctica emergente para un diseño instruccional en un mundo de datos interconectados. TEXTOS Revista Internacional de Aprendizaje y CiberSociedad, 17.
- [3] Suthers, D. & Rosen, D. (2011). A unified framework for multi-level analysis of distributed learning. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge. (pp. 64-74). NY, USA: ACM New York.
- [4]-Dietz-Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. Journal of interactive online learning, 12(1), 17-26.

MATERIALES EDUCATIVOS DIGITALES PARA LA CONSTRUCCIÓN DE CONOCIMIENTOS EN LA UNIVERSIDAD

- González, Liliana Mirna, lilianamirna@gmail.com

Universidad Nacional de San Juan
Facultad de Ciencias Exactas, Físicas y Naturales
Gabinete de Tecnología e Innovación Educativa.
Dpto. de Informática

- Cuadros Patricia, pcuadrosan@hotmail.com

Universidad Nacional de San Juan
Facultad de Ingeniería. Dpto. de Matemáticas

- Pontoriero Francisco ruffopontoriero@gmail.com

Universidad Nacional de San Juan
Facultad de Ciencias Exactas, Físicas y Naturales
Gabinete de Tecnología e Innovación Educativa.
Dpto. de Informática

- Gallardo Vanesa vanesagallardol@gmail.com

Universidad Nacional de San Juan
Facultad de Ciencias Exactas, Físicas y Naturales
Gabinete de Tecnología e Innovación Educativa.
Dpto. de Geofísica y Astronomía y Dpto. de Geología

Resumen

Los jóvenes tienen en las yemas de los dedos las tecnologías; ellas atraviesan sus modos de pensar, aprender y conocer, de ahí porque debemos identificar y reconocer las tendencias culturales.

Maggio (2012), es necesario reconocer el valor de los “usos culturales y formas cognitivas del mismo modo como deberíamos tener en cuenta los estilos cognitivos propios de los sujetos culturales que son nuestros alumnos” (p.153)

Cabe preguntarnos, ¿Qué materiales educativos digitales prefieren lo hipertextual, el texto plano, lo multimedial?

¿Cuándo utilizamos tecnologías priorizamos el “efecto de la tecnología”?

La oralidad y la escritura han sido y seguirán prevaleciendo como las formas de narrar. Así lo

visual, lo audiovisual, el sonido y otras formas de representación van imprimiendo su sello como complementos de lo oral y lo escrito.

Con el avance de las tecnologías, el lenguaje multimedial trae consigo un nuevo tipo de narración, aquella en la que se interrelacionan distintos medios como el vídeo, el texto, la fotografía, la infografía, la animación y la ilustración; son las nuevas narrativas digitales.

La situación de pandemia ha acelerado la necesidad del diseño e implementación de materiales educativos digitales. A su vez ha contribuido a consolidar la opción pedagógica a distancia.

Palabras clave: MATERIALES-EDUCATIVOS- DIGITALES- EFECTOS-COGNICIÓN- UNIVERSIDAD

Contexto

El Proyecto mencionado –en proceso de evaluación por CICITCA- tiene como objetivo principal generar conocimientos que aporten a la problemática: qué efectos cognitivos dejan en alumnos universitarios los materiales educativos digitales. Se encuentra enmarcado en el “Gabinete de Tecnología e Innovación Educativa creado por Resolución 03/18, dependiente de Decanato de la Facultad de Ciencias Exactas, Físicas y Naturales. El Gabinete se constituye en un espacio académico y de investigación cuyo es generar estrategias, metodologías, medios y materiales, así como instrumentos tecnológicos, que contribuyan a la inserción de innovaciones tecnológicas en pos de la mejora de los procesos educativos de la Facultad de Ciencias Exactas, Físicas y Naturales y de la Facultad de Ingeniería de la Universidad Nacional de San Juan

1. Introducción

La intencionalidad pedagógica de las prácticas de enseñanza es el eje vertebrador de los procesos de aprendizajes que desarrollan los alumnos. El docente es quién define cómo desarrollar e implementar las prácticas de enseñanza contemplando la realidad socio-educativa, el contexto socio-histórico en el que se da la interacción entre los componentes del triángulo interactivo –docente/alumnos/contenidos-

El alumno realiza los procesos de construcción de conocimiento mediados por la ayuda, guía y orientación del docente y de sus compañeros. Estos procesos están mediados por la acción del docente, de otros alumnos, el contenido y -en la actualidad, en muchos casos- por las nuevas tecnologías. La concepción constructivista considera que los procesos de construcción de conocimientos no se realizan en el vacío sino en un contexto socio-cultural determinado y en un

tiempo particular, orientado por objetivos plasmados en un diseño educativo.

Salomon, Perkins y Globerson (1992) distinguen dos tipos de efectos cognitivos:

los efectos que se obtienen en conjunción CON la tecnología en el curso de la colaboración intelectual con ella, y los efectos procedentes DE la tecnología, en términos del residuo cognitivo transferible dejado por la colaboración, tras la forma de un mayor dominio de habilidades y estrategias. (1992:7)

Siguiendo a los autores mencionados, los efectos CON la tecnología “influyen en lo que los alumnos hacen, en la calidad y cuando lo hacen”. En cambio, refieren a efectos DE la tecnología cuando se dan cambios duraderos en las capacidades cognitivas generales de los alumnos como resultado de haber operado con tecnologías; cambios en el “dominio del conocimiento, de la habilidad, o bien de la profundidad de la comprensión” después de dejar el uso de la tecnología. (Salomon et al,1992:8)

El conocimiento es el resultado de un proceso permanente de construcción a través de la interacción y del diálogo con “otros”, mediado por materiales educativos digitales y situado en un contexto socio-histórico. El aprendizaje colaborativo da sustento a la interacción entre los diferentes elementos del Triángulo didáctico mediado por las tecnologías.

El sistema educativo ha privilegiado la oralidad y la escritura en sus formas de narrar. Así lo visual, lo audiovisual, el sonido y otras formas de representación se fueron incorporando paulatinamente como complementos de lo oral y lo escrito.

En la actualidad con el avance de las tecnologías en sus diferentes formatos y soportes, el lenguaje multimedial trae consigo un nuevo tipo de narración, aquella en la que se interrelacionan distintos medios como el vídeo,

el texto, la fotografía, la infografía, la animación y la ilustración; son las nuevas narrativas digitales.

La lógica de la narración unidireccional fue tomando otros formatos, aunque persiste, dando lugar a la lógica de lo hipertextual y lo multimedial y así a la hipermedia.

Ambrosino (2017) expresa:

Las narraciones establecen un contexto las mismas nos otorgan una estructura que posibilita construir una trama más elaborada. [...] son fundamentalmente un lenguaje que se construye a partir de relatos que ofrecen significación y realización. Se refieren al tejido de la acción e intencionalidad docente. [...]...la narrativa es la que otorga sentido y significación al contenido, la sociabilidad y estructura los escenarios de aprendizaje para la construcción de conocimiento.

Teniendo en cuentas estas conceptualizaciones nos propusimos llevar adelante una investigación de “diseño” que nos permitió construir materiales educativos digitales que orientan a nuestros alumnos a desarrollar conocimiento como sujetos culturales en un paradigma socio-tecnológico.

2. Líneas de Investigación, Desarrollo e Innovación

- Efectos cognitivos que dejan los materiales educativos digitales en estudiantes universitarios

3. Resultados Obtenidos / Esperados

Objetivos

- Generar conocimientos acerca de qué efectos cognitivos dejan en alumnos universitarios los materiales educativos digitales.

Objetivos específicos

- Identificar características de macro y micro narrativas como materiales educativos digitales en carreras universitarias.
- Diseñar materiales educativos digitales y e-actividades que favorezcan la construcción de conocimiento para diferentes áreas disciplinares.

Resultados Obtenidos

En el marco de una investigación de “diseño” se logró construir narrativas que orienten a nuestros alumnos a desarrollar conocimiento como sujetos culturales.

❖ Diseño e Implementación de Materiales Educativos

- **Acerca de la Ética y la Moral: Exelearning.**

- **Derivada de un cociente:** El video elaborado explica cómo derivar funciones que involucren un cociente entre funciones. <https://www.youtube.com/watch?v=PRxl7K7OtLo>

- **Proceso de Transformación de conocimientos.** El video elaborado explica conceptos, relaciones de conceptos y ejemplifica.

- **Función Diferenciable. Cálculo II** <https://youtu.be/uyS9zufCPkI>

- **Función Diferenciable. Ejercicios. Cálculo II** <https://youtu.be/Z0JZ7VzJzFk>

- **Derivada de un producto:** El video elaborado explica cómo derivar funciones que involucren un producto entre funciones. https://www.youtube.com/watch?v=BiZgyaKok_o

- **Tecnologías y algo más:** texto plano acerca de conceptos generales del tema para un “Curso de capacitación de Diseño Básico de Aula Virtuales en Moodle”. SIED-UNSJ

- **Tecnología:** El video elaborado explica conceptos, relaciones de conceptos y ejemplifica.

<https://www.youtube.com/watch?v=v82ZddoBPgo&t=9s>

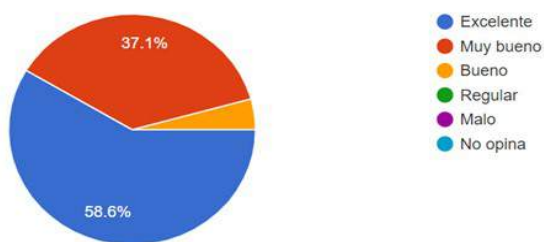
❖ Entono tecno-pedagógico:

- Aspectos Profesionales y Sociales. Licenciatura en Ciencias de la Computación. FCEFYN. SIED UNSJ <https://campusvirtual.unsj.edu.ar/course/view.php?id=6>
- Ciencia, Tecnología y Sociedad. Licenciatura en Sistemas de la Información. FCEFYN. SIED UNSJ <https://campusvirtual.unsj.edu.ar/course/view.php?id=7>

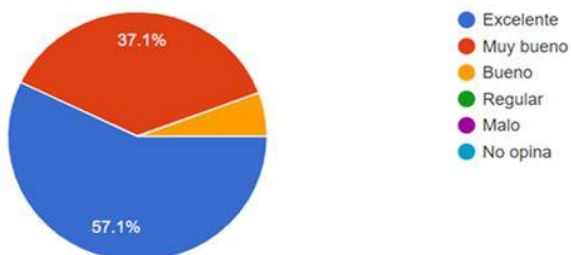
❖ Datos preliminares encuesta a estudiantes

Encuesta a los alumnos de la asignatura cálculo II – especialidades Ing. Industrial – Ing. Química – Ing. en Alimentos – Año 2020.

- Claridad y comprensión del material educativo (videos, documentos, ejercitación)



- Los materiales multimedia favorecieron la construcción de conocimientos.



4. Formación de Recursos Humanos

El equipo de trabajo está conformado de la siguiente manera:

- Directora: González, Liliana Mirna
- Co-directora: Ureta Laura
- Investigadores: Pontoriero, Francisco, Cuadros Patricia, Margarit Viviana, Gallardo Vanesa, Marcovecchio, María J., Rossetti Beiram, Gabriela.

Se propone volcar los resultados en actividades de Postgrado, tales como:

- Diseño e implementación de la Diplomatura “Educación y nuevas tecnologías en tiempos de convergencia tecnológica”. Aprobada por Ordenanza N° 5 CD-FCEFN, Prof. Responsable: Mgter Liliana Mirna González.
- Co-dirección/Asesoramiento de Tesis de Maestría en Informática. F.C.E.FyN de la UNSJ.
- Docente de la Maestría en Informática. F.C.E.FyN de la UNSJ. Materia: Educación y Tecnologías.
- BECARIOS
Puigdengolas Márquez, Fabricio Agustín Facultad de Ingeniería. Universidad Nacional de San Juan. Tareas de investigación y desarrollo. Co-director o co-tutor: Cuadros, Patricia Del Valle

5. Referencias bibliográficas

- Ambrosino, María Alejandra (2017) Docencia y NARRATIVAS Transmedia en la educación superior. Trayectorias Universitarias. Volumen 3, N° 4. Recuperado de <http://revistas.unlp.edu.ar/TrayectoriasUniversitarias>
- Lion Carina (13 de setiembre de 2012) Las narrativas docentes en entornos digitales. Ibertí. OEI. Recuperado de <https://www.youtube.com/watch?v=8TJQFeP51rg>

- Maggio M. (2018) Reinventar la clase en la Universidad. Paidós. Argentina
- Salomon, Perkins y Globerson (1992) Coparticipando en el conocimiento: la ampliación de la inteligencia humana con las tecnologías inteligentes. En Revista Comunicación, lenguaje y educación. CL&E, n° 13, Madrid.
- Scolari, C. (2014). Narrativas transmedia: nuevas formas de comunicar en la era digital. Anuario AC/E de cultura digital. Recuperado de https://www.accioncultural.es/media/Default%20Files/activ/2014/Adj/Anuario_ACE_2014/6Transmedia_CScolari.pdf
- Odetti, V. (2017). El diseño de materiales didácticos hipermediales: El caso del PENT FLACSO. Buenos Aires: Teseo.
- — (2016). Materiales didácticos hipermediales: lecciones aprendidas y desafíos pendientes. En García, J. M., y Báez Sus, M. (Comp.) Educación y tecnologías en perspectiva. Diez años de FLACSO Uruguay. Disponible en http://flacso.edu.uy/publicaciones/libro_educacion_tecnologia_2016/Baez_Garcia_Educacion_y_tecnologias_en_perspectiva.pdf
- — (2018) Narrativas Trasmmedia. Educación y TIC. EL ABROJO. Montevideo. Uruguay.

NetOS-Lab: Laboratorio portátil de Redes y Sistemas Operativos

Mercedes Barrionuevo⁽¹⁾, Carlos Taffernaberry^(1,2)

⁽¹⁾ Univ. Nacional de San Luis, San Luis

⁽²⁾ Universidad Tecnológica Nacional, Facultad Regional Mendoza
Argentina
{mbarrio, ctaffer}@unsl.edu.ar

RESUMEN

Las redes de computadores tienen un gran impacto en el mundo profesional, el conocimiento y el esparcimiento a nivel mundial. En el ámbito laboral, por ejemplo, la finalidad principal para la creación de una red de computadoras es compartir los recursos e información a distancia, asegurando la confiabilidad y la disponibilidad de la misma.

En la actualidad la complejidad y heterogeneidad que presentan las redes de computadoras, e Internet mismo, requieren de profesionales cada vez capacitados en todo lo referido a estas temáticas para poder administrarlas correctamente, diagnosticar y reparar sus fallos.

Nuestra principal motivación es desarrollar herramientas que ayuden a formar profesionales aptos para dicho mundo laboral, como así también, dejarlas disponibles de forma libre para su uso y aportes en futuras mejoras y/o actualizaciones.

En este sentido, NetOS-Lab es una herramienta con fines educativos que permitirá ensayar servicios de red (DNS, VPN, SMTP, firewall, etc.) y simular distintos aspectos de Sistemas Operativos de Red, como planificación de procesos y gestión de memoria.

Palabras claves: Redes de computadoras, Virtualización, Simuladores, Educación, Laboratorios portátiles.

CONTEXTO

Esta propuesta se desarrolla dentro del Laboratorio de Redes de Computadoras del Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales (UNSL).

Actualmente este proyecto ha sido preseleccionado junto a otros cinco en la convocatoria Programa de Diseño de Material Didáctico Digital para la Innovación Educativa de la Universidad Nacional de San Luis. El objetivo de dicho programa es la formulación de proyectos innovadores en el área educativa y se destinarán fondos para subsidiar la ejecución de los mismos.

1. INTRODUCCIÓN

En la actualidad, nuestra sociedad experimenta Internet a través del uso del protocolo World Wide Web, los correos electrónicos, programas que permiten compartir información, etc., sin la necesidad de conocer *cómo* es que estas aplicaciones funcionan realmente. Sin embargo, a medida que avanzamos en el estudio de las redes de datos, se vuelve cada vez más importante conocer cómo un servicio de red es capaz de transmitir e interpretar los mensajes enviados y recibidos a través de la red.

Un aspecto fundamental para todo estudiante de las Tecnologías de la Información y la Comunicación (TIC) es la posibilidad de

configurar y probar el correcto funcionamiento de los servicios más comunes utilizados en Internet. Realizar estas pruebas en un ambiente seguro, donde se pueda volver a una configuración inicial y estable es un objetivo de toda cátedra dedicada a la puesta a punto de los servicios de red.

Estos objetivos pueden ser logrados mediante el uso de técnicas de virtualización, las cuales permiten a diversas aplicaciones ejecutarse en una red virtual exactamente igual que en una red física. Otra ventaja de los entornos de trabajos virtualizados es la posibilidad de realizar diversas pruebas de configuración y ejecución de comandos o de aplicaciones pudiendo guardar el estado actual y recuperar el trabajo realizado luego de un tiempo.

Por otra parte, contar con una herramienta que se pueda personalizar agregando aplicaciones didácticas o cualquier servicio necesario durante el desarrollo de los prácticos es una mejora sustancial para el dictado de toda materia.

En base a lo mencionado anteriormente se propone la creación de NetOS-Lab (Laboratorio portátil de Redes y Sistemas Operativos), una herramienta didáctica fundamental para el desarrollo de las prácticas en las materias afines a los Servicios en SO de Redes y Sistemas Operativos de Red, buscando ayudar a los profesores y estudiantes en el proceso de enseñanza-aprendizaje fundamentalmente no presencial.

El objetivo principal de este proyecto consiste en construir una herramienta de software libre, autocontenida, personalizable y portátil que incorpore funcionalidades tales como: diseño de diversas topologías de red LAN, MAN o WAN de manera gráfica; despliegue de redes y host virtuales para ensayos de servicios de red como HTTP, DNS, SMTP, entre otros [1]; simulación de planificación de procesos y de gestión de memoria principal y secundaria en Sistemas Operativos [2]; otras funcionalidades

afines a los Sistemas Operativos cuya práctica sea difícil de llevar a cabo en ambientes no simulados.

NetOS-Lab surge como una integración de varios trabajos finales y de aplicación de alumnos de la carrera Tecnicatura Universitaria en Redes, el cual está en constante actualización buscando mejorar dicha herramienta en base a nuevas necesidades.

Los trabajos finales conforman la arquitectura de NetOS-Lab formada por 3 módulos independientes y bien diferenciados: LPR_1.0 [3], el cual hace uso de la plataforma docker [4] y el sistema operativo GNU/Linux para virtualizar servicios de red pertenecientes a la capa de aplicación pero con una topología de red fija y predefinida; EliaNS [5], una herramienta basada en kivaNS [6] que permite desplegar gráficamente diversas topologías de red; y el módulo de herramientas didácticas conformado por dos aplicaciones: una diseñada para configurar y visualizar la planificación de procesos dependiendo del algoritmo seleccionado [7] y otra basada en Proyecto2SO [8], la cual visualiza la gestión de la memoria principal y secundaria teniendo en cuenta algún algoritmo de reemplazo de páginas. Los módulos mencionados y la arquitectura de NetOS-Lab se muestran en la figura 1.

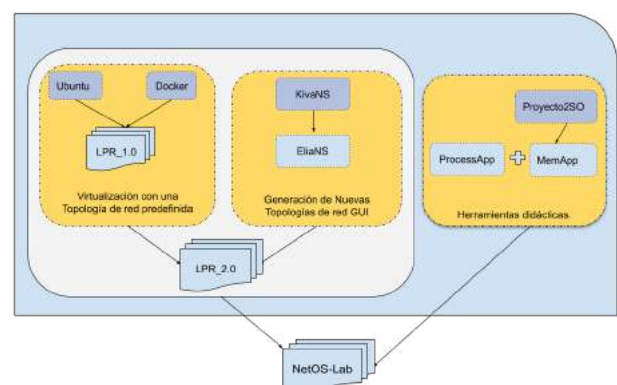


Figura 1: Arquitectura de NetOS-Lab

En la próxima sección presentamos las dos líneas diferentes de trabajo e investigación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La motivación del presente trabajo surge de las siguientes situaciones:

- Imposibilidad de acceso presencial al laboratorio de redes de la UNSL, debido a la pandemia durante el año 2020.
- Dificultad de crear más de una red física con los recursos disponibles en el hogar de los estudiantes.
- Ausencia de un software para simular redes y servicios de capa 7 al mismo tiempo.
- Alto costo económico para desplegar topologías de redes en nubes públicas.
- Necesidad de enseñar instalación y configuración de servicios en Sistema Operativos Posix.
- Dificultad en la enseñanza de conceptos teóricos, como planificación de procesos o gestión de memoria, cuya práctica es compleja por ser mecanismos ocultados por el sistema operativo a los usuarios.

En base a estas realidades surgieron dos diferentes líneas de trabajo, *Virtualización* de recursos para ensayar servicios de red y *Creación* de herramientas didácticas para la comprensión de conceptos de Sistemas Operativos. A continuación se describen algunos aspectos fundamentales de cada una de ellas:

- **Virtualización:** Cuando hablamos de virtualización hacemos referencia al proceso de reemplazar dispositivos físicos por dispositivos virtuales disponibles mediante el uso de un software. Se pueden virtualizar servidores, estaciones de trabajo, redes y aplicaciones. Para ello, el software de virtualización administra los recursos físicos de esa máquina: memoria, CPU, almacenamiento y ancho de banda de la red, entre los aspectos más relevantes. Junto con

la disminución del uso de equipos físicos, la virtualización trae como beneficios la reducción de costos de mantenimiento y consumo energético. Esto deriva una consolidación de servidores, optimizando el uso del espacio físico y posibilitando el despliegue de servicios de capa 7 sin diferencia alguna con ambientes reales. Existen varias técnicas de virtualización [9], siendo las más utilizadas: Paravirtualización, Nivel de SO, Virtualización completa, Nivel de Kernel, entre otros. En la figura 2 se pueden observar los tipos y tecnologías de virtualización y el lugar que ocupa NetOS-Lab dentro de esta clasificación.

- **Herramientas didácticas:** Las herramientas didácticas que forman parte de NetOS-Lab se centran en el uso de simuladores [10], los cuales permiten configurar y ejecutar algoritmos de planificación de procesos y administrar la memoria principal como secundaria de cualquier sistema operativo. En base a varias encuestas, y exámenes realizados a los estudiantes por parte de la cátedra, se puede observar que el uso de estos simuladores presenta mejoras en la calidad formativa del futuro profesional, optimizando su capacitación, permitiéndole poner en práctica los conocimientos teóricos, comprendiendo el funcionamiento de un sistema de interés, generando capacidad de análisis, y evaluando diferentes escenarios.



Figura 2: Tipos de virtualización

Ambas líneas de investigación están en continuo desarrollo dado que tienen como objetivo final formar un único proyecto integrador a ser usado en las prácticas del laboratorio de redes.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos y/o esperados de nuestro trabajo se visualiza en la figura 3, la cual muestra los distintos estadios del proyecto, desde los inicios hasta la versión actual.

La *figura 3 (A)* representa la primera versión del trabajo, denominado LPR 1.0. Este primer prototipo desarrollado está compuesto por el sistema operativo GNU/Linux distribución Ubuntu 18.04 y la herramienta de virtualización *docker*. LPR 1.0 permite, en base a una única topología de red, crear maquinas clientes y servidores, en las cuales se pueden configurar, iniciar, detener y pausar diversos servicios de red como así también, verificar el correcto funcionamiento de cada uno de ellos. La *figura 3 (B)* muestra la incorporación, al prototipo anterior, del módulo llamado EliaNS y la actualización del sistema operativo a la distribución Ubuntu 20.04. Estas modificaciones forman el segundo prototipo creado denominado LPR 2.0, cuyo principal aporte es la ejecución de los servicios de red mencionados anteriormente pero para cualquier topología de red creada de manera gráfica con la herramienta EliaNS. Finalmente, en la *figura 3 (C)* se detalla el trabajo que se está llevando a cabo actualmente correspondiente a la incorporación de las herramientas de simulación de procesos, de gestión de memoria y de mecanismos de automatización, conformando NetOS-Lab.

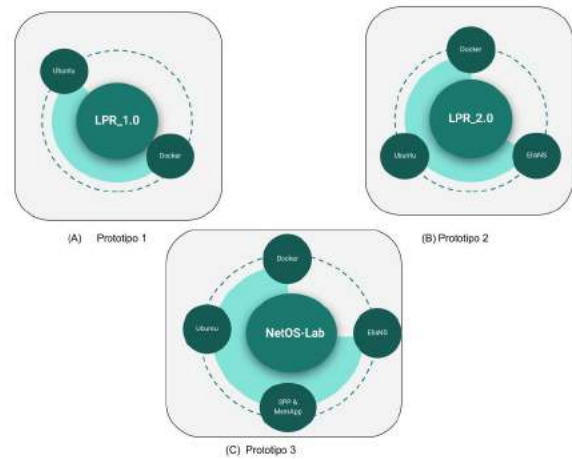


Figura 3: Evolución de NetOS-Lab.

Por lo tanto, se espera obtener dos resultados finales:

- 1) La construcción de una herramienta de *software libre, autocontenida y portátil* que integre funcionalidades como: creación de servicios de red, simulación de planificación de procesos y gestión de memoria en Sistemas Operativos, entre otras.
Al ser de software libre, la herramienta tendrá libre acceso y uso. Quedará disponible a toda la comunidad, permitiendo aporte de mejoras por parte de cualquier estudiante o desarrollador que lo desee.
Al ser autocontenida, la herramienta permitirá la ejecución de todas sus funcionalidades, sin necesidad de acceso a Internet o instalar o actualizar software alguno en la computadora donde se ejecutará.
Al ser portátil, la herramienta permitirá ser ejecutada en cualquier computadora, guardar su estado actual en un dispositivo USB y continuar trabajando más adelante en otra computadora.
- 2) Mejoras en el proceso de enseñanza / aprendizaje fundamentalmente en ambientes no presenciales.

Los resultados parciales obtenidos hasta el momento se consideran satisfactorios y es por ello que estas líneas de investigación están en constante desarrollo y actualización, teniendo en cuenta que las necesidades pedagógicas y tecnologías futuras pueden ir cambiando.

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados obtenidos respecto a la formación de recursos humanos son hasta el momento el desarrollo de tres trabajos de fin de carrera correspondientes a la carrera Tecnicatura Universitaria en Redes de Computadoras.

Con el Programa de Diseño de Material Didáctico Digital de la UNSL se espera formar recursos humanos en la orquestación y automatización de servicios para construir NetOS-Lab y sus actualizaciones.

5. BIBLIOGRAFÍA

- Fabiana, Apolloni, Rubén.
<http://sedici.unlp.edu.ar/handle/10915/26781>
- [8] Proyecto2SO. Simulación de Paginación y Segmentación. Web site: <https://github.com/randymorales/Proyecto2-SO>
- [9] Fernando Rodríguez-Haro, Felix Freitag, Leandro Navarro, A summary of virtualization techniques, Procedia Technology, Volume 3, 2012, Pages 267-272, ISSN 2212-0173.
- [10] SiSO: Un Simulador integral del Sistema Operativo. Barrionuevo Mercedes, Apolloni Rubén, Piccoli María Fabiana. 2010. <http://sedici.unlp.edu.ar/handle/10915/19640>
- [1] Kurose, James F., Ross, Keith W. Redes de computadores: un enfoque descendente basado en Internet 2da. Edición ISBN: 8478290613. 2004.
- [2] Tanenbaum, Andrew S. Sistemas operativos modernos. Pearson Educación, 2003. ISBN: 9786074420463.
- [3] Virtualización en la educación: Laboratorio Portátil de Redes. Barrionuevo Mercedes, Gil Cristian, Giribaldi Matías, Suarez Christopher, Taffernaberry Carlos. ISBN: 978-950-34-1539-9. Páginas: 465-474. XXIII Congreso Argentino de Ciencias de la Computación (La Plata). 2017
- [4] Docker Enterprise Edition. Web site: <https://www.docker.com/>
- [5] LPR_2.0: mejoras a un ambiente de aprendizaje de servicios de red. Yrigaray Elian, Barrionuevo Mercedes, Taffernaberry Carlos. ISBN: 978-987-733-196-7. Páginas: 105-114. XIV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET). 2019.
- [6] KivaNS. Web site: <http://www.aurova.ua.es/kiva/>
- [7] Una herramienta de simulación para la planificación de procesos. Barrionuevo, Mercedes, Piccoli, María

Realidad virtual, aprendizaje inmersivo y realidad aumentada: Casos de Estudio en Carreras de Ingeniería

Santiago Pérez, Ana Muñoz, Maria Eugenia Stefanoni, Daniela Carbonari
CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería)
Facultad Regional Mendoza, Universidad Tecnológica Nacional
Rodriguez 273, Mendoza, Argentina - 0261-5244576
(santiagocp, maria.stefanoni,dcarbonari)@frm.utn.edu.ar;

Ana Muñoz

Instituto Francés de Investigación en Ciencias y Tecnologías Digitales INRIA Chile
Av. Apoquindo 2827, piso 12, Las Condes, Región Metropolitana Chile - +56 2 2584 7277
ana.munoz@inria.cl

RESUMEN

La realidad virtual es una experiencia simulada, similar, o completamente diferente, al mundo real. Puede usarse en aplicaciones de entretenimiento, educativos, o industriales. En el contexto de la realidad virtual, la inmersión se produce cuando el usuario, en distintos grados de profundidad, se olvida de que está en un mundo artificial. Mientras que la realidad aumentada ofrece una percepción e interacción con el mundo real, brindando al usuario un escenario real, aumentado con información adicional generada por computación. De este modo, la realidad física se combina con elementos virtuales, disponiéndose de una realidad mixta en tiempo real. Este trabajo tiene como objeto presentar una línea de investigación en el marco de la realidad virtual, el aprendizaje inmersivo y la realidad aumentada. Para tal fin, es de gran importancia efectuar un estudio, análisis y evaluación del estado del arte y de casos de implementación de realidad virtual, niveles de inmersión y realidad aumentada, pero específicamente en los procesos de enseñanza-aprendizaje de los estudiantes universitarios de ingeniería, con énfasis en aquellos que propician un mayor nivel de inmersión y realidad aumentada. Y para contraste, se realizará un relevamiento con las

implementaciones de estos aspectos en Cátedras seleccionadas al efecto, de las Carreras de Ingeniería en Electrónica y de Ingeniería en Sistemas de Información, de la UTN Regional Mendoza. Como resultado de una compilación y estudio preliminar previo sobre la temática, se considera que hay muchas posibilidades de aplicaciones en esta línea de investigación, en ambientes universitarios, con recursos y plataformas de libre disponibilidad, especialmente en el caso de las Cátedras que hacen uso intensivo de prácticas de taller.

Los integrantes del Proyecto tienen formación de posgrado, han dirigido tesis de posgrado y/o han formado parte de grupos y proyectos de Investigación en Educación.

Palabras clave: realidad virtual, aprendizaje inmersivo, realidad aumentada,

CONTEXTO

La línea de investigación está inserta en el ámbito del Centro UTN CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería), de la Facultad Regional Mendoza, de la Universidad Tecnológica Nacional, y del Instituto Francés de Investigación en

Ciencias y Tecnologías Digitales INRIA, de Chile.

1. INTRODUCCIÓN

Fundamentos

Se conoce con el nombre de Realidad Virtual (VR, Virtual Reality) al uso de la modelación y la simulación por computadora, para que una persona pueda interactuar con un entorno visual o sensorial tridimensional artificial (3-D). Los usos avanzados de realidad virtual sumergen al usuario en un ambiente que simula la realidad mediante el uso de complejos dispositivos electrónicos, eventualmente interactivos, que envían y reciben información. Por ejemplo, un usuario que lleva un casco con una pantalla estereoscópica podría ver imágenes animadas de un entorno simulado, que captan los movimientos del usuario y ajustan la vista en la pantalla en consecuencia, generalmente, en tiempo real, o podría recorrer un conjunto simulado de habitaciones, experimentando puntos de vista y perspectivas cambiantes, que están convincentemente relacionados con sus propios giros y pasos de cabeza. El usuario podría incluso tomar y manipular objetos que ve en el entorno virtual.

Jaron Lanier (EEUU), es quien introdujo la expresión “realidad virtual” en 1987 y su investigación e ingeniería contribuyó con varias innovaciones y productos a la naciente industria de la realidad virtual.

La realidad virtual y la inmersión

“El concepto de inmersión consiste en lograr la percepción de estar físicamente presente en un mundo no físico, ya sea que se trate de la réplica de un contexto realista (como visitar una plataforma petrolera), o bien una experiencia que rompe las reglas o estructuras cotidianas

de la realidad física (volar en un entorno fantástico)” (Crespo, et al.; 2013).

Cada una de las experiencias de VR presentan distintos niveles de involucramiento dentro de esa realidad virtual, y es por ello que se identifican tres tipos de realidad virtual dependiendo del nivel de inmersión y consciencia del usuario: de baja inmersión, semiinmersiva y altamente inmersiva.

La realidad virtual con bajo nivel de inmersión es cuando se presenta un entorno virtual en tercera dimensión, a través de un monitor de computadora y la interacción de los usuarios es a través de dispositivos como el teclado, el mouse o un joystick. Mientras que es semiinmersiva, si el entorno virtual 3D se despliega en una gran pantalla, y el usuario puede visualizar el entorno a través de lentes 3D. También se suele hablar de sistemas semi-inmersivos o inmersivos de proyección caracterizados por ser 4 pantallas en forma de cubo, las cuales rodean al observador, el usuario usa lentes y un dispositivo de seguimiento de movimientos de la cabeza. Este tipo de sistemas son usados principalmente para visualizaciones, donde se requiere que el usuario se mantenga en contacto con elementos del mundo real. Finalmente, se habla de realidad virtual altamente inmersiva, si el usuario interactúa con el entorno virtual a través dispositivos especiales, que le permiten experimentar esa realidad virtual: guantes, lentes o cascos, botas, y todo tipo de sensores que permiten que el usuario perciba “estar” en esa realidad virtual.

La realidad aumentada

La Realidad Aumentada, (AR, Augmented Reality) superpone a una imagen real, obtenida a través de una pantalla, imágenes, modelos 3D u otro tipo de información generada computacionalmente.

Otros autores definen el concepto con más precisión. Por ejemplo, De Pedro (2011) explica la AR como «aquella tecnología capaz de complementar la percepción e interacción con el mundo real, brindando al usuario un escenario real aumentado con información adicional generada por ordenador. De este modo, la realidad física se combina con elementos virtuales disponiéndose de una realidad mixta en tiempo real». Se observan los conceptos de interacción, realidad mixta y tiempo real.

Esto hace a la AR una tecnología con un potencial enorme, y la convierte en una de las principales para los próximos años. Sus características principales son:

- Permite la combinación del mundo real y el mundo virtual,
- Es interactiva en tiempo real,
- Depende del contexto, y
- Utiliza las tres dimensiones.

En el mundo de la educación y la formación es probablemente donde la realidad aumentada cobra su máximo sentido. Con la AR se puede visualizar lo que un capacitador o docente está explicando, de manera directa. Por ejemplo, se podría contemplar el cuerpo humano y sus diferentes capas en una clase de anatomía, o las partes más internas de un motor de combustión.

La realidad mixta

La realidad mixta MR (Mixed Reality) es una mezcla entre la realidad virtual y la aumentada. La VR permite la inmersión en mundos completamente digitales, mientras que la AR permite gestionar objetos digitales en entornos reales.

Existen 5 principales diferencias entre ellas:

1. La VR es una realidad distinta y la AR no. La VR "traslada" al usuario a una realidad diferente de la suya. La AR se

aprovecha del entorno más cercano del usuario para ofrecerle información digital sobre el mundo real, complementando lo que percibe a través de sus sentidos.

2. La realidad virtual requiere siempre de un dispositivo. La tecnología sobre la que se basa la VR es la antigua metodología de la estereoscopia, ofreciendo así una sensación tridimensional. Se requiere un dispositivo que aisle visualmente al usuario del mundo real. La AR no necesita de un dispositivo llevado por el usuario. En ocasiones puede ser el teléfono móvil o una pantalla interactiva los que ofrezcan la posibilidad de iniciar la realidad aumentada en el lugar donde se encuentre el usuario.

3. La realidad aumentada ha encontrado usos más diversos. La AR ha encontrado aplicaciones en campos como el deporte, la medicina y la información. En el campo médico, permite por ejemplo visualizar los vasos sanguíneos, proyectándolos sobre la piel para facilitar el trabajo del personal sanitario. La AR tiene además usos educativos: WordLens reconoce mediante la cámara del móvil un texto para traducirlo al idioma requerido por el usuario.

4. La realidad virtual es ahora más compleja que la aumentada... pero la tendencia se invertirá. La VR está limitada actualmente por el dispositivo que requiere para la inmersión del usuario, que aún está en desarrollo.

La realidad virtual en la capacitación

Un área importante de aplicación para los sistemas de realidad virtual siempre ha sido la capacitación para actividades de la vida real. El atractivo de simulaciones es que pueden proporcionar una formación igual o casi igual a la práctica con sistemas reales, pero a un costo reducido y con mayor seguridad. Este fue particularmente el caso del entrenamiento militar y la primera aplicación

significativa de Simuladores fue la formación de pilotos durante la Segunda Guerra Mundial. Los simuladores de vuelo se basan en la retroalimentación visual y de movimiento, para aumentar la sensación de volar sentado en un sistema mecánico cerrado en el suelo.

Estas tecnologías son obviamente prometedoras, aunque no se puede perder de vista que es diferente la aplicación de las innovaciones tecnológicas en el entorno real de los ambientes educativos. Sin embargo, el elemento motivacional, tan importante en la educación parece garantizado, como dice Reinoso (2012): «numerosas han sido las investigaciones que sugieren que la AR refuerza el aprendizaje e incrementa la motivación por aprender».

En lo que hace a educación, nuestro interés está focalizado hacia las aplicaciones en los procesos de enseñanza-aprendizaje universitaria de ingeniería.

2. OBJETIVOS

El objetivo de la investigación es estudiar, analizar y evaluar el estado del arte de estos aspectos temáticos de virtualidad, y contrastar con la tipología de las actividades que se realizan en centros educativos universitarios, fundamentalmente de la UTN, conociendo las tecnologías virtuales que utilizan, y el software y la metodología aplicada, todo ello en orden a analizar las experiencias relevantes en este campo

3. METODOLOGÍA

A partir de la experiencia preliminar de compilación de trabajos afines, se sabe que hay una gran cantidad de ellos referidos a realidad virtual, niveles de inmersión y realidad aumentada, pero el número disminuye drásticamente cuando

se habla específicamente de su aplicación en el ambiente universitario de ingeniería. Por otro lado, hay abundante bibliografía que aún en estos casos adolece de estudios de investigación sobre la idoneidad de estas tecnologías en la educación, las metodologías que se han seguido, los resultados obtenidos de estudios cuantitativos y cualitativos, y las conclusiones basadas en fuentes primarias sobre el objeto de estudio.

Se pretende una investigación documental en la línea de Arias (2006) como «proceso basado en la búsqueda, recuperación, análisis, crítica e interpretación de datos secundarios, es decir, los obtenidos y registrados por otros investigadores en fuentes documentales: impresas, audiovisuales o electrónicas».

Posteriormente, se pretende hacer un trabajo similar, exploratorio, en algunas Cátedras Universitarias de Ingeniería, de las Carreras de Ingeniería en Electrónica y de Ingeniería en Sistemas de Información, de la Universidad Tecnológica Nacional, para evaluar el estado actual de aplicación de estas tecnologías, para ponerlas en valor respecto a otros contextos de aplicación.

4. ESTADO DE AVANCE

La realidad virtual, entendida como las aplicaciones que sumergen al usuario en un entorno que simula la realidad mediante el uso de dispositivos interactivos, que envían y reciben información, han tenido un gran desarrollo con los modernos dispositivos que hacen uso de las capacidades de cómputo actuales.

La realidad virtual inmersiva, aquella que permite que el usuario perciba estar dentro de ella sin la interrupción de los estímulos externos, ha demostrado ser una potente herramienta para las

capacitaciones laborales, especialmente en entornos riesgosos, como por ejemplo la fuerza aérea donde se entrena a los pilotos, o en medicina donde los médicos pueden practicar operaciones complicadas para poder prever posibles complicaciones, que de otro modo se verían en el momento de la cirugía.

Estas situaciones, unido a nuestro conocimiento de la aplicación de realidad virtual en niveles previos al universitario, nos orientaron para realizar una compilación y revisión preliminar bibliográfica, en busca de casos de implementación de la realidad virtual altamente inmersiva en carreras de ingeniería (Stefanoni y otros 2020). Este estudio nos ha mostrado un aspecto no esperado: no son tan numerosas las implementaciones en carreras de ingeniería, aunque los casos analizados han sido enriquecedores, por su capacidad de potenciar el proceso de enseñanza aprendizaje de los estudiantes de ingeniería (Díaz y otros 2018; Piscitelli 2017).

5. ESTADO DEL ARTE

Díaz y otros describen la Realidad Virtual en el ámbito educativo soportado por las tecnologías móviles. Indican que la tendencia va en aumento con el desarrollo tecnológico, el abaratamiento de los dispositivos visuales y la universalización del smartphone, y concluyen que su integración en la educación debe ser objeto de estudio. Esta afirmación está basada en una revisión bibliográfica, donde encuentran que desde el 2015 existe un gran interés en el tema, y clasificaron la producción científica en el área de la siguiente manera: educación primaria, secundaria, museística y universitaria. Encontraron que la mayor producción se encuentra en la etapa de educación primaria y secundaria por lo

que en la educación universitaria existe muy poca producción. Concluyen que la Realidad Virtual abre la puerta a nuevas posibilidades educativas y a un aprendizaje experiencial y significativo desde las siguientes tres características:

- Facilidad para el aprendizaje desde el constructivismo.
- Proporcionar formas alternativas de aprendizaje.
- Posibilitar la colaboración entre estudiantes más allá del espacio físico.

Mientras que Pisticelli presenta el estado del arte con el propósito de evaluar una serie de proyectos de realidad aumentada y realidad virtual en términos de la reducción de materiales utilizados y el uso de estos en sus formas más eficaces. Analiza treinta iniciativas de realidad virtual y realidad aumentada en la educación, y concluyen que en la mayor parte de los casos el valor agregado del proyecto se relaciona con el acceso a contenidos y experiencias que de otro modo estarían negados para el usuario. Sin embargo, se plantean la interrogante del valor instrumental de las soluciones: al igual que la visita a una biblioteca análoga no garantiza el involucramiento alumno-docente ni el aprendizaje significativo, la inclusión de realidad virtual o realidad aumentada en la experiencia educativa no garantiza el interés, la motivación o la comprensión del contenido, a menos que estos recursos se utilicen en un entorno educativo adecuado.

6. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por docentes investigadores, becarios graduados y alumnos del Centro UTN CeReCoN (Centro de Investigación y

Desarrollo en Computación y Neuroingeniería) de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional, y docentes-investigadores del Instituto Francés de Investigación en Ciencias y Tecnologías Digitales INRIA, Chile. Entre estos se encuentra una tesista de Maestría, que presentará su tesis de Maestría afín a este proyecto. Las actividades se llevan a cabo en el ámbito de las instalaciones de dichos entes, que cuentan con sus propias áreas de trabajo.

- Reinoso, R. (2012). "Posibilidades de la realidad aumentada en educación". En Hernández, J.; Pennesi, M.; Sobrino, D. y Vázquez, A. (Coords.) "Tendencias emergentes en educación con TIC". Barcelona: Editorial Espiral. pp.357-400.
- Stefanoni y otros (2020). Aprendizaje inmersivo y realidad virtual: Casos de estudio en ingeniería. ECEFI 2020.

7. BIBLIOGRAFÍA Y REFERENCIAS

- Arias, F. (2006). El proyecto de Investigación. Introducción a la Metodología Científica. Caracas. Editorial Episteme, C.A.
- Crespo, R.; Riestra, E.; Gánem, R.; Cárdenas, D. (2013). "Realidad Virtual como herramienta para aprendizaje inmersivo en ingeniería". Instituto Tecnológico y de Estudios Superiores de Monterrey. México. Disponible en: <https://repositorio.tec.mx/handle/11285/621360>
- De Pedro Carracedo, J.; Martínez Méndez, C. (2012). Realidad Aumentada: Una Alternativa Metodológica en la Educación Primaria Nicaragüense. Disponible en: <http://rita.det.uvigo.es/201205/uploads/IE-EE-RITA.2012.V7.N2.A9.pdf>.
- Díaz, I. A., Rodríguez, J. M. R., & García, A. M. R. (2018). La tecnología móvil de Realidad Virtual en educación: una revisión del estado de la literatura científica en España. EDMETIC, 7(1), 256-274. <https://blog.rtve.es/webdocs/page/44/>
- Piscitelli, A. (2017). Realidad virtual y realidad aumentada en la educación, una instantánea nacional e internacional. Economía creativa.

Relación entre Analítica del Aprendizaje y Diseño de Aprendizaje

Marisa Panizzi¹, Iris Sattolo¹

¹Universidad de Morón

Cabildo 134, Buenos Aires, Argentina

marisapanizzi@outlook.com, isattolo@unimoron.edu.ar

Resumen

La línea de investigación se centra en la aplicación de la Analítica del Aprendizaje asociado al Diseño de Aprendizaje en cursos de e-learning.

La Minería de Datos Educativos propone métodos para extraer información útil a partir de los datos generados en los entornos educativos. Como nuevas disciplinas surgen la Analítica del Aprendizaje y el Diseño de Aprendizaje, las que conducen a nuevas ideas sobre el comportamiento de los alumnos, interacciones y rutas de aprendizaje. La información que se puede obtener puede ser insumo para tomar decisiones sobre las prácticas educativas.

Se presentan los interrogantes para llevar a cabo la aplicación de Analítica del Aprendizaje sobre un curso desarrollado a distancia con técnicas de Diseño del Aprendizaje en el contexto de la Universidad de Morón.

Palabras clave: Minería de datos Educativa, Analítica del Aprendizaje, Diseños de Aprendizaje.

Contexto

La línea de investigación que se reporta en este artículo es financiada parcialmente por un proyecto de investigación titulado “Aplicación de Analítica del Aprendizaje sobre un curso a distancia desarrollado con técnicas de Diseño del Aprendizaje” (Código 80020190300011

UM) de la Secretaría de Ciencia y Tecnología de la Universidad de Morón.

Introducción

En nuestra institución el desarrollo de la opción pedagógico-didáctica a distancia se viene realizando a través de un esfuerzo sostenido a lo largo del tiempo, alineada con los valores institucionales y acorde a sus posibilidades reales de expansión [1]. La Universidad de Morón cuenta con un sistema de gestión del conocimiento (Learning Management System, LMS), también llamados Virtual Learning Environment (VLE) o Entornos Virtuales de Aprendizaje (EVA) [2], que ofrece carreras a distancia como también diversos cursos de capacitación. La plataforma utilizada desde sus comienzos para su gestión fue Moodle. A partir del año 2018 se incorporó como apoyo a la presencialidad la plataforma Blackboard [3] para los primeros años. A raíz del surgimiento de la pandemia generada por el SAR-COV2 se implementó para toda la comunidad (clases, administrativos, etc.). Estas plataformas generan datos, los que se pueden utilizar para mejorar dichos entornos.

La minería de datos (en inglés Data Mining o DM) es una subdisciplina de la informática que permite descubrir información nueva y potencialmente útil en grandes cantidades de datos [4]. En la última década, la DM se ha aplicado en el ámbito educativo y esto dio lugar a la aparición de la Minería de Datos Educativos (en inglés, Educational Data Mining o EDM) la cual propone métodos para extraer

información útil a partir de los datos generados en los entornos educativos [5].

La Sociedad Internacional de Minería de Datos Educativos define EDM de la siguiente manera: “La Minería de Datos Educativos es una disciplina emergente, preocupada por desarrollar métodos para explorar los tipos únicos de datos que provienen de entornos educativos, y usar esos métodos para comprender mejor a los estudiantes y los entornos que ellos aprenden” [6].

Los datos que ofrecen las plataformas de aprendizaje posibilitan la búsqueda de patrones, dando lugar a una analítica del aprendizaje que supone un paso más allá de la recolección, ya que sirve para analizar e interpretar esos datos. Se usan técnicas pedagógicas y algoritmos propios de la Minería de Datos, con la finalidad de obtener información para mejorar la práctica educativa, optimizando el rendimiento de alumnos y profesores, así como el modelo educativo.

En los últimos diez años, se gestó la Analítica del Aprendizaje como un campo separado. Se debe a la tendencia creciente en la digitalización en el campo de la educación, la aparición de entornos de aprendizaje distribuidos y la mayor participación en las experiencias de aprendizaje en línea [7].

La Analítica del Aprendizaje, o Learning Analytics (LA), es la denominación de esta reciente disciplina con potencial transformador (relacionado con el aprendizaje personalizado y adaptativo) y con incidencia en todas las disciplinas educativas, particularmente en e-Learning [8]. Esta disciplina ha evolucionado y persigue los siguientes objetivos:

- explicar comportamientos de aprendizaje inesperados,
- identificar patrones de aprendizaje exitosos,
- detectar conceptos erróneos y esfuerzos fuera de lugar,
- introducir intervenciones apropiadas y
- aumentar la conciencia de los usuarios

sobre sus propias acciones y progreso [7].

Para que los sistemas de análisis de aprendizaje orientados al estudiante funcionen de manera efectiva, deben integrarse con enfoques pedagógicos y con el diseño de aprendizaje. El Diseño de Aprendizaje es el proceso en el cual se realizan las siguientes actividades:

- 1 definir las tareas,
- 2 proporcionar los contextos y recursos para realizar dichas tareas,
- 3 apoyar al alumno durante el desempeño de la tarea y
- 4 proporcionar retroalimentación sobre los resultados.

Estas son las actividades que se deben llevar a cabo en cualquier proceso de aprendizaje.

A esta secuencia de actividades, se le llama “Diseño de Aprendizaje” [8]. La investigación se ha centrado en “conceptualizar los principios del diseño del aprendizaje, sin evaluar lo que sucede después del proceso de diseño” [9].

En la literatura existente se observa que la Analítica del Aprendizaje es un campo interdisciplinario que abarca métodos y enfoques de diversas disciplinas, pero carece de un modelo consolidado para sistematizar cómo se fusionan esas disciplinas. Además, se evidencia la ausencia de investigaciones para medir qué decisiones de Diseño de Aprendizaje estimulan entornos de aprendizaje productivos y cuáles generan conocimientos de diseño procesables [10].

Líneas de Investigación, Desarrollo e Innovación

Dentro del campo de la Minería de datos educativos, Analítica del Aprendizaje, y Diseños de Aprendizaje, orientamos los esfuerzos en responder a estas preguntas de investigación:

PI 1: ¿Cuál fue el impacto originado al abordar el e-learning en el año 2020 en el dictado de las clases de los docentes de la Universidad de Morón?

PI 2: ¿Qué se debe tener en cuenta al realizar el diseño de un curso a dictarse a distancia, para que éste pueda ofrecer datos valiosos recuperables a través de los LMS?

PI 3: ¿Qué se debe tener en cuenta cuando utilizamos datos de los estudiantes para que no se viole la privacidad del usuario?

PI 4: Una vez realizado el análisis de los datos, luego de aplicar LA, los resultados obtenidos, ¿contribuyen a los resultados esperados?

Para dar respuesta a las preguntas de investigación planteadas nos proponemos el siguiente plan de trabajo:

- 1) Elaboración del estado de arte de la Minería de Datos Educacional y Analítica del Aprendizaje.
- 2) Elaboración del estado del arte del Diseño de Aprendizaje aplicado al e-Learning.
- 3) Construir un instrumento de recolección de datos.
- 4) Analizar el conjunto de datos que proveen las plataformas existentes en nuestra institución.
- 5) Proporcionar una metodología para diseñar un proyecto de curso virtual, evaluarlo y seguir su realización.
- 6) Aplicar técnicas de EDM y LA.
- 7) Aplicación de los resultados del punto 5 al curso elegido.
- 8) Puesta en marcha del curso y evaluación.

La construcción del estado de arte sobre Minería de Datos Educacional, y Analítica del Aprendizaje y Diseño de Aprendizaje aplicado al e-Learning (puntos 1 y 2 del plan de trabajo) se desarrollará mediante el método de investigación, mapeo sistemático de la literatura (en inglés, Systematic Mapped Study) [11].

Para el diseño del instrumento de recolección de datos (punto 3 del plan de trabajo) se

aplicará las directrices del proceso de encuesta propuesto en [12].

Para el proceso de minería de datos (punto 6 del plan de trabajo) se aplicará el proceso KDD (en inglés Knowledge Discovery in Databases) “Descubrimiento de conocimiento en bases de datos” [4].

Objetivos y Resultados Esperados

En esta línea de investigación en progreso se han logrado una serie de **resultados** que se detallan a continuación:

- a) Proyectos de Investigación: ha finalizado el PIO/17-01-MP-001 titulado: “Aplicación de tecnologías inteligentes de explotación de información para el análisis de perfiles de tesis de grado de carreras informáticas de la UM” (período 2017-2019). Proyecto en ejecución titulado: “Aplicación de Analítica del Aprendizaje sobre un curso a distancia desarrollado con técnicas de Diseño del Aprendizaje” (Código 80020190300011 UM) (período 2020-2022).
- b) Resultados Académicos, se ha logrado una tesis de grado.
- c) Producción Científica: se ha presentado la línea de investigación en eventos científicos de alcance nacional (TE&ET¹ ediciones 2018 y 2019, CACIC² ediciones 2018 y 2019 y en el ámbito internacional, JENUI³ 2018 e InNGENIO⁴ 2019). Además, se logró una publicación en Springer.
- d) Formación en investigación: el grupo de investigación se encuentra en un proceso de aprendizaje constante de métodos de investigación de ingeniería de software

1 Congreso de “Tecnología en Educación” y “Educación en Tecnología” (TE&ET).

2 Congreso Argentino de Ciencias de la Computación (CACIC).

3 Jornadas sobre la Enseñanza Universitaria de la Informática (JENUI 2018). Link: <http://jenui2018.uoc.edu/>

4 Congreso Latinoamericano de Ingeniería. Link: <http://fundacioniai.org/ingenio/>

experimental, revisiones sistemáticas [11] y encuestas [12], Además, se experimenta con procesos, técnicas y algoritmos de DM.

El grupo de investigación se plantea como **próximos objetivos**:

Analizar los resultados del SMS con el propósito de identificar las contribuciones realizadas sobre Analítica del Aprendizaje y Diseños de Aprendizaje aplicadas a cursos de e-Learning en la educación superior.

Avanzar con el diseño de la encuesta para la recolección de los datos que dará respuesta a nuestra pregunta de investigación: *¿Cuál fue el impacto originado al abordar el e-learning en el año 2020 en el dictado de las clases de los docentes de la Universidad de Morón?*

Presentar los avances realizados en InNGENIO 2021, TEyET 2021 y CACIC 2021.

Formación de Recursos Humanos

El grupo se encuentra conformado por un Director, un Co-Director, dos tesistas de grado, y dos docentes-investigadores en proceso de formación.

Se estiman dos tesis de la carrera Licenciatura en Sistemas en el marco de la línea de investigación.

Bibliografía

- [1] <https://www.unimoron.edu.ar/area/distancia/stream/a94733556-estudia-a-distancia>
- [2] Belloc Consuelo. *Entornos virtuales de Aprendizaje*
http://www.formaciondocente.com.mx/04_RinconTecnologia/03_AmbientesVirtuales/Entornos%20Virtuales%20de%20Aprendizaje%203.pdf
- [3] <https://www.blackboard.com/>
- [4] Hernández Orallo José, Ramírez Quintana María José, Ferri Ramírez César. (2014). *Introducción a la Minería de Datos*. Ed.

Pearson Educación S.A. Madrid.

[5] A Peña Ayala. Educational Data Mining. A Survey and a data mining-based analysis of recent works. *Expert Systems with applications*, vol. 41, nro 4, Mar., pp. 1432-1446, 2014.

[6] Siemens y Baker *Learning Analytics and Educational Data Mining: Towards Communication and Collaboration*.

[7] Jones, K. M. L. (2019). *Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy*. International Journal of Educational Technology in Higher Education, 16(1) doi:10.1186/s41239-019-0155-0.

[8] Koper R. & Bennett, S. (2008). Learning Design: Concepts. In H. Adelsberger, J. M. Pawlowski, Kinshuk & D. Sampson (Eds.), *Handbook on Information Technologies for Education and Training*. Heidelberg: Springer.

[9] Lockyer, L., Heathcote, E., Dawson, S. (2013). *Informing pedagogical action: aligning learning analytics with learning design*. Am. Behav. Sci. 57(10), 1439–1459

[10] Berland, M., Davis, D., Smith, C.P. (2015). AMOEBA: *Designing for collaboration in computer science classrooms through live learning analytics*. Int. J. Comput. Supported Collaborative Learn. 10(4), 425–447.

[11] Kitchenham, B. y Chartes, S. (2007). *Guidelines for performing systematic literature reviews in Software engineering*, Keele University, EBSE-2007-01.

[12] Jefferson Seide Molléri, Kai Petersen, Emilia Mendes (2020). *An empirically evaluated checklist for surveys in software engineering*. Information and Software Technology 119.

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

Martín LOBOS¹, Silvia BAST¹, Gustavo ASTUDILLO¹,
¹ Departamento de Matemática/FCEyN/UNLPam
lobmar146@gmail.com, silviabast@exactas.unlpam.edu.ar,
astudillo@exactas.unlpam.edu.ar

RESUMEN

Como parte del trabajo del proyecto de investigación “Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática” en los años 2019 y 2020, nos hemos propuesto identificar, categorizar y evaluar distintos Entornos de Desarrollo Integrado (EDI) para la enseñanza de la robótica educativa y el aprendizaje de las nociones básicas de programación, con el objetivo final, de seleccionar el que mejor se adapte a las exigencias establecidas por el cuerpo docente que se desempeña en el Taller de Introducción a la Programación (TIP). El mencionado taller se lleva a cabo todos los años en el marco de los cursos de ambientación para ingresantes

Para lograr tal objetivo, se llevó adelante la búsqueda de los EDI en distintos repositorios, para someterlos posteriormente a un proceso de medición y evaluación de calidad, haciendo uso de la estrategia denominada *Goal-Oriented Context-Aware Measurement and Evaluation* (GOCAME) y de estándares internacionales ISO 25010:2011.

Palabras clave: entorno de desarrollo integrado, medición, evaluación, robótica educativa, programación

CONTEXTO

El grupo de investigación GrIDIE¹ (Grupo de Investigación y Desarrollo en Innovación

Educativa) enfoca, desde 2005, su investigación en tecnologías informáticas aplicadas en educación. Desde 2018, el grupo impulsa dos proyectos de investigación: “Aprendizaje de las ciencias con tecnologías educativas” e “Incorporación de estrategias innovadoras en los procesos de enseñanza y de aprendizajes de informática”. Este último cuenta con cuatro líneas de investigación: (i) la medición y evaluación de EDIs para robótica educativa, (ii) la definición de criterios de evaluación que permitan identificar las posibilidades de la programación tangible para el aprendizaje de nociones básicas de programación, (iii) la revisión/evaluación de propuestas didácticas para la enseñanza/aprendizaje de la programación en función de su enfoque pedagógico/didáctico, y (iv) la implementación de una propuesta didáctica para el aprendizaje de la programación en el ámbito Universitario.

Los proyectos se desarrollan y son financiados por la Facultad de Ciencias Exactas y Naturales - UNLPam y fueron aprobados por resolución 27/18 CD-FCEyN.

1. INTRODUCCIÓN

En este trabajo presentamos los avances realizados en 2019 y 2020 en la línea de investigación de medición y evaluación de EDIs para robótica educativa.

El Taller de Introducción a la Programación (TIP)

¹ Facultad de Ciencias Exactas y Naturales
UNLPam

La asignatura Introducción a la Computación (IC), es, en la mayoría de los casos, el primer contacto de los estudiantes con los conceptos de programación. Por tal razón, los integrantes de la cátedra IC en colaboración con el grupo de investigación GrIDIE, vienen impulsando, desde 2010, distintas propuestas para ingresantes y estudiantes que cursarán la asignatura.

Las actividades incluidas en el TIP fueron diseñadas para desarrollarse durante el período de ambientación universitaria que se lleva a cabo, las semanas previas al inicio del primer cuatrimestre, de forma semipresencial.

El TIP se desarrolla sobre el entorno virtual de enseñanza aprendizaje (EVEA) Moodle y presenta una impronta lúdica (Deterding, Khaled, Nacke, & Dixon, 2011), Kapp (2012), Astudillo et al, 2016). Sus actividades se estructuran por niveles, de forma tal que, los desafíos y propuestas del nivel siguiente se ponen visibles una vez que el estudiante ha completado el nivel anterior.

La primera semana se desarrolla de forma virtual asincrónica, focalizando en la construcción del concepto de algoritmo y secuencia. En la etapa presencial se trabaja haciendo uso de un EDI de robótica, para fortalecer el concepto de secuencia, e incorporar los conceptos de alternativa condicional, repetición y nociones de variables. El objetivo principal del taller es que los estudiantes logren apropiarse de los principales conceptos haciendo uso de la programación en bloques, para luego poder transferir los mismos al contexto de la cursada donde se usa el lenguaje de programación Pascal.

La secuencia didáctica que se aplica en el TIP

La propuesta diseñada desde el grupo de investigación hace uso de la robótica educativa y se sustenta en los principios del construccionismo (Papert, 1990), el buen aprendizaje (Pozo, 2008) y aprendizaje basado en problemas (Barrows, 1986).

Para la especificación de la secuencia fue necesaria la definición de kit (sensores y actuadores), la selección de un EDI, un simulador (para tareas extra-clase) y de un conjunto de ejercicios y problemas que hacen uso de los mismos. La secuencia abarca los conceptos de estructuras de control (secuencia, repetición y selección) y la noción de variable. Asimismo, se desarrollaron un conjunto de *concept cards* que complementan la secuencia en busca de la transferencia de los aprendizajes a partir del planteo de problemas.

Robótica Educativa

La presente investigación comienza a partir de la decisión de incluir robótica educativa en el TIP, la cual, fue tomada teniendo en cuenta que los robots resultan la conexión ideal entre la programación con una impronta lúdica y la representación de las instrucciones en un contexto real. Tal como afirma Vaillant (2013) “La robótica forma parte de un enfoque pedagógico centrado en el alumno, que le permite construir objetos tangibles de su propio diseño y con sentido para él.” (p. 38). Para Monsalves González (2011) y Ruiz-Velasco (2007) se trata de una disciplina que tiene como objetivo generar entornos de aprendizaje heurístico poniendo el foco en la participación activa, donde los aprendizajes se construyen a partir de la experiencia del estudiante durante el proceso de armado y programación de los robots.

El primer EDI de robótica que se aplicó en la etapa presencial del TIP fue Visualino², junto con un simulador on-line denominado TinKerCad³. Si bien Visualino permitía el desarrollo de las actividades del taller, presentaba algunas debilidades tales como: condiciona la propuesta didáctica al grupo de sensores y actuadores que propone, falta de inclusión de extensiones, y la forma en la que construye el código fuente.

Para subsanar esas desventajas se realizó una búsqueda que se fue ampliando y definió la pregunta que guía la investigación ¿Qué EDIs, para robótica educativa, cuentan con la

² Visualino (<http://www.visualino.net/index.es.html>)

³ TinKerCad (<https://www.tinkercad.com/>)

mayor flexibilidad para adaptarse a la propuesta didáctica?

De la búsqueda de EDIs de robótica que se ajustaran de la mejor manera a la secuencia didáctica propuesta, surgió una amplia oferta de entornos basados en Google Blockly que se presentaron como posibles candidatos. Frente a esta oferta, se comenzó el análisis para seleccionar el EDI más adecuado a través de metodologías de evaluación de calidad para productos de software.

El proceso de búsqueda, medición y evaluación

La búsqueda y los EDI seleccionados.

La búsqueda de los EDI comenzó en repositorios tales como Github en proyectos derivados, similares o relacionados a los entornos ya conocidos (Visualino y Mixly). Para la selección de EDIs a evaluar se tuvo en cuenta como criterios iniciales la posibilidad de trabajar con placas Arduino UNO ya que es el Hardware utilizado para el taller. También, la utilización de Google Blockly para la programación en bloques por la familiaridad con las actividades ya desarrolladas para el TIP.

En total se encontraron diecisiete (17) EDIs, pero teniendo en cuenta los criterios mencionados anteriormente se seleccionaron nueve (9) entornos candidatos para su evaluación:

1. Mixly
2. Visualino
3. Ardublockly
4. BlocklyDuino
5. Blockly@rduino
6. Blockly4Arduino
7. ArduinoBlocks
8. Facilino (Offline)
9. Facilino (Online)

Evaluación de Calidad de los EDIs

Dado que los EDIs son desarrollos de software, debe usarse una metodología de evaluación de calidad para tales productos, aplicando criterios establecidos por estándares

internacionales, tal como la ISO 25010:2011 (International & Organization for Standardization, 2011).

Según David Garvin (Garvin, 1984) la calidad es un concepto complejo y multifacético. Tal como afirman (Pfleeger & Kitchenham, 1996) para poder medir la calidad habitualmente se construyen modelos que relacionan las características de la calidad, descomponiendo cada una de ellas en diversos factores o relacionándolas de forma jerárquica.

En este caso, el modelo de calidad está representado por una estructura jerárquica que define un conjunto de características y las relaciones entre ellas y proporciona las bases para especificar los requerimientos y evaluarlos.

Se aplica el proceso de medición que propone la estrategia denominada Goal-Oriented Context-Aware Measurement and Evaluation (GOCAME) (Olsina et al., 2008) que usa el marco conceptual denominado Contextual-Information Need, Concept model, Attribute, Metric and Indicator (C-INCAMI), que sigue el proceso que se especifica a continuación.

Para comenzar, se establece la necesidad de información, y a partir de la misma se definen los requerimientos, generando el modelo de calidad que tomará sus bases del estándar ISO 25010:2011 (International & Organization for Standardization, 2011), al que se agregan un conjunto de subconceptos y atributos que surgen de los requerimientos, establecidos por el grupo de investigación. Seguidamente, para cada uno de los atributos del modelo se definen las métricas, que se aplican en el proceso de medición.

Luego se lleva a cabo el proceso de medición aplicando las métricas definidas para cada atributo.

Posteriormente, se definen los indicadores, y se aplican en el proceso de evaluación. Los mismos ofrecen información contextualizada para la toma de decisiones.

Finalmente, se analizan los resultados y se realizan las recomendaciones de acuerdo a la necesidad de información establecida al inicio del proceso.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Este proyecto se desarrolla bajo la hipótesis de que es posible definir estrategias innovadoras para la enseñanza y el aprendizaje de temas de Ciencias de la Computación, utilizando TIC, que impacten positivamente en la motivación de los estudiantes y en los diseños didácticos de los docentes.

Esto implica, por una parte, indagar y evaluar distintas tecnologías que permitan innovar y que faciliten el diseño de actividades y recursos educativos; y por otro, identificar las bases pedagógico-didácticas que sustenten y den sentido a la innovación.

Se trabaja actualmente en la definición de criterios de evaluación que permitan identificar las posibilidades de la programación tangible para el aprendizaje de nociones básicas de programación, la revisión desde propuestas didácticas para enseñanza de la programación con el fin de identificar las bases pedagógicas que las sustentan, la implementación y evaluación de una propuesta de enseñanza de la programación basada en la didáctica por indagación y la resolución de problemas y tal como se detalla en el presente artículo en un proceso de medición y evaluación de la calidad de EDIs que permitan la implementación de actividades con robótica educativa.

3. RESULTADOS OBTENIDOS/ESPERADOS

Durante los años 2019-2020 se realizaron los siguientes avances:

El relevamiento de 17 EDIs candidatos que aplican para el aprendizaje de la programación y robótica educativa.

El análisis de los EDIs candidatos y la selección de 9 para su evaluación.

Se definió un modelo de calidad y se desarrollaron las métricas que se aplicarán en el proceso de medición.

En cuanto a los trabajos futuros:

Se espera completar las mediciones, definir los indicadores, realizar el proceso de evaluación y finalmente seleccionar el EDI más acorde a la secuencia didáctica.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan un investigador formado y tres investigadores en formación. Dos de ellos avanzando sobre sus tesis de doctorado y un tercero en su tesis para alcanzar el grado de magíster.

El proyecto cuenta con tres graduados que se inician en la investigación. Uno de ellos definiendo su proyecto de trabajo final de especialización, para ser presentado para su aprobación en el primer trimestre de 2021. Otro de los graduados se encuentra desarrollando su trabajo en el marco de una beca Estímulo a las Vocaciones Científicas-Convocatoria 2019, cuyo desarrollo retroalimenta el presente artículo.

5. BIBLIOGRAFÍA

Astudillo, G. J., Bast, S. G., & Willging, P. A. (2016). *Enfoque basado en gamificación para el aprendizaje de un lenguaje de programación. Virtualidad, Educación y Ciencia*, 7(12), 125–142.

Barrows, H. S. (1986). *A taxonomy of problem-based learning methods*. *Medical education*, 20(6), 481-486.

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011). *Gamification: Toward a definition*. En *CHI 2011 Gamification Workshop Proceedings*. Vancouver, BC, Canada.

Garvin, D.A. (1984), *What Does 'Product Quality' Really Mean?* *Sloan Management Review*, 26 (1), pp. 25-43

International Standar Organization. (2011). *ISO/IEC 25010 Systems and software engineering. Systems and software Quality*

Requirements and Evaluation (SQuaRE) System and software quality models.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Kitchenham, B., & Pfleeger, S. (1996). *Software Quality: The Elusive Target*. *IEEE Softw.*, 13, 12-21.

Monsalves González, S. (2011). *Estudio sobre la utilidad de la robótica educativa desde la perspectiva del docente*. *Revista de Pedagogía*, 32(90), 81-117.

Papert, S. (1990). *A critique of technocentrism in thinking about the school of the future*. *Epistemology and Learning Group*, MIT Media Laboratory. Recuperado de <http://cursa.ihmc.us/rid=1N5PWKQGN-H47QOR-37ZW/Papert%20critique%20of%20technocentrism.pdf>

Pozo, J. I. (2008). Capítulo 4. *Los rasgos de un buen aprendizaje*. En *Aprendices y maestros: la psicología cognitiva del aprendizaje* (pp. 159-175). Madrid, España: Alianza.

Pozo, J. I., & Pérez Echeverría, M. del P. (1994). *Aprender a resolver problemas y resolver problemas para aprender*. En *La solución de problemas*. Santillana Madrid.

Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon*, 9(5).

Ruiz-Velasco, E. (2007). *Educatrónica: Innovación en el aprendizaje de las ciencias y la tecnología*. Madrid: Díaz de Santos.

Vaillant, D. (2013). *Integración de TIC en los sistemas de formación docente inicial y continua para la Educación Básica en América Latina*. Argentina: UNICEF Argentina. Recuperado de https://www.unicef.org/argentina/spanish/educacion_Integracion_TIC_sistemas_formacion_docente.pdf

Olsina, L., Papa, F., Molina, H. (2011). *Modelling and Implementing Web Applications*, Rossi G., Pastor O., Schwabe D. & Olsina L.—Cap. 13 - *How to Measure and Evaluate Web Applications in a Consistent Way*. Springer -Verlag HCIS, London, 2008, 385 -420

Tecnologías Digitales y Colaborativas en Educación

**Adair Martins, Carina Fracchia, Claudia Allan, Susana Parra,
Natalia Baeza, Carolina Celeste, Nahuel Mamani, Kevin Isaias Pascual,
Ana Alonso de Armiño, Betina Molero, Norberto Larrosa, Roberto Laurent**

Departamento de Computación Aplicada / Facultad de Informática
Departamento de Electrotecnia / Facultad de Ingeniería
Universidad Nacional del Comahue

Dirección: Buenos Aires 1400, 8300 – Neuquén

Teléfono: 0299 - 4490300 int. 429

e-mails: {adair.martins, carina.fracchia, claudia.allan, susana.parra}@fi.uncoma.edu.ar,
{baeza.natalia, celeste.carolina.s, anaalonso}@gmail.com,
{nahuel.mamani, kevin.isaia, norberto.larrosa}@est.fi.uncoma.edu.ar,
{betinamolero}@hotmail.com,
{roberto_laurent}@yahoo.com.ar

Resumen

La constante evolución de las nuevas Tecnologías de la Información y Comunicación (TIC) en las últimas décadas, ha impulsado en forma significativa la utilización de herramientas y recursos digitales en la resolución de problemas de distintos campos del conocimiento. En el ámbito educativo facilita la modificación de las prácticas docentes propiciando contextos de aprendizajes más ricos e interactivos. Esto nos ha motivado a continuar avanzando en el desarrollo de herramientas basadas en software libre, además de la utilización de aplicaciones de Realidad Aumentada (RA) y Realidad Virtual (RV). El objetivo principal es fortalecer los procesos colaborativos de enseñanza y aprendizaje en los distintos ámbitos, tratando de lograr el acercamiento de la universidad al medio. En este trabajo se presentan los avances alcanzados hasta el momento en las principales líneas de investigación.

Palabras Claves: Tecnologías digitales, Realidad Aumentada y Realidad Virtual, Juegos educativos, Aplicación web, Grafos.

Contexto

Las líneas de investigación pertenecen al Proyecto de Investigación 04/F016: “Computación Aplicada a las Ciencias y Educación” de la Facultad de Informática (FAIF), Universidad Nacional del Comahue (UNCo). Participan docentes y estudiantes avanzados de la carrera de Licenciatura en Ciencias de la Computación (FAIF), docentes de la Facultad de Ingeniería (FI) e investigadores de la Universidad Católica de Brasilia (UCB), Brasil.

1. Introducción

El grupo de investigación viene desde proyectos anteriores trabajando en el desarrollo y uso de herramientas digitales. Particularmente el año 2020 mostró la importancia de tener docentes formados en el uso de TIC, además de contar con infraestructura tecnológica adecuada para la formación en estos nuevos escenarios digitales. En la actualidad se cuenta con un

gran número de herramientas y recursos educativos abiertos, que han posibilitado enriquecer los materiales que los docentes elaboran o seleccionan para llevar adelante su práctica pedagógica. En este sentido las capacitaciones realizadas a docentes del nivel secundario de diferentes establecimientos educativos localizados en las provincias de Neuquén y Río Negro, mostraron el potencial que ofrecen los materiales educativos lúdicos para favorecer la motivación, colaboración y participación de estudiantes en la construcción de sus aprendizajes. Se presentó la experiencia y resultados obtenidos al trabajar con competencias de juegos educativos en el XV Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET) realizado en la Facultad de Informática de la Universidad Nacional del Comahue entre los días 6 y 7 de julio de 2020, Neuquén [1].

Adicionalmente, siguiendo con la línea de investigación “Desarrollo y Uso de Recursos TIC” se propuso el desarrollo de una herramienta web interactiva, que favorezca la experimentación, para ser utilizada en la enseñanza y aprendizaje de algoritmos sobre grafos, principalmente para estudiantes de la licenciatura en Ciencias de la Computación. La teoría de grafos y sus innumerables aplicaciones forman parte hoy en día de diversos estudios de matemática discreta, informática, investigación operativa y de diversas especialidades de ingeniería. Los algoritmos de grafos resultan especialmente útiles para hallar soluciones a numerosos problemas de la vida real. Los grafos y sus algoritmos son de naturaleza visual y dinámica, motivo por el cual un entorno gráfico y dinámico representa una gran ayuda en la comprensión del funcionamiento de estos algoritmos [2]. Se ha realizado un breve análisis sobre las características de algunas herramientas para grafos, tales como: IAGraph [3], VisuAlgo [4], Data Structure Visualizations [5] y Algorithm Visualizer [6] entre otras. Si bien

todas permiten la visualización de grafos y la aplicación de algoritmos sobre estos, presentan diferencias respecto a la forma del ingreso de los datos de entrada. IAGraph y VisuAlgo disponen de un editor visual de grafos, lo que facilita un modelado más intuitivo y fácil. Sin embargo, algunas requieren de instalación previa de ciertos programas o paquetes para su funcionamiento. Es recomendable que todo software educativo sea de fácil acceso y puesta en marcha, como así también utilizable en distintas plataformas. Por ejemplo VisuAlgo, Algorithm Visualizer y Data Structure Visualizations cumplen con tales requerimientos, dado que son aplicaciones web que solo requieren un navegador web para funcionar. Estas, además se destacan por permitir la visualización de la ejecución de los algoritmos y sus resultados. Esto facilita la comprensión de su funcionamiento y el análisis de los valores obtenidos. Por otro lado, Algorithm Visualizer a diferencia de las otras herramientas, no limita al usuario a experimentar con un conjunto limitado de algoritmos, ya que posee un editor de código fuente que permite al usuario escribir sus propios algoritmos. Para el desarrollo e implementación de la herramienta propuesta se utilizó una arquitectura cliente-servidor y diferentes tecnologías de código libre. Para mostrar sus funcionalidades se presenta un caso de aplicación del algoritmo de PageRank [7]. Este estudio dio origen al trabajo publicado en el XV Congreso TE&ET 2020 [8].

Por otra parte, en la línea de investigación de “Realidad Aumentada y Realidad Virtual” se continúa con el uso y desarrollo de recursos digitales más visuales e interactivos aplicados a distintas áreas. La Psicoprofilaxis Quirúrgica es una técnica psicoterapéutica que tiene como finalidad, preparar emocionalmente al paciente y su familia para afrontar una intervención quirúrgica con los recursos adecuados para evitar que se convierta en una experiencia traumática. Las

tareas que se realizan abarcan desde ofrecer información y contención, además de favorecer el apoyo entre los pacientes y el equipo de salud [9]. En el área de pediatría del Hospital Castro Rendón de la ciudad de Neuquén, desde la década del 80, en forma semanal e ininterrumpida, esta tarea está a cargo de un grupo conformado por personal del hospital mencionado y voluntarios conocidos como “Cuentacuentos”, los cuales a través de la narración de un cuento, que incluye los procedimientos, algunos dolorosos, que debe soportar el pequeño paciente, se les trata de brindar una aproximación de lo que sucederá, lo que ayuda a liberar tensiones y angustias. Estos encuentros son posibles sólo para los pacientes que pueden movilizarse hacia el hospital, pero para los casos donde estos no pueden concurrir, la situación para ellos se torna más traumática. En función de esto, se propuso la elaboración de un cuento aumentado que incorpore las tecnologías informáticas RA y RV para poder acercar el hospital a la comunidad.

2. Líneas de Investigación y Desarrollo

Se está trabajando en las siguientes líneas de investigación: “Uso y Desarrollo de Recursos TIC”, “Realidad Aumentada y Realidad Virtual”. Estas están interrelacionadas, persiguiendo como propósito general acercar la Universidad al medio.

3. Resultados Obtenidos/Esperados

A continuación se mencionan los avances de los resultados obtenidos hasta el momento. En la línea “Uso y Desarrollo de Recursos TIC”. Con participación de docentes del departamento de Programación y del departamento de Computación Aplicada de la FAIF se desarrolló un taller denominado

“Herramientas para la producción de recursos digitales” destinado a la capacitación de docentes del nivel medio para la creación de materiales educativos lúdicos. Para estructurar la capacitación se trabajó con una propuesta metodológica en la cual se combinaron instancias de trabajo colaborativo y competitivo. Inicialmente se realizaron encuestas a los docentes para determinar el nivel de conocimiento y manejo de las herramientas TIC, principalmente respecto a la utilización de herramientas de edición de imagen, y de herramientas que soportan colaboración. En la capacitación se trabajó con tecnologías emergentes RA y RV, códigos QR, repositorios libres, derechos de autor, entre otras, y las herramientas de software libre para diseño Gimp e Inkscape. Para el diseño del juego se especificó para cada institución el uso por lo menos de 5 áreas temáticas. Se dispuso también que debería haber un mínimo de 15 preguntas contemplando diferentes niveles de complejidad. La condición inicial fue que las preguntas sean desarrolladas por los alumnos, además deberían ser verificadas y validadas por sus respectivos docentes de acuerdo a los temas definidos. Se utilizó la generación de códigos QR en las distintas partes del juego, y en las respuestas válidas. Los docentes participaron en una instancia de competencia desarrollada en el marco del evento PROGRAMATE organizado por la Facultad de Informática en Octubre del año 2019 y fue dirigido a estudiantes del nivel medio. En la competencia se organizó el tiempo de entrega del juego, sus características, diseño y estructura. Las premiaciones consideraron el diseño más motivador, mejor producción y mejor diseño educativo. La elección del diseño más motivador se basó en las estrategias y reglas que resultaran más atractivas para los estudiantes. En el caso de mejor producción se consideró la calidad del diseño gráfico, y por último, respecto al mejor diseño educativo se tuvo en cuenta al juego más didáctico y que incorporara más conocimiento de las áreas involucradas.

Como conclusión la experiencia fue muy positiva, permitiendo a los docentes incorporar los conocimientos de las tecnologías informáticas y a los estudiantes fortalecer el conocimiento de los contenidos en las distintas disciplinas y áreas. Dio origen al trabajo publicado en [1,12].

Continuando con la línea “Uso y Desarrollo de Recursos TIC”, se presentó una propuesta para el desarrollo de una herramienta web para ser utilizada en el aprendizaje de grafos, principalmente para estudiantes de las carreras de Ciencias de la Computación. Este software ha sido diseñado teniendo en cuenta la comparativa realizada sobre las herramientas mencionadas anteriormente. Consiste en una aplicación web basada en una arquitectura cliente-servidor, compuesta de un editor visual de grafos y de un editor de código fuente con el que es posible escribir algoritmos que se ejecutarán sobre el grafo modelado, y cuyos resultados pueden ser visualizados sobre el mismo. Una particularidad de esta herramienta es que otorga al usuario la libertad de elegir el lenguaje de programación con el cual escribir sus algoritmos e incluye algunos ya implementados, listos para su ejecución. La finalidad de la herramienta propuesta es apoyar y fortalecer el aprendizaje del funcionamiento de los algoritmos sobre grafos, y está dirigida principalmente a estudiantes de las carreras en de Ciencias de la Computación. Para mostrar sus funcionalidades se presentó un caso de aplicación del algoritmo de Pagerank [7]. Algoritmo creado por Google para establecer un ranking de importancia de las páginas web en base a la estructura de grafo de la web. Este estudio dio origen al trabajo presentado en el Congreso TE&ET 2020 [8].

Combinar las ventajas asociadas a los aprendizajes cooperativo y competitivo es todo un desafío, que también puede ser beneficioso en la enseñanza y aprendizaje de la programación. En este contexto de Covid-19, los torneos de programación surgen como

un recurso didáctico capaz de promover una motivación positiva y mejorar el rendimiento académico en el área de programación. Un estudio realizado entre los años 2014 y 2017 en el marco de una tesis doctoral de una integrante del grupo de investigación: “Propuestas de torneos en los niveles medio y universitario en el ámbito de la enseñanza de la programación: análisis de su impacto en la motivación y el rendimiento académico de los estudiantes”, donde se realizaron diferentes torneos de programación en los niveles secundario y universitario, facilitó la generación de contribuciones teóricas y metodológicas relacionadas a su uso en contextos educativos. La administración de diferentes encuestas y aplicación de métodos estadísticos permitió conocer indicadores de aspectos motivacionales y del uso de estrategias de aprendizaje, que tienen relación con los torneos y su impacto, tanto en el aprendizaje de programación como en el rendimiento académico. Se presentan los indicadores que hacen referencia principalmente al trabajo cooperativo. Este estudio dio origen al trabajo publicado en la revista TE&ET 2021 [10].

De acuerdo a lo mencionado anteriormente y utilizando las tecnologías emergentes de RA y RV se realizó el cuento aumentado titulado “Una Aventura Distinta” [11]. La secuencia narrativa de este cuento, se ha inspirado en las indicaciones que los profesionales (Pediatras, Cirujanos y Psicólogos) del Hospital Castro Rendón de la ciudad de Neuquén, presentan ante las familias de los pacientes pediátricos que van a ser operados. Esta primera versión del cuento es digital, y está acompañada de dos aplicaciones: una empleando tecnología RA y otra empleando RV. Para el desarrollo de las mismas se realizaron visitas al hospital para entrevistar al personal médico y se realizaron grabaciones 360° de diferentes áreas del hospital (quirófano, sala de preparación del paciente, sala de padres, entre otras). La aplicación RV se está desarrollando en el marco de una tesis de

Licenciatura en Ciencias de la Computación de una integrante del proyecto. En este contexto de pandemia, este cuento resulta un aporte valioso y motivador, mediante los recursos tecnológicos desarrollados se ayuda a que pacientes y sus familiares puedan familiarizarse con las instalaciones del hospital, personal médico y procedimientos que se llevarán a cabo. Se ofrece un campo referencial, donde es posible identificarse con los personajes, con las emociones generadas por el cuento y así al transitar una ficción cercana a la realidad, puedan volver de otra manera a la difícil situación que deben atravesar. Este estudio dio origen a la publicación del libro [11].

4. Formación de Recursos Humanos

El grupo de investigación viene trabajando y formándose en proyectos anteriores a través de la realización de cursos de postgrado, extensión, entre otros. Se encuentran en desarrollo dos tesis de grado de la carrera de Licenciatura en Ciencias de la Computación. Dos docentes investigadores realizaron todos los cursos de las maestrías: “Tecnología Aplicada en Educación”, UNLP, y “Enseñanza de las Ciencias Exactas y Naturales”, UNCo. Ambas se encuentran en etapa de tesis.

5. Bibliografía

[1] Martínez Carod, N., Fracchia, C. C. (2020). Competencia y creatividad: una experiencia en la elaboración de Juegos de mesa en colegios secundarios. XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020), REDUNCI, ISBN: 978-987-604-553-7, pp. 75-83.

[2] Sánchez, M. A., (2017) Visualización de datos utilizando grafos: ejemplos de aplicación, Revista CEA, vol. 1, N°1.

[3] Nuñez, J. C. D., (2012), “IAGraph: “Aplicación Integral de Grafos: Búsquedas”, Universidad Politécnica de Madrid.

<http://www.dma.fi.upm.es/personal/gregorio/grafos/web/iagraph>. [Consultado: 11-abr-2020].

[4] Halim, S., Halim, F., “VisuAlgo”, (2011), Visualización de estructuras de datos y algoritmos mediante animación. National University of Singapore (NUS). <https://visualgo.net/es>. [Consultado: 29-abr-2020].

[5] Galles D., (2011). Data Structure Visualization. Computer Science, University of San Francisco: <https://www.cs.usfca.edu/~galles/visualization/Algorithms.html>. [Consultado: 29-abr-2020].

[6] “Algorithm Visualizer”, Algorithm Visualizer. [Offline]. <https://algorithm-visualizer.org/>. [Consultado: 25-feb-2021].

[7] Page, L. , Brin S. , Motwani, R., y Winograd, T. ,(1998), “The PageRank citation ranking: Bringing order to the Web”, en 7th International World Wide Web Conference, Brisbane, Australia, pp. 161–172.

[8] Mamani, N., Martins, A., (2020). “Propuesta de una herramienta web para el aprendizaje de grafos: aplicación del algoritmo de PageRank TE&ET 2020. XV Congreso de Tecnología en Educación y Educación en Tecnología – REDUNCI, ISBN: 978-987-604-553-7, pp. 66-74.

[9] Mandelbaum, S. (2011). Psicoprofilaxis Quirúrgica en la actualidad. Congreso del Centenario de la Sociedad Argentina de Pediatría.

[10] Fracchia, C. C., Bramardi, S., (2021). Torneos de programación: combinando los aprendizajes competitivo y cooperativo. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET), (27), e6.

<https://doi.org/10.24215/18509959.27.e6>.

[11] Álvarez, A. L., Fracchia, C. C., Alonso, A., Una aventura distinta. Editorial EDUCO Universidad Nacional del Comahue Neuquén 2020. ISBN:978-987-604-560-5

<http://rdi.uncoma.edu.ar/handle/123456789/1603>

5.

[12] Fracchia C., A. Martins, Realidad Aumentada en la Enseñanza Primaria: Diseño de Juegos de Mesa para las Áreas Ciencias Sociales y Matemáticas, (2017), III Congreso Internacional de Investigación y Docencia de la Creatividad. ISBN: 978-84-697-5239-5, Universidad de Granada, España.

Tecnologías Emergentes

Claudia Russo¹, Mónica Sarobe¹, Hugo Ramón¹, Carlos Di Cicco¹, Mariana Adó¹, Tamara Ahmad¹,
Leonardo Esnaola¹, Paula Lencina¹, Pablo Luengo¹, María Rosana Piergallini¹,
Marina Rodríguez¹, Eliana Serrano¹. Sandra Serafino¹, Lucas Benjamín Cicerchia¹, Javier Charne¹
Gabriel Pérez¹, Marcelo Guiguet¹, Damián Montes de Oca¹, Sebastián Adorno¹

Instituto de Investigación y Transferencia en Tecnología (ITT), UNNOBA-CIC.

Sarmiento N° 1169, 2^{do} Piso, Junín (B) – Te: (236) 4477-050 INT 11610.

{claudia.russo, monica.sarobe, hugo.ramon, carlos.dicicco, mariana.ado, tamara.ahmad,
leonardo.esnaola, paula.lencina, pablo.luengo, rosana.piergallini,
marina.rodriguez, eliana.serrano, sandra.serafino, lucas.cicerchia, javier.charne, gabriel.perez,
marcelo.guiguet, damian.montesdeoca, sebastian.adorno}@itt.unnoba.edu.ar

Resumen

En pos del bienestar general, el desarrollo humano se ve modelado por innovaciones tecnológicas que inciden directa e indirectamente en la forma de vida. Las *Tecnologías Emergentes* crecen de manera acelerada permitiendo avances en el campo de la comunicación, la medicina, la agricultura, el comercio y la educación; entre tantos otros.

En el ámbito educativo particularmente, las *Tecnologías Emergentes* han generado un nuevo escenario para el desarrollo del proceso de enseñanza-aprendizaje, facilitando posibilidades múltiples y planteando nuevos desafíos en todos los niveles de formación. Actualmente, el espacio de educación formal trasciende el aula, gracias a la existencia de múltiples herramientas disponibles tanto en dispositivos móviles como en computadoras ocasionando la necesidad de repensar la educación. Asimismo, las *Tecnologías Emergentes* también permiten generar

indicadores dinámicos que proporcionan nuevas métricas para el análisis del proceso de aprendizaje y la gestión educativa.

El presente artículo sintetiza el plan de trabajo 2021 del ITT-UNNOBA-CIC en relación al desarrollo e impacto de las *Tecnologías Emergentes* al servicio de la Educación, guardando un especial hincapié en la accesibilidad a entornos virtuales de enseñanza-aprendizaje (EVEA), la inteligencia artificial y la informática en salud.

Palabras claves: Tecnologías emergentes; Educación; Accesibilidad web; Brecha digital; Entornos Virtuales de Enseñanza-Aprendizaje; Inteligencia artificial; Salud; Realidad Virtual; Realidad Aumentada; Gamificación.

Contexto

¹ Investigador ITT-UNNOBA-CIC.

Las líneas de investigación a describir se enmarcan en el proyecto de investigación: *Informática y Tecnologías Emergentes*, 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) 2019, aprobado y financiado por la Secretaría de Investigación de la Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA). Su objetivo es estudiar cómo la informática impacta en el desarrollo de tecnologías emergentes a fin de analizar, definir y crear herramientas y estrategias innovadoras, capaces de incidir en el desarrollo de la sociedad.

1. Introducción

Desde hace una década, la UNNOBA trabaja en la consecución de mejoras en la calidad de los métodos y las técnicas de enseñanza-aprendizaje. En este sentido, la institución transita un profundo proceso en el que intenta definir una metodología que permita explotar el potencial de las herramientas informáticas aplicadas al campo educativo en pos de reducir la brecha digital, tanto para el público general como para las personas con discapacidad.

La introducción de las TIC en las universidades y la implementación de los EVEA, abrió una amplia variedad de posibilidades de acceso al conocimiento para una gran cantidad de personas; sin embargo, paralelamente, aumentó la brecha digital. El presente trabajo se enmarca en un proyecto mayor que tiene como finalidad investigar de qué manera la informática impacta en el desarrollo de *Tecnologías Emergentes*. Para ello se propone analizar, definir y desarrollar herramientas y estrategias innovadoras que

repercutan responsablemente en el desarrollo de la sociedad.

2. Líneas de investigación y desarrollo

El equipo de trabajo se centra en las *Tecnologías Emergentes* a partir de seis líneas de investigación:

Línea 1. *Definición de indicadores para evaluar la calidad de cursos virtuales.*

Línea 2. *Accesibilidad en entornos virtuales de enseñanza-aprendizaje y su impacto en la brecha digital.*

Línea 3. *Inteligencia artificial aplicada a la enseñanza de la Programación.*

Línea 4. *Aplicaciones de realidad virtual, realidad aumentada y estrategias de gamificación en propuestas pedagógicas a distancia.*

Línea 5. *Informática en Salud preventiva en Centros de Atención Primaria.*

Línea 6. *Tecnología GPS aplicada al deporte.*

Línea 7. *Estrategias de optimización de performance en sistemas de almacenamiento definido por software.*

Línea 8. *Procesamiento de imágenes aplicada al diagnóstico por imágenes y agronomía.*

Dentro de la línea 1, *Definición de indicadores para evaluar la calidad de cursos virtuales*, se ha creado un modelo para la evaluación de la calidad de los cursos con horas virtualizadas. El trabajo fue realizado atendiendo a la necesidad de evaluar cómo se desarrolla el sector *e-Learning* a nivel mundial y poniendo particular atención en medir aquellas variables que influyen

directamente sobre el proceso de enseñanza-aprendizaje.

Para ello se tuvieron en cuenta los factores que afectan a la gestión educativa construyendo una mirada global de su calidad. Estos factores fueron entendidos como el conjunto de variables que influyen en la calidad de la educación superior y operacionalizados en *dimensiones*. A su vez, el término *dimensión* fue definido como “cada uno de los elementos macros que permiten delimitar el conjunto en evaluación, es decir, el sistema educativo virtual implementado a través de la plataforma Educación Digital de UNNOBA, un EVEA basado en Moodle” [1]. Constituido por cuatro dimensiones observadas y medidas mediante ocho, diez, cuatro y tres indicadores respectivamente, el modelo de evaluación propuesto fue administrado por la red de tutores de Educación Digital UNNOBA que recorrió las aulas virtuales al finalizar el primer y el segundo cuatrimestre del año 2020. Estas mediciones permiten realizar un análisis comparativo de la calidad de dichas aulas en ambas instancias y corroborar la virtud anticipatoria e innovadora de los indicadores.

En la línea 2, *Accesibilidad en entornos virtuales de enseñanza-aprendizaje (EVEA) y su impacto en la brecha digital*, se han definido lineamientos de un modelo institucional favorable a la implantación en la UNNOBA de propuestas educativas formales accesibles. Esto se ha logrado a partir de una evaluación conforme a la metodología ESVIAL y a un hacer accesible el sitio web de Educación Digital UNNOBA, considerando siempre la normativa vigente. Se define a las *personas con discapacidad* como aquellos “grupos tradicionalmente nominados como lisiados, impedidos, minusválidos, deficientes, discapacitados o personas con discapacidad, según diversos

momentos históricos y perspectivas teóricas de análisis” [2]. Asimismo, se precisa la *accesibilidad web (AW)* como la “posibilidad de que la información de la página web pueda ser comprendida y consultada por personas con discapacidad y por usuarios que posean diversas configuraciones en su equipamiento o en sus programas” [3].

La línea 3, *Inteligencia artificial aplicada a la enseñanza de la Programación*, se centra en el desarrollo de un método que, empleando técnicas de *Procesamiento de Lenguaje Natural (PLN)* conjuntamente con otras herramientas y métodos propios del desarrollo de algoritmos, permite guiar al estudiante en la tarea de programación. El PLN es un área de investigación importante dentro de la *Inteligencia Artificial (IA)* [4]. Se define a la IA como una rama de las Ciencias de la Computación que se ocupa de simular las capacidades de inteligencia del cerebro humano mediante métodos, técnicas y herramientas destinadas a modelizar y resolver problemas siguiendo un proceder acorde a los sujetos cognoscentes. En este sentido, la IA es una “ciencia que se orienta a la búsqueda de la comprensión profunda de la inteligencia, teniendo en cuenta la delimitación de la misma, sus posibilidades y caracterizándose como un desafío de enorme complejidad” [5]. El método propuesto parte de enunciados escritos en lenguaje natural y pretende asistir a los estudiantes en el reconocimiento de las estructuras de datos y de control necesarias para conseguir obtener una especificación concreta, escrita en pseudocódigo, que resuelva el problema presentado. En este proceso los estudiantes deberán aplicar las buenas prácticas de programación que se procuran enseñar.

En la línea 4, *Aplicaciones de realidad virtual, realidad aumentada y estrategias de gamificación en propuestas pedagógicas a distancia*, se propone analizar las diferentes herramientas de realidad virtual (RV) y

aumentada (RA) e incorporarlas en el dictado de cursos a distancia, acompañadas de una adecuada fundamentación que brinde marco pedagógico. En esta misma línea de investigación, se sugiere estudiar las técnicas y los elementos utilizados actualmente para favorecer dinámicas de juego dentro de propuestas educativas con el objetivo de fortalecer los procesos de enseñanza y aprendizaje, haciendo uso del juego como estrategia motivacional que permita enriquecer los espacios educativos. Las neurociencias han demostrado que el juego puede ser utilizado como un recurso de gran utilidad para el aprendizaje a cualquier edad, en tanto propicia las formas diferentes de aprender y compromete distintos factores que reportan beneficios en lo personal y lo social y, por ende, beneficios a nivel educativo [6]. Por esta razón, se propone aplicar la *gamificación*, es decir, la “*metodología emergente que consiste en utilizar juegos o mecánicas de juego en contextos no lúdicos, proporcionando en educación la motivación de los alumnos*” [7], como estrategia para enriquecer el aprendizaje.

La línea 5, *Informática en Salud preventiva en Centros de Atención Primaria*, se ocupa de evaluar el uso de las Tecnologías de la Información y Comunicación (TIC) en la Atención Primaria de la Salud dentro de los Centros de Atención Primaria de Salud (CAPS). En 1978 se definió la *Atención Primaria de Salud (APS)* como “*la asistencia sanitaria esencial basada en métodos y tecnologías prácticos, científicamente fundados y socialmente aceptables, puesta al alcance de todos los individuos y familias de la comunidad mediante su plena participación y a un costo que la comunidad y el país puedan soportar, en todas y cada una de las etapas de su desarrollo con un espíritu de autorresponsabilidad y autodeterminación. La atención primaria forma parte integrante tanto del sistema nacional de salud, del que constituye la función central y el núcleo principal, como del desarrollo social y económico global de la comunidad.*” [8]. Desde entonces la Organización Mundial de la Salud (OMS) considera a la Atención Primaria en Salud como una estrategia clave

para alcanzar los objetivos de equidad y acceso universal a los servicios de salud y reconoce el potencial que las TIC para lograr una mayor eficacia de los servicios de salud y un mejor acceso a la atención médica. En vinculación con el Instituto Académico de Desarrollo Humano (IADH) de la UNNOBA, el ITT se propuso lograr un uso más eficiente de los bienes y servicios sociales, así como también adecuar la organización y los recursos de los CAPS a las necesidades de su comunidad.

La línea 6, *Tecnología GPS aplicada al deporte*, procura generar un sistema de entrenamiento que permitiese el seguimiento de aspectos del rendimiento físico del deportista en un contexto real, facilitando la mejora de la práctica deportiva mediante el análisis científico de los datos obtenidos. Con tal objetivo, el ITT-UNNOBA se ha asociado con la empresa Silamberts SRL para desarrollar un prototipo económico de GPS – accesible para las instituciones deportivas y profesionales– que permita recopilar los datos necesarios en el cálculo de los parámetros de rendimiento físico de un deportista. El desarrollo de esta nanotecnología cuenta con el financiamiento tanto de la universidad como de la Fundación Dr. Manuel Sadosky [9].

La línea 7, *Estrategias de optimización de performance en sistemas de almacenamiento definido por software*, se ocupa de estudiar los principales sistemas de almacenamiento distribuido basados en software libre, procurando analizar su desempeño frente a diversos requerimientos, como son mayor cantidad y simultaneidad de clientes conectados, número creciente de datos accedidos de manera concurrente, usuarios distribuidos geográficamente, tiempos de respuesta más acotados, throughput que se amplía de forma exponencial y todo esto con capacidad para recuperarse ante fallas que deben suponerse la regla más que la excepción.

En la Definición de la Línea 8, *Procesamiento de imágenes aplicada al diagnóstico por imágenes y agronomía*, donde se trabaja en la resolución de problemáticas de detección de

patologías en imágenes de cardiología y se comenzará a trabajar con imágenes de microscopía con el fin de aportar un valor agregado a los laboratorios de la universidad, ambas con la participación de Becarios doctorales que trabajan en la temática. Además, sumado a ello se continuará trabajando en el procesamiento de imágenes ligado a la agronomía, por un lado en el desarrollo de un prototipo de sensado de limones para una máquina cosechadora [10]. Y por otro en el desarrollo de un sistema ciberfísico para una plataforma de sensado a campo que permita obtener imágenes que luego serán procesadas para obtener información del cultivo, trabajo del cual se desprende una beca doctoral. Y por último el procesamiento de imágenes satelitales de campos cultivados en la región del noroeste de la provincia de buenos aires, con el fin de aportar información relevante a los productores agrícolas[11], y de la que se desprenden dos doctorados. Todas estas temáticas se abordan desde el procesamiento digital de imágenes acompañado además de técnicas de inteligencia artificial.

3. Resultados obtenidos/esperados

Dentro de la línea 1, *Definición de indicadores para evaluar la calidad de cursos virtuales*, se espera que tras concluir el primer cuatrimestre del 2021 pueda realizarse un análisis comparativo de los datos obtenidos durante la primera mitad de los ciclos lectivos 2020 y 2021. Asimismo, podrán compararse los resultados de la segunda mitad de los ciclos 2020 y 2021 al finalizar el año, logrando confirmar o rectificar los resultados antes obtenidos. Esto permitirá ratificar o desestimar el carácter transformativo del modelo de evaluación creado para medir la calidad del *e-Learning*. Hasta el momento el modelo ha funcionado como una herramienta capaz de guiar al docente en la mejora permanente del diseño de su aula virtual. Su mayor virtud es hacer eficiente la evaluación de la calidad del aula virtual cuantificando sus

características, mediante los valores específicos asignados a cada indicador, y cualificándolas, gracias a la valoración general de cada una de las cuatro dimensiones. Por otro lado, se estima la posibilidad de que los estudiantes accedan a esta herramienta en forma de encuesta final del curso, permitiéndoles brindar una devolución detallada y sistematizada de su experiencia en el entorno de enseñanza-aprendizaje. De esa manera se capitalizaría la comunicación estudiante-docente a través de un censado estudiantil permanente sobre la calidad de las aulas virtuales.

En la línea 2, *Accesibilidad en entornos virtuales de enseñanza-aprendizaje (EVEA) y su impacto en la brecha digital*, se espera: (a) analizar el marco y el contexto para la implementación de un proyecto educativo accesible e inclusivo en la UNNOBA; (b) evaluar las competencias e idoneidad del personal docente responsable de la acción formativa, como así también la accesibilidad de los programas curriculares; (c) describir los recursos técnicos disponibles, las condiciones de infraestructura y las barreras existentes para la impartición de la formación accesible; (d) caracterizar el perfil del grupo objetivo de estudiantes en función de variables demográficas, sociales y culturales, como así también de sus estilos y preferencias de aprendizaje; y (e) definir los modelos didácticos, sus objetivos y las actividades a realizar durante la enseñanza aplicando principios de accesibilidad. Para ello se trabajará en la confección de los cuestionarios auto administrados que se usarán para evaluar las competencias docentes y el perfil de los y las estudiantes. También se describirán en función de la información recopilada los recursos técnicos disponibles y las condiciones de infraestructura y barreras existentes para la impartición de la formación accesible.

Particularmente, para hacer accesible el sitio web de Educación Digital UNNOBA considerando la normativa vigente, se espera: (a) profundizar el estudio del estado del arte y la legislación actual en relación a la *accesibilidad web*; (b) buscar, comparar y seleccionar distintas herramientas gratuitas y pagas que permitan evaluar la accesibilidad de un sitio web; y (c) documentar el trabajo de implementación de mejoras, para que el sitio siga creciendo de forma accesible en base al estudio realizado. Para lograrlo se compararán y evaluarán distintas herramientas de revisión automática y manual, gratuitas y pagas, online e instalables que permitan analizar la accesibilidad del sitio web de Educación Digital UNNOBA. Se pondrá especial énfasis en aquellas que sean Open Source y no instalables. Una vez analizadas y seleccionadas las herramientas, se evaluará el nivel de accesibilidad del sitio conforme las metodologías y normas vigentes, se realizará un diagnóstico en base al análisis efectuado, se implantarán las mejoras propuestas en el análisis efectuado y se documentará el proceso de trabajo llevado a cabo durante la implementación de estas mejoras. El enfoque de la investigación es básicamente cuantitativo por lo que se recolectarán datos respecto del cumplimiento de la *Accesibilidad Web* en el sitio de Educación Digital para posteriormente ser analizados.

En la línea 3, *Inteligencia artificial* se espera desarrollar una aplicación que, a través del uso de técnicas de IA, asista al estudiante en el desarrollo de algoritmos partiendo de enunciados escritos en lenguaje natural.

En la línea 4, *Aplicaciones de realidad virtual, realidad aumentada y estrategias de gamificación en propuestas pedagógicas a distancia*, se espera la realidad virtual, la realidad aumentada y la gamificación, sean alternativas reales y concretas de las cuales

puedan valerse los docentes a la hora de formular propuestas, involucrando un rol más activo de los estudiantes e invitando a desarrollar las diferentes competencias propiciadas. Tras la definición de identificadores que permitan medir la incidencia de la implementación de estas tecnologías en propuestas pedagógicas a distancia, se podrá realizar un análisis de resultados.

En la línea 5, *Informática en Salud preventiva en Centros de Atención Primaria*, se espera que el uso de las TICs como herramienta de apoyo de los Centros de Atención Primaria de Salud (CAPS) facilite: (a) la identificación de pacientes por afección y clasificación de riesgo, (b) la planificación de la atención médica, (c) el seguimiento activo atento a las necesidades de cada una de las personas, (d) la generación de un mapa de salud que permita visualizar la distribución geográfica de los diferentes tipos de afecciones y (e) la evaluación/involución del tipo de patologías detectadas en el tiempo para generar acciones de prevención.

En la línea 6, *Tecnología GPS aplicada al deporte*, se espera obtener una segunda versión del prototipo con sensor cardíaco incorporado que permita la medición del pulso cardíaco y transferencia de datos a la nube, en tiempo real para una rápida toma de decisiones. Otro punto a desarrollar es la definición de pruebas adicionales que validen los prototipos en diferentes entornos y condiciones, y con la retroalimentación de deportistas, directores técnicos y médicos especialistas, entre otros.

Finalmente, en la línea 7, *Estrategias de optimización de performance en sistemas de almacenamiento definido por software*, se implementó un cluster de almacenamiento basado en Ceph [11] sobre hardware físico, se definieron estrategias de almacenamiento de acuerdo a diferentes patrones de acceso, se

estudió la incidencia de capas de caché y su impacto en la performance general del sistema. Se espera poder definir esquemas de configuración que resulten óptimos para cada patrón de acceso definido, y se espera establecer valores de rendimiento base que puedan utilizarse a modo de referencia para realizar monitoreos continuos y evaluar los ajustes futuros sobre parámetros que afectan la performance.

En el caso de la línea 8, Procesamiento de imágenes aplicada al diagnóstico por imágenes y agronomía, se espera continuar consolidando el equipo de trabajo en las diferentes temáticas abordadas. Continuar avanzando en lo que a procesamiento de imágenes en el área de diagnóstico por imágenes e incursionar en lo que a imágenes de microscopía refiere, de las cuales se desprenderán dos doctorados en la temática. Se espera además, terminar el prototipo de sensado de limones con el fin de llevarlo a la máquina cosechadora para que pueda ser probado en campo en busca de mejoras y con el objetivo de una versión final del producto. En cuanto a las imágenes agrícolas, se espera continuar trabajando en los sistemas ciberfísico en la búsqueda de una arquitectura de sistema que sea óptima para la problemática abordada e implementando nuevas técnicas de procesamiento de imágenes esperando mejorar los resultados obtenidos hasta el momento, que redundará en futuras nuevas publicaciones. Y por último en cuanto al área de imágenes satelitales lograr concluir con los doctorados en los cuales se está trabajando y que permitan brindar un valor agregado al sector. Una de estas tesis se desarrolla de manera conjunta con investigadores de la Universidad Nacional de La Plata.

4. Formación de recursos humanos

El equipo de trabajo está compuesto por docentes e investigadores formados y en formación pertenecientes a la Universidad Nacional del Noroeste de la Provincia de Buenos Aires, quienes durante el transcurso del 2020 dirigieron becas CIN, becas CIC y diversas tesinas de grado. Entre estas últimas se destaca la basada en el uso de redes neuronales convolucionales para la detección de objetos de interés en imágenes, cual trabajo final a defender entre abril y mayo del corriente año. Asimismo, el director de esta tesina de grado se encuentra pronto a la conclusión y defensa de su tesis doctoral. Por otro lado, en agosto y septiembre de 2020 respectivamente, dos investigadores de este equipo defendieron sus tesis de maestría obteniendo el título de *Magíster en Tecnologías aplicadas a Educación* otorgado por la Universidad Nacional de La Plata. A su vez, otras tres integrantes se hallan próximas a finalizar la *Maestría en Educación en Entornos Virtuales* (Universidad Nacional de la Patagonia Austral), la *Maestría en Ingeniería en Calidad* (Universidad Tecnológica Nacional) y la *Maestría Internacional en Bioinformática* (Esneca Business School). Sumado a ello, 2 integrantes se encuentran en proceso de finalización de doctorados y se incorporan 3 becarios doctorales con becas UNNOBA, CIC y CONICET. Por último, tanto la presentación a congresos como CLEI / LACLO 2021 y CACIC 2021 como la realización de múltiples cursos de posgrado y actualización profesional aseguran una formación permanente que repercute en el dictado de materias en carreras de grado, ofertas de posgrado y cursos de extensión universitaria.

5. Bibliografía y referencias

- [1] Russo, Claudia; Sarobe, Mónica; Ahmad, Tamara (2021). "Definición de indicadores. Calidad en cursos virtuales", artículo en evaluación para ser publicado en *TE&ET. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, N° 28, Enero-Junio 2021, La Plata: Facultad de Informática, Universidad Nacional de La Plata.
- [2] Mareño Sempertegui, M. (2010). "Inclusión Educativa en la Universidad Nacional de Córdoba: el desafío de aplicar los principios del Diseño Universal en la gestión de políticas universitarias", *Ponencia presentada en: Primeras Jornadas Nacionales de Accesibilidad y Diseño Universal*, Córdoba, Argentina: Universidad Nacional de Córdoba.
- [3] Ley Nacional N. 26.653 (2010). Accesibilidad de la Información en las Páginas Web. Autoridad de Aplicación. Plazos. Reglamentación. Recuperado de: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/175000-179999/175694/norma.htm>
- [4] Aggarwal, M. (2011). Information retrieval and question answering nlp approach: an artificial intelligence application. *International Journal of Soft Computing and Engineering (IJSCE)*, 1 (NCAI2011).
- [5] Ocaña-Fernández, Yolvi; Valenzuela-Fernández, Luis Alex; Garro-Aburto, Luzmila Lourdes (2019). "Inteligencia artificial y sus implicaciones en la educación superior", *Propósitos y Representaciones*, 7(2), pp. 536-568. doi: <http://dx.doi.org/10.20511/pyr2019.v7n2.274>
- [6] Fernández, Héctor (s/f). "¿Qué es la gamificación? Aplicaciones y ejemplos reales", *Economía TIC*, <https://economytic.com/gamificacion/>
- [7] Parra-González, María Elena; Segura-Robles, Adrián (2019). "Producción científica sobre gamificación en educación un análisis cuantitativo", *Revista de educación*, N° 386, octubre-diciembre 2019, pp. 113-136, Madrid: Ministerio de Educación y Formación Profesional. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=7464667>
- [8] OMS-UNICEF (1978). "Almá-Atá. Atención Primaria de Salud", *Conferencia Internacional sobre Atención Primaria de Salud*, Kazajistán. Recuperado de <https://apps.who.int/iris/bitstream/handle/10665/39244/9243541358.pdf;jsessionid=96EC3180E69DF3FF63E494202B9C8E16?sequence=1>
- [9] Santos, Luisina; Guiguet, Marcelo; Luengo, Pablo; Álvarez, Eduardo; Di Cicco, Carlos; Useglio, Gustavo; Gómez, Federico; Capelli, Matías (2020). "Dispositivo GPS para monitoreo del rendimiento deportivo", *XXVI Congreso Argentino de Ciencias de la Computación - CACIC 2020. Libro de actas*, Red de Universidades con Carreras en Informática, La Plata: Universidad Nacional de La Plata. Disponible en <http://sedici.unlp.edu.ar/handle/10915/114088>
- [10] Sandra Serafino, Lucas Benjamin Cicerchia, Gabriel Pérez, Sebastian Adorno & Agustin Balmer (2020), "Detección y conteo de limones mediante técnicas de visión artificial y tracking para la estimación de cosecha en tiempo real," 2020 XLVI Latin American Computer Conference (CLEI), Loja, Ecuador
- [11] Cicerchia L.B., Abasolo M.J., Russo C.C. (2020) Classification of Summer Crops Using Active Learning Techniques on Landsat Images in the Northwest of the Province of Buenos Aires. 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_10

[12] Weil, Sage (2007). CEPH: Reliable,
Scalable and High-Performance Distributed
Storage. Obtenido el 10 de mayo de 2020 de
<https://ceph.com/wp-content/uploads/2016/08/weil-thesis.pdf>

Una aproximación para el Diseño de una Aplicación destinada a la confección de Diagramas de clases: Caso estudiantes universitarios con disminución visual

Nicolas Gabriel del Valle^{1,2}, Ing. Laura del Carmen Ligorria^{1,3}, Ing. Roxana María Manera^{1,4}, Ing. Lorena Daniela Peralta^{1,5}, Ing. Alejandra Di Gionantonio^{1,6}, Juan Gabriel Moreno^{1,7}

¹UTN Facultad Regional Córdoba

{²dv.nico13, ³liuniversidad, ⁴roxanamanera, ⁵peralta.lorena.d, ⁶ing.alejandrardg}@gmail.com, ⁷gabriel_2008@live.com.ar

RESUMEN

Ante la necesidad de actualizar las estrategias utilizadas en la enseñanza de nivel superior que permitan la transmisión del conocimiento, se analiza diseñar una aplicación que permita la conversión de un lenguaje formal basado en gramática libre de contexto hacia un diagrama gráfico. En la presente investigación se procederá a la construcción de diagramas de clases, permitiendo la confección de los mismos por parte de estudiantes universitarios con disminución visual. Esto constituye una primera etapa con miras a la generación de módulos destinados a abarcar diferentes diagramas UML.

Palabras clave: UML, aplicación, disminución visual.

CONTEXTO

El presente trabajo se realiza en el Área de Investigación del Departamento de Ingeniería en Sistemas de Información de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional.

1. INTRODUCCIÓN

El avance de las herramientas didácticas para considerar a estudiantes con capacidades heterogéneas ha sido tenida en cuenta por las instituciones. Los esfuerzos actuales realizados por parte de las mismas con respecto a la flexibilidad de las herramientas pedagógicas de enseñanza para así beneficiar a personas con capacidades heterogéneas existen, pero en su mayoría están enfocadas a

la educación preescolar, primaria y secundaria, no así en niveles superiores [1]. En este escenario, Aquino, García & Izquierdo, en su investigación revelaron que el grado de visión es un factor relevante en la trayectoria académica de estudiantes con disminución visual, así como la necesidad de apoyos institucionales en tres categorías: accesibilidad arquitectónica, tecnológica y de personal especializado [2].

En el caso específico del diseño de diagramas UML, se han encontrado en el mercado diferentes aplicaciones que realizan la traducción de texto a diagramas gráficos [3] similares al planteado en el presente trabajo de campo. Sin embargo, no se han observado herramientas enfocadas para estudiantes con capacidad visual disminuida, ni aquellas que realicen la conversión inversa, es decir, de diagrama a texto.

La aplicación va a permitir la generación y lectura de los elementos que forman parte del lenguaje de modelado UML 2.0.

2. OBJETIVOS Y LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se inscribe dentro de los lineamientos de investigación en Ciencias de la Computación y la Informática.

El objetivo de este proyecto de investigación es: “La producción de una herramienta de software que permita la conversión de una técnica estándar a un modelado gráfico”.

3. MATERIALES Y MÉTODOS

Para la construcción del prototipo de software se emplearon tecnologías web. Como base arquitectónica se utilizó la estructura provista

por el framework Angular [4], pero debido a la subutilización de las herramientas provistas por dicho framework, se prevé migrar a la librería ReactJS [5]. Dicha librería implementa patrones que nos permitieron construir dicho prototipo teniendo en cuenta la calidad del sistema. Se tuvieron en cuenta características como la extensión, la flexibilidad y la comprensión de los componentes de software. En un nivel más bajo de abstracción, se implementó el patrón de diseño “Strategy”, el cual permite la mantenibilidad de la lógica involucrada en la interpretación sintáctica y generación de los modelos gráficos. De tal forma que, si se considera otro algoritmo, el mismo podrá incorporarse al flujo de control del software sin involucrar cambios significativos.

En referencia a la sintaxis perteneciente al lenguaje textual estructurado, se definió su gramática, parsers, semántica y gestión de notificación de errores por medio de la herramienta ohmjs [6].

Para adaptar la ejecución del software a un entorno de escritorio se utilizó la librería ElectronJS [7]. Dicha librería permite la generación de artefactos ejecutables pertenecientes a los sistemas operativos utilizados por la industria (Linux, Windows y Mac OS).

4. RESULTADOS Y AVANCES

Lo logrado hasta el momento constituye un prototipo. El mismo permite reducir la incertidumbre propia al desarrollo de una aplicación, la visualización de las bondades de esta herramienta y su impacto en el desarrollo de las estrategias pedagógicas llevadas a cabo por las materias afines a la programación y diseño de software.

Se estima en un futuro la implementación de una transformación inversa, es decir aquella que tiene como entrada un diagrama gráfico, generando como salida una descripción textual formalizada. Además, también se establecen para las futuras iteraciones la adaptabilidad de la interfaz de usuario [8] y la generación y lecturas de varios tipos de diagramas UML.

A medida que se vayan concretando los distintos módulos, estarán disponibles para que la comunidad pueda probarla y enviar sus sugerencias de manera que se optimice la herramienta concebida.

Al plantear como objetivo la generación de un grafo conexo por parte de un estudiante con disminución visual, una de las ideas que surgen es establecer como punto de partida algo que sea familiar para el mismo, y luego formular un sistema de transformación para obtener como resultado el gráfico deseado.

Una de las herramientas más usadas en la transferencia de conocimientos por parte de este grupo poblacional es el uso de texto, ya sea por medio del lenguaje Braille o mediante el uso del teclado y lector de pantalla.

Así surge la iniciativa de implementar una aplicación que mediante el ingreso de un texto con sintaxis definida permita la generación del grafo deseado. Ver Figura 1.

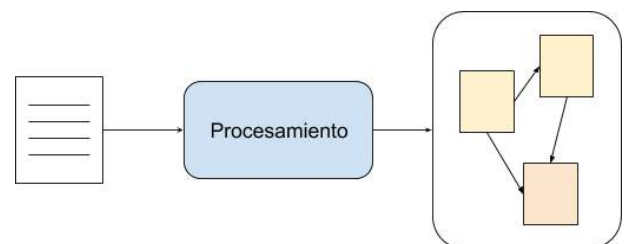


Figura 1. Esquema general del sistema.

Un primer aspecto a considerar es la sintaxis del lenguaje perteneciente al texto estructurado utilizado como entrada del procesamiento. El mismo puede ser definido utilizando un número de tokens escaso (siendo un token un componente léxico formado por una cadena de caracteres que tiene significado coherente en cierto lenguaje) ya que, al generarse un diagrama estructurado, la cantidad de relaciones propiedades de los nodos son limitados. Esto tiene como consecuencia una reducción de la curva de aprendizaje involucrada a la hora de memorizar la sintaxis de dicho lenguaje.

Cuando se considera la lectura de texto digitalizado, de las herramientas utilizadas, se destaca el uso de lectores de pantalla. El mismo posee un sintetizador de voz encargado de la generación de los sonidos propios de un idioma previamente

configurado. Dichos sintetizadores no son infalibles y suelen fallar a la hora de pronunciar palabras pertenecientes a una lengua foránea, por lo que es necesario como requisito que los tokens pertenecientes al lenguaje estructurado estén en una lengua de común uso en la región considerada, para que de esta manera el sintetizador de voz pueda pronunciarlos sin dificultad permitiendo su interpretación de manera inequívoca. Por el momento se tuvo en cuenta sólo el español ya que es la lengua madre perteneciente a nuestra región.

Cuando se diseñan diagramas como artefactos del proceso de modelación de software, los mismos pueden ser muy sencillos con pocos componentes y adornos, hasta muy complejos dependiendo de los objetivos propios que se quieran alcanzar y de la complejidad inherente del sistema en construcción. Como consecuencia, dicha variación en la complejidad de las abstracciones debe ser soportada por el lenguaje requerido como entrada del sistema, es decir se debe permitir la generación de bocetos rápidos hasta la confección de modelos complejos y extensos, teniendo en cuenta en el último caso aspectos tales como la navegabilidad y la reducción en la carga de la memoria a corto plazo, etc.

Para definir el lenguaje formal se construyó una gramática libre de contexto, la cual posee reglas de producción de la forma $V \rightarrow w$, donde V es un símbolo no terminal y w es una cadena de terminales o no terminales. Dicha gramática se puede observar en la figura 2.

```

1  const MyGrammar =
2  ClassDiagramEntities {
3    classDiagramEntities = class*
4
5    class = ( stereotypedClass myspaces className
6             | "clase" myspaces className
7             | "interfaz" myspaces className ) eol classDetail
8
9    className = alnum+
10
11   stereotypedClass = "clase" myspaces stereotype
12
13   stereotype = alnum+
14
15   classDetail = attributes? operations?
16
17   attributes = "atributos" eol attribute+
18
19   operations = "operaciones" eol operation+
20
21   attribute = (visibility myspaces)? (static myspaces)? attributeName eol
22
23   operation = (visibility myspaces)? (static myspaces)? operationName eol
24
25   visibility = "privado" | "paquete" | "protegido" | "publico" ...

```

Figura 2. Fragmento gramática libre de contexto de estructura de diagrama de clases.

Una vez introducido el lenguaje formal en la aplicación, la misma generará el diagrama requerido, como se muestra en la figura 3.

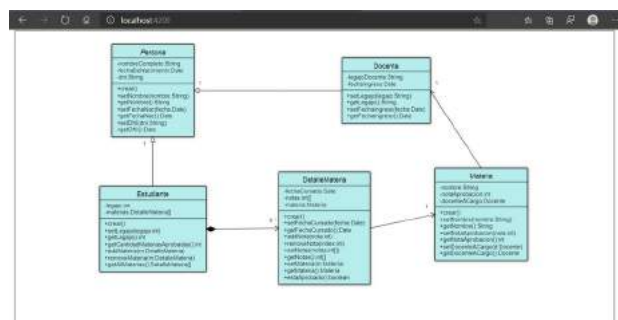


Figura 3. Diagrama de clases generado por la aplicación.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por docentes-investigadores pertenecientes a la carrera de grado de Ingeniería en Sistemas de Información.

El grupo está compuesto por una Directora, tres ingenieras investigadoras de apoyo, una ingeniera y dos estudiantes aspirantes a incorporarse a la carrera de investigador.

Este proyecto contribuirá a la formación y crecimiento de la carrera de investigador de los integrantes del mismo.

6. BIBLIOGRAFÍA

- [1] Fernández Morales, F. H.; Duarte, J. E. and Gutiérrez, G. J. Estrategia pedagógica para la formación de ingenieros con discapacidad visual. Papeles de trabajo - Centro de Estudios Interdisciplinarios en Etnolingüística y Antropología Socio-Cultural. Versión On-line ISSN 1852-4508. no.29 Rosario jun. (2015).
- [2] Aquino, S. P., García, V. & Izquierdo, J. La inclusión educativa de ciegos y baja visión en el nivel superior: Un estudio de caso. *Sinéctica*, 39, 1-21 (2012).
- [3] PlantText, The expert's design tool, <https://www.planttext.com>
- [4] Google. Introduction to the Angular Docs, <https://angular.io/docs>
- [5] Getting Started ReactJS, <https://reactjs.org/docs/getting-started.html>
- [6] GitHub - harc/ohm: A library and language for building parsers, interpreters, compilers, etc. <https://github.com/harc/ohm>
- [7] GitHub. (n.d.). Documentación de Electron, <https://electronjs.org/docs>
- [8] Web Content Accessibility Guidelines (WCAG) Overview <https://www.w3.org/WAI/standards-guidelines/wcag/>

Una propuesta para la mejora de funciones ejecutivas en niños con autismo a través del desarrollo de habilidades de Pensamiento Computacional

Marcelo Uva, Marcela Daniele, Gisela Montanari y Francisco Bavera

Departamento de Computación, Facultad de Ciencias Exactas, Físico-Químicas y Naturales,
Universidad Nacional de Río Cuarto

Ruta 36 Km. 601 –CP 5800 - Río Cuarto – Córdoba - Argentina Tel. (0358) 4676235
{uva, marcela.pancho}@dc.exa.unrc.edu.ar, gisemontanari@gmail.com

Resumen

El Pensamiento Computacional es definido como un proceso mental utilizado para formular problemas y soluciones de manera que puedan ser llevadas adelante por un agente de proceso de información. Estudios y tendencias a nivel mundial dan cuenta de la importancia de incorporar y desarrollar el Pensamiento Computacional (PC) en el sistema educativo obligatorio. El PC define un proceso de resolución de problemas fortaleciendo capacidades de organización lógica, abstracción, secuenciamiento, verificación de errores, entre otras.

Por otro lado, los Trastornos del Espectro Autista (TEA) se definen como una disfunción neurológica del desarrollo que afecta principalmente capacidades sociales y ejecutivas. La función ejecutiva puede ser definida como el proceso que interviene en el control y monitoreo de los pensamientos y acciones, tomando dentro de esto la planeación, autorregulación, organización del pensamiento, flexibilidad cognitiva, detección y corrección de errores, inhibición y resistencia a la interferencia.

En este artículo se presenta una línea de investigación cuyo objetivo es analizar, diseñar e implementar actividades de aprendizaje para niños con TEA destinadas a

mejorar de sus funciones ejecutivas a través del desarrollo de sus habilidades de PC. Se tiene planificado la implementación de un framework colaborativo para la creación de estas actividades. Este framework deberá ser lo suficientemente flexible como para permitir adaptar actividades según las capacidades de cada niño/a con TEA.

Contexto

En el marco del Proyecto de Investigación y Desarrollo (PID 2018) "La construcción del pensamiento computacional: estudio del impacto desde la formación de formadores" aprobado por el Ministerio de Ciencia y Tecnología de Córdoba se realizaron acciones en pos de caracterizar la construcción del Pensamiento Computacional y contribuir a la formación docente continua y el impacto sobre sus prácticas. La enseñanza de las Ciencias de la Computación promueve la construcción del Pensamiento Computacional (PC) e implica el desarrollo de capacidades y competencias útiles para la búsqueda de soluciones a diversos tipo de problemas. Estudios y tendencias a nivel mundial dan cuenta de la importancia de introducir y desarrollar el PC en el sistema educativo obligatorio [1].

En esta misma línea de investigación fue presentado el programa “Estudios Interdisciplinarios en Evaluación de Procesos de Software y sus aportes al Desarrollo del Pensamiento Computacional en Prácticas Educativas de Ciencias Naturales” (convocatoria de Proyectos y Programas de Investigación PPI 2020-2022 Universidad Nacional de Río Cuarto). En ese programa se incluyó el proyecto “El pensamiento computacional y las prácticas docentes en ciencias”.

1. Introducción

El Pensamiento Computacional es definido como un proceso mental utilizado para formular problemas y soluciones de manera que puedan ser llevadas adelante por un agente de proceso de información [3]. Estudios y tendencias a nivel mundial dan cuenta de la importancia de introducir y desarrollar el Pensamiento Computacional (PC) en el sistema educativo obligatorio [2, 3]. Estonia, Reino Unido, Finlandia, Francia y Australia han sido pioneros en esta decisión. Incluso en algunas regiones de España y de otros países europeos están siendo introducidos en la actualidad. También en países de América Latina como Costa Rica, Perú, Colombia y República Dominicana, en otros países de la región se ha situado el debate, se han generado diversas propuestas y se han incluido las Ciencias de la Computación en sus sistemas educativos. En la Argentina, muy recientemente, el Consejo Federal de Educación, por resolución CFE No 343/18 de fecha 12 de septiembre de 2018, aprueba los NAP (Núcleos de aprendizaje prioritarios para educación digital, programación y robótica) y establece que las jurisdicciones llevarán adelante su implementación e inclusión en sus documentos curriculares.

La Asociación de Docentes en Ciencias de la Computación (CSTA) y la Sociedad Internacional para la Tecnología en Educación (ISTE) plantean una definición abierta, enumerando algunas de las características principales del PC que conjuga tanto habilidades cognitivas como actitudinales. El PC define un proceso de resolución de problemas [4] que incluyen las siguientes características:

- Formulación de problemas de manera que permitan utilizar computadoras y otras herramientas para solucionarlos.
- Organización y análisis de datos de manera lógica.
- Representación de datos mediante abstracciones, como modelos y simulaciones.
- Automatización de soluciones mediante pensamiento algorítmico (pasos ordenados).
- Identificación, análisis e implementación de soluciones eficientes con un orden de pasos y recursos.
- Generalización y transferencia del proceso de resolución de problemas.

Estas habilidades se apoyan y acrecientan mediante una serie de actitudes que son dimensiones esenciales del PC. Estas disposiciones incluyen:

- Confianza en el manejo de la complejidad.
- Persistencia en el trabajo con problemas difíciles.
- Tolerancia a la ambigüedad.
- Habilidad para lidiar con problemas no estructurados.
- Habilidad para comunicarse y trabajar con otros para alcanzar una meta o solución común.

Por otro lado, el trastorno del espectro autista (TEA) es un trastorno neurobiológico del

desarrollo, que se manifiesta durante los tres primeros años de vida y que perdura a lo largo de todo el ciclo vital. Según Martos Pérez [11] el trastorno se precipita tras un periodo de aparente normalidad, que se extiende aproximadamente hasta el final del primer año de vida. A partir de aquí emergen una serie de manifestaciones características que evocan una especie de regresión en el desarrollo. Esto se produce en el momento evolutivo que tiene lugar la construcción de las funciones psicológicas superiores elementales que son la base para la apropiación del conocimiento y la cultura humana. Los niños con autismo plantean un reto para sus familias y requieren una atención especial por parte de las personas de su entorno, por lo cual además de abordar los déficits inherentes a dicho trastorno deben hacer frente a otras dificultades que generalmente aparecen asociadas al autismo. Como alteraciones del sueño problemas en la alimentación, escasez de juego y las limitadas habilidades de la autonomía personal.

Algunas características en niños con TEA son deficiencias persistentes en la comunicación y en interacción sociales junto a patrones restrictivos y repetitivos de comportamiento, intereses o actividades. Los niños con autismo de distintas edades y niveles cognitivos pueden mostrar alteraciones en el funcionamiento ejecutivo [7,12] que afectan al uso de habilidades de planificación, flexibilidad e inhibición de respuestas. Algunas dificultades vinculadas con el déficit ejecutivo son:

- Anticipar situaciones.
- Organizar, secuenciar y/o mantener información.
- Producir conductas con una meta definida..

- Iniciar una actividad sin ayuda (dependencia ambiental).
- Monitorear la propia ejecución de una tarea o actividad (tiempo, atención).
- Autocorregir errores.
- Generar diferentes soluciones para un mismo problema.
- Generalizar lo aprendido.
- Autorregular las emociones.
- Cambiar el foco de atención de un estímulo a otro.

El desarrollo y potenciación de habilidades de PC [5, 6] en niños diagnosticados con TEA puede contribuir al fortalecimiento de las capacidades ejecutivas en estos propios de estos casos. En este sentido es posible mencionar algunas experiencias como [8, 9, 10] en donde a través de un ambiente de aprendizaje basado en juegos se busca potenciar habilidades sociales en grupos de niños que trabajando en equipo resuelven problemas de programación simples. En [9] por ejemplo se describe la experiencia desarrollada en un workshop de programación de juegos para niños-adolescentes con TEA.

Las líneas de trabajo propuestas en este artículo están direccionadas a potenciar habilidades de PC en niños con TEA en beneficio de sus funciones ejecutivas. Se tiene planificado el desarrollo de un framework informático mediante el cual se puedan desarrollar actividades específicas para niños con TEA potenciando sus habilidades de PC

2. Línea de Investigación y Desarrollo

La propuesta de trabajo presentada en este artículo tiene como objetivos: estudiar, analizar, diseñar e implementar actividades orientadas a potenciar las habilidades de PC en niños con TEA con el objetivo de

contribuir al beneficio de sus capacidades ejecutivas y sociales.

Se tiene previsto trabajar con las habilidades de PC vinculadas en resolución de problemas, secuenciamiento, abstracción y testing. Se plantea el desarrollo de un framework informático que permita la definición de actividades específicas utilizando pictogramas. Los pictogramas son recursos muy utilizados por los profesionales abocados a la educación de niños con TEA. Del mismo modo, los dispositivos tecnológicos tales como celulares, tablets, notebooks son, en general, amigables para los niños/as con TEA.

El framework proveerá una estructura sobre la cual un profesional o la propia familia a cargo de un niño con TEA podrá definir determinado tipo de actividades. Por ejemplo, vinculado a las habilidades de PC de resolución de problemas, secuenciamiento y abstracción se podrán establecer las condiciones iniciales, el objetivo final de la tarea y las acciones (pictogramas) para llevar adelante el proceso. El niño deberá ordenar (programar) y ejecutar el la solución.

Los lineamientos de la propuesta fueron definidos por un grupo interdisciplinario de profesionales conformado por psicólogos, psicopedagogos, acompañantes terapéuticos, maestros, profesores y profesionales informáticos.

3. Resultados Obtenidos/Esperados

Durante 2019 y 2020, se lograron importantes avances en la investigación propuesta, hacia el cumplimiento del objetivo general planteado. Se divulgaron resultados en eventos científicos nacionales e internacionales, compartiendo e intercambiando con otros investigadores, en torno a la apropiación de habilidades de pensamiento computacional y la modificación

de sus prácticas docentes, por parte de docentes de primaria a partir de propuestas formativas recibidas. Se consolidó la conformación de un equipo de investigación integrado por docentes investigadores con trayectoria en didáctica de las ciencias, de Computación, Matemática y Física, Fac Cs Exactas, UNRC.

Como resultado de este trabajo se espera contribuir a la mejora de las funciones ejecutivas en niños con TEA potenciando sus habilidades de PC. El objetivo a corto plazo es el desarrollo de un framework Open Source en donde profesionales y niños puedan trabajar y compartir recursos con otros niños con las mismas problemáticas.

4. Formación de recursos humanos

Durante el 2020 una estudiante de Licenciatura en Ciencias de la Computación comenzó a desarrollar su trabajo de tesis dentro de la línea de investigación propuesta en este trabajo. Actualmente se ha comenzado a implementar un prototipo de framework el cual está siendo evaluado por el equipo de trabajo interdisciplinario.

Bibliografía

- [1] Robins et al., 2003. A. Robins, J. Rountree, N. Rountree. Learning and teaching programming: A review and discussion. *Computer Science Education*, 13 (2) (2003), pp. 137-172.
- [2] K. Boyer, P. S. Buffum et al. "Engage: A game-based learning environment for middle school computational thinking," in *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE15. ACM, 2015, pp. 440-440.

- [3] J. M. Wing, "Computational thinking," *Communications of the ACM*, vol. 49, no. 3, p. 33, mar 2006.
- [4] R. E. Mayer, "Teaching for transfer of problem-solving skills to computer programming," in *Computer-Based Learning Environments and Problem Solving*, E. De Corte, M. C. Linn, H. Mandl, and L. Verschaffel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 193–206.
- [5] M. Román-González, J. Moreno-León, and G. Robles, "Complementary tools for computational thinking assessment," in *CTE 2017: International Conference on Computational Thinking Education 2017*, July 2017, pp. 154-159.
- [6] S. Fridenson-Hayo, S. Berggren, A. Lassalle et al. "'emotiplay': a serious game for learning about emotions in children with autism: results of a cross-cultural evaluation," *European Child & Adolescent Psychiatry*, vol. 26, no. 8, pp. 979–992, mar 2017.
- [7] E.L.Hill, "Evaluating the theory of executive dysfunction in autism," *Developmental Review*, vol. 24, no.2, pp. 189–233, 2004.
- [8] M. Schmidt and D. Beck, "Computational thinking and social skills in Virtuoso: An immersive, digital game-based learning environment for youth with autism spectrum disorder," in *Immersive Learning Research Network*. Eds. Cham: Springer International Publishing, 2016, pp. 113-121.
- [9] R. Munoz, R. Villarroel, T. S. Barcelos, F. Riquelme, Á. Quezada and P. Bustos-Valenzuela, "Developing Computational Thinking Skills in Adolescents With Autism Spectrum Disorder Through Digital Game Programming," in *IEEE Access*, vol. 6, pp. 63880-63889, 2018.
- [10] R. Munoz, T. S. Barcelos et al., R. Villarroel, and I. F. Silveira, "Game design workshop to develop computational thinking skills in teenagers with autism spectrum disorders," in *2016 11th Iberian Conference on Information Systems and Technologies (CISTI)*, June 2016, pp. 1-4.
- [11] Juan Martos Perez, Maria Llorente Comi, Ana Gonzalez Navarro, Raquel Ayuda Pascual, Sandra Freire. "Los niños pequeños con Autismo, soluciones prácticas para problemas cotidianos". CEPE. CIENCIAS DE LA EDUCACIÓN PREESCOLAR Y ESPECIAL. España 2001.
- [12] S. Ozonoff, B. F. Pennington, and S. J. Rogers, "Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind," *Journal of Child Psychology and Psychiatry*, vol. 32, no. 7, pp. 1081–1105, nov 1991.

Tesis de Doctorado

Coordinación de dispositivos en ambientes ubicuos mediante coreografías

Tesis doctoral desarrollada en la Universidad Nacional de San Luis - San Luis - Argentina

Defendida el 17 de Noviembre de 2020

Oscar A. Testa¹, Germán A. Montejano (director)², and Oscar Dieste (codirector)³

¹Universidad Nacional de La Pampa, Facultad de Ciencias Exactas y Naturales, La Pampa - Argentina, otesta@exactas.unlpam.edu.ar

²Universidad Nacional de San Luis, Facultad de Ciencias Físico, Matemáticas y Naturales, San Luis - Argentina, gmonte@unsl.edu.ar

³Universidad Politécnica de Madrid, Escuela Técnica Superior de Ing. Informática, Madrid - España, odieste@fi.ump.es

Resumen

Introducción: Actualmente nos encontramos involucrados en ambientes donde los dispositivos ubicuos forman parte de nuestra vida cotidiana y de nuestras tareas diarias. De forma permanente estamos interactuando con dichos dispositivos y más aún, con los servicios que ellos nos brindan. En casi todos los casos, los dispositivos ubicuos no proporcionan servicios de forma aislada, sino que deben cooperar con otros dispositivos. **Objetivo:** El objetivo principal de esta investigación es el de definir un mecanismo de coordinación de dispositivos ubicuos que garantice su interoperabilidad independientemente del modelo y fabricante del mismo; utilizando los estándares de SOA y de coreografías para la composición de servicios. **Metodología:** Se ha utilizado como metodología de investigación design science, ya que es la que mejor se adapta a la naturaleza del problema, planteando como uno de sus lineamientos la construcción de artefactos y su posterior evaluación. **Resultados:** Se ha obtenido un framework de coordinación de dispositivos a través de la utilización de coreografías, que funciona de manera correcta y dando soporte a las características distintivas de los dispositivos ubicuos. La solución planteada es simple, interoperable y extensible.

1. Introducción

Dispositivos ubicuos son aquellos dispositivos electrónicos que tienen capacidad de procesamiento y comunicación, y pueden ser encontrados en cualquier lugar: la oficina, el auto, la casa, o la misma ropa con la que vestimos [1].

La computación ubicua es un paradigma tecnológico que pretende que las computadoras no se perciban como objetos diferenciados; así como también, que su utilización sea lo más transparente y cómoda posible para las personas, en las diversas situaciones en que existe interacción [1].

Los avances tecnológicos (especialmente de hardware y comunicaciones) han permitido que los dispositivos ubicuos sean al mismo tiempo generadores y consumidores de servicios; es decir, de acuerdo a las capacidades del dispositivo, éste pueda no solo obtener, sino también ofrecer su funcionalidad; lo que se traduce en un ambiente de cooperación entre dispositivos, permitiendo componer funcionalidades más complejas. Por composición, entendemos la forma en que los dispositivos ubicuos se pueden combinar para realizar una tarea determinada. La composición implica que los dispositivos deben comunicarse entre ellos con la finalidad de obtener un servicio con valor agregado, lo que a

su vez conlleva desafíos tales como: tolerancia a fallas, escaso nivel de procesamiento, problemas de conectividad, por mencionar algunos ejemplos [2].

Si bien hoy en día podemos decir que distintos dispositivos se pueden comunicar entre ellos, compartiendo de alguna manera sus servicios, generalmente lo realizan a partir de protocolos propietarios y sin seguir definiciones estándar, provocando que otros dispositivos no puedan ser integrados para la comunicación. Esto representa una importante limitación en la composición de dispositivos ubicuos [2].

La composición de dispositivos ubicuos presenta desafíos adicionales tales como: la heterogeneidad de los mismos, las contingencias de los dispositivos, y la personalización de los mismos (ej: provisión de servicios de acuerdo a las preferencias del usuario). Dado que los dispositivos ubicuos poseen limitaciones de recursos (ej. poca memoria y batería), se deben hacer consideraciones especiales respecto a la eficiencia y rendimiento de la composición de dispositivos. Todas estas dificultades hacen que la composición de dispositivos ubicuos constituya un área de investigación importante, donde los avances no son claros al día de hoy [2].

Esta situación nos motivó a plantear una propuesta que aglutine las especificaciones y estandarizaciones existentes en SOA para la coordinación de dispositivos ubicuos. Si pensamos que cada dispositivo ubicuo es proveedor o consumidor de un servicio, encaja perfectamente en la arquitectura de servicios.

Para poder llevar adelante este tipo de investigación, y de acuerdo a la naturaleza misma de la solución que intentamos encontrar, nos basamos en la metodología de investigación Design Science, ya que la misma se basa en la construcción y evaluación de un artefacto con la intención de dar solución a problemas debidamente identificados [3].

En este caso, nuestro objetivo principal es la definición de un mecanismo de coordinación de dispositivos ubicuos que garantice su interoperabilidad independientemente del modelo y fabricante del mismo; utilizando los estándares de SOA y de coreografías para la composición de servicios. Para poder alcanzar este objetivo principal hemos planteado otros objetivos más específicos, de los cuales se desprende la principal contribución de este trabajo de Tesis.

La contribución consiste entonces, en la construcción de un framework de ejecución de coreografías en ambientes pervasivos a través de la utilización de dispositivos ubicuos. Este framework, por lo tanto, al estar basado en SOA es estandarizado, lo que conlleva a que se evite el problema en la heterogeneidad de dispositivos, además de permitir la utilización de características más avanzadas (dependiendo de la potencia y capacidad que posea el dispositivo) como es el caso de WS-Transaction, WS-Security, etc. Además, permitirá la interconectividad entre dispositivos ubicuos y servidores arbitrarios para llevar adelante la composición.

2. Motivación

Los mecanismos de composición de servicios, como las orquestaciones y coreografías, son aspectos bien conocidos en SOA (Service Oriented Architecture) que permiten construir sistemas de negocio complejos y aplicaciones a partir de una gran cantidad de servicios heterogéneos, simples y distribuidos. Estos conceptos podrían ser aplicables a ambientes ubicuos, en especial las coreografías, las cuales pueden fácilmente interpretarse en términos de dispositivos y ambientes ubicuos. Sin embargo, en determinados ambientes donde los servicios son dinámicos, móviles, menos fiables y dependientes del dispositivo, los mecanismos de composición establecidos para servicios web no son directamente aplicables [16].

Adicionalmente, la composición de múltiples dispositivos ubicuos presenta nuevos desafíos que no son compatibles con la composición de servicios web. En particular, los mecanismos de composición en ambientes masivos ¹ necesitan hacer frente a distintas contingencias que pueden ocurrir con estos elementos, 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 encuentren en un momento determinado, etc.

¹Por ejemplo el de dispositivos móviles

En ambientes ubicuos, la disponibilidad y confiabilidad de los dispositivos no puede ser garantizada. No obstante, a pesar de todas las dificultades, existen ventajas que podría brindar la adaptación de los conceptos de SOA para la composición de dispositivos ubicuos, como son la estandarización tanto de los protocolos como de los mecanismos de comunicación entre los dispositivos, y la compatibilidad con otras plataformas de servicios ya existentes.

Las similitudes entre la composición de servicios web y la coordinación de dispositivos ubicuos es sorprendente. Si pensamos que cada dispositivo ubicuo en un ambiente pervasivo² es proveedor, o, consumidor de un servicio, la coordinación de dispositivos encaja perfectamente con la composición de servicios. Es también sorprendente que esta similitud no haya sido apenas explorada con anterioridad, salvo en el trabajo de Sheng [2] se hace mención a ello.

A finales de la década de los 80 y principios de los 90, Mark Weiser introdujo el término de computación ubicua (también utilizó el acrónimo “ubicomp”) [1]. La teoría de Weiser postula que la computación ubicua tiene como propósito mejorar el uso de las computadoras, haciéndolas disponibles en el entorno físico, a través de una multiplicidad de elementos. El problema de esta teoría fue que los protocolos de comunicación existentes, no se alineaban con ella y debieron ser mejorados, en especial para permitir la movilidad de los dispositivos [1].

Se han realizado diversos proyectos de entornos ubicuos como **Aura** [4] de la Universidad de Carnegie Mellon, cuyo objetivo principal es el de proveer a cada usuario con un halo invisible de servicios de información, más allá del lugar que se encuentre; **Gaia** [5] de la Universidad de Illinois donde se plantean que los espacios físicos se convierten en espacios activos a partir de la utilización de dispositivos ubicuos y proponen un sistema operativo para manejar todos estos elementos en conjunto; **Oxigen** [6] perteneciente al MIT (Massachusetts Institute of Technology) donde se trabaja para que los dispositivos sean incorporados en la vida humana y cotidiana de manera natural e imperceptible.

Si bien estos proyectos han sido un avance para la integración de los dispositivos, no se ha logrado hacer que los mismos trabajen de forma independiente y colaborativa para la obtención o realización de una tarea específica, sin necesidad de ser coordinados a través de un nodo central.

Para la coordinación de dispositivos (o sistemas) ubicuos existen algunos mecanismos, como por ejemplo los frameworks de desarrollo. Si bien estos frameworks han presentado un avance respecto de la integración de los dispositivos ubicuos con otros elementos de computación, no son suficientes para que los mismos puedan interactuar de manera independiente.

Existen actualmente también trabajos de investigación relacionados con la coordinación de dispositivos ubicuos en ambientes pervasivos, aunque los mismos no han logrado trabajar sobre el tema de esta tesis en profundidad.

Todas estas características hacen que la composición de dispositivos ubicuos se configure en un área de investigación muy importante donde los avances han sido limitados al día de hoy.

3. Metodología de Investigación

La computación orientada a servicios, y en particular los servicios web, proporcionan mecanismos para la composición de servicios. Las orquestaciones, por ejemplo, son mecanismos bien conocidos que permiten construir sistemas de negocio complejos a partir de una gran cantidad de servicios heterogéneos, simples y distribuidos.

Nuestro principal objetivo de investigación es *Definir un mecanismo de coordinación de dispositivos ubicuos que garantice su interoperabilidad independientemente del modelo y fabricante del mismo; utilizando los estándares de SOA y de coreografías para la composición de servicios*. Si pensamos que cada dispositivo ubicuo en un ambiente pervasivo es proveedor, o consumidor de un servicio, la coordinación de dispositivos parece encajar perfectamente con la composición de servicios. Sin embargo, en contextos tales como el de Internet de las cosas (IoT de las siglas del inglés Internet of

²Ambientes pervasivos son entornos poblados por varios dispositivos (sensores, actuadores, etc) y aplicaciones de software integrados de forma transparente [1]

Things), donde los servicios son dinámicos, móviles, menos fiables y dependientes del dispositivo, los mecanismos de composición establecidos para servicios web no son directamente aplicables [16].

Por otra parte, la composición en ambientes pervasivos, como es el caso de las redes de sensores, dispositivos *wearables*, etc., necesita hacer frente a las distintas contingencias que pueden ocurrir con estos elementos, sin descuidar la heterogeneidad de los mismos. No debemos olvidar, que los dispositivos en mención 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. Por tanto, en ambientes pervasivos, la disponibilidad y confiabilidad de los dispositivos no puede ser garantizada.

Si bien las semejanzas entre la tecnología SOA y su aplicabilidad a sistemas ubicuos es muy importante, como mencionábamos, la mera traslación de los conceptos no alcanza. Para ello es necesario adaptar los conceptos de SOA para poder dar lugar a soportar las características propias de los dispositivos ubicuos. Es sorprendente que esta similitud no haya sido apenas explorada con anterioridad. Únicamente en el trabajo de Sheng [2] se hace mención a la necesidad de más investigación en esta área. En miras de alcanzar este propósito, nos hemos planteado las siguientes preguntas de investigación:

1. ¿Es SOA capaz de actuar como mecanismo de coordinación de sistemas ubicuos?
2. ¿Es posible desarrollar un framework que permita incorporar SOA a dispositivos ubicuos?
3. ¿Es posible demostrar la aplicabilidad de SOA en ambientes ubicuos mediante una prueba de concepto?

De acuerdo a la naturaleza del problema y en base al objetivo planteado en el presente trabajo de investigación, la metodología seleccionada para llevar adelante el proyecto es “Design Science”. Design Science crea y evalúa artefactos de tecnologías de la información con la intención de dar solución a problemas debidamente identificados, tal como es expresado en [3]. Los artefactos que se crean o evalúan a partir de Design Science van desde software, lógica formal y matemática rigurosa, hasta descripciones informales en lenguaje natural.

En la Figura 1 podemos apreciar las fases que plantea la metodología Design Science asociados a las preguntas de investigación planteadas.

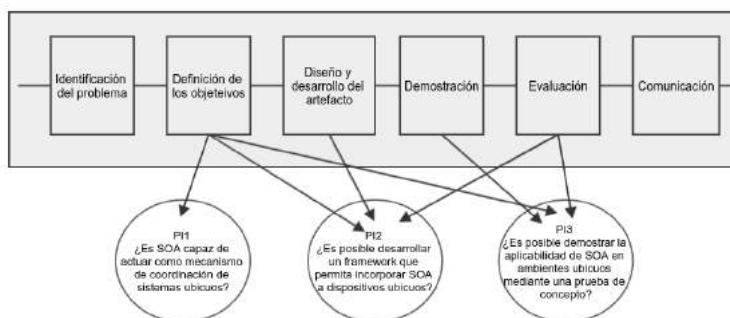


Figura 1: Metodología de investigación

4. Contribución

La principal contribución de este trabajo es el framework de ejecución de coreografías, incluyendo a los dispositivos ubicuos como parte importante de él, funciona de manera correcta y dando soporte a las características propias de los dispositivos ubicuos. La solución planteada es simple, interoperable y extensible. Hemos logrado trasladar, a un espacio o entorno donde las soluciones realizadas eran ad-hoc, conceptos de una teoría que existía únicamente en SOA, cuyo propósito es el de realizar sistemas

interoperables de forma estandarizada y transparente. Esto permite que los dispositivos ubicuos puedan cooperar a la hora de realizar sus tareas con ordenadores, teléfonos móviles, etc. que implementen la interfaz de SOA de forma transparente.

El framework de ejecución de coreografías³ se construyó en base a los lenguajes de programación PHP y C++. Estos lenguajes fueron seleccionados en función de los dispositivos que se utilizaron para la prueba de concepto. La base de la ejecución de coreografías se centra en una serie de clases que, en principio, leen la descripción en WS-CDL (del inglés Web Service Choreography Description Language) de la coreografía, y en base al dispositivo que lo ejecuta; determina en primer lugar en qué posición de la ejecución de la coreografía se encuentra para luego poder determinar cuál o cuáles son los pasos siguientes en la ejecución. Una vez determinados los pasos que se deben ejecutar, se hacen las invocaciones a otros dispositivos, teniendo en cuenta las actividades descritas en la definición de la coreografía en el lenguaje WS-CDL. Esta ejecución se hace de manera controlada, en el sentido de examinar que no se produzcan contingencias provenientes de las características de los dispositivos (desapariciones, latencia en la respuesta por falta de capacidad de procesamiento, etc.). En base a esta verificación se realizan las tareas correctivas correspondientes, según se han expresado en la definición de la coreografía. La estructura de ejecución del framework la podemos apreciar en la Figura 2. Allí se presentan los pasos que se llevan adelante dentro o desde un dispositivo que compone la coreografía, donde en primer lugar se determina cuál o cuáles son los pasos siguientes y los va ejecutando según corresponda. Como vemos la secuencia de ejecución es simple, clara y aplicable a cualquier dispositivo.

Para comprender mejor la estructura del framework desarrollado, mostraremos como ejemplo, un código escrito en el lenguaje de especificación de coreografías WS-CDL donde se aprecia la interacción entre dos dispositivos que forman parte de la coreografía⁴. Desde el primer dispositivo denominado VehiculoAccidentadoRole se produce una comunicación con otro dispositivo BalizaRole, éste último a su vez se relaciona con otro dispositivo CentralBalizaRole. Esta coordinación se pueden apreciar en las dos interacciones definidas. Si observamos en la Figura 2, el paso *Determinar siguiente* hace la tarea de fijarse en la definición XML de la coreografía a qué dispositivo debe llamar, en el ejemplo que mostramos debería leer la interacción que corresponde y fijarse cuál es el dispositivo que figura como *toRole*; si esto fuese el caso de la primer interacción del listado, debería invocar o coordinar con el dispositivo que se denomina BalizaRole.

```
<interaction name="reportarAccidente" operation="informarIncidente" >
  <participate relationshipType="tns:Vehiculo_Baliza" fromRole="tns:VehiculoAccidentadoRole" toRole="
    tns:BalizaRole" />
  <exchange action="request" name="informarIncidente" informationType="tns:avisoIncidenteType">
    <send variable="cdl:getVariable(tns:DatosIncidente , VehiculoAccidentadoRole)" />
    <receive variable="cdl:getVariable(tns:DatosIncidente , BalizaRole)" />
  </exchange>
</interaction>

<interaction name="publicarAccidente" operation="publicarIncidente">
  <participate relationshipType="tns:Baliza_CentralBaliza" fromRole="tns:BalizaRole" toRole="
    tns:CentralBalizasRole" />
  <exchange action="request" name="informarIncidente" informationType="tns:avisoIncidenteType">
    <send variable="cdl:getVariable(tns:DatosIncidente , BalizaRole)" />
    <receive variable="cdl:getVariable(tns:DatosIncidente , CentralBalizasRole)" />
  </exchange>
</interaction>
```

A continuación mostramos cómo se implementan los pasos de la ejecución de la coreografía, descritos en la figura 2, pero ya en el código propiamente dicho. En la figura 3 se pueden apreciar los diagramas de clases en C++; éstas clases son las encargadas de llevar adelante la ejecución de la coreografía. En este diagrama podemos apreciar también que se encuentran ya implementadas las clases que dan manejo a una transacción dentro de la ejecución.

Para poder hacer uso del framework para un caso específico se deben incluir las librerías que implementan las clases dentro del código específico de la coreografía que se desea llevar adelante. Además, se deben generar clases que heredan de las clases preexistentes en el framework. A continuación se muestra un template de una clase que implemente y haga uso del framework. Más específicamente,

³El código fuente del framework desarrollado, puede ser accedido a través de esta dirección: https://github.com/GRISE-UPM/ml_server_rest.

⁴Solamente se muestra un trozo de la totalidad de la especificación XML, a los fines prácticos

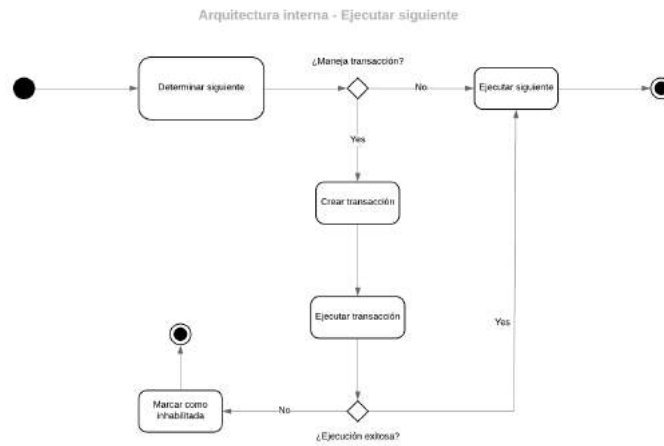


Figura 2: Diagrama de arquitectura de ejecución

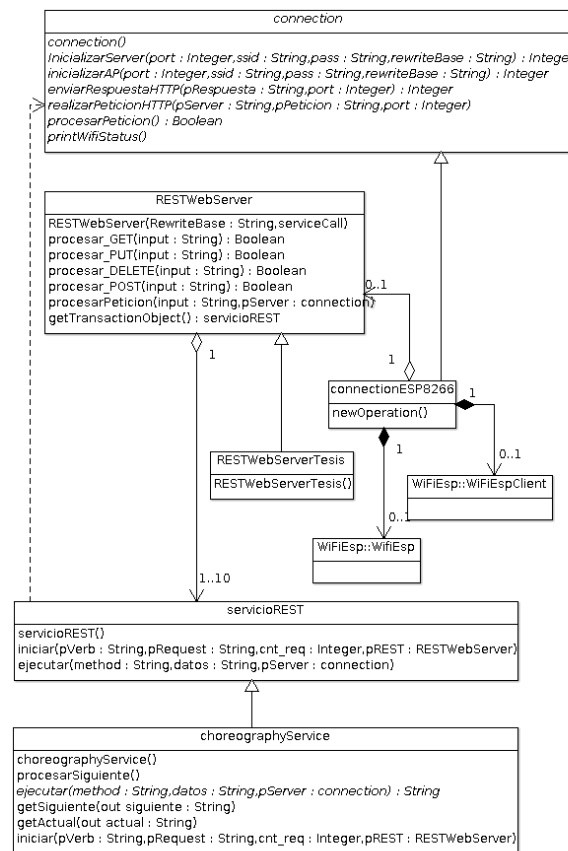


Figura 3: Diagrama de clases para ejecución de coreografías

Se deben incluir las librerías `connection` y `RESTWebServer`, y se deben implementar las clases constructoras, iniciarlas y ejecutarlas. En este caso mostramos el código que implementa el dispositivo `BalizaRole`, el cual es parte de las interacciones mostradas en la definición previa.

```
#include "ClaseX.h"
#include <connection.h>
#include <RESTWebServer.h>

BalizaRole::BalizaRole() : choreographyService(){
    chor_act_role = "BalizaRole";
}

return;

int BalizaRole::iniciar(){
    return choreographyService::iniciar();
}

void BalizaRole::ejecutar(){
    if (strcmp(method,"informarIncidente") == 0){
        if (this->_verb != 'O'){
            pServer->enviarRespuestaHTTP("405_Method_not_allowed","");
            return;
        }
        String respuesta = this->informarIncidente(method,datosIncidente);
        strcpy(sRespuesta, "{ \"resultado\": \"\" }");
        strcat(sRespuesta, respuesta.c_str());
        strcat(sRespuesta, \" \", \"token\": \"\");
        strcat(sRespuesta, _token);
        strcat(sRespuesta, \" } \");
        pServer->enviarRespuestaHTTP("200_OK",sRespuesta);

        // Llamo a que se ejecute lo que tenga que seguir de la coreografia
        choreographyService::ejecutar();
    }
}
```

El template mostrado corresponde a la programación en lenguaje C++; no obstante, la programación en PHP es muy similar a la presentada, incluso al ser un lenguaje de más alto nivel, su implementación es mucho más simple.

Para validar la funcionalidad del framework de coordinación, se ha seleccionado una prueba de concepto basada en un escenario de trabajo específico; no obstante, creemos que esta solución podría ser extrapolada hacia prácticamente cualquier ambiente de trabajo.

Se planteó un ambiente inteligente relacionado directamente con los problemas referentes al tráfico vehicular a lo largo de las autopistas, rutas o carreteras. Simulamos vehículos que transitan por una carretera, los cuales se comunican con balizas distribuidas a lo largo de la vía. Esta comunicación puede ser bi direccional, cuando el vehículo informa de un problema encontrado o bien que la baliza da aviso de posibles problemas a lo largo de la carretera. A su vez las balizas pueden comunicarse con centrales de coordinación o de emergencia según corresponda. Una aplicación real de este escenario podría ser que un ómnibus con pasajeros se traslada desde una ciudad a otra a través de una carretera normal (con dos carriles), el conductor del vehículo cuenta con un sensor que detecta la posibilidad de que esté por sufrir un ataque cardíaco, o más aún, que ya lo esté sufriendo en ese momento (este dispositivo puede ser una placa de desarrollo Arduino). A partir de ese instante, el vehículo le avisa al conductor (ya sea visualmente como auditivamente), a su vez intenta comunicarse con una baliza (la cual puede ser representada por una placa Arduino) de la carretera para que la ayuda llegue lo antes posible al lugar del evento. También podría disparar alertas a los vehículos cercanos, para que los mismos puedan tomar acciones preventivas, además de avisar al pasaje del ómnibus y detener la marcha en caso de ser necesario.

Los dispositivos utilizado para la construcción de la prueba de concepto son los siguientes: Equipo servidor, Equipo Laptop, Equipo RaspberryPi B+ y Placas de desarrollo Arduino Mega 2560 y Arduino Nano V3.

El diagrama de secuencia de la Fig. 4 representa la ejecución de la coreografía, en base a la prueba de concepto definida. Esta figura se genera automáticamente desde el mismo framework ya que por cada ejecución que se hace de un caso de prueba, se genera un código de identificación único que es informado en la ejecución para luego poder obtener el gráfico correspondiente. En el diagrama se pueden observar, como etiquetas en la secuencia de mensajes, el formato de información que viaja desde un dispositivo hacia otro, el cual está definido dentro de la definición de la coreografía realizada en WS-CDL. De acuerdo al extracto en XML que mostramos previamente, las dos interacciones que allí presentamos

se ven representadas en este diagrama como las interacciones marcadas con las numeraciones 1, 2, 3, 7 y 10 respectivamente.

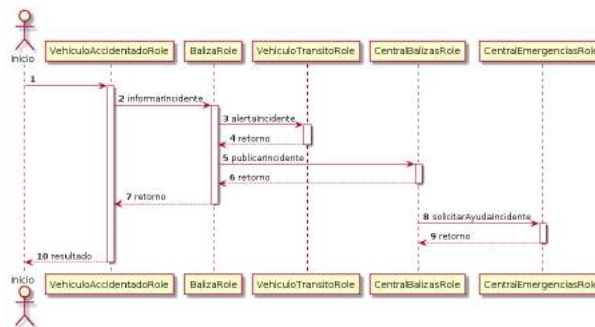


Figura 4: Diagrama de secuencia de ejecución de la coreografía

Este escenario que se plantea puede ser fácilmente abstraído a otras situaciones como puede ser un sistema de alarma hogareño, coordinación de una flota de distribución de mercadería, provisión de servicios a pasajeros en tránsito dentro de un aeropuerto, etc.

5. Conclusiones y Trabajo Futuro

La solución planteada, esto es, la utilización del estándar WSCDL de coreografías para SOA, junto con el framework desarrollado, ha resuelto satisfactoriamente el problema planteado.

El principal problema que la presente tesis ha afrontado, el cual está descrito en detalle en la Sección de Planteamiento del problema, consistía en que la composición de dispositivos se realizaba a través de protocolos propietarios y sin seguir definiciones estándares. Esto ha provocado que dispositivos de distintos proveedores no puedan ser utilizados en una composición, presentando limitaciones importantes a la hora de realizar composiciones con estos dispositivos.

La solución planteada permite realizar composiciones con dispositivos ubicuos de una manera abierta, basada en estándares y escalable. Esto implica que diversos dispositivos, de diversos fabricantes, y con distintas capacidades, pueden ser incorporados fácilmente al framework y, por consiguiente, participar en coreografías con otros dispositivos de forma sencilla.

El framework desarrollado no posee gran complejidad. Es bastante simple desde el punto de vista de diseño, y no demasiado exigente en términos de procesamiento y memoria. Por ello, creemos que la solución planteada puede aplicarse en sectores comerciales como, por ejemplo, el de sistemas de seguridad hogareña, donde se podría utilizar el framework desarrollado para implementar sistemas más sofisticados que los actuales, con la posibilidad de manejar una mayor cantidad de dispositivos, realizando tareas más complejas que la generación de una simple señal de activación, como ocurre hoy en día. Otro sector en donde podría aplicarse el framework desarrollado es en la atención domiciliar de pacientes con enfermedades que deban ser monitorizados a distancia, incluso con posibilidad de poder aplicar medicación en caso de ser necesario.

La solución planteada también es aplicable a IoT, donde los dispositivos no solamente estarían conectados entre sí, sino que también podrían utilizar otros protocolos de comunicación más allá de TCP/IP.

La solución propuesta permite, adicionalmente que los dispositivos ubicuos pueden interactuar no sólo con otros dispositivos ubicuos, sino también con aplicaciones basadas en SOA. De esta manera, los dispositivos pueden ser integrados en sistemas existentes ampliando enormemente sus capacidades.

Finalmente, la solución propuesta no es, ni mucho menos, completa. Es posible, y deseable, realizar toda una serie de mejoras y ampliaciones que describimos en la continuación.

De la mano con los hallazgos alcanzados, se han abierto tanto nuevas líneas de investigación como también posibilidades de producción de software de uso industrial. A continuación reportamos las futuras opciones de investigación y desarrollo que esta tesis ha propiciado.

5.1. Líneas futuras de investigación

- Descubrimiento de servicios: Durante la presente investigación se ha trabajado sobre servicios y dispositivos ubicuos estáticos, es decir, preestablecidos dentro de la coreografía y presentes en la red de comunicaciones antes del inicio de cualquier interacción. Resultaría una línea de investigación muy interesante la de poder coordinar a través de coreografías dispositivos ubicuos descubiertos “on the fly”.
- Introducción de capas de seguridad: Una línea de investigación que también resultaría en un notable progreso a este trabajo de tesis es la relacionada con la introducción dentro del framework de coordinación de las capas de seguridad existentes en servicios web como es el caso de WS-SECURITY. La inclusión de esta capa de seguridad significa un desafío importante ya que se debe estudiar de qué manera es posible incorporarla dentro de dispositivos con poca capacidad de procesamiento y escasa memoria.
- Implementación de transacciones de acuerdo a estándares: Otra línea de investigación, relacionada en cierta forma con la anterior, es la de implementar la totalidad de los estándares de transacciones disponibles en SOA dentro del framework de coordinación propuesto en esta tesis. En esta investigación se realizó la implementación de una transacción distribuida, pero sin ajustarse en su totalidad a los estándares WS-TRANSACTION y WS-COORDINATION.
- Convergencia con microservicios: Hemos mencionado durante la sección de Discusión que existe una relación bastante estrecha entre esta tesis y la tecnología de microservicios. En microservicios se discute sobre la forma en que se pueden coordinar los servicios para llevar adelante tareas en conjunto, donde una de las formas presentadas es la de coreografías. Por lo tanto, una futura línea de investigación consistiría en la convergencia entre los conceptos emanados de este trabajo de Tesis y la tecnología de microservicios.
- Convergencia con IoT: Como hemos indicado en la sección de Discusión, IoT utiliza únicamente el protocolo TCP/IP para la interconectividad de los distintos componentes. También hemos mencionado que es una limitación que no se ha afrontado en el framework desarrollado en esta tesis. Por lo tanto, una futura línea de investigación consistiría en la convergencia entre ambas tecnologías ampliándolas a la utilización de otros protocolos como puede ser Bluetooth o RFID, por mencionar dos ejemplos destacados.

5.2. Líneas futuras de desarrollo

- Implementación del framework de ejecución de coreografías sobre dispositivos ubicuos con calidad industrial: Para poder llevar adelante la investigación fue necesario implementar un framework de coordinación de dispositivos ubicuos tal y como exige la metodología de “Design science” elegida. Si bien ello fue suficiente para los fines de la presente investigación, el prototipo desarrollado no alcanza la calidad necesaria para ser aplicado efectivamente sobre problemas de coordinación de dispositivos ubicuos en la industria.

Aunque sí deja las bases sentadas para que el desarrollo siga evolucionando hasta poder producir un framework sólido y de características industriales para que el mismo pueda ser utilizado en entornos reales.

- Ampliación de las pilas de protocolos: El objetivo en este caso sería mejorar la conectividad de los dispositivos que forman parte de una coreografía. Para ello se debería extender el framework para utilizar otros protocolos de comunicaciones, como puede ser el caso de la utilización de protocolos Bluetooth, RFID, etc. Esta mejora permitiría ejecutar coreografías no sólo con una mejor diversidad de dispositivos ubicuos sino también probablemente de forma más versátil, aprovechando las posibilidades que dichos protocolos ofrecen, ej: la comunicación a través de redes de sensores.

Referencias

- [1] M. Weiser, “Hot topics-ubiquitous computing,” *Computer*, vol. 26, pp. 71–72, Oct 1993.
- [2] 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.
- [3] S. T. March and G. F. Smith, “Design and Natural Science Research on Information Technology,” *Decis. Support Syst.*, vol. 15, pp. 251–266, Dec. 1995.
- [4] J. Harkes, T. Farbacher, and N. Miller, “Project Aura Distraction-free Ubiquitous Computing.” Available at <http://www.cs.cmu.edu/~./aura/people.html>, Last visited: Aug 15th, 2002.
- [5] R. H. Campbell, D. M. Mickunas, D. Reed, and K. Nahrstedt, “Active Spaces for Ubiquitous Computing.” Available at <http://gaia.cs.illinois.edu/>, Last visited: Aug 15th, 2002.
- [6] M. L. f. C. Science and M. A. I. Laboratory, “Pervasive Human-Centered Computing.” Available at <http://oxygen.csail.mit.edu/>, Last visited: Aug 15th, 2002.
- [7] “Amigo Project.” Available at <http://gforge.inria.fr/projects/amigo/>.
- [8] 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.
- [9] S. Najjar, 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).
- [10] 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.
- [11] 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.
- [12] S. Cherrier, Y. M. Ghamri-Doudane, S. Lohier, and G. Roussel, “Services collaboration in wireless sensor and actuator networks: Orchestration versus choreography,” in *2012 IEEE Symposium on Computers and Communications (ISCC)*, pp. 000411–000418, July 2012.
- [13] C. Duhart, P. Sauvage, and C. Bertelle, “Emma: A resource oriented framework for service choreography over wireless sensor and actor networks,” *International Journal of Wireless Information Networks*, 06 2015.
- [14] L. Mostarda, S. Marinovic, and N. Dulay, “Distributed orchestration of pervasive services,” in *2010 24th IEEE International Conference on Advanced Information Networking and Applications*, pp. 166–173, April 2010.
- [15] Z. Zhou, D. Zhao, L. Liu, and P. C. Hung, “Energy-aware composition for wireless sensor networks as a service,” *Future Generation Computer Systems*, vol. 80, pp. 299 – 310, 2018.
- [16] 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.

Framework para la evaluación de calidad de procesos ágiles

Resumen de Tesis presentada para obtener el Doctorado en Ciencias Informáticas
en la Facultad de Informática de UNLP

Dra. Noelia Pinto

ns.pinto@ca.fre.utn.edu.ar

Centro de Investigación Aplicada a TIC (CInApTIC). UTN FRRe

Director Dr. César Acuña

Centro de Investigación Aplicada a TIC (CInApTIC). UTN FRRe

Codirector Dr. Gustavo Rossi

Centro de Investigación LIFIA Facultad de Informática UNLP

19 de Noviembre de 2020

Resumen

La industria del software requiere de productos y servicios de alta calidad, puede lograrse mediante la aplicación de modelos y metodologías de calidad reconocidos internacionalmente. Sin embargo, estos modelos en pequeñas y medianas empresas (PYMES) son muy difíciles de implementar ya que ello implica una gran inversión en dinero, tiempo y recursos.

Por ello, resulta necesario establecer estrategias que permitan la automatización del ciclo de desarrollo de software, de forma tal de promover la reducción de costos y la eficiencia en los procesos de obtención de productos finales. En este caso, la filosofía ágil resulta ser el enfoque más adecuado para los entornos de desarrollo actuales, y se posicionan como alternativa a los procesos de desarrollo con alto costo en documentación y procesos excesivamente prolongados.

Con el objetivo de facilitar la adopción de prácticas ágiles en las PYMES, promoviendo el aseguramiento de la calidad de los procesos de desarrollo de software, se presenta Agile Quality Framework (AQF), un framework cuya propuesta integra el diseño y desarrollo de un modelo que permita evaluar la calidad en procesos ágiles de software, y una plataforma que permitirá el seguimiento de proyectos de software ágiles junto a la evaluación sistemática de calidad del proceso de desarrollo.

AQF surge, como una plataforma que contribuye con los equipos de desarrollo de software a partir de la evaluación de calidad en proyectos ágiles, considerando como objeto de la medición al proceso de desarrollo independientemente del enfoque ágil seleccionado.

Palabras Claves: *Calidad de Software; Procesos ágiles de Desarrollo de Software; Evaluación de Calidad*

1. Introducción

La mejora e innovación de los procesos de software, con el objetivo de incrementar la calidad de sus productos y servicios, se ha convertido, en los últimos años, en el elemento diferenciador que las empresas necesitan para mejorar sus niveles de competitividad en la Industria del Software. Sin embargo, diversos estudios coinciden en la dificultad de las PYMES (Mas., Amengual., 2005)(Pasini, Esponda, Bertone & Pesado, 2008) (Pfleger, 2002) para implementar programas de Mejoras de Proceso de Software (Software Process Improvement - SPI), fundamentalmente porque la aplicación de estos modelos resulta costosa en términos económicos y de esfuerzo, pues requieren una gran inversión en dinero, tiempo y recursos, sus recomendaciones son complejas de aplicar y el retorno de la inversión se produce a muy largo plazo (Garzías, Fernández & Piattini, 2009). De esta forma, los

parámetros de tiempos de desarrollo y costo de soluciones afectarán directamente al trabajo que se realice, siendo la calidad la primera variable de ajuste disponible.

Dado que la calidad del producto de software desarrollado está estrechamente relacionada con la calidad del proceso utilizado, las PyMEs necesitan implementar proyectos para la mejora de sus procesos que le permitan incrementar la calidad de sus productos. De esta forma, analizando la situación de la Industria del Software en el NEA (Región Nordeste de Argentina) respecto a la adopción del ciclo de vida que guíe los procesos de desarrollo de las empresas, surge la necesidad de proporcionar un marco de trabajo que permita evaluar la calidad cuando optan por trabajar con enfoques ágiles (Acuña, Cuenca Pletsch, Tomaselli, Pinto & Tortosa; 2015).

Habiendo realizado un análisis de la literatura disponible se observa que, en su mayoría, los estudios se centran en analizar la implementación de prácticas ágiles, sin evaluar la relación entre el proceso de desarrollo de software y la calidad. Así, y teniendo en cuenta las particularidades de organizaciones tipo PYMES, resultados de estudios previos realizados (Rujana, Romero Franco, Tortosa, Tomaselli, & Pinto, 2016), demuestran la ausencia de estrategias que permitan a las empresas integrar agilidad a sus ciclos de desarrollo sin dejar de lado aspectos relacionados a la calidad de software.

En términos prácticos, si bien existen numerosos estudios que demuestran una creciente adopción de estos enfoques en empresas de software (Version One 12th Annual State of Agile™ Report, 2018), el rápido crecimiento de la agilidad ha generado confusión, malas interpretaciones e incluso efectos negativos en el desarrollo de algunos proyectos de software (Bollati, V. A., Gaona, G., Pletsch, L. C., Gonnet, S., & Leone, H., 2017). Este escenario implica, por lo tanto, la búsqueda de estrategias que permitan no solo guiar a las empresas en la adopción de las prácticas ágiles sino también en cuestiones relacionadas a la gestión de la calidad que contribuyan en la generación de valor en la producción de software.

En este sentido, y como primera alternativa, se ha implementado y validado un marco de trabajo que permite evaluar la calidad cuando se opta por trabajar con procesos ágiles de desarrollo de software. Dicho framework se denomina AQF (Agile Quality Framework) y su versión actual está formada por un modelo, QuAM (Quality Agile Model), compuesto por métricas, atributos y criterios que permiten medir los niveles de calidad asociados a las prácticas ágiles, y por QuAGI (Quality AGile), una herramienta de software que brinda soporte a dicho modelo a través de la automatización del seguimiento de proyectos y la visualización de diferentes informes (Pinto, Acuña, Tortosa & Cabas Geat; 2018).

Cabe destacar que este desarrollo constituye uno de los objetivos de la tesis doctoral denominada “Framework para la evaluación de calidad de procesos ágiles”, realizada en el marco del proyecto “Evaluación de Calidad en Procesos Ágiles de Desarrollo de Software”, financiado por la UTN y ejecutado en el Centro de Investigación aplicada en TICS (CInApTIC) de la Facultad Regional Resistencia, con el código IAI4445TC.

1.1. Hipótesis y Objetivos de la Tesis

Se ha definido para la tesis la Hipótesis de investigación como sigue:

“Será factible mejorar la calidad de los procesos de desarrollo de software guiados por prácticas ágiles a través de la evaluación automática de la calidad de dichos procesos utilizando un

framework compuesto por un modelo de calidad y una herramienta que automatice la gestión de dicho modelo”

El Objetivo General de esta tesis doctoral ha consistido en: “*Proponer un framework que facilite el seguimiento de proyectos y la evaluación de calidad de procesos en empresas PYMES que implementen prácticas ágiles en el desarrollo de software”*.”

Para conseguirlo, se establecieron los siguientes objetivos específicos:

1. Realizar un estudio de trabajos previos relacionados a modelos de calidad y herramientas de software aplicables a procesos ágiles en contextos de empresas PYMES.
2. Proponer un modelo de evaluación de calidad sobre procesos de desarrollo de software guiados por prácticas ágiles, en base a la definición de un conjunto de componentes.
3. Diseñar e implementar herramientas de software que soporten la gestión de componentes del modelo de calidad e integren un framework que evalúe la calidad en procesos ágiles.
4. Analizar y estudiar el comportamiento del framework propuesto utilizando un método de validación sobre entornos reales de producción de software.

1.2. Enfoque Metodológico

Teniendo en cuenta las características de esta tesis, luego de analizar modelos tradicionales y otros más específicos, el método de investigación que se ha seguido constituyó una adaptación del propuesto por Marcos y Marcos (Marcos, E., & Marcos, A., 1999) para la investigación en Ingeniería del Software. El método genérico es un método general de trabajo, basado en los pasos a seguir, según Bunge (Bunge, M, 1989) en cada investigación científica. Aunque estos pasos se basan en el método deductivo hipotético, debido a su generalidad, son aplicables, con ciertas modificaciones, a cualquier tipo de investigación e incluso a la resolución de cualquier tipo de problema que intentemos resolver.

Particularmente, durante la tesis las fases de *Resolución* y *Validación* se han llevado a cabo siguiendo la teoría de la Ingeniería de Software Empírica, es decir la rama de la Ingeniería de Software que se basa en la experimentación como método para corresponder ideas o teorías con la realidad, la cual refiere a mostrar con hechos las especulaciones, suposiciones y creencias sobre la construcción de software (Wohlin, C et al, 2000). Por un lado, la *Resolución* comprende todas las actividades que conducen a lograr el producto final que, para la tesis, ha significado la obtención del framework AQF. Y por otro lado la *Validación* implica el proceso del estudio, análisis y observación de la implementación del framework sobre proyectos ágiles reales.

Tanto para el proceso de Resolución como el de Validación, se han utilizado dos de las técnicas más relevantes de la Ingeniería de Software Empírica: la encuesta y el estudio de casos.

Las encuestas se han realizado sobre una muestra representativa de la población, y luego los resultados han sido generalizados al resto de la población. Su utilización ha permitido obtener datos, primero para el diagnóstico de la situación actual y a posteriori para el estudio de la reacción de la población analizada al implementar el framework AQF en sus proyectos.

Durante la etapa de *Validación* se han realizado estudios de casos, un instrumento adecuado para muchos tipos de investigación de ingeniería de software, ya que los objetos de estudio (en este caso proyectos ágiles de software), son fenómenos contemporáneos, que son difíciles de estudiar de forma aislada (Kitchenham, B., Pickard, L., & Pfleeger, S. L., 1995) Este método ha permitido la

observación de diversas variables y su comportamiento en cada uno de los equipos donde se han puesto los proyectos ágiles bajo evaluación del framework AQF.

Cabe destacar también que los estudios de caso no se realizaron en simultáneo, por lo que el carácter iterativo del proceso de *Validación* ha permitido que los resultados de cada uno de los estudios de casos sirvan como punto de control de lo realizado hasta el momento, pudiendo realizar los ajustes que se consideren necesario tanto sobre el modelo como sobre la herramienta que conforman el framework AQF para un próximo estudio.

2. AQF: Desarrollo de la propuesta

El aporte significativo de AQF se centra en que el nivel de calidad de un proyecto ágil no sea definido únicamente a través de un valor numérico, sino que sea el resultado también de la evaluación de otros aspectos, en base al contexto en el que se desarrolla el proceso y los factores que lo impactan.

Aquí, es necesario aclarar que, cuando se hace referencia al término “Proyecto Ágil”, se incluye a aquéllos que implementan prácticas ágiles para el proceso de desarrollo del producto software en base a cualquier enfoque tal como Scrum, Kanban, Lean o híbridos.

El framework AQF está compuesto por:

- Un *modelo de calidad*, denominado QuAM (Quality Agile Model) que consta de:
 - Una estructura formal definida en función a componentes, atributos, métricas y criterios relacionados entre sí.
 - Una estrategia de medición que establece el nivel de calidad asociado a cada componente.
- Una herramienta de software, QuAGI (Quality AGile) que da soporte al modelo, constituyendo un medio para su validación, y se compone de:
 - Una arquitectura que define cada uno de los componentes que forman la herramienta.
 - Una representación lógica que permite implementar el proceso de medición sobre cada uno de los componentes del modelo QuAM.
 - Un conjunto de interfaces que facilitan la interacción de actores externos (tales como administrador del proyecto, equipo de desarrollo, etc) con la herramienta.

2.1. QuAM: El modelo de calidad

QuAM (Quality Agile Model), es un modelo conceptual cuyo objetivo es describir el conjunto de componentes que influye en la calidad de proyectos ágiles junto a un procedimiento de evaluación de calidad que los integra. QuAM permite asociar los resultados de la medición con un perfil de calidad asociado al proyecto ágil evaluado (Pinto, Acuña, Cuenca Pletsch; 2016). El modelo QuAM surge de un relevamiento sistemático de la literatura y del análisis del contexto actual a partir del relevamiento de información realizado en la Industria del software en el nordeste argentino (NEA) del cual se han obtenido factores de relevancia que fueron incluidos al esquema de componentes de QuAM. Tal como se observa en la Figura 1, la estructura del modelo incluye 4 componentes, formados por atributos, cada uno de los cuales se mide a través de métricas directas e indirectas.



Figura 1. Componentes y atributos del modelo de calidad QuAM

Sin embargo, para la evaluación de calidad de proyectos ágiles que propone AQF no basta con la definición de un modelo, sino que resulta necesario automatizar la gestión de los componentes de QuAM a partir de una herramienta que permita el seguimiento de proyectos acompañado por la obtención de informes parciales respecto de la calidad del proceso ágil que lo caracteriza. Por ello, a fin de complementar el framework AQF, durante la tesis se abordó el desarrollo de QuAGI, una aplicación web que permite la gestión de los componentes de QuAM a través de la medición y obtención automática de los valores correspondientes a cada una de las métricas a lo largo del proceso de desarrollo. El objetivo central de esta herramienta es permitir el seguimiento online de proyectos basados en prácticas ágiles junto a la posibilidad de realizar evaluaciones continuas respecto al nivel de calidad que se esté logrando en el proceso.

2.2. QuAGI: Plataforma web para seguimiento de proyectos ágiles

QuAGI fue diseñada como una aplicación web escalable de forma tal que, permita añadir nuevas funcionalidades de manera rápida y sencilla a partir de reportes de feedback que surjan luego de escenarios de validación que se describirán en el capítulo siguiente de esta tesis.

Por lo tanto, QuAGI permite administrar proyectos ágiles, realizar el seguimiento de cada uno de ellos, obtener diversos informes y constituye una herramienta de comunicación interna entre los integrantes del equipo. Por último, da soporte a los procesos de toma de decisiones asistiendo a los responsables mediante reportes que informen sobre evaluación de calidad del proyecto y recomendaciones de ajustes para la mejora continua (Pinto, Tortosa, Cabas Geat, Ibáñez & Acuña; 2018).

La arquitectura propuesta para QuAGI consiste en una plataforma basada en componentes reutilizables, permitiendo así que las aplicaciones que se integren a la plataforma hagan uso de dichos componentes. Tal como se observa en la Figura 2, cada uno de los componentes de QuAGI representa una vista particular de su arquitectura, y esta configuración propuesta permite asegurar el mantenimiento de la trazabilidad entre los diferentes artefactos desarrollados, la reutilización de estos y el mejor control de su evolución a la hora de incorporar nuevas funcionalidades.



Figura 2. Arquitectura conceptual de QuAGI

Considerando que existen actualmente en el mercado muchas herramientas que permiten gestionar proyectos de software (por ejemplo, Trello, Taiga.io, Jira, etc), QuAGI, a diferencia de las existentes, además forma parte del framework AQF, por lo que no solo permite el seguimiento online de los proyectos de software, sino que también da soporte a lo establecido en un modelo de calidad propio y permite, entonces, realizar evaluación de calidad del proyecto en cuestión. Así, actualmente QuAGI permite: Configuración inicial del proyecto ágil, Gestión del Product Backlog, Evaluación de calidad y Emisión de reportes.

Cabe destacar que QuAGI almacena constantemente la interacción de los usuarios con la herramienta; por lo que, cuando se requiere conocer sobre el nivel de calidad del proyecto, QuAGI recupera dicha información, realiza las mediciones de todos los componentes, definidos por QuAM, y obtiene la evaluación de forma automática. Una de las contribuciones más importantes derivadas de esta tesis se enfoca en la posibilidad de incorporar, al seguimiento de proyectos ágiles, el proceso de evaluación de calidad asociado al ciclo de desarrollo que se pretende gestionar.

Finalmente, resulta importante destacar que, como resultado del desarrollo de esta plataforma web QuAGI la misma ha sido registrada, durante el desarrollo de la tesis, como Obra Inédita de Software ante la Dirección Nacional de Derecho de Autor dependiente del Ministerio de Justicia y Derechos Humanos de la República Argentina¹.

2.3. Valor agregado de AQF

La propuesta del framework AQF se destaca por ciertos aspectos que no solo benefician a la calidad de los procesos de desarrollo, sino que, asimismo, ayudan a tomar decisiones que añaden valor agregado a los proyectos de software.

- **Propuesta integral**

Un aspecto clave de AQF, es que representa una propuesta que integra el seguimiento de un proyecto de software y la evaluación de calidad del proceso ágil asociado, en una única plataforma.

¹ El Registro se ha publicado en el Boletín Oficial del Poder Ejecutivo Nacional el 25 de junio de 2019, con el Número RE-2019-57162220-APN-DNDA-#MJ

A diferencia de trabajos relacionados, que se enfocan unilateralmente en algunos aspectos del ciclo de desarrollo de software, AQF observa el proceso ágil desde 4 factores: producción de entregables, equipo de trabajo, comunicación con el cliente y gestión de requerimientos, coincidiendo con los valores de la filosofía ágil en la gestión de proyectos. AQF, obtiene una visión integral de la implementación del proceso ágil, a lo largo de todo el ciclo de desarrollo, reforzando la información de la evaluación de calidad obtenida con informes útiles para la toma de decisiones.

- **Procedimiento de Evaluación**

AQF ofrece la definición detallada de un procedimiento de evaluación integral, que basándose en QuAM, permite comprender el significado del nivel de calidad del proceso ágil asociado a su capacidad de producción de entregables del equipo, a la evaluación del equipo de trabajo, a la comunicación con el cliente y a la gestión de requerimientos. De hecho, el aspecto diferenciador de este procedimiento de evaluación es que la calidad de un proyecto ágil no es un valor binario, sino que puede darse en un espectro de valores que seguramente estén condicionados por condiciones del contexto donde el proyecto se lleva a cabo.

- **Visibilidad del trabajo de equipo**

Otro beneficio de AQF, frente a otras propuestas disponibles en la literatura, es que ofrece mayor visibilidad del trabajo asignado al equipo. Esta característica exalta uno de los principios ágiles: la transparencia, haciendo hincapié en la importancia del acceso y conocimiento sobre la información que se maneja de cada proyecto. QuAGI propone, entonces, una visión holista de cada proyecto, mostrando, no solamente la actividad y quién tiene asignado cada tarea, sino promoviendo que cada integrante del equipo conozca todo el trabajo no finalizado en el que está participando.

- **Gestión del conocimiento**

AQF, a partir de la interacción de usuarios con QuAGI, permite generar conocimiento útil que se transformarán, luego, en lecciones aprendidas para toma de decisiones en proyectos actuales y futuros proyectos ágiles. A partir de los datos y de las evidencias generadas por el uso de QuAGI, se podrían analizar indicadores que permitan exponer las bondades y debilidades de los elementos intervinientes en los procesos ágiles de desarrollo de software de los proyectos evaluados. Este aporte de conocimiento permite el monitoreo y evaluación de los logros obtenidos por la ejecución de determinadas prácticas ágiles; fomentando la creación de una cultura orientada al autoaprendizaje y socialización, para permitir que las buenas prácticas y las lecciones aprendidas sean de total conocimiento de los participantes en la mejora continua de los procesos de desarrollo de software.

3. Resultados obtenidos

Para validar la propuesta presentada en la tesis, se ha elegido trabajar con estudios de caso para que la validación de la propuesta se realice sobre entornos reales de proyectos ágiles en ejecución, de forma tal de obtener retroalimentación de los potenciales usuarios de QuAGI y al mismo tiempo poner en práctica toda la definición del método de evaluación propuesto por QuAM, lo que en sí mismo ya es un mecanismo de mejora. Además, los estudios de casos son una herramienta

empírica que provee de cierto grado de libertad para definir contextos que puedan adecuarse a las necesidades de validación que fueran surgiendo.

Se han ejecutado 3 estudios de casos para validar QuAGI a partir del seguimiento de proyectos ágiles reales y su evaluación de calidad en base a los componentes establecidos por QuAM².

A modo de resumen, se destaca que en el estudio de caso N° 1 se llevó adelante la validación de la herramienta en un primer sprint de proyectos ágiles en 20 equipos de desarrollo de empresas PYMES de software. En función al uso de la herramienta QuAGI se realizó una encuesta a los equipos de forma tal de validar la estructura de QuAM analizando la utilidad de cada componente en la evaluación de calidad de procesos ágiles. Del estudio se concluyó que un 70% de los equipos participantes consideran útil la herramienta QuAGI para la evaluación de calidad de procesos ágiles gracias a que facilita el seguimiento de la estructura del modelo tal como se define en QuAM. En el caso del 30% que consideró de poca o nula ayuda utilizar QuAGI, se pudo relevar que se debió a la falta de informes representativos del proceso asociado al trabajo en equipo, factor fundamental para el éxito de proyectos ágiles de acuerdo con su perspectiva.

Los resultados del estudio de caso N° 1 dieron lugar al diseño del estudio de caso N° 2, el cual tuvo por objetivo validar la información que QuAGI ofrece para mejorar la calidad de trabajo en equipo en proyectos ágiles. El estudio de caso N° 2 estuvo constituido por 90 personas pertenecientes a 20 equipos, y se enfocó en los resultados proporcionados por el componente N° 2 de QuAM “Evaluación del trabajo en equipo”. Los resultados que se lograron, luego de analizar los informes proporcionados por QuAGI y recomendaciones realizadas durante reuniones posteriores a la implementación, fueron muy superiores a los obtenidos en un primer control sin aplicar las mejoras indicadas por la herramienta. No solo se observaron mejoras a nivel del trabajo en equipo, sino que esto impactó favorablemente a los demás componentes considerados al evaluar la calidad del proceso asociado al proyecto ágil correspondiente. La experiencia aportó además mejoras que eran necesarias a nivel de interfaz para continuar añadiendo valor a QuAGI como herramienta de seguimiento.

Finalmente, se diseñó un nuevo estudio de caso que facilite la validación integral de QuAGI teniendo en cuenta la evaluación de todos los componentes de QuAM. El estudio de caso N° 3 se caracterizó por analizar el rendimiento completo de QuAGI en equipos sin experiencia previa con otras herramientas de seguimiento en proyectos ágiles. Además, se incluyó en el estudio el análisis de la calidad de proyectos ágiles en equipos que utilizaron otra herramienta diferente con el objetivo de comparar resultados. Durante la experiencia se realizaron observaciones en forma gradual y en diferentes momentos acordados desde el inicio, registrándose cada caso para su posterior análisis. Concluyendo el estudio de caso, los resultados que se obtuvieron, han sido mucho mejores para la calidad de los procesos en los equipos que utilizaron QuAGI en sus proyectos ágiles no solo a nivel de gestión sino también relacionados a curva de aprendizaje, experiencia de usuario y utilidad.

Además, el proceso de validación de la propuesta de la tesis ha permitido obtener conclusiones y adquirir experiencias de las que se desprenden ciertas lecciones aprendidas. En primer lugar, luego de haberse ejecutado todos los estudios de caso se destaca cómo las recomendaciones ofrecidas por QuAGI en conjunto con el seguimiento de los componentes definidos en QuAM permite mejorar notoriamente la calidad de los procesos ágiles de cada proyecto. Del análisis se destaca

² Y las características y particularidades de cada caso se detallan en profundidad en la tesis disponible en <http://sedici.unlp.edu.ar/handle/10915/109769>

también que los participantes subrayan entre los beneficios de QuAGI que sea una herramienta fácil de usar, útil y, sobre todo, manifiestan intención de usarla en el futuro pues la información obtenida resulta de vital importancia para el éxito de sus proyectos. Todo esto no hubiera sido posible sin analizar la aplicabilidad real de la evaluación de calidad de proyectos ágiles mediante el estudio de componentes definidos en QuAM y el seguimiento automático a través de QuAGI, logrado gracias al proceso de ejecución de estudios de casos.

En segundo lugar, la realización de esta experiencia de validación ha permitido, además, analizar el grado de adecuación, completitud y corrección de los resultados en los informes de calidad aportados por QuAGI; logrando modificaciones que aporten valor en cada instancia de validación y favorezca la obtención de la mejor versión de la herramienta.

Y, por último, pero no menos importante, el estudio de casos sobre proyectos ágiles reales ha posibilitado obtener realimentación de los participantes que aplicaron el modelo y usaron la herramienta. Esto no solo colaboró con mejoras a nivel de interfaz y usabilidad de la herramienta QuAGI, sino que también permitió añadir valor a través de mejoras en la configuración de los componentes de QuAM.

4. Conclusiones y Trabajos Futuros

La tesis doctoral que se presenta en este artículo tiene como objetivo principal el diseño, desarrollo e implementación del framework, AQF, para la evaluación de calidad en procesos ágiles de desarrollo de software.

El trabajo de esta tesis ha permitido integrar en AQF la presentación de un nuevo Modelo de Calidad, QuAM; y la implementación de una herramienta de software, QuAGI, que da soporte automático a lo establecido en QuAM. Y, mejor aún, la propuesta ha sido adaptada, en base a estudios de casos de validación, para que pueda utilizarse en entornos reales de producción con proyectos de software cuyo desarrollo es guiado por prácticas ágiles.

Como trabajos futuros se pretende llevar a cabo más casos de estudio que permitan lograr la versión que se adecúe, en mayor medida, a la realidad de las PyMES, incorporando sus prácticas más comunes y permitiendo obtener el nivel de calidad más representativo a cada proyecto ágil que se evalúe a través de AQF. Además, resulta necesario incorporar al estudio de validación el resto de los componentes para analizar el funcionamiento integral del framework.

Además, a pedido de las empresas, se continuará expandiendo AQF mediante la incorporación de un Asistente Virtual que actúe en función a eventos, libere de trabajo de monitorización a los Administradores de Proyecto y de soporte a la toma de decisiones de directivos de las empresas.

Finalmente, cabe destacar que, como resultado de realización de la tesis, se han generado 2 proyectos de investigación, “Evaluación del impacto de las emociones en la calidad de software desde el punto de vista del usuario”, y otro, denominado “iQuAGI: Un enfoque inteligente para la evaluación de calidad de procesos de software ágiles”, ambos pertenecientes al CInApTIC y en ejecución actualmente.

Referencias

Acuña, C., Cuenca Pletsch, L., Tomaselli, G., Pinto, N., Tortosa, N. (2015) “Calidad de Software y Metodologías Ágiles en las PYMES de la Industria del Software”. Publicado en Memorias de 3er Congreso Nacional de Ingeniería Informática/Sistemas de Información (CONAIISI 2015). Noviembre, 2015. ISBN 978-987-1896-47-9.

Bollati, V. A., Gaona, G., Pletsch, L. C., Gonnet, S., & Leone, H. (2017). "The state of agile development adoption in Argentine software companies". In Computer Conference (CLEI), 2017 XLIII Latin American (pp. 1-10). IEEE.

Bunge, M. (1989). *La investigación científica* (No. 330.018 BUN).

Garzás, Fernández & Piattini (2009). "Una aplicación de la Norma ISO/IEC 15504 para la evaluación por nivel de madurez de Pymes y pequeños equipos de desarrollo". *Revista Española de Innovación, Calidad e Ingeniería del Software*, Vol.5, No. 2.

Kitchenham, B., Pickard, L., & Pfleeger, S. L. (1995). Case studies for method and tool evaluation. *IEEE software*, 12(4), 52-62.

Mas A., Amengual E. (2005). "Las mejoras de los procesos de Software en las pequeñas y medianas empresas (pymes). Un nuevo modelo y su aplicación a un caso real". *Revista Española de Innovación, Calidad e Ingeniería del Software*, Vol.1, No. 2

Marcos, E., & Marcos, A. (1999). *An Aristotelian Approach to the Methodological Research: a Method for Data Model Construction. Information Systems. The Next Generation.* McGraw Hill, 532-543.

Pasini, A. C., Esponda, S., Bertone, R. A., & Pesado, P. (2008). "Aseguramiento de Calidad en PYMES que desarrollan software." XIV Congreso Argentino de Ciencias de la Computación.

Pfleger, S. (2002) "Ingeniería de Software. Teoría y Práctica." Pearson Education.

Pinto, N., Tortosa, N., Cabas Geat, B., Ibáñez, L., & Acuña, C. (2018, November). Validación de la reingeniería aplicada sobre la primera versión de Agile Quality Framework. In XIX Simposio Argentino de Ingeniería de Software (ASSE)-JAIIO 47 (CABA, 2018).

Pinto, N., Acuña, C., & Cuenca Pletsch, L. R. (2016). Quality Evaluation in Agile Process: A First Approach. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)...

Rujana, M., Romero Franco, N., Tortosa, N., Tomaselli, G., & Pinto, N. (2016). "Análisis sobre adopción de metodologías ágiles en los equipos de desarrollo en pymes del NEA". WICC 2016

Version One 12th Annual State of Agile™ Report (2018) Available at <https://explore.versionone.com/state-of-agile/versionone-12th-annual-state-of-agile-report>

C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in software engineering: an introduction.* Kluwer Academic Publishers, Norwell, MA, USA, 2000.

Resumen de Tesis Doctoral: Integración escalable de realidad aumentada basada en imágenes y rostros

Programa: Doctorado en Ciencias Informáticas
Facultad de Informática - UNLP
Tesisista: Nahuel Adiel Mangiarua
Director: Jorge Ierache
Codirector: María José Abásolo
Fecha de defensa: 18/12/2020

Grupo de Investigación en Realidad Aumentada Aplicada
Departamento de Ingeniería, UNLaM, Buenos Aires, Argentina
nmangiarua@unlam.edu.ar

Resumen La realidad aumentada permite la adición de elementos digitales sobre la visión natural de las personas. La tesis aquí resumida se enfoca en desarrollar una arquitectura que permita la integración de realidad aumentada basada en imágenes arbitrarias y la realidad aumentada basada en rostros de forma escalable. Se estudiaron los procesos involucrados y los algoritmos necesarios para solventar cuellos de botella. Los resultados obtenidos sugirieron la utilización de múltiples bucles con esquemas de ejecución asíncrono y paralelos así como al algoritmo de búsqueda aproximada de vecinos más cercanos HNSW como el más apropiado para este contexto particular.

Palabras Claves: *Realidad Aumentada; Reconocimiento Facial; Escalabilidad; Arquitectura de Integración; Inferencia biométrica; Approximate Nearest Neighbor Search*

1. Temas de investigación

La Realidad Aumentada (RA) se encuentra inmersa en lo que se ha llamado el continuo de la virtualidad, introducido por (Milgram y Kishino 1994). Este continuo que se extiende desde el ambiente real que nos rodea hasta un ambiente completamente virtual, artificial. La RA consiste en la creación de un entorno en el que la información y los objetos virtuales se fusionan con la realidad, ofreciendo al usuario una experiencia enriquecida, incrementando su percepción natural. La RA puede ser usada para expandir nuestros sentidos, define una visión directa o indirecta de un entorno físico del mundo real, cuyos elementos se combinan con elementos virtuales, como pueden ser textos, imágenes, audio o videos para la creación de una realidad mixta en tiempo real (Abásolo Guerrero et al. 2011).

Una aplicación de RA, en la mayoría de los casos, está conformada mínimamente por un conjunto de cuatro elementos básicos que necesitan estar vinculados de cierta manera para que la aplicación logre su cometido. Los elementos en cuestión son:

- Un elemento que capture las imágenes de la realidad
- Un elemento sobre el que proyectar la integración de las imágenes reales con los contenidos virtuales.
- Una arquitectura de software, embebida o no en un marco de trabajo
- Un disparador o elemento que actúe como desencadenante de RA.

RA basada en imágenes

En particular, la RA basada en el reconocimiento de imágenes planas consiste en utilizar como disparador y anclaje al mundo físico, una imagen plana arbitraria que puede ser detectada, permitiendo la incorporación de información virtual de forma espacial y

contextualmente coherente con la realidad. A diferencia de las imágenes denominadas marcadores, que poseen un patrón geométrico dominante para facilitar su detección y un código visual embebido para diferenciar uno de otro, la imagen arbitraria puede ser entendida como una fotografía, que captura y representa la realidad de un instante congelado en el tiempo. Si bien la realidad física que nos rodea se encuentra en continuo cambio, muchos de estos cambios son lo suficientemente graduales como para considerarlos constantes durante el tiempo de vida útil de una determinada aplicación que utilice esta tecnología.

Escalabilidad y búsqueda de imágenes

La escalabilidad en términos de RA puede naturalmente relacionarse a múltiples aspectos de su ejecución. La acepción utilizada en el contexto particular del trabajo de tesis, hace referencia específicamente a la cantidad de objetivos de aumentación que una arquitectura, u aplicación que la utilice, pueda contener y aumentar. En este sentido no se considera la cantidad de objetivos (imágenes o rostros) que se estén aumentando en simultáneo sobre un mismo cuadro de vídeo o imagen, sino al tamaño del repositorio de objetivos que un software puede reconocer y aumentar individualmente, sin depender de procesamiento externo.

Detección y reconocimiento de rostros e inferencia biométrica

La detección de rostros en tiempo real en un flujo de vídeo, aunque aún un área de investigación activa, resulta posible y aplicable desde los aportes introducidos por el famoso trabajo de Viola y Jones (Viola y Jones 2001). Siguiendo la tendencia general en el área de visión de computador, en los años posteriores muchos otros mecanismos basados en características (usualmente llamadas con el término en inglés, features) complejas o esquemas de aprendizaje de máquina simples se han propuesto, no sólo para aumentar el desempeño sino para incrementar el rango de ángulos y de variación lumínica aceptables, como los presentados en (Jegou, Douze, y Schmid 2008; Wang, Han, y Yan 2009; Waring y Liu 2005; Wu y Nevatia 2007). En la actualidad, el problema de la detección de rostros es abordado mayormente mediante la aplicación de técnicas avanzadas de aprendizaje de máquina como presenta (Dantone et al. 2012), en particular destacan las redes neuronales convolucionales y sus arquitecturas derivadas como presentan (Girshick 2015; Li et al. 2015).

Por su parte, la inferencia de parámetros biométricos a partir de imágenes es un campo con numerosas especializaciones. En particular, para en la detección de emociones a través del rostro encontramos un gran número de técnicas tanto a nivel experimental como a nivel productivo. (Corneanu et al. 2016) nos ofrece una extensiva introducción y recorrido por los más importantes aportes hasta la actualidad a la vez que (Valstar et al. 2017) describe los esfuerzos por definir un marco de trabajo y evaluación estandarizado mediante las competencias FERA (Facial Expression Recognition and Analysis).

2. Motivación y objetivos

Tanto la RA basada en imágenes arbitrarias como el reconocimiento facial son utilizados por un número de sistemas, aplicaciones o frameworks en diversos campos de aplicación. Sin embargo, no existen en la actualidad frameworks cuyas arquitecturas integren la capacidad de reconocer imágenes y rostros de manera simultánea y escalable, es decir con un número elevado de objetivos a aumentar. A su vez, ningún framework integra la capacidad de efectuar inferencia biométrica de información a partir de las imágenes percibidas, en particular la de rostros humanos.

Se planteó como objeto principal del trabajo de tesis el diseñar una arquitectura escalable de RA basada en el reconocimiento visual monocular de imágenes y rostros humanos, con capacidad de inferencia de datos biométricos, que no haga uso de servicios externos para su etapa de explotación.

En este orden se plantearon como objetivos particulares:

- Establecer los procesos y sus pasos necesarios para efectuar aumentación de imágenes, detección y reconocimiento de rostros e inferencia de información biométrica.
- Analizar comparativamente la complejidad computacional teórica y la carga de procesamiento empírica de cada paso de los procesos de RA, considerando en particular distintas variaciones de algoritmos disponibles para la búsqueda y descripción de POI.
- Diseñar una arquitectura que integre los procesos descritos, contemplando la ejecución paralela y/o asíncrona, identificando el o los pasos que resulten en el principal cuello de botella con respecto a la escalabilidad.
- Analizar comparativamente la velocidad y precisión de los algoritmos aplicables para aliviar o solventar el o los cuellos de botella detectados.
- Diseñar criterios de evaluación y conjuntos (sets) de datos de prueba para los algoritmos aplicables a los cuellos de botella que sean representativos del dominio de explotación propuesto.
- Incorporar al diseño un mecanismo de integración abierto que facilite el agregado futuro de algoritmos de inferencia biométrica adicionales en la arquitectura propuesta.

Mientras que no se busca competir con sistemas existentes en términos del refinamiento y calidad que estos han logrado con años de desarrollo continuo, se desarrollo una integración de las tecnologías propuestas con una escalabilidad superadora.

3. Aportes y Conclusiones

El trabajo de tesis propone una arquitectura escalable que integra la RA basada en imágenes arbitrarias con la detección y reconocimiento de rostros humanos junto con la inferencia de datos biométricos a partir de ellos.

Partiendo desde la definición de los procesos para RA basada en imágenes y rostros, se determina cuáles son los pasos necesarios para lograr respectivamente y se estudia la complejidad computacional teórica de cada uno junto con la distribución de carga de procesamiento relativa. Estableciendo el algoritmo ORB o la combinación de ORB y FREAK como alternativas viables para la detección y descripción de POI en imágenes, se identifican los pasos de búsqueda de correspondencias entre descriptores, tanto de imágenes (POI) como de rostros como los cuellos de botella de cada proceso. Se selecciona para la descripción de rostros algoritmos de redes neuronales convolucionales entrenadas con el error por tripletas que producen descriptores continuos de 128 dimensiones y se establece que este paso debe ser implementado de forma asíncrona debido a su tiempo de ejecución.

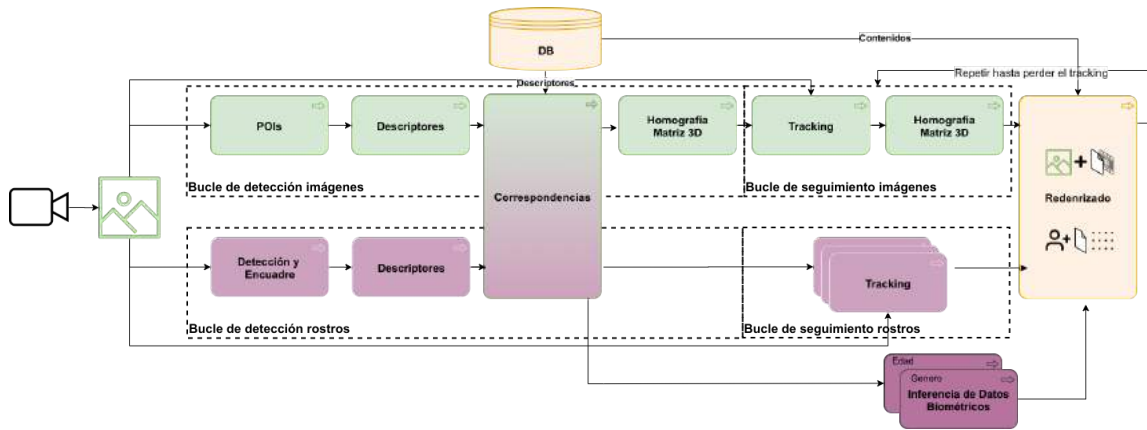
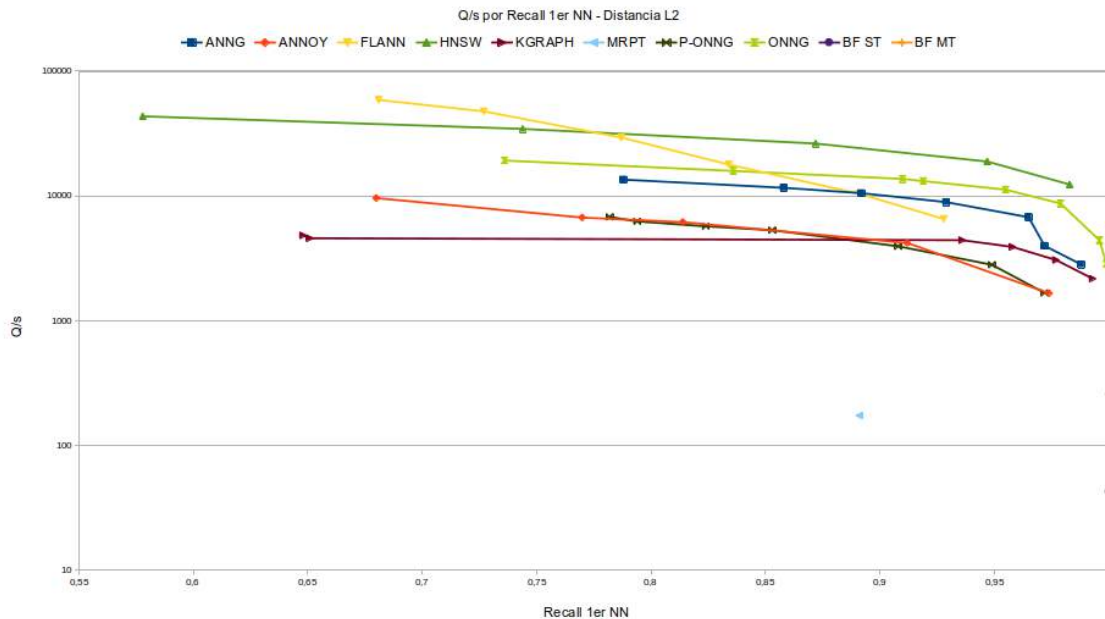


Figura 1: Diagrama conceptual de la arquitectura propuesta. Los recuadros de línea punteada delimitan cada uno de los 4 bucles.

Se presenta el diseño de una arquitectura integrada, compuesta por cuatro bucles en un flujo de ejecución alternante con derivación de tareas asíncronas en un esquema de ejecución paralelo para las tareas de descripción de rostros e inferencia de información biométrica, como se aprecia en la figura 1. Los algoritmos seleccionados permiten considerar el cuello de botella de cada proceso integrado como un mismo problema el cual se propone aliviar con el uso de algoritmos de búsqueda ANN.

Para lograr la escalabilidad de la arquitectura integrada, se realizan una serie de experimentos para analizar comparativamente la velocidad, precisión y estabilidad de distintos algoritmos de búsqueda ANN, estableciendo un marco de evaluación y conjuntos de datos específicos para el contexto de la RA basada en imágenes y reconocimiento de rostros. En la figura 2 se observa particularmente la medida de queries (operaciones de búsqueda) por segundo de cada algoritmo de acuerdo a su precisión (recall). Se establece y valida la superioridad del algoritmo HNSW para la tarea en este contexto particular, difiriendo de los resultados provistos por otros autores para contextos generales.



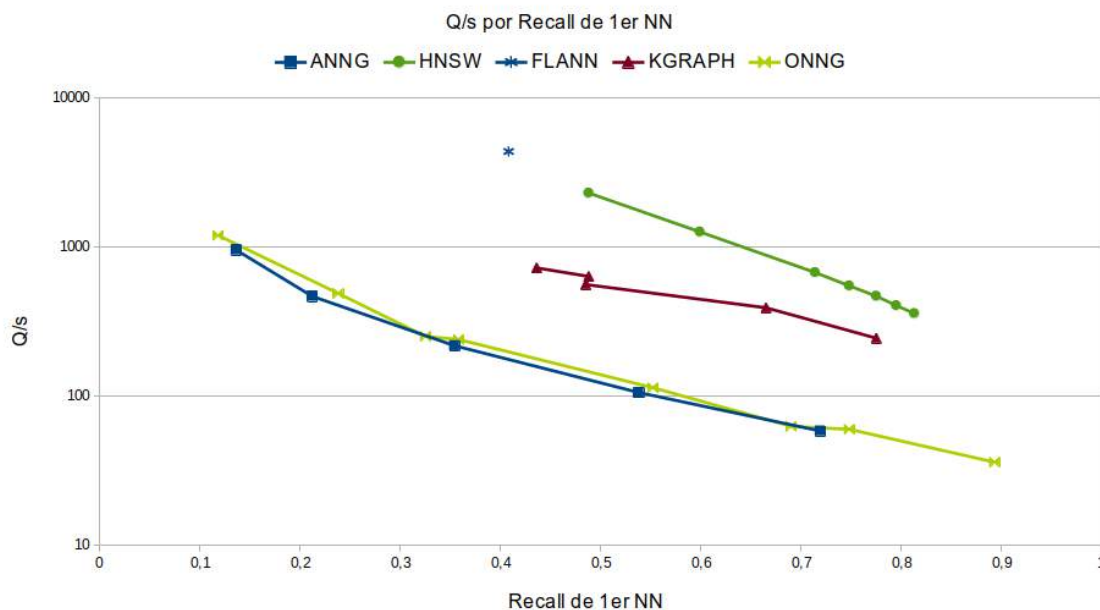


Figura 2: Gráficos de cantidad de operaciones por segundo según precisión (recall) para múltiples algoritmos de búsqueda ANN probados. Mientras más arriba a la derecha mejor es el algoritmo.

También se obtiene como aporte secundario un nuevo esquema de evaluación de algoritmos de búsqueda de vecinos más cercanos aproximados específico para el contexto de la RA. El mismo, implica la evaluación de esta familia de algoritmos utilizando sets de datos donde la variación entre los elementos query y los elementos de entrenamiento, base u originales sea reducida. Se establece un porcentaje de variación de entre el 5% y el 15% de acuerdo al tipo función de distancia utilizada pero se recomienda ajustar estos valores según el algoritmo específico que genera los elementos. Con estos sets de datos se propone un esquema donde los algoritmos de búsqueda de ANN sean comparados por su recall en los dos primeros vecinos.

Se crea un prototipo demostrador experimental que implementa la arquitectura propuesta en C++, compatible con cualquier plataforma para la que se lo compile. La misma optimiza la utilización de múltiples hilos de procesamiento y permite la posibilidad de integración con sistemas de más alto nivel para la generación y explotación de aplicaciones de RA. En la misma se implementa una interfaz abstracta que permite la incorporación dinámica de algoritmos de inferencia biométrica que serán automáticamente ejecutados de manera paralela y asíncrona.

Contexto

La tesis para el doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP se radicó en el Grupo de Realidad Aumentada Aplicada (GRAA) del Departamento de Ingeniería de la UNLaM, fue desarrollada en el marco del proyecto PROINCE C-231. La dirección estuvo a cargo de Jorge Ierache (GRAA, UNLaM) y la co-dirección de Maria Jose Abasolo (LIDI, UNLP).

Futuros trabajos

Si bien el marco de evaluación de algoritmos de ANN para el contexto de la RA propuesto predice con mayor precisión el desempeño de los algoritmos, se requiere continuar sobre esta

línea para establecer condiciones aún más específicas y con el fin de desarrollar una métrica de comparación concreta más precisa.

Se plantea abordar la compilación cruzada del prototipo de implementación a plataformas ARM, así como la integración con sistemas de alto nivel, específicamente el motor de gráficos Unity3D.

También se plantea la integración continua de algoritmos de inferencia de datos biométricos a medida que sean desarrollados y publicados por sus respectivos autores.

Finalmente se planea, continuando con la línea de investigación bajo el marco del proyecto PROINCE C-231 Comandos de Voz y Reconocimiento Facial para Aplicaciones de Realidad Aumentada, el desarrollo de un prototipo demostrador en el contexto de la emergentología.

Publicaciones relacionadas con la tesis

Revistas, series internacionales y capítulos de libro

Mangiarua, Nahuel, Jorge Ierache, and María José Abasolo. "Scalable Integration of Image and Face Based Augmented Reality." In *Augmented Reality, Virtual Reality, and Computer Graphics*, edited by Lucio Tommaso De Paolis and Patrick Bourdot. Cham: Springer International Publishing, https://doi.org/10.1007/978-3-030-58465-8_18

En esta publicación se comunican los principales elementos de diseño y la arquitectura integrada objeto de este trabajo de tesis.

Mangiarua, Nahuel, Jorge Ierache, and María José Abasolo. "Implementation of an Open Source Based Augmented Reality Engine for Cloud Authoring Frameworks." *Journal of Computer Science and Technology* 19, no. 2 (October 2019): e16. <https://doi.org/10.24215/16666038.19.e16>.

En esta publicación se comunica el diseño del proceso basado en imágenes y su análisis particular así como consideraciones para el aprovechamiento de capacidades de la nube para herramientas de autor.

Mangiarua, Nahuel, Jorge Ierache, Martín Becerra, Hernán Maurice, Santiago Igarza, and Osvaldo Sposito. "Templates Framework for the Augmented Catalog System." In *Computer Science – CACIC 2018*, edited by Patricia Pesado and Claudio Aciti, 267–276. Cham Springer International Publishing: Springer International Publishing, 2019. <http://sedici.unlp.edu.ar/handle/10915/73478>.

En esta publicación se comunican los principales elementos de diseño y la implementación del concepto de template para sistemas de autor que facilita el trabajo a usuarios finales frente a un incremento de escala en el sistema.

Ierache, Jorge, Nahuel Mangiarua, Martín Ezequiel Becerra, and Santiago Igarza. "Framework for the Development of Augmented Reality Applications Applied to Education Games." In *Proceedings of the 5th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics*, edited by Lucio Tommaso De Paolis and Patrick Bourdot, 340–350. Cham Springer International Publishing: Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-95270-3_28.

Esta publicación complementaria comunica un framework especializado desarrollado sobre el Sistema de Catálogos Virtuales Aumentados, el cual constituye un caso de aplicación potencial para la implementación de la arquitectura propuesta.

Congresos internacionales

Montalvo, C., F. Petrolo, D. Sanz, Nahuel Mangiarua, N. Verdicchio, S. Igarza, and J. Ierache. “Knowledge Based Augmented Card System for Medical Assistance Over Mobile Devices.” In Selected Papers of the XXI Argentine Congress of Computer Science, 257–65, 2017.

Ierache, Jorge, Nicolás Nazareno Verdicchio, N. Duarte, and Nahuel Mangiarua. “Augmented Reality Card System for Emergency Medical Services.” In Proceedings of the International Work-Conference on Bioinformatics and Biomedical Engineering, 487–94. Granada, Spain, 2016.

Ierache, Jorge, Nahuel Mangiarua, N. Verdicchio, and D. Sanz. “Augmented Card System Based on Knowledge for Medical Emergency Assistance.” In 2016 IEEE Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI), 1–3, 2016. <https://doi.org/10.1109/CACIDI.2016.7785979>.

Ierache, Jorge, Nahuel Mangiarua, S. Bevacqua, N. Verdicchio, and M. Becerra. “Development of a Catalogs System for Augmented Reality Applications.” In World Academy of Science, Engineering and Technology, International Science Index, 9:1–7, 2015.
<http://waset.org/Publications/development-of-a-catalogs-system-for-augmentedreality-applications/10000077>.

Congresos nacionales y workshops

Mangiarua, Nahuel, Jorge Ierache, and María José Abásolo Guerrero. “Avances en Línea de Investigación Doctoral: Integración Escalable de Realidad Aumentada Basada en Imágenes y Rostros.” In Libro de actas del XXII Workshop de Investigadores en Ciencias de la Computación. Santa Cruz, Argentina, 2020.

Mangiarua, Nahuel, Jorge Ierache, Martín Ezequiel Becerra, Hernán Maurice, Santiago Igarza, and Osvaldo Mario Sposito. “Framework para la generación de templates en sistema de catálogos de RA.” In Libro de actas del XXIV Congreso Argentino de Ciencias de la Computación. La Plata, Argentina, 2018.
<http://sedici.unlp.edu.ar/handle/10915/73478>.

Ierache, Jorge, Javier Urien, Nahuel Mangiarua, Martín Ezequiel Becerra, Paola Pezoimburg, Emidio Bueno, Hernán Maurice, and Eric S. Auchterberge. “Desarrollo de un prototipo de RA de tipo geoespacial.” In Libro de actas del XII Jornadas de Vinculación Universidad-Industria, 5–8, 2018.
<http://sedici.unlp.edu.ar/handle/10915/71289>.

Becerra, Martín Ezequiel, Nahuel Mangiarua, Santiago Igarza, Jorge Ierache, and María José Abásolo Guerrero. “Líneas de investigación del grupo de RA aplicada: templates de catálogos aumentados integración escalable de RA basada en imágenes y rostros aumentación de sistemas SCADA en el contexto de la industria 4.0.” In Libro de actas del XX Workshop de Investigadores en Ciencias de la Computación, 2018.
<http://sedici.unlp.edu.ar/handle/10915/67455>.

Mangiarua, Nahuel, Cristian Montalvo, Facundo Petrolo, Diego Rubén Sanz, Nicolás Nazareno Verdicchio, E. Lobatto, A. Rosenthal, Martín Ezequiel Becerra, Santiago Igarza, and Jorge Ierache. “Framework para la generación de templates en sistemas de catálogos de RA.” In Libro de actas del XIX Workshop de Investigadores en Ciencias de la Computación, 2017. <http://sedici.unlp.edu.ar/handle/10915/61824>.

Verdicchio, Nicolás Nazareno, Diego Rubén Sanz, Cristian Montalvo, Facundo Petrolo, Nahuel Mangiarua, Santiago Igarza, and Jorge Ierache. “Sistema de catálogo virtual aumentado: integración de framework especializado orientado a juegos didácticos.” In Libro de actas del XI Congreso de Tecnología en Educación y Educación en Tecnología, 2016. <http://hdl.handle.net/10915/54651>.

Trabajos previos

Bevacqua, Sebastián Ariel, Santiago Igarza, Nahuel Mangiarua, Martín Ezequiel Becerra, Nicolás Nazareno Verdicchio, Fernando Martín Ortiz, Diego Rubén Sanz, Nicolás Daniel Duarte, Matías Ezequiel Sena, and Jorge Ierache. “Líneas de investigación del grupo de Realidad Aumentada Aplicada de UNLaM.” In Libro de actas del XVII Workshop de Investigadores en Ciencias de la Computación. Salta, Argentina, 2015. <http://sedici.unlp.edu.ar/handle/10915/45657>.

Mangiarua, Nahuel, Jorge Ierache, and Martin Becerra. “Herramienta de Realidad Aumentada Para La Explotación de Material Didáctico Tradicional.” In Libro de Actas Del IX Congreso de Tecnología En Educación En Tecnología, 240–54. Chilecito, Argentina, 2014.

Ierache, Jorge, Santiago Igarza, Nahuel Mangiarua, Martín Ezequiel Becerra, Sebastián Ariel Bevacqua, Nicolás Nazareno Verdicchio, Fernando Martín Ortiz, Diego Rubén Sanz, Nicolás Daniel Duarte, and Esteban de la Llave. “Realidad Aumentada (RA) en el contexto de usuarios finales.” In Libro de actas del XVI Workshop de Investigadores en Ciencias de la Computación, 2014. <http://sedici.unlp.edu.ar/handle/10915/41253>.

Ierache, Jorge, Nahuel Mangiarua, and Sebastián Ariel Bevacqua. “Sistema de catálogo para la asistencia a la creación, publicación, gestión y explotación de contenidos multimedia y aplicaciones de RA.” In Libro de actas del XX Congreso Argentino de Ciencias de la Computación. Buenos Aires, Argentina, 2014. <http://sedici.unlp.edu.ar/handle/10915/42339>.

Ierache, Jorge, Nahuel Mangiarua, and Martin Becerra. “Herramienta de Realidad Aumentada Para Facilitar La Enseñanza En Contextos Educativos Mediante El Uso de Las TICs,” *Revista Latinoamericana de Ingeniería de Software*, 1 (2014): 1–3.

Esta tesis doctoral

Nahuel A. Mangiarua. “Integración Escalable de Realidad Aumentada Basada en Imágenes y Rostros”. Tesis doctoral, Universidad Nacional de La Plata, La Plata, Argentina, 2020. <http://sedici.unlp.edu.ar/handle/10915/111301>

Bibliografía

Abásolo Guerrero, María José, Cristina Manresa Yee, Ramón Más Sansó, y Marcelo Vénere. 2011. *Realidad virtual y realidad aumentada*. Editorial de la Universidad Nacional de La Plata (EDULP). <http://sedici.unlp.edu.ar/handle/10915/18399>.

Corneanu, C.A., M.O. Simón, J.F. Cohn, y S.E. Guerrero. 2016. «Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications». *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38. <https://doi.org/10.1109/TPAMI.2016.2515606>.

Dantone, M., J. Gall, G. Fanelli, y L. Van Gool. 2012. «Real-Time Facial Feature Detection Using Conditional Regression Forests». En *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2578–85. <https://doi.org/10.1109/CVPR.2012.6247976>.

Girshick, Ross B. 2015. «Fast R-CNN». *CoRR Abs/1504.08083*. <http://arxiv.org/abs/1504.08083>.

Jegou, Herve, Matthijs Douze, y Cordelia Schmid. 2008. «Hamming Embedding and Weak Geometric Consistency for Large Scale Image Search». En *Computer Vision – ECCV 2008: 10th European Conference on Computer Vision*, editado por David Forsyth, Philip Torr, y Andrew Zisserman, 304–317. Marseille, France: Springer. https://doi.org/10.1007/978-3-540-88682-2_24.

Li, H., Z. Lin, X. Shen, J. Brandt, y G. Hua. 2015. «A Convolutional Neural Network Cascade for Face Detection». En *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5325–34. <https://doi.org/10.1109/CVPR.2015.7299170>.

Milgram, Paul, y Fumio Kishino. 1994. «A Taxonomy of Mixed Reality Visual Displays». *IEICE Trans. Information Systems* E77-D, n.º 12: 1321-29.

Valstar, Michel F., Enrique Sánchez-Lozano, Jeffrey F. Cohn, László A. Jeni, Jeffrey M. Girard, Zheng Zhang, Lijun Yin, y Maja Pantic. 2017. «FERA 2017 - Addressing Head Pose in the Third Facial Expression Recognition and Analysis Challenge». *CoRR Abs/1702.04174*. <http://arxiv.org/abs/1702.04174>.

Viola, P., y M. Jones. 2001. «Rapid Object Detection Using a Boosted Cascade of Simple Features». En *1:1-511-1-518*. IEEE Comput. Soc. <https://doi.org/10.1109/CVPR.2001.990517>.

Wang, X., T.X. Han, y S. Yan. 2009. «An HOG-LBP Human Detector with Partial Occlusion Handling». En *2009 IEEE 12th International Conference on Computer Vision*, 32–39. <https://doi.org/10.1109/ICCV.2009.5459207>.

Waring, C.A., y Xiuwen Liu. 2005. «Face Detection Using Spectral Histograms and SVMs». *IEEE Transactions on Systems*. <https://doi.org/10.1109/TSMCB.2005.846655>.

Wu, Bo, y Ram Nevatia. 2007. «Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet Based Part Detectors». *International Journal of Computer Vision* 75 (2). <https://doi.org/10.1007/s11263-006-0027-7>.

Interacción Tangible en escenarios educativos. Diseño de una herramienta de autor para la creación de actividades educativas basadas en Interacción Tangible

**Tesis de Doctorado en Ciencias Informáticas
Facultad de Informática, Universidad Nacional de La Plata**

Tesista: Verónica Artola

wartola@lidi.info.unlp.edu.ar

Instituto de Investigación en Informática LIDI (III LIDI), Facultad de Informática, UNLP

Directora: Cecilia Sanz (III LIDI, CIC, Facultad de Informática, UNLP)

Co-Directora: Patricia Pesado (III LIDI, Facultad de Informática, UNLP)

Fecha de exposición: 30 de marzo de 2020

CONTEXTO

El plan de trabajo se enmarca en las investigaciones que se desarrollan en el Proyecto 11-F023, 'Metodologías, Técnicas y Herramientas de Ingeniería de Software en Escenarios Híbridos. Mejora de Proceso'. El trabajo se realiza en el marco del Doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP.

La tesista contó con una beca Doctoral CONICET. Al mismo tiempo, esta línea de I+D forma parte de la agenda de cooperación con la Universidad de Zaragoza, específicamente con el Grupo de Informática Gráfica Avanzada GIGA Affective Lab, con el que se realizan acciones conjuntas sobre nuevos paradigmas de interacción persona-ordenador en escenarios educativos. Esta vinculación ha sido de impacto para el desarrollo de este trabajo.

MOTIVACIÓN

El área de interacción persona-ordenador ha despertado el interés de los investigadores y convoca la mirada de diferentes disciplinas tales como la Psicología Cognitiva, la Educación, la Comunicación, el Diseño, y las Ciencias de la Computación. Se encuentra en continua evolución, teniendo como una de sus metas buscar las mejores alternativas para la interacción entre las personas y los entornos digitales (Artola, Sanz, Baldassarri, y cols., 2016). Así, las interfaces de líneas de comando (Command Line Interface-CLI), dieron paso a las interfaces gráficas de usuario (Graphical User Interface-GUI) y hoy día ya se habla de un conjunto de paradigmas y estilos de interacción que han sido agrupados dentro de la categoría de Interfaces Naturales (Natural User Interface-NUI). Las ideas de Weiser (1999) sobre la Computación Ubicua, las de Ishii y Ullmer (1997) sobre la Interacción Tangible (IT) y el surgimiento de las NUI han dado lugar a que el usuario pueda controlar una aplicación de cualquier tipo, a través de una interfaz invisible que resulte intuitiva, fácil de aprender y que permite ser accedida a través de la voz, gestos, movimientos corporales, o movimientos de objetos del entorno que son interpretados como comandos de entrada para la aplicación (Diego-Mendoza, Márquez-Domínguez, y Sabino-Moxo, 2014).

Esta tesis propone ser un aporte al área específica de interacción persona-ordenador, y en particular para la investigación y prácticas vinculadas a sistemas basados en interacción tangible, considerando específicamente sus posibilidades para escenarios educativos.

El término Interfaz de Usuario Tangible fue acuñado por Hiroshi Ishii, científico y programador, director del proyecto 'Bits Tangibles' (*Tangible Bits*), en funcionamiento desde el año 1995, en el Massachusetts Institute of Technology (MIT). En 1995, junto a Fitzmaurice (Fitzmaurice, Ishii, y Buxton, 1995) introdujeron la noción de una interfaz aprehensible. Unos años más tarde, en 1997, Ishii y sus estudiantes presentaron una visión más completa en su trabajo *Tangible Bits* (Ishii y Ullmer, 1997). El proyecto *Tangible Bits* busca volver tangible la información digital, asociando los bits con objetos físicos cotidianos y superficies arquitectónicas. El objetivo del proyecto consiste en reducir la brecha existente entre las personas, la información digital y el entorno, así como abordar las nociones de centro y periferia en las actividades humanas. El resultado es un nuevo tipo de interfaz que se encarna en el ambiente (Ishii y Ullmer, 1997).

Las interfaces basadas en IT permiten que, a través del uso de objetos cotidianos, se interactúe con una aplicación informática. Estas interfaces hacen hincapié en la percepción multisensorial y en volver tangible la información digital.

A lo largo de los años, diferentes marcos de diseño se han aplicado al desarrollo de sistemas basados en IT. De esta forma se ha ido abriendo la puerta a nuevas formas de relacionarlos con los procesos de enseñanza y aprendizaje

(Marshall, Price, y Rogers, 2003; Price, 2008). Las teorías del aprendizaje cognitivistas y constructivistas permiten dar un sustento y un valor a la IT en escenarios educativos. Esta justificación abarca conceptos teóricos y prácticas de participación, de exploración y construcción de modelos, la actividad colaborativa entre estudiantes, entre otros. Entre los diversos argumentos que se han planteado se puede nombrar: la generación de metáforas (Antle, Droumeva, y Corness, 2008; Bakker, Antle, y Van Den Hoven, 2012), la posibilidad de centrar la atención en la tarea (Marshall, 2007), la inclusión de un canal adicional (táctil) para transmitir/percibir información (Dillenbourg, 2016), el razonamiento del mundo a través del descubrimiento y la participación (Zuckerman y cols., 2005; Marshall, 2007), la mejora de la memoria a través de la acción física, la incitación a la interacción social y a la colaboración (Hornecker, 2002), por mencionar algunos. El aprendizaje mediado por objetos tangibles además tiene el potencial para permitir que los niños combinen y recombinen lo conocido y familiar en formas nuevas y desconocidas (Antle, 2007; Antle y Wise, 2013; Price, 2008; Manches, O'Malley, y Benford, 2010; Price, Rogers, Scaife, Stanton, y Neale, 2003; Marshall, 2007; Martinez-Maldonado y cols., 2013)

Un gran potencial de la IT en el ámbito educativo puede encontrarse en las *tabletops*. Estas arquitecturas son superficies horizontales (mesas) aumentadas computacionalmente que permiten apoyar objetos físicos. Sus principales beneficios se encuentran relacionados con la motivación, el disfrute, el aprendizaje, el involucramiento, la tangibilidad, la multimodalidad y la posibilidad de realizar y enriquecer tareas colaborativas. Desde sus inicios se han realizado numerosas experiencias que muestran buenos resultados en diversos dominios.

La creación de aplicaciones basadas en IT enfrentan al momento múltiples desafíos. Uno de estos desafíos es facilitar la participación de expertos no técnicos en el desarrollo de este tipo de aplicaciones. Es así, que han comenzado a gestarse herramientas que facilitan el proceso de creación de actividades basadas en IT para usuarios sin conocimientos en programación, de manera tal de que puedan tomar decisiones de diseño vinculadas a las interacciones entre objetos físicos e información digital. Estas herramientas varían en sus objetivos, en el grupo destinatario y en las posibilidades que ofrece cada una. Este escenario ha dado pie a una de las motivaciones principales de este trabajo. Así, esta tesis se enfoca en el área de la Interacción Tangible aplicada al escenario educativo, y en particular en poder acercar estas tecnologías a los docentes para el diseño de actividades educativas basadas en IT.

Las motivaciones que impulsan este trabajo se vinculan entonces con:

- La importancia de los sistemas IT, como eslabón evolutivo en el campo de la interacción persona-ordenador, y sus posibilidades ya evidenciadas en situaciones educativas, lo que invita a profundizar la investigación en el área.
- La necesidad de involucrar aún más a los expertos en el dominio en el contexto de creación de actividades basada en IT.
- La creencia de que una herramienta de autor, que se oriente al diseño y desarrollo de actividades IT, especialmente educativas, constituye un aporte a la comunidad y al área, ya que posibilita que los expertos en el dominio puedan apropiarse de estas tecnologías.

OBJETIVOS

Este trabajo ha tenido como objetivo investigar sobre la Interacción Tangible en el escenario educativo y desarrollar tecnología basada en esta forma de interacción para su aplicación en procesos de enseñanza y aprendizaje.

Para tal fin, se elaboraron una serie de objetivos específicos que se listan a continuación:

1. Estudiar el concepto de IT, sus características, y marcos teóricos y descriptivos existentes, en general y en relación al ámbito educativo.
2. Componer un estado del arte en relación a las interfaces de interacción tangible en el escenario educativo. Estudiar buenas prácticas por parte de diseñadores y docentes que estén trabajando en esta temática.
3. Diseñar y desarrollar una herramienta de autor que permita la construcción de actividades educativas digitales basadas en el paradigma de interacción tangible.
4. Generar un estudio de caso que permita relevar las opiniones y aceptación de los docentes en el uso de este tipo de herramienta de autor.
5. Analizar los resultados obtenidos del estudio de caso, de manera tal de lograr el uso efectivo de la herramienta de autor en escenarios educativos específicos, mediante un plan de difusión que se conformará como parte de este trabajo.

Para guiar esta investigación se plantearon doce (12) preguntas de investigación:

- P1. ¿En qué contexto surge la propuesta de IT?
- P2. ¿Cuáles son las principales motivaciones para el desarrollo de la IT?
- P3. ¿Cómo propone la relación entre el mundo físico y el mundo digital?
- P4. ¿Qué relaciones se establecen entre objetos físicos y digitales en los diferentes marcos conceptuales de IT?
- P5. ¿Cuáles son los marcos que orientan diseño aplicaciones basadas IT?
- P6. ¿Cuáles son los marcos teóricos que vinculan la IT con el área educativa?
- P7. ¿Qué aportes realiza la IT en los procesos de enseñar y aprender?
- P8. ¿Qué tipos de experiencias educativas se están desarrollando con IT, y en particular, sobre *tabletops*?
- P9. ¿Qué consideraciones son necesarias al momento de diseñar y desarrollar aplicaciones basadas en IT, y en particular sobre *tabletops*?
- P10. ¿Qué herramientas se encuentran disponibles para que personas no expertas puedan desarrollar aplicaciones basadas en IT?
- P11. ¿Cuáles son las necesidades aún no atendidas por este tipo de herramientas para el desarrollo de actividades educativas basadas en IT?
- P12. ¿Qué aspectos constituyen una barrera para la inclusión de aplicaciones en procesos educativos?

APORTES

A continuación se mencionan los aportes relacionados a los objetivos específicos planteados.

1. Estudiar el concepto de IT, sus características, y marcos teóricos y descriptivos existentes, en general y en relación al ámbito educativo.
 - Se realizó una indagación sobre diferentes paradigmas IPO. Se puso mayor énfasis en el estudio de la IT, por ser el tema de interés en este trabajo. Se recuperaron conceptos e ideas clave que han motivado el surgimiento de la IT y su evolución.
 - Se abordó el estudio de 20 marcos que proponen categorías, pautas y orientaciones de diseño de aplicaciones basadas en este paradigma, y que permiten definir a la IT y analizarla desde diversos puntos de vista. Se presentaron estos marcos agrupados por diferentes categorías. A partir del análisis se generó un mapa conceptual que busca organizar lo presentado por diferentes autores de referencia en el área. El mapa y la categorización propuesta resultan en un aporte.
2. Componer un estado del arte en relación a las interfaces de interacción tangible en el escenario educativo. Estudiar buenas prácticas por parte de diseñadores y docentes que estén trabajando en esta temática.
 - Se elaboró un estado del arte, a partir de un proceso de revisión sistemática de bibliografía, recuperando experiencias de actividades basadas en IT sobre *tabletops* de los últimos 10 años. Esta revisión permitió conocer distintos dominios en los cuales la IT se ha aplicado y los beneficios encontrados en cada una, destacándose algunas buenas prácticas encontradas. Se analizó así un total de 63 experiencias que, en general, han demostrado éxito al alcanzar los objetivos planteados por los investigadores de cada una.
3. Diseñar y desarrollar una herramienta de autor que permita la construcción de actividades educativas digitales basadas en el paradigma de interacción tangible.
 - Se abordó un estudio que recupera un estado del arte de las herramientas para la creación de aplicaciones basadas en IT con la finalidad de conocer las características y descubrir vacancias en relación a los aspectos deseados. Se consideraron 7 herramientas para su análisis siguiendo una serie de criterios retomados de (Tetteroo y cols., 2013; Sanz, Artola, y cols., 2017). Entre las aplicaciones analizadas pudo notarse algunas vacancias relacionadas con el empaquetamiento de diferentes tipos de actividades secuenciadas en forma de proyecto y con la posibilidad de exportar las aplicaciones con estándares que faciliten su acceso y su reutilización por parte de la comunidad educativa.
 - Se presentó una descripción del diseño y el desarrollo de la herramienta de autor, EDIT, propuesta en el marco de esta tesis como uno de los principales aportes que realiza. La herramienta propone atender aspectos no abordados por las herramientas analizadas.

4. Generar un estudio de caso que permita relevar las opiniones y aceptación de los docentes en el uso de este tipo de herramienta de autor.
 5. Analizar los resultados obtenidos del estudio de caso, de manera tal de lograr el uso efectivo de la herramienta de autor en escenarios educativos específicos, mediante un plan de difusión que se conformará como parte de este trabajo.
- Se llevó a cabo un estudio de caso para validar la aceptación de la EDIT por parte de los docentes, sus opiniones, percepciones y reflexiones acerca de la IT en contextos educativos. Se trabajó con una metodología, en la cual se invitó a docentes a utilizar EDIT. El objetivo del estudio fue analizar la aceptación de EDIT por parte de los docentes como una tecnología para la creación de actividades educativas basadas en IT. La aceptación de EDIT se estudió en base a una adaptación del Modelo de Aceptación de la Tecnología (TAM) realizada por Teo (2009). Se trabajó con una muestra de 38 docentes que se desempeñan en diversos niveles educativos. Los resultados mostraron una alta aceptación por parte de los docentes que participaron de las experiencias en relación a incorporar EDIT en actividades educativas. Estos resultados permitieron arribar a una serie de conclusiones.

CONCLUSIONES

El presente trabajo ha tenido como objetivo investigar sobre la Interacción Tangible en el escenario educativo y desarrollar tecnología basada en esta forma de interacción para su aplicación en procesos de enseñanza y aprendizaje. Para tal fin, se elaboraron una serie de objetivos específicos, mencionados previamente.

El primer objetivo se trabajó mediante una indagación acerca de los diferentes paradigmas de interacción, que han surgido a lo largo de la evolución de las interfaces. Se puso mayor énfasis en el estudio de la IT, por ser el tema de interés en este trabajo. Se realizó una introducción a la evolución de las interfaces, y luego, un estudio de diversos marcos que proponen categorías, pautas y orientaciones de diseño, y que permiten definir a la IT y analizarla desde diversos puntos de vista. El segundo objetivo se aborda con una revisión sistemática de bibliografía, recuperando experiencias de actividades basadas en IT sobre tabletops de los últimos 10 años. Esta revisión permite conocer distintos dominios en los cuales la IT se ha aplicado y los beneficios encontrados en cada una, destacándose algunas buenas prácticas encontradas. Para el tercer objetivo fue necesario realizar un estudio acerca del estado del arte de las herramientas de autor para la creación de aplicaciones basadas en IT. Esto tuvo como finalidad conocer las características de las herramientas desarrolladas y descubrir vacancias en relación a los aspectos deseados. En base a estas vacancias se diseñó y desarrolló una herramienta de autor, EDIT. Finalmente, se realizó un estudio de caso para validar la aceptación de EDIT por parte de los docentes, sus opiniones, percepciones y reflexiones acerca de la IT en contextos educativos.

Las respuestas a las preguntas de investigación presentadas anteriormente permiten arribar a algunas conclusiones que se presentan a continuación.

El surgimiento de la IT se relaciona con la evolución de las interfaces, que ha avanzado desde las basadas en comando (CLI), a las GUI, y luego a las naturales (NUI). Esta evolución ha motivado el auge de paradigmas de realidad híbrida o combinada. La IT nace del trabajo de Ishii y Ullmer (1997) y presenta sus bases en las tendencias de la Computación Ubicua y la RA. Promueve un importante cambio en el diseño de interfaces de usuario, y el modo de interactuar con sistemas informáticos. El objetivo de este nuevo tipo de interfaz fue potenciar la colaboración, el aprendizaje y la toma de decisiones a través de la tecnología digital, y aprovechar las habilidades naturales de los usuarios para manipular objetos y materiales físicos.

La IT permite crear relaciones entre el mundo físico y el mundo digital que puede abordarse desde distintos puntos de vista, de acuerdo al marco conceptual que se utilice. Se observó que los primeros trabajos se han centrado en el desarrollo técnico, las descripciones taxonómicas y los marcos teóricos. Con el tiempo, se visualiza una tendencia a avanzar en el campo y la exploración de la aplicación de ideas teóricas en diseños concretos y estudios empíricos que investigan sistemáticamente los supuestos beneficios de tales diseños. Algunos de los marcos se relacionan con las interfaces tangibles para escenarios educativos, relacionándolos con otros previos, y proponiendo categorías, pautas y orientaciones para el diseño de actividades educativas. Estos marcos atienden a las operaciones cognitivas involucradas, tipos de feedback según grupos destinatarios, objetivos que pueden alcanzarse, entre otros. A partir del análisis realizado, se generó un mapa conceptual que busca organizar todos los aportes realizados por diferentes autores de referencia en el área. Se distinguen 4 categorías de interés: la relación entre los objetos físicos y el entorno (Ullmer y Ishii, 2001; Shaer y cols., 2004), las relaciones entre objetos físicos e información digital (Shaer y cols., 2004; Ullmer y Ishii, 2001; Van Den Hoven y Eggen, 2004; Koleva y cols., 2003), las conceptualización de la entrada y la salida (Fishkin, 2004), y los tipos de objetos físicos de acuerdo a su metáfora en el entorno digital (Underkoffler y Ishii, 1999; Holmquist y cols., 1999; Van Den Hoven y Eggen, 2004; Antle, 2007). Respecto al diseño de

aplicaciones basadas en IT, 7 de los marcos estudiados ofrece guías, heurísticas y/o buenas prácticas de diseño. Tal es el caso de (Dourish, 2001; Hornecker y Buur, 2006; Sharlin y cols., 2004; Jacob y cols., 2008; Rogers, Scaife, Harris, y cols., 2002; Benford y cols., 2005; Wensveen y cols., 2004). Algunos de ellos han considerado la relación acción/efecto, y han resultado de interés para la creación de la herramienta de autor propuesta en esta tesis. Se puso especial énfasis en la revisión de trabajos enfocados en aplicaciones basadas en IT para el área educativa, de acuerdo a las preguntas de investigación P6, p7, P8 y P9. Se analizaron y categorizaron 6 marcos que aportan fundamentos y guías de diseño para la integración de la IT en experiencias educativas.

Las 63 experiencias educativas basadas en IT sobre *tabletops* estudiadas han demostrado éxito al alcanzar los objetivos planteados por los investigadores de cada una. Una de las conclusiones a las que se arribó, es que las aplicaciones abarcan dominios muy diversos como la Biología, la Química, la Programación, las Matemáticas, la Terapia ocupacional, las Artes, etc. Entre los resultados reportados por los autores de estos trabajos se manifiesta que han favorecido múltiples aspectos tales como la colaboración, la motivación, entre otros. Cada uno de estos aspectos fue fomentado a través de distintas actividades que fueron diseñadas en función de las metas planteadas (por ejemplo, el uso de objetos físicos específicos para fomentar la colaboración). Un aspecto a destacar es que el diseño y el desarrollo de estas experiencias, en muchos casos, involucraron la participación de los expertos en dominios específicos y a los usuarios finales, llevando adelante estrategias de diseño participativo. Esta situación ha llevado a investigar un conjunto de trabajos y proyectos en los que se proponen herramientas para la creación de aplicaciones IT, orientadas a usuarios sin conocimientos en programación (expertos en el dominio de la aplicación).

Las herramientas de autor analizadas permitieron conocer su disponibilidad, tecnología subyacente, modo en el que permite la tarea de creación, y los aspectos de personalización que posibilita. Las principales vacancias observadas se relacionan con la posibilidad de crear proyectos que incluyan una secuencia de actividades, y exportar dichos proyectos siguiendo estándares para poder encontrarlos y compartirlos con la comunidad.

A partir de los resultados del análisis de las herramientas de autor se realizó el aporte central de esta tesis que consiste en una herramienta de autor EDIT, que posibilita crear actividades educativas basadas en IT, secuenciadas según las necesidades del docente e integradas en un proyecto. EDIT fundamenta su diseño en varios de los marcos revisados y en modelos de herramientas de autor orientadas a la creación de materiales educativos. Para evaluar EDIT se consideró la aceptación de esta herramienta tecnológica por parte de los docentes como un indicador de su intención de uso. Al mismo tiempo, se analizaron las opiniones, percepciones y reflexiones de los docentes en relación a la IT en el campo educativo.

Los resultados permitieron afirmar que los docentes intervinientes, se mostraron interesados y motivados por las posibilidades de la IT en sus contextos educativos. Además, como resultados de las sesiones del estudio de caso, los docentes muestran un alto nivel de aceptación de EDIT. Una conclusión de interés a la que se arriba es que los participantes requirieron de la visualización de ejemplos para poder imaginar sus propias actividades educativas con IT. Se encontró una tendencia a creer que las aplicaciones basadas en IT favorecen situaciones de aprendizaje con niños, y/o en el ámbito de educación especial y no tanto con adolescentes o adultos. En general, luego de conocer aplicaciones basadas en IT para adolescentes o adultos, se notó un cambio en la opinión de los docentes manifestando la necesidad de contar con ejemplos. Los docentes finalmente indicaron creer que esta tecnología puede aportar en escenarios educativos de nivel primario y secundario. Resulta una necesidad así, establecer un espacio donde los docentes puedan compartir aplicaciones IT creadas y al cual puedan acceder, tanto para reusar aplicaciones, pudiendo adaptarlas a sus contextos, como para encontrar inspiración para crear sus propias aplicaciones IT. Además, entre los resultados obtenidos se puede mencionar que los docentes perciben a la IT como interesante, motivadora, y como un nuevo desafío en el diseño de actividades para sus ámbitos educativos. Otro aspecto que fue considerado por los docentes en el desarrollo de este tipo de actividades educativas fue la importancia de contemplar la dinámica de las actividades con IT cuando se trabaja con grupos numerosos. Para profundizar el estudio acerca de la opinión y percepción de la IT por parte de docentes, se realizó una encuesta a 12 docentes que participaron de un curso sobre interfaces de RA, RV e IT. Estos datos enriquecieron los resultados previos respecto al interés y motivación por las posibilidades de la IT en los contextos educativos.

Finalmente se pudo concluir que todos los docentes valoraron a EDIT como un puente para el acercamiento de la IT a las aulas.

De esta manera, se considera que esta tesis ha logrado abordar el total de las preguntas de investigación planteadas al inicio de este recorrido y constituye un aporte al área de Interacción Tangible, en particular aplicada en el ámbito de la educación.

LÍNEAS DE TRABAJO FUTURO

Se planea evolucionar hacia la siguiente versión de EDIT incorporando nuevas plantillas que habiliten la creación de otros tipos de actividades basadas en IT. También se planea atender a las propuestas realizadas por los docentes participantes de las sesiones del estudio de caso, tales como incorporar más guías y ayudas, y generar algunas alertas adicionales en la plantilla de asociación, entre otras.

Un aspecto central es la creación de un entorno o comunidad donde los docentes puedan compartir las aplicaciones IT, para fomentar la divulgación de esta tecnología. Al mismo tiempo, dejar disponible EDIT y su código fuente para socializar este proyecto y continuar con el objetivo de acercar la IT a los contextos educativos. En este sentido, cabe resaltar que la necesidad de contar con una tabletop se vuelve un aspecto de limitación en algunos casos, aunque en otros se ha visto predisposición y entusiasmo por construir una para la institución educativa.

Una línea que se deja abierta para su profundización es la de integrar como parte de la herramienta EDIT diferentes niveles de usuarios y posibilitar para aquellos más expertos el diseño del comportamiento de objetos activos.

Finalmente, se propone avanzar en las tareas de difusión de la IT para Educación, a través de talleres y espacios de divulgación.

En relación a los aspectos teóricos, se abre la puerta para trabajar en la profundización de los marcos de diseño de aplicaciones IT, especialmente, relacionados con el ámbito educativo. Además, de considerarlos en el desarrollo de nuevas aplicaciones.

Se espera así que esta tesis pueda ser continuada y profundizada por otros investigadores y sea un aporte para la comunidad.

BIBLIOGRAFÍA

Antle, A. N. (2007). Designing tangibles for children: What designers need to know. En *Chi '07 extended abstracts on human factors in computing systems* (pp. 2243–2248). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/1240866.1240988> doi: 10.1145/1240866.1240988

Antle, A. N., Droumeva, M., y Corness, G. (2008). Playing with the sound maker: do embodied metaphors help children learn? En *Proceedings of the 7th international conference on interaction design and children* (pp. 178–185).

Antle, A. N., y Wise, A. F. (2013). Getting down to details: Using theories of cognition and learning to inform tangible user interface design. *Interacting with Computers*, 25(1), 1–20.

Artola, V., Sanz, C., Baldassarri, S., Zangara, A., Guisen, A., y Martorelli, S. (2016). Capítulo v. interacción tangible. aplicaciones en el escenario educativo. el caso de itcol. En H. G. S. Catalina, S. C. Verónica, C. G. Cornelio, y N. Marcelo (Eds.), *Aplicación de la tecnología en el aprendizaje. casos de argentina y méxico* (p. 175–208). Guadalajara. México: Universidad de Guadalajara.

Bakker, S., Antle, A. N., y Van Den Hoven, E. (2012). Embodied metaphors in tangible interaction design. *Personal and Ubiquitous Computing*, 16(4), 433–449.

Benford, S., Schnädelbach, H., Koleva, B., Anastasi, R., Greenhalgh, C., Rodden, T., ... others (2005). Expected, sensed, and desired: A framework for designing sensing-based interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(1), 3–30.

Diego-Mendoza, J., Márquez-Domínguez, J. A., y Sabino-Moxo, B. A. (2014). Desarrollo de una interfaz natural de usuario para rehabilitación motriz. *Revista Salud y Administración*, 1(3), 3–15.

Dillenbourg, P. (2016). The evolution of research on digital education. *International Journal of Artificial Intelligence in Education*, 26(2), 544–560.

Dourish, P. (2001). *Where the action is: Where the action is: The foundations of embodied interaction*. MIT Press.

Fitzmaurice, G. W., Ishii, H., & Buxton, W. A. (1995, May). Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 442–449).

- Fishkin, K. P. (2004, septiembre). A taxonomy for and analysis of tangible interfaces. *Personal Ubiquitous Comput.*, 8(5), 347–358. Descargado de <http://dx.doi.org/10.1007/s00779-004-0297-4> doi: 10.1007/s00779-004-0297-4
- Holmquist, L. E., Redström, J., y Ljungstrand, P. (1999). Token-based access to digital information. En *Proceedings of the 1st international symposium on handheld and ubiquitous computing* (pp. 234–245). London, UK, UK: Springer-Verlag. Descargado de <http://dl.acm.org/citation.cfm?id=647985.743869>
- Hornecker, E. (2002). Understanding the benefits of graspable interfaces for cooperative use. En *Coop* (pp. 71–87).
- Hornecker, E., y Buur, J. (2006). Getting a grip on tangible interaction: A framework on physical space and social interaction. En *Proceedings of the sigchi conference on human factors in computing systems* (pp. 437–446). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/1124772.1124838> doi:10.1145/1124772.1124838
- Ishii, H., y Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. En *Proceedings of the acm sigchi conference on human factors in computing systems* (pp. 234–241). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/258549.258715> doi:10.1145/258549.258715
- Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., y Zigelbaum, J. (2008). Reality-based interaction: A framework for post-wimp interfaces. En *Proceedings of the sigchi conference on human factors in computing systems* (pp. 201–210). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/1357054.1357089> doi: 10.1145/1357054.1357089
- Koleva, B., Benford, S., Hui Ng, K., y Rodden, T. (2003, 09). A framework for tangible user interfaces. En *Proceedings of physical interaction (pi03) - workshop on real world user interfaces*.
- Manches, A., O'Malley, C., y Benford, S. (2010). The role of physical representations in solving number problems: A comparison of young children's use of physical and virtual materials. *Computers & Education*, 54(3), 622–640.
- Martinez-Maldonado, R., Dimitriadis, Y., Clayphan, A., Muñoz-Cristóbal, J. A., Prieto, L. P., Rodríguez-Triana, M. J., y Kay, J. (2013). Integrating orchestration of ubiquitous and pervasive learning environments. En *Proceedings of the 25th Australian computer-human interaction conference: Augmentation, application, innovation, collaboration* (pp. 189–192).
- Marshall, P., Price, S., y Rogers, Y. (2003). Conceptualising tangibles to support learning. En *Proceedings of the 2003 conference on interaction design and children* (pp. 101–109). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/953536.953551> doi: 10.1145/953536.953551
- Marshall, P. (2007). Do tangible interfaces enhance learning? En *Proceedings of the 1st international conference on tangible and embedded interaction* (pp. 163–170). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/1347390.1347425> doi:10.1145/1347390.1347425
- Price, S. (2008). A representation approach to conceptualizing tangible learning environments. En *Proceedings of the 2nd international conference on tangible and embedded interaction* (pp. 151–158). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/1347390.1347425> doi:10.1145/1347390.1347425
- Price, S., Rogers, Y., Scaife, M., Stanton, D., y Neale, H. (2003). Using 'tangibles' to promote novel forms of playful learning. *Interacting with computers*, 15(2), 169–185.
- Rogers, Y., Scaife, M., Gabrielli, S., Smith, H., y Harris, E. (2002). A conceptual framework for mixed reality environments: designing novel learning activities for young children. *Presence: Teleoperators & Virtual Environments*, 11(6), 677–686.
- Sanz, C., Artola, V., Guisen, A., Javier, M., Eva, C., y Sandra, B. (2017). Shortages and challenges in augmentative communication through tangible interaction using a user-centered design and assessment process. *J. UCS*, 23(10), 992–1016.
- Shaer, O., Leland, N., Calvillo-Gamez, E. H., & Jacob, R. J. (2004). The TAC paradigm: specifying tangible user interfaces. *Personal and Ubiquitous Computing*, 8(5), 359–369.

Sharlin, E., Watson, B., Kitamura, Y., Kishino, F., y Itoh, Y. (2004). On tangible user interfaces, humans and spatiality. *Personal and Ubiquitous Computing*, 8(5), 338–346.

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, 52(2), 302–312.

Tetteroo, D., Soute, I., y Markopoulos, P. (2013). Five key challenges in end-user development for tangible and embodied interaction. En *Proceedings of the 15th acm on international conference on multimodal interaction* (pp. 247–254). New York, NY, USA: ACM. Descarga- [dodehttp://doi.acm.org/10.1145/2522848.2522887](http://doi.acm.org/10.1145/2522848.2522887) doi:10.1145/2522848.2522887

Ullmer, B., y Ishii, H. (2001, 01). Emerging frameworks for tangible user interfaces. En *Human-computer interaction in the new millenium* (p. 579 - 601).

Underkoffler, J., y Ishii, H. (1999). Urp: A luminous-tangible workbench for urban planning and design. En *Proceedings of the sigchi conference on human factors in computing systems* (pp. 386–393). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/302979.303114> doi:10.1145/302979.303114

Van Den Hoven, E., y Eggen, B. (2004). Tangible computing in everyday life: Extending current frameworks for tangible user interfaces with personal objects. En *European symposium on ambient intelligence* (pp. 230–242).

Weiser, M. (1999, jul). The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3), 3–11. Descargado de <http://doi.acm.org/10.1145/329124.329126> doi: 10.1145/329124.329126

Wensveen, S. A., Djajadiningrat, J. P., y Overbeeke, C. (2004). Interaction frogger: a design framework to couple action and function through feedback and feedforward. En *Proceedings of the 5th conference on designing interactive systems: processes, practices, methods, and techniques* (pp. 177–184).

Zuckerman, O., Arida, S., y Resnick, M. (2005). Extending tangible interfaces for education: Digital montessori-inspired manipulatives. En *Proceedings of the sigchi conference on human factors in computing systems* (pp. 859–868). New York, NY, USA: ACM. Descargado de <http://doi.acm.org/10.1145/1054972.1055093> doi:10.1145/1054972.1055093

Método de Reducción de Incertidumbre Basado en Algoritmos Evolutivos y Paralelismo Orientado a la Predicción y Prevención de Desastres Naturales

Doctorado en Ciencias de la Computación,
Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional
de San Luis

Autor: Miguel Méndez-Garabetti

LICPaD, Laboratorio de Investigación en Cómputo Paralelo/Distribuido, Departamento de Ingeniería en Sistemas de Información, Facultad Regional Mendoza, Universidad Tecnológica Nacional, Mendoza, Argentina.

Fecha de defensa: 20 de julio de 2020.

Director:

Dr. Bianchini Germán

LICPaD, Laboratorio de Investigación en Cómputo Paralelo/Distribuido, Departamento de Ingeniería en Sistemas de Información, UTN-FRM.

gbianchini@frm.utn.edu.ar

Codirectora:

Dra. Verónica Gil Costa

Laboratorio de Investigación y Desarrollo Científico (LIDIC), Departamento de Informática, Facultad de Ciencias Físico Matemáticas y Naturales, Universidad Nacional de San Luis, Argentina.

ggvcosta@gmail.com

Codirectora Beca Doctoral CONICET:

Dra. Caymes-Scutari Paola

CONICET, LICPaD, Laboratorio de Investigación en Cómputo Paralelo/Distribuido, Departamento de Ingeniería en Sistemas de Información, UTN-FRM.

pcaymesscutari@frm.utn.edu.ar

Resumen

La presente tesis doctoral aborda la problemática de la incertidumbre existente en todo sistema de predicción, focalizando en el desarrollo de métodos de reducción de incertidumbre aplicados a la predicción de fenómenos naturales. Debido a que estos fenómenos suelen causar gran impacto en las comunidades, la flora y la fauna, el ecosistema, entre otros, los sistemas de predicción deben proporcionar respuesta en el menor tiempo posible. Por estos motivos, los métodos propuestos han sido desarrollados utilizando capacidades de alto rendimiento. El primer método desarrollado en esta tesis (ESS-IM), comenzó con el objetivo de lograr una mejora a una metodología previamente desarrollada denominada ESS (Sistema Estadístico Evolutivo). Específicamente se trabajó en el incremento del paralelismo de la metaheurística interna, incorporando una arquitectura basada en modelo de islas bajo un esquema de migración. Este desarrollo logró incrementar la capacidad de búsqueda de la metaheurística interna, impactando de forma directa en un incremento en la calidad de predicción del método. En la validación, ESS-IM fue aplicado en una serie de casos de quemas controladas e incendios forestales. Es importante destacar que, en forma conjunta al desarrollo de la tesis, se llevaron a cabo diferentes investigaciones complementarias, tales como: estudios de sintonización de parámetros, desarrollo de un sistema de generación de mapas de incendios forestales a partir de imágenes satelitales, diseño de una red inalámbrica de sensores como sistema de alerta temprana, entre otros. Finalmente, en la última etapa de la tesis, se implementó una versión híbrida basada en metaheurísticas evolutivas bajo una estrategia colaborativa

basada en islas. El método HESS-IM, se implementó de forma heterogénea (a nivel de hardware), logrando que los resultados obtenidos incrementen la calidad de predicción y eficiencia del método.

Palabras clave: reducción de incertidumbre, paralelismo, computación de alto rendimiento, metaheurísticas evolutivas híbridas poblacionales.

1. Introducción

Los modelos se han convertido en uno de los principales instrumentos de la ciencia moderna, éstos suelen utilizarse en la representación de gran cantidad de sistemas físicos (Stanford University & Center for the Study of Language and Information (U.S.), 1997), inclusive en el ámbito de la ciencia ambiental para efectuar predicciones. La modelización puede ser vista como el establecimiento de relaciones semánticas entre la teoría y los fenómenos u objetos. Generalmente, los modelos representan alguna situación real, pero de manera incompleta e inexacta; esto permite realizar estudios de sistemas complejos de forma aproximada. Usualmente, los modelos deben ser alimentados con parámetros de entrada, los cuales brindan la información necesaria con la cual es posible determinar las condiciones iniciales del fenómeno bajo estudio, y de esta manera estimar sus posibles variaciones en el tiempo. Esto se conoce como “salida del modelo”, la cual representa la evolución del sistema a lo largo del tiempo.

Cuando se lleva a cabo el estudio de un fenómeno complejo mediante la simplificación de la realidad, usando modelos, éstos pueden verse afectados debido a la calidad de la información con que dicho modelo es alimentado. Es decir, el mero hecho de simplificar un fenómeno físico para poder representarlo mediante un modelo implica pérdida de precisión, exactitud y certeza, respecto del comportamiento real del fenómeno. Si junto a esto se considera que dicho modelo “simplificado” es alimentado con información incompleta o de baja calidad, evidentemente tendremos resultados no del todo satisfactorios y distantes de la realidad.

Esta problemática, que surge a partir de la falta de precisión o exactitud en los parámetros de entrada de un modelo, se conoce como incertidumbre. Como bien se menciona en (Ishigami & Homma, 1990): “cuando se analizan sistemas complejos, los resultados calculados incluyen incertidumbres debidas a los propios modelos informáticos, como así también a las existentes en los valores de las variables de entrada o los parámetros utilizados en el análisis. Debido a esto, es importante no solo cuantificar las incertidumbres de salida sino también identificar las variables de entrada dominantes que contribuyen a las incertidumbres de salida”.

1.1. Motivación

En base a esto, se puede decir que en cualquier ámbito donde se realice algún tipo de predicción existirá mayor o menor grado de incertidumbre (tanto se trate de un modelo económico, meteorológico o natural) y con ello el esfuerzo por intentar minimizar sus efectos negativos. Dicho en otras palabras, ante la presencia de incertidumbre (principal fuente de imprecisión en todo sistema de predicción), es necesario intentar eliminarla, o en el mejor de los casos lograr reducirla. En la literatura se utiliza el término Reducción de Incertidumbre o Métodos de Reducción de Incertidumbre para hacer referencia a aquellas técnicas o herramientas que de alguna manera permiten reducir los efectos causados por la falta de precisión en cualquier variable o factor de un sistema de predicción. El concepto de Método de Reducción de Incertidumbre (MRI o Uncertainty Reduction Method, URM) ha sido acuñado en el trabajo de Ishigami y Homma en 1990 (Ishigami & Homma, 1990).

El concepto de MRI se utiliza en diversas áreas de la ciencia, algunos ejemplos son (Bartezzaghi, Verganti, & Zotteri, 1999; Hu, Mahadevan, & Ao, 2018; Kalyuzhnaya & Boukhanovsky, 2015; Li, Hamel, & Azarm, 2010; Purica & Bae, 2016; Qian, Li, & Yang, 2013; Sano, Takeda, & Yamasaki, 2009; Taghribi & Sharifian, 2017). En el contexto de la presente tesis se ha definido un MRI de la siguiente manera: “sistema computacional cuyo objetivo consiste en mejorar la calidad de salida de un modelo determinado, de forma independiente a los datos con que éste haya sido alimentado”. Esto se logra mediante algún tipo de pre-procesamiento y/o pos-procesamiento de los datos de entrada y/o salida del modelo. Las funciones llevadas a cabo por un MRI no deben interferir con las operaciones internas del modelo bajo análisis, es decir, este último debe ser tratado como una caja negra. En este sentido podemos mencionar que la reducción de incertidumbre corresponde al proceso que permite minimizar los efectos negativos causados por la falta de precisión en las variables de entrada de un modelo determinado, permitiendo optimizar el funcionamiento de éste e impactando de forma positiva en la calidad de la salida del mismo. La reducción de incertidumbre puede lograrse de diferentes maneras en función del fenómeno con el que se esté tratando, el modelo utilizado, la implementación realizada, etc. Por ejemplo, en la predicción meteorológica, la reducción de incertidumbre puede llevarse a cabo mediante la medición de variables atmosféricas en tiempo real (Doeswijk, 2007; Farguell et al., 2016).

En esta tesis se abordó el desarrollo de métodos generales de reducción de incertidumbre que puedan ser aplicados en la predicción de diferentes fenómenos naturales. Puntualmente se utilizó como caso de estudio a los incendios forestales dado que para este tipo de fenómenos ha sido factible conseguir datos de casos de quemas, y además desde el contexto local, éste es uno de los fenómenos naturales que ocurren con mayor frecuencia.

2. La problemática de los incendios forestales

Cuando se habla de incendios forestales, comúnmente se connota con los aspectos negativos que éstos poseen. Sin embargo, éstos juegan un rol fundamental en los cambios naturales que ocurren en los ecosistemas de nuestro planeta. Los efectos del fuego, sobre bosques, praderas, etc., propician la diversidad de la vida vegetal y animal. Incluso ciertos tipos de plantas no tienen posibilidad de reproducirse sin el fuego. El fuego inicia procesos naturales que permiten convertir la materia orgánica en nutrientes que posteriormente la lluvia entrega al suelo, proporcionando un semillero fértil rejuvenecido para las plantas (Graham, 1999). Sin embargo, los incendios forestales fuera de control suelen generar grandes pérdidas y daños, destacándose entre ellos: pérdidas de vidas humanas, daños a la flora, la fauna y al suelo (Morgan et al., 2001). Las pérdidas económicas son difíciles de estimar debido a la gran cantidad de sectores que pueden verse afectados a corto, mediano o largo plazo, pero en general los recursos ambientales de toda zona afectada suelen ser considerablemente damnificados.

Los incendios forestales afectan millones de hectáreas año tras año haciendo de estos fenómenos uno de los principales agentes de perturbación de los ecosistemas del planeta. En Argentina, según el último informe publicado por el Ministerio de Ambiente y Desarrollo Sustentable de la Presidencia de la Nación Argentina, año 2016 (*Estadísticas de Incendios Forestales 2016*, 2016), durante 2016 se consumieron 1.072.642 hectáreas. De este total un 54,74% corresponde a superficie de pastizal, un 28,04% de arbustal, un 16,47% de bosque nativo y el resto corresponde a bosque cultivado y superficie sin determinar. Además, los incendios forestales ocurren con mayor frecuencia en época

estival, donde se incrementa la sequedad, las temperaturas y/o el viento, lo que en conjunto facilita la propagación del fuego destruyendo así extensas superficies.

Debido a las pérdidas y daños que pueden producir, desde diferentes áreas de la ciencia se trabaja constantemente en el desarrollo de herramientas, sistemas, estrategias, etc., que permitan reducir los efectos negativos que éstos ocasionan. Entre estos se pueden encontrar: estrategias de evacuación, tecnologías para la extinción del fuego, cortafuegos, modelos de comportamiento, simuladores de incendios, sistemas de predicción, entre otros. Estas herramientas suelen utilizarse en distintas etapas del proceso de incendios forestales: prevención, predicción, detección y monitorización.

3. Predicción de comportamiento de incendios forestales

Tal como se mencionó previamente, en el contexto de esta tesis, se ha trabajado en el desarrollo de MRI aplicados a la predicción del comportamiento de los incendios forestales. Es decir, el proceso que consiste en intentar pronosticar el comportamiento de la propagación de un incendio forestal ya iniciado, con el objetivo de conocer las zonas que presentan mayor riesgo de ser alcanzadas por el fuego y tomar decisiones o medidas de acuerdo a ello. La simulación de la propagación de incendios forestales es una tarea difícil desde el punto de vista computacional, tanto por la complejidad involucrada en los modelos utilizados, como así también por la incertidumbre que afecta a este proceso. Más allá de la imprevisibilidad del comportamiento de los incendios forestales, la incertidumbre proviene de diferentes fuentes tales como: datos inexactos y/o faltantes, comprensión científica incompleta de la respuesta ecológica al fuego y de la respuesta del comportamiento del fuego a los tratamientos de mitigación, entre otros (Thompson & Calkin, 2011).

El concepto general de predicción del comportamiento de incendios forestales consiste en un simulador de incendios (basado en algún modelo de comportamiento de incendios forestales), tratado como una caja negra, alimentado con ciertos parámetros de entrada (**PE**) que representan condiciones actuales del incendio, tales como: velocidad y dirección del viento, vegetación, características de la superficie, entre otros. El simulador necesita, además de los **PE**, la Línea de Fuego Real (**LFR**) del incendio en el instante de tiempo t_n (**LFR_n**). Una vez que el simulador es ejecutado, éste genera en su salida la predicción de la línea de fuego o Línea de Fuego Predicha (**LFP**) para el instante de tiempo posterior, t_{n+1} . Por supuesto, se espera que dicha predicción tenga un alto grado de coincidencia con el estado real del incendio para dicho instante de tiempo (**LFR_{n+1}**) de la mejor manera posible. Este enfoque también es conocido como predicción de un paso adelante (Nowak, 2002). Sin embargo, debido a la complejidad e incertidumbre del modelo de comportamiento del fuego, la incertidumbre en los parámetros de entrada, y a que la predicción está basada en una única simulación, esta metodología no suele ofrecer una predicción lo suficientemente aproximada a la realidad como para ser considerada una herramienta confiable para la toma de decisiones. Algunos ejemplos de predicción clásica son (Andrews et al., 2008; DX, PR, & L, 1998; Finney, 1998; Heinsch & Andrews, 2010; Lopes, Cruz, & Viegas, 2002; Wallace, 1993). Las limitaciones presentes en el enfoque de predicción clásica han hecho evidente la necesidad de incorporar técnicas de reducción de incertidumbre que permitan mejorar el rendimiento (en calidad y tiempo de respuesta) de los sistemas de predicción actuales. En este contexto, los desarrollos de los MRI cobran vital importancia ya que han demostrado, tras ser aplicados en diferentes áreas de la ciencia, efectividad a la hora de minimizar los efectos negativos causados por las diferentes incertidumbres permitiendo mejorar la calidad de salida de los modelos tratados. Además, debido a que estos sistemas suelen trabajar con grandes conjuntos de

datos sobre los cuales es necesario realizar complejas operaciones, éstos sistemas suelen requerir de sistemas de computación de alto rendimiento (Buyya, 1999), los cuales permiten reducir el tiempo de procesamiento al utilizar múltiples unidades de cómputo en simultáneo.

4. Contribuciones de la tesis

La presente tesis contribuye al desarrollo de métodos de reducción de incertidumbre capaces de asistir y participar en el proceso de resolución de problemas de alta magnitud y complejidad computacional, tal como la predicción y prevención de desastres naturales. Este tipo de problemas puede ser clasificado en la categoría de Grand Challenge Problems (Wilkinson & Allen, 1999), es decir, problemas de alta complejidad, y a la vez importantes (desde el punto de vista social y económico) donde se requieren varios órdenes de magnitud de recursos para resolverlos.

Los métodos desarrollados, si bien han sido concebidos como generales (ya que podrían ser aplicados a diferentes fenómenos naturales con características de propagación), fueron validados tras su aplicación en la predicción del comportamiento de incendios forestales. Los métodos desarrollados pertenecen a la categoría de Métodos Guiados por Datos con Solución Múltiple Solapada (DDM-MOS, Data Driven Methods with Multiple Overlapping Solutions) (Germán. Bianchini, 2006), a la cual también pertenecen los métodos: S²F²M (Germán Bianchini, Denham, Cortés, Margalef, & Luque, 2010), ESS (G. Bianchini, Caymes-Scutari, & Méndez-Garabetti, 2015) y ESSIM-DE (Tardivo, Caymes-Scutari, Méndez-Garabetti, & Bianchini, 2018). Es importante resaltar que esta tesis ha permitido incrementar dos de las características principales de este tipo de metodologías, a saber: a) la calidad de predicción, y b) el tiempo de respuesta, esto mediante la utilización de metaheurísticas híbridas paralelas (Alba, Luque, & Nesmachnow, 2013) en conjunto con técnicas de análisis estadístico y computación de alto rendimiento.

Métodos desarrollados:

- ESS-IM (M. Méndez-Garabetti, Bianchini, Caymes-Scutari, & Tardivo, 2016): Desarrollo del método de reducción de incertidumbre denominado Sistema Estadístico Evolutivo con Modelo de Islas (Evolutionary Statistical System with Island Model, ESS-IM). Este método ha sido aplicado a la predicción del comportamiento de incendios forestales, logrando incrementar considerablemente la calidad de predicción, respecto de metodologías previas. Dicha mejora es lograda mediante el incremento del paralelismo de la metaheurística interna.
- HESS-IM (Miguel Méndez-Garabetti, Bianchini, Tardivo, Caymes Scutari, & Gil Costa, 2017): Desarrollo del método de reducción de incertidumbre denominado Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (Hybrid Evolutionary Statistical System with Island Model, ESS-IM). Este desarrollo, también aplicado a la predicción del comportamiento de incendios forestales, ha permitido incrementar tanto la calidad de predicción como el tiempo de respuesta. Esto ha sido posible mediante la utilización de un esquema híbrido a nivel de la técnica de optimización basada en metaheurísticas, y uno heterogéneo a nivel de utilización de hardware. Para ello el método implementa un framework capaz de operar con diferentes metaheurísticas de forma colaborativa (conformando así una metaheurística híbrida) y el paralelismo necesario ha sido implementado basado en una arquitectura heterogénea CPU-GPU.

5. Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (HESS-IM)

HESS-IM es un método general de reducción de incertidumbre que ha sido planteado como un framework adaptativo para la implementación de técnicas de optimización colaborativa basada en diferentes metaheurísticas evolutivas poblacionales. De esta manera, HESS-IM permite vincular n metaheurísticas evolutivas para resolver de forma colaborativa la optimización de soluciones del método de predicción. HESS-IM implementa una metaheurísticas híbrida basada en tres metaheurísticas poblacionales evolutivas: a) Algoritmos Evolutivos, b) Evolución Diferencial y c) Optimización por Cúmulo de Partículas, bajo un esquema de colaboración basado en migraciones. Desde el punto de vista de paralelismo se utiliza MPI conjuntamente con CUDA (David B. Kirk, 2013) para lograr reducir el tiempo de procesamiento.

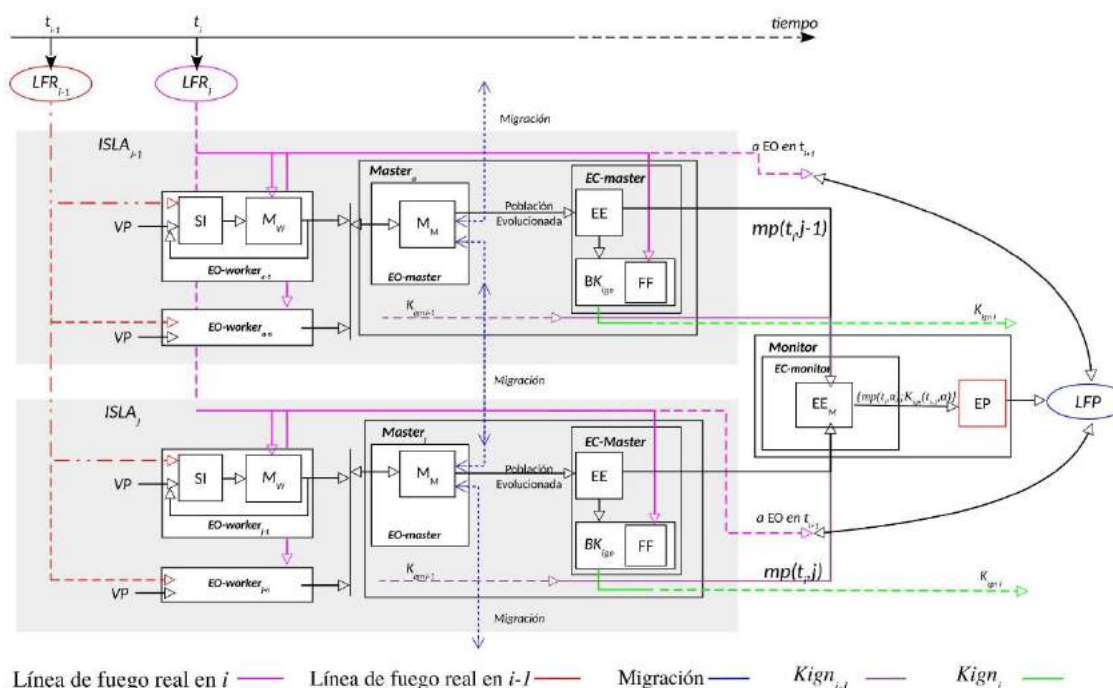


Figura 1 - HESS-IM: Sistema Estadístico Híbrido Evolutivo con Modelo de Islas. SI: simulador de incendios; M_M : etapa metaheurística en master; M_W : etapa metaheurística en worker (evaluación de fitness); EO: etapa de optimización; EE: etapa estadística; BK_{ign} : búsqueda de K_{ign} ; K_{ign} : valor clave de ignición; FF: función de fitness; EC: etapa de calibración; EP: etapa de predicción; LFP: línea de fuego predicha; LFR_x : línea de fuego real en tiempo x ; VP: vector de parámetros de entrada; EEM: etapa estadística en proceso monitor; mp : mapa de probabilidad.

La técnica de optimización de HESS-IM se implementa en un esquema de paralelización basado en islas con doble jerarquía master-worker. Este esquema permite a cada isla operar con una metaheurística diferente sobre el mismo espacio de búsqueda, intercambiando entre sí soluciones candidatas cada cierto intervalo de tiempo, o a demanda de determinados eventos.

Un esquema general del funcionamiento de HESS-IM se observa en la Fig. 1. La cadena de predicción comienza en el proceso Monitor, éste envía a cada una de las n islas dos conjuntos de datos iniciales: a) conjunto de datos del incendio: un mapa con el frente de fuego real, intervalos de tiempo a considerar para cada paso de predicción, valores y rangos para cada parámetro de entrada del modelo, y b) conjunto de datos de las metaheurísticas: cantidad de islas, metaheurística que cada isla deberá utilizar, cantidad

de workers por isla, parámetros de migración, parámetros propios de cada metaheurística, etc.

Dichos conjuntos de datos son recibidos por el proceso master de cada isla, donde se instancia la Etapa de Optimización $\mathbf{EO}_{\text{master}}$, que da inicio a la metaheurística requerida por el proceso monitor, ya sea: EA, DE o PSO. Dicho procesamiento se realiza en la Etapa Metaheurística, la cual se encuentra dividida en dos subetapas: por un lado, la Etapa Metaheurística, del proceso master \mathbf{MM} , y por el otro, la Etapa Metaheurística de los procesos workers \mathbf{MW} . La etapa \mathbf{MM} realiza diferentes operaciones dependiendo del momento en que se encuentre operando el método y en función de la metaheurística configurada en dicha isla. Sin embargo, existen ciertas funciones que todas las islas realizan en común tales como:

1. Inicialización de la población: al inicio del proceso de predicción, esta etapa se encarga de inicializar la población de individuos y de enviar los mismos a los workers para que éstos sean tratados.
2. Migración de individuos: mientras los workers procesan a los individuos, la etapa \mathbf{MM} lleva a cabo la migración de individuos hacia las islas vecinas. La migración se realiza en función de la topología de comunicación y tipo de migración configurada. Esta última puede ser: a) con mayor frecuencia entre islas que operen con la misma metaheurística, o b) con un único valor de frecuencia para todas las islas.
3. Evolución de la población: cuando se alcanza la cantidad máxima de iteraciones o el umbral de aptitud, dentro de cada iteración, se realiza la evolución de la población hacia la próxima generación evolutiva.
4. Finalización y envío: cuando la población ha evolucionado, cumpliendo la condición de finalización, se envía, desde la etapa \mathbf{MM} la población final de individuos hacia la Etapa de Calibración ($\mathbf{EC}_{\text{master}}$).

Por su lado, los workers tienen como función principal evaluar la aptitud de cada individuo mediante la utilización del simulador de comportamiento de incendios forestales (\mathbf{SI}). Dicho simulador está basado en el modelo definido por Rothermel (Rothermel, 1972) e implementado mediante la librería fireLib (Collin D. Bevins, 1996) para CPU y CUDA-FGM (Sousa, dos Reis, & Pereira, 2012) para GPU, la cual corresponde a una implementación de firelib para GPU (Kindratenko, Enos, & Shi, 2009; Tsutsui & Collet, 2013).

El cálculo de la evaluación de aptitud en el instante t_i necesita contar con la línea de fuego real del incendio (\mathbf{LFR}) en t_{i-1} (es decir, \mathbf{LFR}_{i-1}) y los valores de los parámetros de entrada, los cuales se almacenan en los vectores de parámetros (\mathbf{VP}) (los individuos de la población). La función de aptitud utilizada para determinar la calidad de cada individuo está basada en el índice de Jaccard (Real & Vargas, 1996). Esta función permite comparar el mapa simulado con el mapa real en t_i (\mathbf{LFR}_i) y, en función de la diferencia entre los dos mapas, se determina el valor de aptitud para cada individuo. La Etapa de Calibración del proceso master ($\mathbf{EC}_{\text{master}}$) calcula un mapa de probabilidad en función de todos los individuos de la población final. Dicho mapa se utiliza para obtener el valor clave de ignición (Key Ignition Value, \mathbf{K}_{ign}), el cual representa el patrón de comportamiento del incendio y es obtenido en la etapa Búsqueda del \mathbf{K}_{ign} (\mathbf{BK}_{ign}).

Para cada instante de tiempo i , cada isla j genera un $\mathbf{K}_{\text{ign}(t_i, j)}$ y un mapa de probabilidad $\mathbf{mp}(t_i, j)$, los cuales son enviados a la etapa de calibración del proceso monitor ($\mathbf{EC}_{\text{Monitor}}$). El primero es enviado a $\mathbf{EC}_{\text{Monitor}}$ en t_{i+1} y el segundo a $\mathbf{EC}_{\text{Monitor}}$ en t_i . Éstos valores finalmente son ingresados en la etapa de predicción (\mathbf{EP}) la cual realiza el cálculo de predicción (línea de fuego predicha, \mathbf{LFP}) para el próximo instante de tiempo \mathbf{LFR}_{i+1} .

Debido a esto HESS-IM no genera predicción en el primer instante de tiempo, ya que en dicho momento se calcula el primer K_{ign} .

Es importante tener en cuenta que la hibridación tiene en consideración las diferentes velocidades de los algoritmos, con el propósito de que todas las metaheurísticas avancen al mismo tiempo. Se estudiaron configuraciones de cada metaheurística con el objetivo de garantizar dicho fin, aunque de igual manera los intercambios de información entre las islas se realizan de manera asincrónica, evitando de este modo cualquier problema por desacoplamiento de velocidades.

6. Resultados obtenidos

La hibridación implementada en HESS-IM utilizando tres metaheurísticas poblacionales evolutivas ha permitido mejorar el optimizador del método incrementando la calidad de predicción en una serie de 6 casos que quemas controladas y 3 incendios forestales reales. La hibridación se ve mejorada utilizando un esquema de migración basado en topología de anillo entre diferentes metaheurísticas utilizando una configuración de parámetros clásica. La metodología ha sido validada aplicando la misma sobre casos de quemas controladas como así también en casos de incendios forestales. La implementación CPU-GPU del método ha permitido mejorar también el rendimiento del método en términos de tiempo de ejecución, siendo este un factor de suma importancia en los sistemas abocados a resolver problemáticas con gran impacto social.

7. Conclusiones y trabajo futuro

En esta tesis se ha trabajado en el tratamiento de la reducción de incertidumbre en relación a los métodos de predicción y prevención de desastres naturales. Puntualmente se ha abordado como caso de estudio la problemática de los incendios forestales, uno de los más grandes problemas que afecta extensas regiones a lo largo del mundo y de nuestro país. El problema en cuestión ha sido tratado desde diferentes perspectivas. Por un lado, el enfoque principal de esta tesis consistió en el desarrollo de métodos de reducción de incertidumbre que puedan ser utilizados para predecir el comportamiento de este fenómeno. Y, por otro lado, se ha avanzado en otras aristas, como el diseño de un sistema integral de detección, alerta temprana y predicción de incendios forestales, el cual podría brindar asistencia en la lucha y mitigación de incendios, como así también en la prevención de los mismos. Este sistema integra los métodos propuestos, con el uso de redes inalámbricas de sensores y un sistema de generación automática de mapas alimentado por imágenes satelitales. Esta línea de trabajo podría hacer posible que este tipo de metodologías llegasen a ser implementadas y adoptadas por entes gubernamentales, permitiendo de esta manera minimizar los daños causados por este fenómeno (esta línea se encuentra dentro del plan de trabajo futuro).

En la presente tesis se han propuesto y desarrollado dos métodos de reducción de incertidumbre capaces de asistir y participar en el proceso de resolución de problemas de alta magnitud y complejidad computacional, tal como la predicción y prevención de desastres naturales. En este sentido, los métodos desarrollados fueron: a) Sistema Estadístico Evolutivo con Modelo de Islas (ESS-IM)¹, y b) Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (HESS-IM), los cuales han logrado mejoras considerables en relación a la calidad de las predicciones generadas y así también en términos de tiempo de respuesta, inclusive tras su aplicación con incendios forestales de gran escala.

8. Bibliografía

Alba, E., Luque, G., & Nesmachnow, S. (2013). Parallel metaheuristics: recent advances and new trends.

¹ Por motivos de espacio en el presente artículo este método no ha sido descrito detalladamente.

- International Transactions in Operational Research*, 20(1), 1–48. <https://doi.org/10.1111/j.1475-3995.2012.00862.x>
- Andrews, P. L., Bevins, C. D., Seli, R. C., Andrews, A., Bevins, P. L., & Seli, C. D. (2008). BehavePlus fire modeling system, version 4.0: User's Guide Revised.
- Bartezzaghi, E., Verganti, R., & Zotteri, G. (1999). A simulation framework for forecasting uncertain lumpy demand. *International Journal of Production Economics*, 59(1–3), 499–510. [https://doi.org/10.1016/S0925-5273\(98\)00012-7](https://doi.org/10.1016/S0925-5273(98)00012-7)
- Bianchini, G., Caymes-Scutari, P., & Méndez-Garabetti, M. (2015). Evolutionary-Statistical System: A parallel method for improving forest fire spread prediction. *Journal of Computational Science*, 6, 58–66. <https://doi.org/10.1016/J.JOCS.2014.12.001>
- Bianchini, Germán. (2006). *Wildland fire prediction based on statistical analysis of multiple solutions. TDX (Tesis Doctorals en Xarxa)*. Universitat Autònoma de Barcelona.
- Bianchini, Germán, Denham, M., Cortés, A., Margalef, T., & Luque, E. (2010). Wildland fire growth prediction method based on Multiple Overlapping Solution. *Journal of Computational Science*, 1(4), 229–237. <https://doi.org/10.1016/j.jocs.2010.07.005>
- Buyya, R. (1999). *High Performance Cluster Computing: Architectures and Systems*. Prentice Hall, PTR, NJ, USA.
- Collin D. Bevins. (1996). *fireLib User Manual and Technical Reference*.
- David B. Kirk, W. W. H. (2013). *Programming Massively Parallel Processors (2nd). Special Edition*. <https://doi.org/10.1016/B978-0-12-381472-2.00001-5>
- Doeswijk, T. G. (2007). *Reducing Prediction Uncertainty of Weather Controlled Systems*. Wageningen University, The Netherlands.
- DX, V., PR, R., & L, M. (1998). An empirical model for the spread of a fireline inclined in relation to the slope gradient or to wind direction. En *Proceedings of the 3rd international conference on forest fire research. 14th Conference on Fire and Forest Meteorology* (pp. 325–342). Luso, Portugal.
- Estadísticas de Incendios Forestales 2016*. (2016). Buenos Aires: Presidencia de la Nación.
- Farguell, A., Moré, J., Cortés, A., Miró, J. R., Margalef, T., & Altava, V. (2016). Reducing Data Uncertainty in Surface Meteorology Using Data Assimilation: A Comparison Study. *Procedia Computer Science*, 80, 1846–1855. <https://doi.org/10.1016/J.PROCS.2016.05.475>
- Finney, M. A. (1998). FARSITE : Fire Area Simulator — Model Development and Evaluation. *USDA Forest Service Research Paper*, (February), 47. <https://doi.org/U.S.ForestServiceResearchPaperRMRS-RP-4Revised>
- Graham, S. (1999, octubre 22). Global Fire Monitoring. Recuperado 1 de mayo de 2018, de https://earthobservatory.nasa.gov/Features/GlobalFire/fire_2.php
- Heinsch, F. A., & Andrews, P. L. (2010). *BehavePlus fire modeling system, version 5.0: Design and Features. USDA Forest Service - General Technical Report RMRS-GTR*.
- Hu, Z., Mahadevan, S., & Ao, D. (2018). Uncertainty aggregation and reduction in structure–material performance prediction. *Computational Mechanics*, 61(1–2), 237–257. <https://doi.org/10.1007/s00466-017-1448-6>
- Ishigami, T., & Homma, T. (1990). An importance quantification technique in uncertainty analysis for computer models. En *Proceedings of the ISUMA 90', First International Symposium on Uncertainty Modelling and Analysis*, (pp. 398–403). IEEE Comput. Soc. Press. <https://doi.org/10.1109/ISUMA.1990.151285>
- Kalyuzhnaya, A. V., & Boukhanovsky, A. V. (2015). Computational Uncertainty Management for Coastal Flood Prevention System. *Procedia Computer Science*, 51, 2317–2326. <https://doi.org/10.1016/J.PROCS.2015.05.397>
- Kindratenko, V. V., Enos, J. J., & Shi, G. (2009). GPU clusters for high- performance computing. *IEEE International Conference on Cluster Computing and Workshops*, 1–8. <https://doi.org/10.1109/CLUSTER.2009.5289128>
- Li, M., Hamel, J., & Azarm, S. (2010). Optimal uncertainty reduction for multi-disciplinary multi-output systems using sensitivity analysis. *Structural and Multidisciplinary Optimization*, 40(1–6), 77–96. <https://doi.org/10.1007/s00158-009-0372-6>
- Lopes, A. M. G., Cruz, M. G., & Viegas, D. X. (2002). FireStation — an integrated software system for the numerical simulation of fire spread on complex topography. *Environmental Modelling & Software*, 17(3), 269–285. [https://doi.org/10.1016/S1364-8152\(01\)00072-X](https://doi.org/10.1016/S1364-8152(01)00072-X)
- Méndez-Garabetti, M., Bianchini, G., Caymes-Scutari, P., & Tardivo, M. L. (2016). Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the

- internal metaheuristic. *Fire Safety Journal*, 82, 49–62. <https://doi.org/10.1016/j.firesaf.2016.03.002>
- Méndez-Garabetti, Miguel, Bianchini, G., Tardivo, M. L., Caymes Scutari, P. G., & Gil Costa, G. V. (2017). Hybrid-Parallel Uncertainty Reduction Method Applied to Forest Fire Spread Prediction. *Journal of Computer Science and Technology, ISSN-e 1666-6038, Vol. 17, N°. 1, 2017, págs. 12-19, 17(1)*, 12–19.
- Morgan, P., Hardy, C. C., Swetnam, T. W., Rollins, M. G., Long, D. G., Morgan, P., & The, M. (2001). Mapping fire regimes across time and space: Understanding coarse and fine-scale fire patterns. *International Journal of Wildland Fire*, 10, 329–342. <https://doi.org/10.1071/WF01032>
- Nowak, R. D. (2002). Nonlinear system identification. *Circuits, Systems, and Signal Processing*, 21(1), 109–122. <https://doi.org/10.1007/BF01211655>
- Purica, I., & Bae, D. (2016). Climate Change events induced risk assessment and mapping and a potential insurance policy. *Journal of Earth Science & Climatic Change*, 07(09). <https://doi.org/10.4172/2157-7617.C1.027>
- Qian, X., Li, W., & Yang, M. (2013). Two-stage nested optimization-based uncertainty propagation method for uncertainty reduction. En *Communications in Computer and Information Science* (Vol. 402, pp. 243–252). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-45037-2_23
- Real, R., & Vargas, J. M. (1996). The Probabilistic Basis of Jaccard's Index of Similarity. *Systematic Biology*, 45(3), 380–385. <https://doi.org/10.1093/sysbio/45.3.380>
- Rothermel, R. C. (1972). *A mathematical model for predicting fire spread in wildland fuels* (Vol. II). Res. Pap. INT-115, US Dept. of Agric., Forest Service, Intermountain Forest and Range Experiment Station. (Ogden, UT.).
- Sano, T., Takeda, T., & Yamasaki, M. (2009). A new uncertainty reduction method for fuel fabrication process with erbia-bearing fuel. *Journal of Nuclear Science and Technology*, 46(3), 226–231. <https://doi.org/10.1080/18811248.2007.9711525>
- Sousa, F. A., dos Reis, R. J. N., & Pereira, J. C. F. (2012). Simulation of surface fire fronts using fireLib and GPUs. *Environmental Modelling & Software*, 38, 167–177. <https://doi.org/10.1016/J.ENVSOFT.2012.06.006>
- Stanford University, & Center for the Study of Language and Information (U.S.). (1997). *Stanford encyclopedia of philosophy*. Stanford University.
- Taghribi, A., & Sharifian, S. (2017). A Metaheuristically Tuned Interval Type 2 Fuzzy System to Reduce Segmentation Uncertainty in Brain MRI Images. *Journal of Medical Systems*, 41(11), 174. <https://doi.org/10.1007/s10916-017-0821-5>
- Tardivo, M. L., Caymes-Scutari, P., Méndez-Garabetti, M., & Bianchini, G. (2018). *Optimization for an uncertainty reduction method applied to forest fires spread prediction*. *Communications in Computer and Information Science* (Vol. 790). https://doi.org/10.1007/978-3-319-75214-3_2
- Thompson, M. P., & Calkin, D. E. (2011). Uncertainty and risk in wildland fire management: A review. *Journal of Environmental Management*, 92(8), 1895–1909. <https://doi.org/10.1016/J.JENVMAN.2011.03.015>
- Tsutsui, S., & Collet, P. (2013). *Massively Parallel Evolutionary Computation on GPGPUs*. Springer Berlin Heidelberg.
- Wallace, G. (1993). A Numerical Fire Simulation-Model. *International Journal of Wildland Fire*, 3(2), 111. <https://doi.org/10.1071/WF9930111>
- Wilkinson, B., & Allen, C. M. (1999). *Parallel Programming: Techniques and Applications Using Networked Workstations and Parallel Computers*. Pearson/Prentice Hall.

Obtención de reglas de clasificación difusas utilizando técnicas de optimización – Caso de estudio Riesgo Crediticio

Tesis doctoral realizada por

Patricia Rosalía Jimbo Santana

Universidad Central del Ecuador

prjimbo@uce.edu.ec

dirigida en cotutela por

Dra. Laura C. Lanzarini

**Instituto de Investigación en
Informática LIDI, Facultad de
Informática,**

**Universidad Nacional de la Plata,
La Plata, Buenos Aires, Argentina**

laural@lidi.info.unlp.edu.ar

Dr. Aurelio Fernández Bariviera

**Departament of Business, Universitat
Rovira i Virgili,
Avenida de la Universitat,1 Reus, Spain**

aurelio.fernandez@urv.cat

Universidades donde se desarrolló la tesis doctoral:

- **Universidad Nacional de la Plata, La Plata, Buenos Aires, Argentina**
- **Universidad Rovira i Virgili, Reus, España**

Fecha de exposición

14 de Julio 2020

RESUMEN

Esta tesis tiene por objetivo principal contribuir al área conformada por la Minería de datos y el Riesgo Financiero especialmente en el área de crédito, disponiendo de estrategias capaces de generar automáticamente reglas de clasificación difusas, resultando de suma utilidad en cualquier proceso de toma de decisiones. El método desarrollado denominado FRvarPSO, es capaz de operar sobre atributos nominales y numéricos, para obtener reglas de clasificación difusas que combina una red neuronal competitiva con una técnica de optimización basada en un cúmulo de partículas de población variable. La función de aptitud que controla el movimiento de las partículas utiliza un criterio de votación que pondera, de manera difusa, la participación de los atributos numéricos. La eficiencia y eficacia de este método se encuentran fuertemente condicionadas por la manera en que se determinen las funciones de pertenencia de cada uno de los conjuntos difusos, se consideran varias alternativas entre las que mencionamos particionando el rango de cada atributo numérico en intervalos de igual longitud, centrando, en cada uno de ellos, una función triangular con un solapamiento adecuado, utilizando el criterio del experto, y utilizando Fuzzy C-Means para la obtención de los conjuntos difusos. Uno de los principales aportes es en el área del riesgo financiero, ya que asocia el peso de la regla difusa al riesgo que puede asumir el usuario en la interpretación de la regla. Los resultados obtenidos fueron comparados con la versión anterior sobre 11 bases de datos del repositorio UCI y 3 bases de datos reales del sistema financiero ecuatoriano, una de las cuales es una cooperativa de ahorro y crédito y las dos restantes pertenecen a bancos encargados de dar crédito de consumo productivo, no productivo y microcrédito. Los resultados obtenidos fueron satisfactorios.

Palabras Claves: Riesgo, Reglas de Clasificación Difusas (Fuzzy Classification Rules), Optimización mediante Cúmulo de partículas tamaño variable (Variable Particle Swarm Optimization), Minería de datos.

1. MOTIVACION

En los últimos años gracias al avance de la tecnología, las organizaciones han almacenado gran cantidad de información. Esto las ha llevado a la necesidad de incorporar técnicas que permitan procesar y obtener información útil de los datos. El proceso de Extracción del Conocimiento, conocido como proceso de KDD (Knowledge Discovery in Databases), está formado por varias fases que van desde la recolección y transformación de los datos hasta la identificación e interpretación de patrones o relaciones subyacentes sumamente útiles en la toma de decisiones.

Dentro del proceso de KDD, la Minería de Datos (Data Mining) es considerada la fase más importante, ya que agrupa a las técnicas capaces de modelizar la información disponible. A partir del uso o comprensión del modelo generado es posible extraer conocimiento. Este conocimiento que se genera resulta de gran interés para las organizaciones, debido a que constituye una herramienta sumamente importante para la toma de decisiones tácticas y estratégicas, lo cual se convierte en una ventaja competitiva.

Según el tipo de tarea que se quiera resolver existe en la literatura una gran cantidad de técnicas disponibles entre las que podemos mencionar: las algebraicas y estadísticas, estocásticas y difusas, técnicas bayesianas, árboles de decisión, redes neuronales artificiales, máquinas de vectores de soporte y técnicas de optimización entre otras. Cada una de ellas tiene características diferentes.

Para poder seleccionar el tipo de técnica es necesario conocer el problema que se va a resolver teniendo en cuenta que no todas ellas pueden operar con datos faltantes, outliers o datos en cualquier formato. Por otra parte, los problemas que se abordan pueden ser de diferente tipología como, por ejemplo: clasificación, categorización, regresión o agrupamiento.

Una característica deseable de los modelos construidos con las técnicas de la Minería de Datos es que el conocimiento que se extrae se exprese en términos comprensibles. En este sentido, las reglas de asociación son consideradas, por quienes deben tomar decisiones, como una de las formas más comprensibles que puede ser utilizada para representar el conocimiento, ya que tienen la capacidad de explicarse por sí mismas. Una regla de asociación es una expresión de la forma "SI condic1 ENTONCES condic2" donde ambas condiciones son conjunciones de proposiciones de la forma (atributo=valor) cuyos atributos/variables pueden ser cuantitativas o cualitativas. Por su estructura condicional, puede afirmarse que el objetivo principal de una regla es identificar relaciones entre los atributos. Cuando el conjunto de reglas de asociación presenta en el consecuente el mismo atributo, se dice que se trata de un conjunto de "reglas de clasificación". Si a esto se le suma que las reglas de clasificación empleen lógica difusa a través de conjuntos difusos para describir los valores de sus atributos, se obtienen "Reglas de Clasificación Difusas", facilitando aún más su comprensión, permitiendo además el manejo de la incertidumbre, aproximándonos cada vez más al razonamiento humano. Tal como se menciona en (Terano, Asai y Sugeno, 1992), *"cuanto más humano debe ser un sistema, más lógica difusa contendrá"*.

La utilización de la lógica difusa para la generación de las reglas ha permitido que los investigadores obtengan formas de mejorar el desempeño de las metaheurísticas, básicamente para acelerar la convergencia y obtener una mejor calidad en las soluciones planteadas.

Las reglas de clasificación difusas son reglas que tienen sentencias lingüísticas que describen la manera en que los sistemas de inferencia difusa toman las decisiones sobre un conjunto de variables de entrada. Estas reglas de clasificación poseen un antecedente formado por variables nominales y/o condiciones difusas, y un consecuente conocido con anticipación denominado clase.

El aporte central de esta tesis es la definición de un nuevo método capaz de generar un conjunto de reglas de clasificación difusas de fácil interpretación, baja cardinalidad y una buena precisión. Estas características ayudan a identificar y comprender las relaciones presentes en los datos facilitando de esta forma la toma de decisiones.

2. FRvarPSO

El conocimiento que se genera utilizando técnicas de minería de datos resulta de gran interés para las organizaciones, ya que les va a permitir la toma de decisiones tácticas y estratégicas, lo que se convierte en una ventaja competitiva. Existe un conjunto de técnicas presentes en la minería de datos, cada una de ellas tiene características diferentes. Para poder seleccionar el tipo de técnica es fundamental tener conocimiento del problema que se va a resolver. En el estado del arte se pueden encontrar referencias que indican que se obtienen mejores resultados cuando se utilizan modelos híbridos, aquí se enmarca el modelo desarrollado.

El nuevo método propuesto en esta investigación se denomina **FRvarPSO (Fuzzy Rules variable Particle Swarm Optimization)**, combina una red neuronal competitiva con una técnica de optimización basada en cúmulo de partículas de población variable para la obtención de reglas de clasificación difusas, capaces de operar sobre atributos nominales y numéricos. Los antecedentes de las reglas están formados por atributos nominales y/o condiciones difusas. La conformación de estas últimas requiere conocer el grado de pertenencia a los conjuntos difusos que definen a cada variable lingüística. Esta tesis propone tres alternativas distintas para resolver este punto.

Con respecto a la forma de obtención de las reglas, el método propuesto utiliza un proceso iterativo por medio del cual se van cubriendo los ejemplos de una clase a la vez hasta lograr la cobertura deseada. Por lo tanto, el consecuente de la regla queda determinado por la clase seleccionada y es el antecedente el que se extrae a través de la técnica de optimización.

Cada partícula de la población utiliza una representación de longitud fija, donde sólo se codifica el antecedente de la regla y dado el enfoque adoptado, se efectuará un proceso iterativo asociando todos los individuos de la población con una clase predeterminada. El hecho de que todos los individuos pertenezcan a la misma clase hace innecesaria la codificación del consecuente dentro de la partícula.

El proceso de extracción de cada regla inicia con la generación del cúmulo. Las posiciones iniciales de las partículas tienen una gran incidencia en la capacidad exploratoria de cada una de ellas. Si se encuentran demasiado próximas, su conocimiento social puede provocar la convergencia prematura evitando que se alcancen buenas soluciones. Por otro lado, si inician muy lejos del óptimo se requerirán un mayor número de iteraciones o pasos de búsqueda para alcanzarlo. Por todo lo anterior, se decidió utilizar una red neuronal competitiva para identificar las zonas más prometedoras. Esta información fue utilizada para inicializar el cúmulo.

En lo que se refiere al tamaño de la población, se optó por un cúmulo de tamaño variable con capacidad para incorporar individuos en las zonas menos pobladas. Como se dijo anteriormente, en PSO las partículas tienden a reunirse alrededor de las buenas posiciones encontradas por sus vecinos esto hace que el cúmulo pierda diversidad y varias zonas del espacio de búsqueda queden despobladas. La técnica para modificar el tamaño de la población trabaja sobre dos conceptos: por un lado, detecta las zonas densas y elimina individuos muy similares y por otro identifica las zonas muy despobladas y agrega partículas adecuadamente. Esto permite iniciar con pocos individuos e ir modificando el tamaño de la población durante el proceso iterativo de búsqueda.

Volviendo a la representación de la partícula, a la información referida al movimiento del individuo se le agregan dos cuestiones referidas a la construcción del antecedente de la regla. En primer lugar, debe indicarse cuáles serán los atributos que formarán parte del antecedente y esto debe hacerse utilizando información binaria. Por otro lado, es preciso representar el grado de pertenencia de los atributos originalmente numéricos en cada conjunto difuso para conocer cómo operar con las correspondientes variables lingüísticas, a través de una representación difusa.

Uno de los aportes de esta tesis radica en la definición de la función de aptitud o fitness de cada partícula basada en un "Criterio de Votación" que pondera de manera difusa la participación de las condiciones difusas en la conformación del antecedente. Su valor se obtiene a partir de los grados de pertenencia de los ejemplos que cumplen con la regla y se utiliza para reforzar el movimiento de la partícula en la

dirección donde se encuentra el valor más alto. Con la utilización de PSO las partículas compiten entre ellas para encontrar a la mejor regla de la clase seleccionada.

La eficiencia y eficacia de FRvarPSO se encuentran fuertemente condicionadas por la manera en que se determinen las funciones de pertenencia de los conjuntos difusos. En el marco de las investigaciones de esta tesis se han utilizado diferentes opciones. Uno de estas opciones fue particionar el rango de cada atributo numérico en intervalos de igual longitud, y centrando en cada uno de ellos una función triangular con un solapamiento adecuado. Otra de las formas para obtener los conjuntos difusos ha sido utilizando el método Fuzzy C-Means. Adicionalmente, se utilizó también como técnica el conocimiento de un experto para la definición de los conjuntos difusos, y su correspondiente valor de pertenencia.

El desempeño del método propuesto FRvarPSO fue comparado con versiones previas del mismo como son SOM + PSO, SOM + varPSO (PSO con población variable), LVQ + PSO, así como otros métodos de extracción de reglas de clasificación como PART y C4.5. Adicionalmente se comparó FRvarPSO con las versiones SOM + PSO Difuso, SOM + varPSO Difuso, LVQ + PSO Difuso, que son las versiones anteriores a las que se les aplicó lógica difusa al momento de construir el antecedente junto con el criterio de votación utilizado en la función fitness.

La medición se realizó sobre doce bases de datos del repositorio UCI (Machine Learning Repository).

3. RIESGO

Además del aporte en el área informática, se realizó una contribución significativa en el área de riesgo financiero ya que se probó el método propuesto FRvarPSO en tres casos reales en el área de crédito del Sistema Financiero del Ecuador asociadas al riesgo crediticio considerando un conjunto de variables micro y macroeconómicas. Dentro de las bases de datos del Sistema Financiero del Ecuador se contó con la información de una cooperativa de ahorro y crédito, y las otras dos bases de datos correspondieron a bancos que otorgan crédito de consumo productivo, no productivo y microcrédito, que se encuentran en el mercado por más de 25 años.

FRvarPSO se empleó en bases de datos reales en el proceso de análisis para la concesión de crédito. Es importante remarcar la ardua tarea realizada en las tareas de limpieza y transformación de la información, antes de la aplicación de los métodos. Se analizó la función de distribución de cada atributo a fin de detectar valores anómalos, faltantes y/o fuera de rango. Se transformaron variables con la finalidad de tener indicadores aplicados al cliente como, por ejemplo: la capacidad de pago, la capacidad de endeudamiento y la calificación crediticia. Esta última tuvo un análisis particular, debido a que se disponía de información de las características generales del cliente, así como del comportamiento de los créditos otorgados en un horizonte de tiempo, con lo cual se obtuvo el nivel de morosidad.

Otro de los aportes de esta tesis fue haber realizado una consideración especial en la morosidad del cliente teniendo en cuenta los días de vencimiento de la cartera otorgada; esto fue posible debido a que se tenía información del cliente en un horizonte de tiempo, una vez que el crédito se había concedido. De esta forma, se consideraron tres casos: la cartera vencida, la que no devenga intereses y la cartera castigada.

Se verificó que con este análisis las reglas difusas obtenidas a través de FRvarPSO permiten que el oficial de crédito de respuesta al cliente en menor tiempo, y principalmente disminuya el riesgo que representa el otorgamiento de crédito para las instituciones financieras. Lo anterior fue posible, debido a que al aplicar una regla difusa se toma el menor grado de pertenencia promedio de las condiciones difusas que forman el antecedente de la regla, con lo que se tiene una métrica proporcional al riesgo de su aplicación. Con esta observación el oficial de crédito puede tomar la decisión de conceder el crédito incrementando la tasa de interés, las garantías y/o colaterales (activo), con la finalidad de disminuir el riesgo asociado.

Los resultados obtenidos fueron comparados mediante tests de diferencia de medias, verificándose que los modelos difusos presentan en la mayor parte de los casos una precisión superior a la del método PART,

pero ligeramente mayor a la alcanzada por el método C4.5. Sin embargo, si se observa la cantidad de reglas generadas para alcanzar dicha precisión, los métodos difusos utilizan una cantidad promedio de reglas mucho menor, que las reglas de C4.5 y PART. Esto último ratifica el énfasis puesto en la sencillez del modelo y en su facilidad de interpretación, gracias a la utilización de la lógica difusa.

4. RESULTADOS

En la figura 1 se ilustra un análisis comparativo de los resultados obtenidos al aplicar los métodos a las 12 bases de datos del repositorio UCI así como a las 3 bases de datos reales del Sistema Financiero Ecuatoriano. Aquí se puede observar la relación que existe entre la precisión de cada uno de los métodos con la mejor solución encontrada. Dicho valor corresponde al cociente entre la precisión promedio obtenida por el método y la mejor precisión obtenida, es decir, la mayor precisión de la base de datos analizada. Aquí se puede observar como el método propuesto tiene una mejor precisión, que los otros métodos analizados.

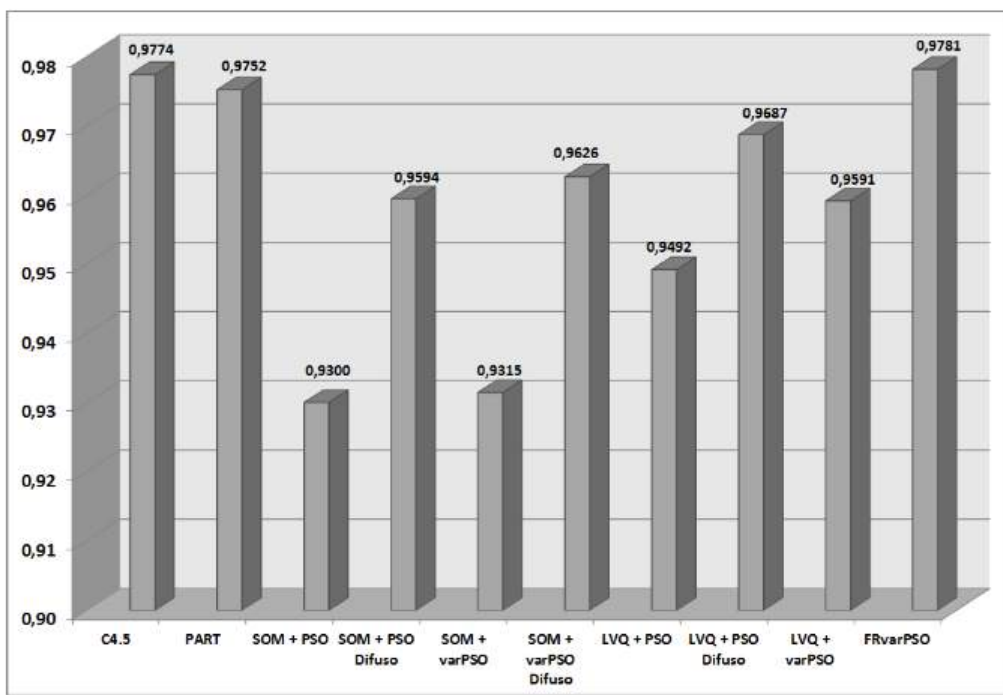


FIGURA 1: Comparación de la Precisión de los métodos

En la figura 2 se puede observar la relación que existe entre la cardinalidad (número de reglas) de cada uno de los métodos con la cardinalidad de la mejor solución encontrada. Dicha relación corresponde al cociente entre el promedio del valor del número de reglas difusas obtenidas por el método y la solución que tiene el menor número de reglas difusas de la base de datos analizada. Los resultados obtenidos permiten afirmar que el método propuesto es el que obtiene el menor número de reglas.

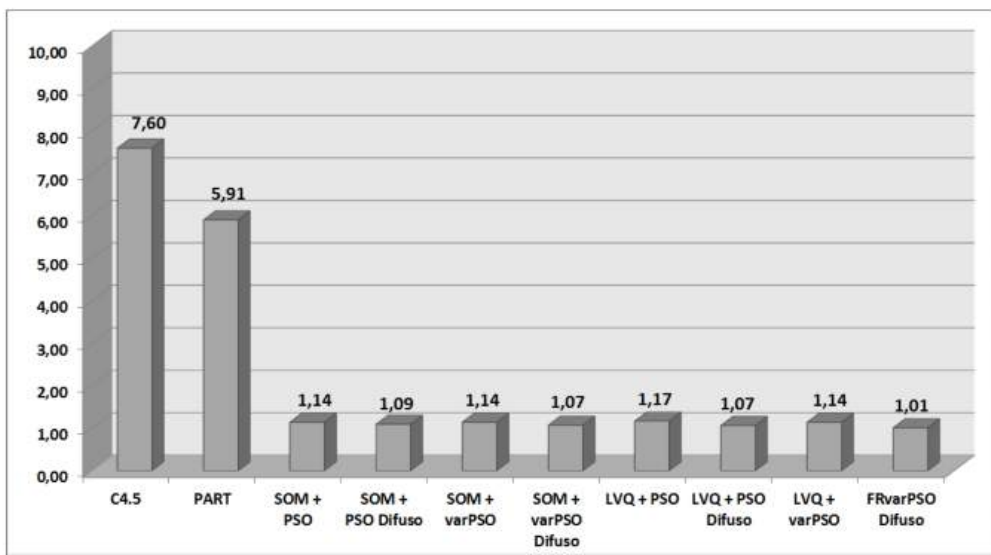


FIGURA 2: Comparación de la Cardinalidad de los métodos

En la figura 3 se puede observar la relación entre la precisión y la cardinalidad de los métodos analizados.

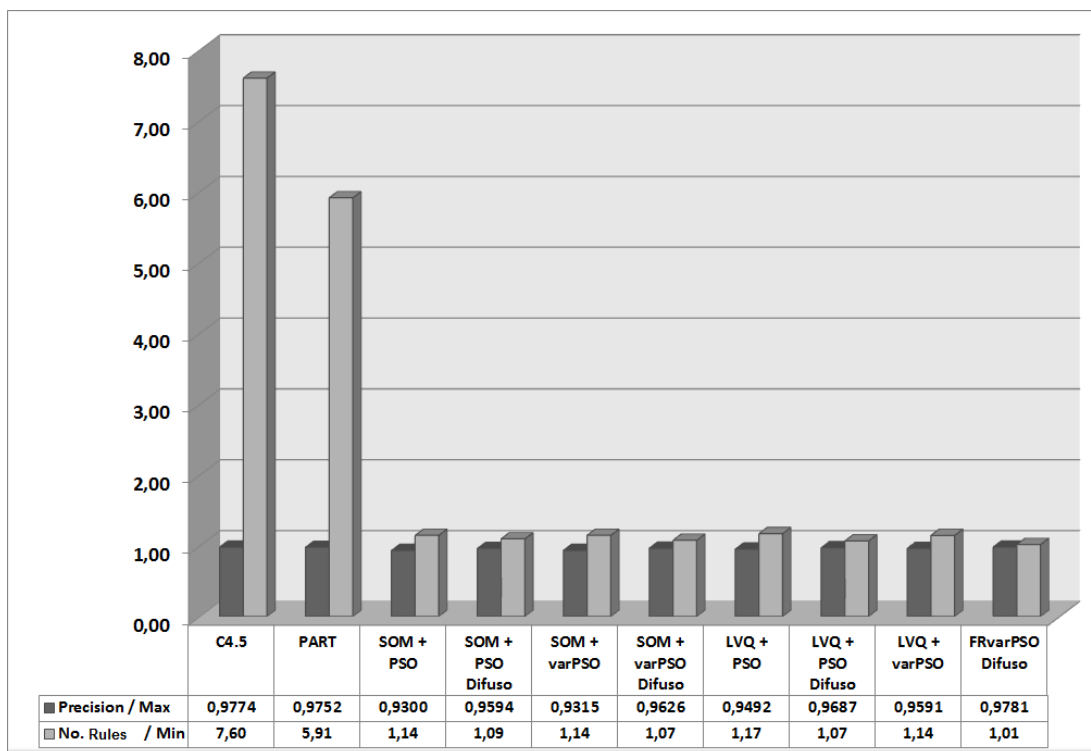


Figura 3: Comparación de la Precisión y Cardinalidad de los métodos

En las figuras 4, 5, 6 se puede observar claramente la relación que existe entre la precisión y el número de reglas de las tres bases de datos reales del Sistema Financiero Ecuatoriano utilizando FRvarPSO y generando los conjuntos difusos por medio de particiones equitativas, Fuzzy C-Means y con conocimiento del experto.

Finalmente, como resultado, se puede indicar que al utilizar FRvarPSO y no contar con el criterio del experto para la determinación de las funciones de pertenencia y el conjunto difuso definido por ella, el utilizar FCM, da como resultado un conjunto de reglas difusas con mayor precisión, menor cardinalidad y fácil interpretación, que cuando se utilizan conjuntos equitativos. A base de los resultados obtenidos se puede decir que el objetivo de la tesis se ha cumplido.

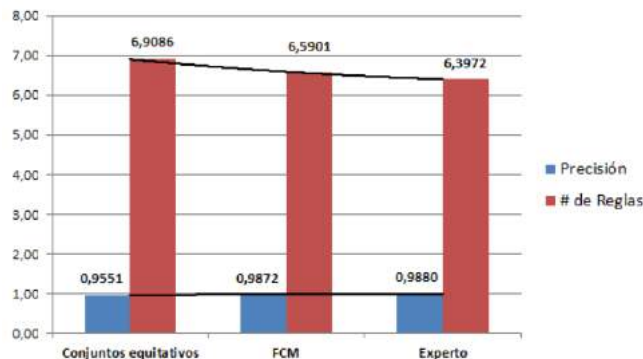


FIGURA 4: Banco encargado de microcrédito

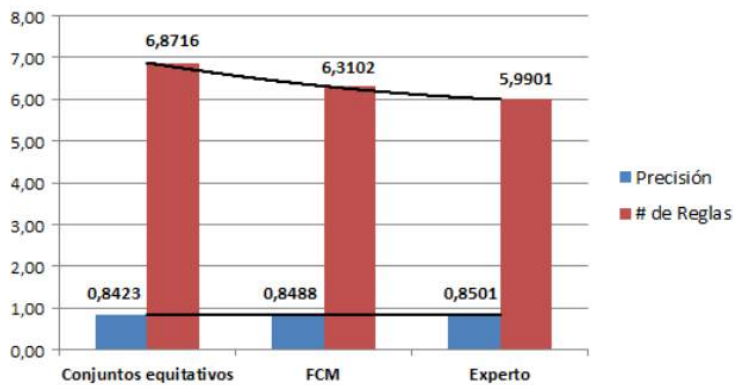


FIGURA 5: Banco encargado de crédito de consumo y crédito productivo empresarial

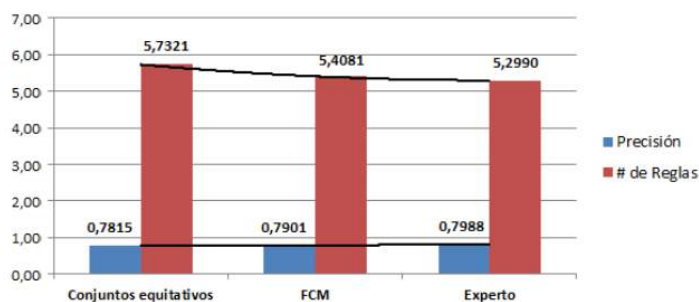


FIGURA 6: Cooperativa de Ahorro y Crédito del Ecuador que encuentra en el segmento 2 dentro de la Superintendencia de Economía Popular y Solidaria, activos mayor a 20'000.000,00 hasta 80'000.000,00

5. CONCLUSIONES Y TRABAJOS FUTUROS

Uno de los objetivos de esta tesis es la revisión de los diferentes métodos que permiten la obtención de reglas de clasificación, presentando un nuevo método que encuentra una mejor solución comparado con otras técnicas. El método propuesto denominado FRvarPSO (Fuzzy Rules Variable Particle Swarm Optimization), permite la representación del conocimiento en términos difusos a través de reglas de clasificación difusas extraídas de un conjunto de datos de entrada.

Las reglas de clasificación difusas obtenidas a través de FRvarPSO se caracterizan por tener cardinalidad baja y precisión adecuada, y especialmente por ser fáciles de interpretar gracias a la incorporación de la lógica difusa. Para conseguir esto se consideraron dos aspectos importantes; el primero tiene que ver con la capacidad del método de operar con variables lingüísticas y el segundo se relaciona con la inserción de información basada en grados de pertenencia tanto en la evaluación de la función fitness, así como en la manera en que se realiza la búsqueda por medio de la técnica de optimización.

El método inicia con la determinación de los conjuntos difusos, y las funciones de pertenencia de cada variable difusa asociada a cada uno de ellos, es aquí donde radica la eficiencia y eficacia de este método. Para esto se utilizaron dos variantes. La primera variante se encarga de particionar el rango de cada atributo numérico en intervalos de igual longitud centrando, en cada uno de ellos, una función triangular con un solapamiento adecuado. En este caso se consideraron particiones difusas uniformes con el mismo número de etiquetas para cada una de las variables y se definieron los conjuntos difusos de tal manera que el valor de una variable difusa sólo puede pertenecer a dos conjuntos de forma simultánea. En la segunda variante se mejora el proceso de obtención de los conjuntos difusos utilizando el algoritmo FCM con tres clusters, donde cada uno de ellos representaba un conjunto difuso.

Posteriormente se utiliza una técnica de optimización de población variable varPSO la misma que emplea una red neuronal competitiva LVQ, con la finalidad de obtener las zonas más prometedoras para realizar de búsqueda. Esta información suministrada por dicha red es utilizada para inicializar el cúmulo. El método utiliza el grado de pertenencia de las variables difusas, incidiendo directamente a través de la técnica de optimización varPSO, en el vector velocidad que controla el movimiento de las partículas. Ninguna de las técnicas por separado logra obtener los resultados que se indican en esta investigación. Los resultados obtenidos permiten afirmar que el modelo híbrido propuesto FRvarPSO tiene un mejor desempeño que sus versiones anteriores a raíz de la incorporación de la lógica difusa.

Una de las ventajas del método propuesto es que además de probarlo con bases de datos del repositorio UCI, se realizaron pruebas con base de datos reales pertenecientes a tres instituciones financieras. Aquí, para la determinación de los conjuntos difusos se utilizó el criterio de un experto en el área de crédito. Adicionalmente, se incorporó en el modelo no solo las variables del sujeto de crédito, si no también aquellas variables que hacen referencia a la economía del país en su conjunto, que son las variables macroeconómicas. Esto da una situación más real para la toma de decisiones, por parte del oficial de crédito en las instituciones financieras.

En este caso FRvarPSO no solo va a influir en la determinación del scoring de crédito, permitiendo analizar el riesgo que representa para la institución financiera la concesión del crédito, si no que en el caso de que el análisis indique que el crédito debe ser negado, este dependiendo del peso que tiene la regla de clasificación difusa puede ser concedido incrementado garantías o aumentando la tasa de interés. Todas estas características hacen que FRvarPSO sea una buena opción para análisis de riesgo crediticio. Es importante indicar que actualmente FRvarPSO está siendo utilizado en instituciones financieras del Ecuador.

FRvarPSO no solo obtiene un modelo más simple ya que utiliza menor cantidad de reglas que otros métodos, si no que presenta una buena precisión, y especialmente gracias a la incorporación de la lógica difusa mejora la interpretabilidad de la regla. Luego, este método aporta al área informática y se ha demostrado que también realiza aportes en el área de la economía, a través del análisis del riesgo crediticio, incorporando para ello variables macro económicas.

Es importante que en investigaciones futuras se consideren:

Una optimización de la función de pertenencia. Dentro de esta una alternativa puede ser la utilización de algoritmos genéticos, o de la misma optimización por cúmulo de partículas, con la finalidad de identificar automáticamente los parámetros de dicha función, sin perder de lado el objetivo que es tener un conjunto de reglas reducidas, pero conservando la interpretabilidad con las variables lingüísticas.

Un segundo aspecto que se debe considerar es la obtención de una nueva forma de representación de los atributos nominales, con el objetivo de reducir el tiempo computacional, y disminuir la longitud de la representación del antecedente dentro de cada partícula.

Otra variante que se debe tener en cuenta es incorporar técnicas para el manejo del desbalance de las clases, ya que en el área de riesgo crediticio la clase que corresponde a los créditos otorgados es la que tiene la mayor cantidad de ejemplos.

Es necesario incorporar al modelo la defuzificación de la variable de salida, siendo esta otra forma de interpretar el riesgo. Este valor crisp que se obtiene de la aplicación de las reglas de clasificación, va a indicar el porcentaje de riesgo que existe. En este caso para las reglas de clasificación aplicables a riesgo crediticio, se va a tener no solamente como salida la concesión o no del crédito, sino también el riesgo que implica el cliente, logrando disminuir considerablemente las consecuencias negativas para la institución financiera. Este mismo criterio puede ser utilizado en otra área de riesgo como el operacional, liquidez entre otros en el área financiera. Incluso el método propuesto puede utilizarse en otras áreas como la medicina, para saber el riesgo que un paciente tiene en contraer una determinada enfermedad.

6. PUBLICACIONES

Publicaciones en Revistas

- [1] **Simpliflyng Credit Scoring Rules using LVQ+PSO.** Lanzarini, Villa Monte, Fernández Bariviera, Jimbo Santana Patricia. *Kybernetes*. Vol.46 No 1. pp.8-16
 - Impact Factor: 0.98. Cuartil 3 Computer Science, Cybernetics
 - SJR 0.29. Cuartil 2 Computer Science (miscellaneous).
 - Número de citas: 11 (Scopus). 7 (Web of Science)
- [2] **Analysis of Methods for Generating Classification Rules Applicable to Credit Risk** Jimbo Santana, Villa Monte, Rucci, Lanzarini, Fernández Bariviera. *Journal of Computer Science & Technology*; vol. 17, No 1. pp.20-28. 2017
 - Indexado en Emerging Sources Citation Index
 - Número de citas: 1 (Web of Science)
- [3] **Data Mining Methods linked to Artificial Intelligence to Credit Risk.** Jimbo Santana, Lanzarini, Villa Monte, Fernández Bariviera. FIGEMPA. Universidad Central del Ecuador: Año V, volumen 2, No 8. 2017
 - Indexada en Latindex
- [4] **Fuzzy Credit Risk Scoring Rules using FRvarPSO,** Jimbo Patricia, Lanzarini Laura, Bariviera Aurelio, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol 26, No 1. pp.39-57. 2018
 - Impact Factor: 1.286. Cuartil 3 Computer Science, Artificial Intelligence
 - SJR 0.39. Cuartil 2 Artificial Intelligence.
 - Número de citas: 2 (Scopus). 1 (Web of Science)
- [5] **Variations of Particle Swarm Optimization for Obtaining Classification Rules Applied to Credit Risk in Financial Institutions of Ecuador.** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez Bariviera. *Risks* Vol 8, No.2. 2020
 - Indexado en Emerging Sources Citation Index
 - Indexado en Scopus. CiteScoreTracker 2019:0.89.
- [6] **Fuzzy Classification Rules with FRvarPSO Using Various Methods for Obtaining Fuzzy Sets.** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez Bariviera. *Journal of Advances in Information Technology* Vol. 11, No. 4, November 2020
 - Indexado en Scopus.

Capítulos de Libro

- [7] **Obtaining Classification Rules Using LVQ+PSO: an application to Credit Risk.** Lanzarini, Villa Monte, Fernández Bariviera, Jimbo Santana. *Advances in Intelligent Systems and Computing*. Springer-Verlag Berlin Heidelberg. Pp 383-391. ISBN 978-3-319-19704-3. DOI 10.1007/978-3-319-19704-3_31
 - Indexado en Scopus
 - Número de citas 7 (Scopus).
- [8] **Extraction of Knowledge with Population-Based Metaheuristics Fuzzy Rules Applied to Credit Risk.** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez-Bariviera. *Advances in Swarm Intelligence*. Springer. pp 153-163. 2018. ISBN 978-3-319-93817-2. DOI 10.1007/978-3-319-93818-9_15
 - Indexado en Scopus
 - Número de citas 4 (Scopus).
- [9] **FRvarPSO: A Method for Obtaining Fuzzy Classification Rules Using Optimization Techniques,** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernández- Bariviera, *Modelling and Simulation in*

Management Sciences, Springer International Publishing, Cham, pp 112–126, 2020 DOI: 10.1007/978-3-030-15413-4_9

- Indexado en Scopus

Congresos con referato internacional

- [10] **An exploratory analysis of methods for extracting credit risk rules.** Jimbo Santana, Villa Monte, Rucci, Lanzarini, Bariviera. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016). pp. 834-841.
- [11] **Fuzzy Classification Rules with FRvarPSO Using Various Methods for Obtaining Fuzzy Sets .** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez Bariviera. The 12th International Conference on Computer Science and Information Technology (ICCSIT December 2019)

Proceso marco orientado a aspectos en las etapas tempranas del ciclo de vida del desarrollo de software para una transición en la industria

Fernando Pincirolí

Instituto de Investigaciones de la Facultad de Informática y Diseño
Universidad Champagnat, Mendoza, Argentina
pincirolifernando@uch.edu.ar

Director: Dr. Raymundo Forradellas, U.N. de San Juan, U.N. de Cuyo, Argentina
Codirector: Dr. José Luis Barros Justo, Universidad de Vigo, España

Doctorado en Ciencias de la Informática
Universidad Nacional de San Juan, Argentina
Fecha de defensa: 19 de noviembre de 2020

Tribunal: Dr. Luis Olsina, Universidad Nacional de La Pampa, Argentina
Dr. Santiago Matalonga, University of the West of Scotland, Escocia
Dr. Leandro Antonelli, Universidad Nacional de La Plata, Argentina

1. Introducción

Por una parte, se halló evidencia de la escasa presencia del paradigma de aspectos en la industria, pero, por otro lado, también se observó que sus beneficios, largamente mencionados en la literatura, sí se pudieron alcanzar en aquellos casos en los que se utilizó la orientación a aspectos en el mundo real, más allá de los ámbitos académicos [1]. Al mismo tiempo, esa evidencia también mostró que las propuestas existentes son incompletas y muy pocas llegan a cubrir tan solo dos fases del ciclo de vida del desarrollo de software (en adelante, SDLC) [1]. Por esto es que surgió la motivación de elaborar una alternativa metodológica que permitiera su aplicación de inmediato en proyectos e iniciativas en el mundo real.

Así, el objetivo de nuestro trabajo consistió en definir un proceso marco para las etapas tempranas del ciclo de vida del desarrollo de software, desde el modelo de negocios hasta la especificación completa de requisitos de software y empleando el paradigma de la orientación a aspectos. A la vez, se buscó propiciar el empleo en la industria de este paradigma para obtener sus beneficios, al aprovechar las herramientas y técnicas estándares disponibles actualmente en el mercado, mientras se siguen desarrollando otras específicas y alcanzan la madurez suficiente. Por tal razón, se decidió llamar a esta propuesta AOP4ST, sigla derivada de Aspect-Oriented Process for a Smooth Transition [2].

Se trata de un proceso marco, no específico, de modo de permitir su empleo con diferentes modelos del ciclo de vida del desarrollo de software a lo largo de sus etapas tempranas y hasta obtener una especificación de requisitos completa y coherente, incluyendo tres vistas: funcional, estática y de estados. Este proceso emplea herramientas y técnicas estándares, de amplia difusión en la industria, para facilitar su adopción inmediata y, también, utiliza notaciones estándares, para permitir elaborar modelos y especificaciones comprensibles y no ambiguas, que puedan contar con el soporte de las herramientas de software actualmente disponibles en el mercado. Se procura que esta alternativa sea completamente orientada a aspectos, que facilite la obtención de las incumbencias en forma progresiva a lo largo de todos los modelos y, al mismo tiempo, las mantenga siempre separadas y asegurando la trazabilidad bidireccional entre ellas. Estas incumbencias deben obtenerse en forma natural a lo largo de todos los

modelos, de manera de no afectar los objetivos propios de cada uno de ellos y, de este modo, potenciar los beneficios que se esperan en cada modelo mediante el empleo del paradigma de aspectos.

En la sección 2 se presenta la motivación de este trabajo, incluyendo el estado de la cuestión y la problemática que se pretende resolver, en la sección 3 se describe la solución diseñada para dar solución a los inconvenientes mencionados y los aportes a la disciplina y, finalmente, la sección 4 presenta las líneas de investigación que quedan abiertas a partir de este trabajo.

2. Motivación

2.1. Estado de la cuestión

A lo largo de la historia del desarrollo de software se produjeron numerosos cambios paradigmáticos que, paulatinamente, fueron adoptados por la industria, por ejemplo, la programación imperativa, la descomposición funcional, la orientación a objetos. Estos paradigmas nacieron en ambientes de investigación y desarrollo, para luego ser ofrecidos a través como productos de mercado [3].

La orientación a aspectos nació en 1996 con la propuesta de Gregor Kiczales, denominada Programación Orientada a Aspectos (*Aspect-Oriented Programming, AOP*) [4]. Su objetivo principal consistió en mejorar la división del código en módulos, para dar solución a los problemas de desparramo (*scattering*) y enredo (*tangling*) de las funciones que se encuentran diseminadas a lo largo de todo un sistema y entremezcladas con otras porciones de código. A estas funciones se las conoce como incumbencias transversales (*crosscutting concerns*) [5]. Con el paso del tiempo aparecieron el diseño orientado a aspectos, el análisis de sistemas orientado a aspectos y, finalmente, se denominó Desarrollo de Software Orientado a Aspectos (*Aspect-Oriented Software Development*) a su empleo en todas las fases del SDLC. Para el caso particular de la aplicación de la orientación a aspectos en las fases tempranas del SDLC, se emplea el nombre de “aspecto temprano” (*early aspect*), que es “una incumbencia que corta transversalmente una descomposición dominante de artefactos o módulos base derivados del criterio de separación dominante de incumbencias, en las primeras etapas del ciclo de vida del software” [6].

La separación de incumbencias es un principio muy importante de la ingeniería de software [7-9]. Esta posibilidad de contar con una mayor modularidad ofrece una serie de beneficios a lo largo de todas las fases del SDLC: comprensibilidad [10-17], reusabilidad [10-14] [17-18], mantenibilidad [10-11] [15-20], reducción de la complejidad [12] [21-25], capacidad de evolución [16] [25-27], extensibilidad [10] [17], flexibilidad [10] [28], reducción en los costos del desarrollo [20] [27], escalabilidad [25] [29], adaptabilidad [13], calidad [26], entre otros.

Estas ventajas son alentadoras e, incluso, aplicables a todas las etapas del SDLC [30-32], pero su explotación en la industria todavía no se evidencia claramente. Aún no existen propuestas lo suficientemente maduras y con una presencia real en el mercado, a través de productos concretos y con soporte de parte de la industria [1]. Esta escasez de propuestas acarrea la falta de aplicación efectiva del paradigma de aspectos en aquel ámbito, aunque en algunas empresas puedan observarse implementaciones en las que se realiza la separación de algunas incumbencias transversales correspondientes a requisitos no funcionales (incumbencias asimétricas) y, en la mayoría de estos casos, sin emplear lenguajes orientados a aspectos [33].

Si bien en el ámbito académico existen numerosas propuestas, su implementación y aplicación están restringidas actualmente al ámbito privado [1]. Por ejemplo, el MIT había pronosticado en 2001 que la programación orientada a aspectos sería una de las diez tecnologías emergentes que cambiarían el mundo y, sin embargo, doce años después prácticamente todavía no se había adoptado [34]. En un estudio de Muñoz et al. [35] se menciona que, casi al momento de cumplirse la previsión del MIT, solamente se

había utilizado programación orientada a aspectos en el 0,5% de los proyectos escritos en Java, subidos al repositorio open source sourceforge.net, entre 2001 y 2008. Por lo tanto, la orientación a aspectos se presenta hoy más como un área de investigación, que necesita ser explotada con mayor profundidad [1]. Esta es la razón que motiva a buscar un enfoque que permita acelerar la adopción del paradigma en el ámbito industrial.

2.2. Dificultades para aplicar las propuestas disponibles

El empleo de la orientación a aspectos en proyectos reales en la industria, tal como se la encuentra disponible actualmente, llevaría a enfrentar algunos obstáculos, impedimentos y dificultades importantes:

1. No existe una propuesta para el SDLC completo. Para cubrirlo totalmente deberían tomarse alternativas de diferentes autores [1] [36-38]. Pero este collage derivaría en un método heterogéneo y prácticamente inaplicable. Además, las alternativas actualmente ofrecidas y más difundidas no abarcan siquiera más de dos fases del SDLC [1].
2. Las propuestas disponibles no cuentan con una difusión suficiente en la industria debido a que en muchos casos se emplean notaciones no estándares [1]; la mayor parte de las propuestas no cuentan con herramientas de modelado que las soporten o, si las poseen, no son asequibles en el mercado ni tienen el soporte técnico necesario [1] [39]; y, por todo esto, se puede concluir que no existe una masa crítica de profesionales formados que puedan incorporarse rápidamente a un equipo de proyecto.
3. Si no hay aplicación en el mundo real, entonces no hay madurez suficiente en técnicas necesarias para conducir adecuadamente los proyectos. Por ejemplo, las técnicas de estimación requieren contar con una base estadística importante para poder asegurar un grado de certeza razonable en sus resultados [40].
4. La aceptación comercial o no de una técnica o producto gravita considerablemente en su posibilidad de uso en el mundo real. Un caso emblemático es el de las bases de datos orientadas a objetos [41], que nunca fueron adoptadas suficientemente por la industria y perteneciente a un paradigma ampliamente difundido y aceptado, no pasaron de ser materia de discusión e implementadas en algunos productos comerciales que no alcanzaron el éxito esperado. Y no hay suficiente evidencia de la aplicación de la orientación a aspectos en la industria [1].

3. Aporte a la disciplina

3.1. Objetivos de la solución

Los objetivos generales de nuestra tesis, junto con sus respectivos objetivos específicos, son:

1. Describir un nuevo proceso marco para las etapas tempranas del SDLC, desde el modelo de negocios hasta la especificación de requisitos de software completa.
 - a. Ofrecer un proceso marco, no específico y liviano, de modo que permita su empleo con diferentes modelos del SDLC.
 - b. Cubrir las etapas tempranas del SDLC, desde el modelado de negocio hasta la obtención de una especificación de requisitos completa y coherente.
 - c. Emplear herramientas y técnicas estándares, de amplia difusión en la industria, para facilitar su adopción inmediata.
 - d. Emplear notaciones estándares, para lograr modelos y especificaciones comprensibles y no ambiguas, que puedan contar con soporte de herramientas de software disponibles en el mercado.
2. Describir ese proceso marco para las etapas tempranas del SDLC empleando el paradigma de la orientación a aspectos.
 - a. Ofrecer una alternativa que sea completamente orientada a aspectos.

- b. Obtener las incumbencias en forma progresiva a lo largo de todos los modelos.
- c. Mantener la separación de incumbencias a lo largo de todos los modelos.
- d. Mantener la trazabilidad bidireccional de las incumbencias de punta a punta.
- e. Obtener las incumbencias en forma natural a lo largo de todos los modelos, de manera de no afectar la obtención de los objetivos de cada modelo.
- f. Potenciar los beneficios que se esperan de cada modelo mediante el empleo del paradigma de aspectos.

3.2. Características de la solución

La propuesta descrita en nuestra tesis doctoral consiste en la definición de un proceso marco genérico e independiente de los modelos del SDLC, que cubre las etapas tempranas de ese ciclo de vida con los siguientes modelos: de procesos de negocio, de requisitos de usuario y de requisitos de software; este último constituido por tres vistas: de procesos, estática y de estados. El modelo de procesos de negocio permite el gobierno de TI con soluciones de Arquitectura Empresarial; el de requisitos de usuario da soporte a las dos partes de la ingeniería de requisitos: el desarrollo y la administración de los requisitos, además de la gestión de la demanda para la etapa de mantenimiento; y el modelo de requisitos de software está constituido por tres vistas complementarias: de procesos, estática y de estados, lo que posibilita una definición completa de los requisitos de software (Figura 1).

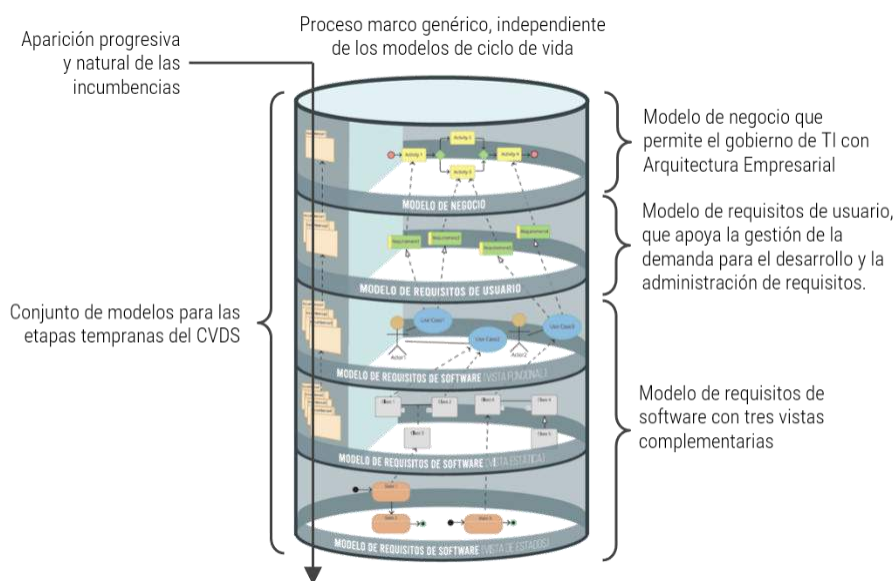


Figura 1. Esquema de AOP4ST: modelos y vistas.

En AOP4ST las incumbencias se detectan en forma progresiva y se mantienen perfectamente separadas a lo largo de los diferentes niveles de abstracción, acrecentándose en número y aumentando en granularidad, además de mantener una trazabilidad bidireccional entre ellas. Esta trazabilidad bidireccional también se mantiene entre los elementos constitutivos de cada uno de los modelos. Se puso especial énfasis en el empleo de notaciones, herramientas y técnicas ampliamente difundidas en la industria [1]. Estas cuestiones permiten que se ponga el foco en los objetivos que se esperan en cada una de las fases del SDLC, con el objetivo de hacer software de calidad, y no distraerse con las exigencias que impone el paradigma de aspectos, sino que, por el contrario, este enfoque sea de verdadera ayuda y un instrumento para alcanzar aquel objetivo (Figura 2).

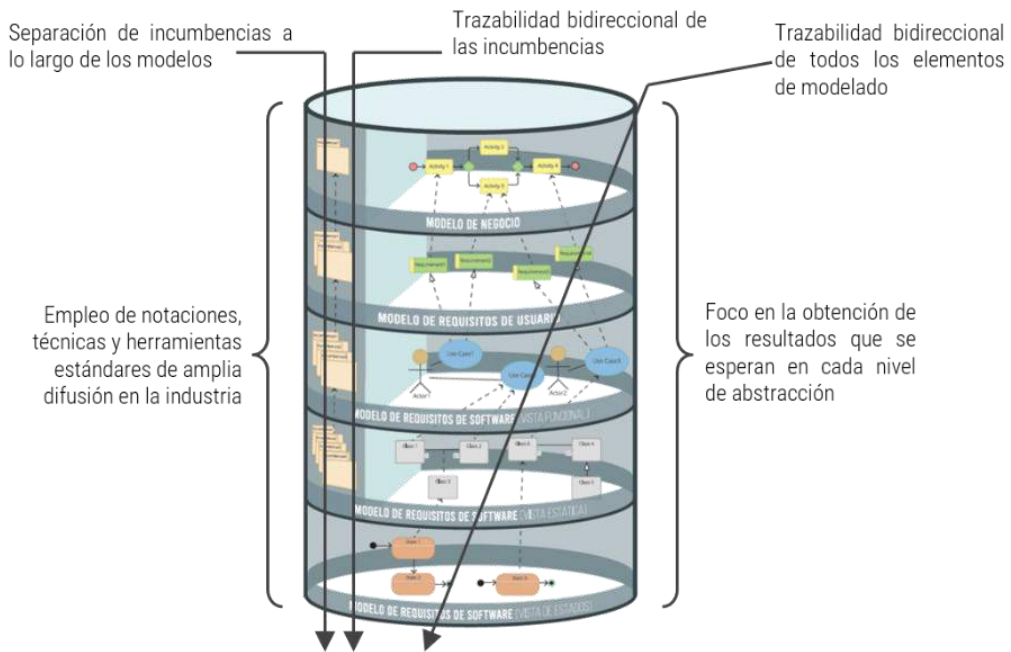


Figura 2. Esquema de AOP4ST: separación de incumbencias y trazabilidad.

Igualmente, la orientación a aspectos exige realizar ciertas actividades para las que se propusieron herramientas y técnicas para reducir su impacto, a la vez que se diseñaron modelos específicos que, además de que se pueden generar en forma automática por medio de herramientas de software, permiten analizar la calidad de los modelos y aportar información de valor para la mejor toma de decisiones (Figura 3).

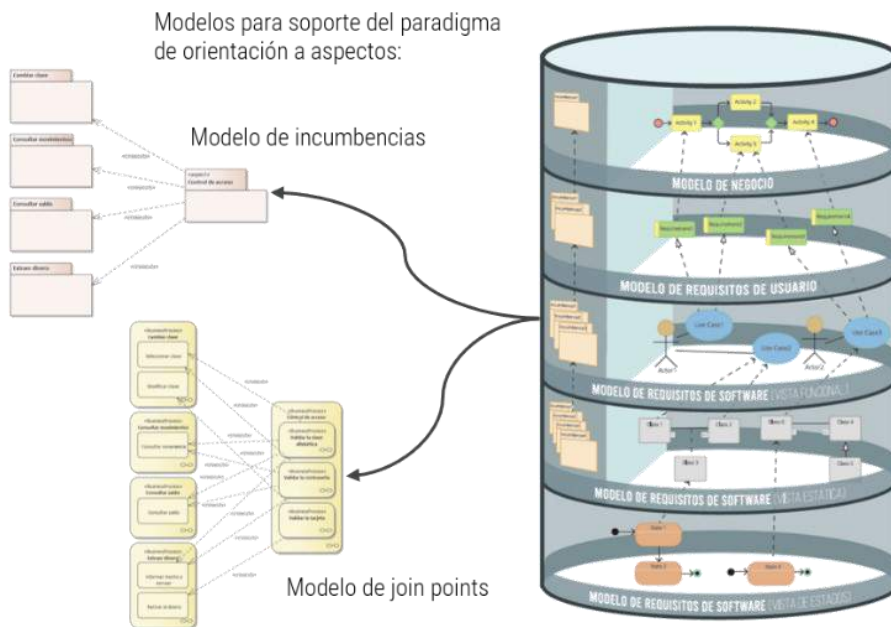


Figura 3. Esquema de AOP4ST: modelos y mecanismos para la administración de los aspectos.

Algunas características destacadas de AOP4ST hacen que ocupe una posición de privilegio en el concierto de las propuestas existentes para el empleo de la orientación a aspectos en el modelo de negocio (Figura 4) según el método de evaluación propuesto por Jalali [42].

Propuesta	Grupos de requisitos						Total general
	Requisitos básicos	Exposición de firma	Composición de reglas	Relaciones de los advices	Patrones de transformación	Soporte de fases	
AOP4ST	3 de 3	3	3,33	4	1	0,8	12,13
Jalali et al.	3 de 3	1	2	2,67	2	3,2	10,87
Cappelli et al.	3 de 3	1	2,67	2	1	0,8	7,47
Jabeen et al.	3 de 3	1	0,67	2,67	1	0,8	6,14
AO4BPMN	3 de 3	1	0,67	2,67	1	0,8	6,14
Patiniotakis et al.	3 de 3	1	0,67	2	1	0,8	5,47
AO4BPEL	2 de 3	0	0	0	0	0	0
Collell	2 de 3	0	0	0	0	0	0
AOBPMN	2 de 3	0	0	0	0	0	0
Wang et al.	1 de 3	0	0	0	0	0	0
Shankardass	1 de 3	0	0	0	0	0	0

Figura 4. Ubicación de AOP4ST con respecto al resto de las propuestas de aspectos existentes para el modelo de negocio.

Todos los modelos contemplan la detección, separación y encapsulamiento de las incumbencias, el establecimiento de las relaciones entre los elementos del mismo nivel de abstracción y con elementos del nivel superior, la posterior composición de las incumbencias y la resolución de conflictos tras la composición, además de prácticas de modelado recomendadas. Además cuenta con un conjunto de reglas de composición de incumbencias en el modelo de negocio; con el empleo de las técnicas de análisis de las relaciones de contribución positivas y negativas entre requisitos de usuario; con la administración de requisitos de software en su vista funcional que permite el empleo de diferentes técnicas de especificación (por ejemplo, casos de uso e historias de usuario); con las técnicas de composición de las clases minimizando los conflictos; y con la composición de las máquinas de estados a partir de la composición de las clases.

3.3. Validaciones

La investigación en el ámbito de la ingeniería de software exige el estudio empírico de los fenómenos que suceden en el mundo real, entre los que se encuentra el desarrollo de nuevos modelos de proceso [43], tema central de este trabajo. De los métodos de investigación empíricos que se emplean en la ingeniería de software fueron utilizados dos métodos cualitativos y uno cuantitativo:

- **Caso de estudio:** en el que se evaluó la aplicación de AOP4ST en un proyecto real completo, pero un tiempo después de la culminación de ese proyecto (post-mortem análisis [44]). Este caso de aplicación fue el remodelado, usando AOP4ST, de todos los procesos del Laboratorio Hidalgo, de la provincia de Buenos Aires, que se habían llevado a cabo con el empleo de un enfoque orientado a objetos y con UML. Los modelos son complejos y, los más de ochenta procesos punta a punta modelados, describen los procesos de funcionamiento del laboratorio para el procesamiento de todas las determinaciones de análisis bioquímicos, tanto para los análisis corrientes como para los correspondientes a la investigación de nuevas drogas para laboratorios de Norteamérica. Los modelos originales fueron auditados por una empresa italiana, que debía asegurar el cumplimiento con las normas establecidas por la Food and Drug Administration (FDA) de los Estados Unidos.
- **Proyectos de investigación-acción:** en la que investigadores y profesionales trabajaron en forma combinada utilizando las técnicas de AOP4ST en tres proyectos reales en la industria, aunque aplicándolas solo a algunos modelos de la propuesta. El primero de los casos fue el modelado de los procesos corporativos de la empresa Prosegur, en donde se detectaron, separaron, encapsularon y compusieron las incumbencias del modelo de negocio del área de Vigilancia en tres países de la compañía: España, Brasil y Argentina. A partir de estos procesos de negocio se detectaron los requisitos de usuario, tal como se propone en este trabajo, y, con ellos, se elaboraron los nuevos procesos corporativos y se estableció el alcance del software a adquirir para dar soporte a tales procesos.

El segundo caso correspondió al modelado de los procesos de la Dirección de Estadísticas e Investigaciones Económicas del Gobierno de la Provincia de Mendoza, ejecutado durante 2019, y en donde nuevamente se realizaron la detección, separación, encapsulamiento y composición de los procesos de negocio.

El tercer proyecto en el que pudieron emplearse las ideas que se presentan en el capítulo dedicado al modelo de negocio de AOP4ST consistió en el modelado de los procesos internos de la empresa de gestión de transacciones bancarias Red Link, en Buenos Aires, proyecto iniciado a finales de 2019 y aún en marcha.

- **Investigación secundaria:** se llevó a cabo un estudio de mapeo sistemático para validar las hipótesis de trabajo y aportar evidencia científica a las afirmaciones vertidas en la tesis, dado que las revisiones sistemáticas son uno de los bloques de construcción clave de la Ingeniería de Software Basada en Evidencia (EBSE) [43]. El protocolo de este estudio fue presentado en un evento en la Argentina en 2018 [45] y el mapeo sistemático fue publicado por la revista científica “Journal of King Saud University – Computer and Information Sciences” [1].

Adicionalmente, las diferentes ideas de AOP4ST fueron desarrolladas a lo largo de cinco proyectos de investigación consecutivos, en los que se produjo una veintena de publicaciones, la mayoría de ellas en congresos y revistas internacionales:

	Proyecto	Institución	Período	Avances y publicaciones
1	“Procesos de desarrollo de software de calidad basados en aspectos”	Universidad Tecnológica Nacional FRM	2010 a 2012	Sentó las bases principales de un proceso marco de desarrollo de software que permita aprovechar los beneficios de la orientación a aspectos [46].
2	“Detección temprana de aspectos en el modelado de negocios y el desarrollo de los requisitos”	Instituto de investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat	2013 a 2015	Se definieron los modelos y vistas de AOP4ST [2].
3	“Modelado de procesos de negocio orientados a aspectos con BPMN”		2015 a 2017	Abordó el modelo de negocio en AOP4ST [47-50] y se obtuvo la principal evidencia con un estudio de mapeo sistemático [1] [45].
4	“Ingeniería de requisitos de usuario en AOP4ST”		2018 a 2020	Se centró en el modelo de requisitos de usuario de AOP4ST [51-57].
5	“Ingeniería de requisitos de software orientada a aspectos en AOP4ST”		2020 a 2022	Se está trabajando sobre los requisitos de software, con énfasis en la vista estática [58].

Las ideas de AOP4ST también fueron aceptadas y expuestas en el “Simposio de Tesis Doctorales del IEEE – 11 Congreso Colombiano de Computación (11 CCC)” llevado a cabo en Popayán, Colombia, en 2016 [59], donde se recibieron opiniones muy enriquecedoras de parte del tribunal, que se plasmaron en la tesis.

4. Líneas de investigación futuras

Se considera de especial interés continuar las investigaciones sobre la detección de patrones, para el análisis de impacto ante cambios y para la resolución de conflictos. La evidencia obtenida de la aplicación de la orientación a aspectos en la industria también da cuenta de esta necesidad [1]. Sería oportuno dar un mayor tratamiento al modelo de requisitos de usuario, para abordar requisitos para sistemas más específicos. En nuestra tesis se trabajó sobre sistemas de propósito general, pero también se entiende que sería de mucho valor el considerar requisitos de sistemas de tiempo real, de uso intensivo de interfaz humano-computador, de computación paralela, de ingeniería, etc.

Un avance importante sería, también, el desarrollo de un soporte más específico de aplicaciones de software. Los lenguajes estándares empleados permitirán la interpretación de los diagramas, ya que en el modelo de negocio se utilizó BPMN que cuenta con BPEL WS-BPEL, y en el resto de los modelos se empleó el UML, que hace uso del OCL, también estándar, para su formalización. Además, se harían grandes aportes con la simulación para analizar la integridad de los modelos tras la composición, especialmente con los diagramas de comportamiento, como son los diagramas de procesos de negocio, los diagramas de actividades y de secuencias, que pueden obtenerse de las especificaciones estructuradas de los casos de uso, y los diagramas de estados, con la inclusión de variables de control para análisis de escenarios. También se observa necesario adaptar a este paradigma las técnicas de estimaciones de software existentes, evaluar el impacto en los roles de los integrantes de los equipos de proyecto a partir de las actividades que surgen del empleo de la orientación a aspectos, especialmente con el empleo de metodologías ágiles para el desarrollo de software [60].

Referencias

- [1] Pinciroli, F., Barros Justo, J.L. and Forradellas, R. (2020). "Systematic Mapping Study: on the coverage of aspect-oriented methodologies for the early phases of the software development life cycle". In: Elsevier, Journal of King Saud University – Computer and Information Sciences.
- [2] Pinciroli, F. (2015). AOP4ST – Aspect-Oriented Process for a Smooth Transition. In: Proceedings of the XVII Workshop de Investigadores en Ciencias de la Computación, WICC 2015, Salta.
- [3] Gabbrielli, M. and Martini, S. (2010). Programming languages: principles and paradigms. Springer, London.
- [4] Kiczales, G., Lamping, J., Mendhekar, A., Maeda, C., Videira Lopes, C., Loingtier, J.M. and Irwin, J. (1997). Aspect-Oriented Programming. In: Proceedings of the European Conference on Object-Oriented Programming, ECOOP, Finland. Springer-Verlag, Lecture Notes in Computer Science, no. 1241.
- [5] Wimmer, M., Schauerhuber, A., Kappel, G., Retschitzegger, W., Schwinger, W. and Kapsammer, E.E. (2011). A survey on UML-based aspect-oriented design modeling. ACM Computing Surveys, vol. 43, no. 4, pp. 1–33.
- [6] Baniassad, E., Clements, P.C., Araújo, J., Moreira, A., Rashid, A. and Tekinerdoğan, B. (2006). Discovering early aspects. In: IEEE Software, vol. 23, pp. 61–70.
- [7] Parnas, DL. (1972). "Information distribution aspects of design methodology". In: Communications of the ACM, vol.15, no.12, pp.1053-1058.
- [8] Dijkstra, E.W. (1982). On the role of scientific thought. In: Selected Writings on Computing: A Personal Perspective, Springer-Verlag, New York, pp. 60–66.
- [9] Ye, S. and He, C. (2013). A comparison of methods for identification of early aspects. In: Proceedings of the 3rd International Conference on Computer Science and Network Technology, ICCSNT 2013, pp. 275-279.
- [10] Mehmood, A. and Jawawi, D.N.A. (2014). A quantitative assessment of aspect design notations with respect to reusability and maintainability of models. In: Proceedings of the 8th Malaysian Software Engineering Conference, MySEC 2014, Langkawi, pp. 136-141.
- [11] Witteborg H., Charfi A., Colomer Collell D., Mezini M. (2014). Weaving Aspects and Business Processes through Model Transformation. In: Villari M., Zimmermann W., Lau KK. (eds) Service-Oriented and Cloud Computing, ESOC 2014, Lecture Notes in Computer Science, vol 8745, Springer, Berlin, Heidelberg.
- [12] Faßbender, S., Heisel, M. and Meis, R. (2015). A Problem-, Quality-, and Aspect-Oriented Requirements Engineering Method. In: Communications in Computer and Information Science, pp. 291–310.
- [13] FanJiang Y.Y., Kuo, J.Y., Ma, S.P., Huang, W.R. (2010). An Aspect-Oriented Approach for Mobile Embedded Software Modeling. In: Taniar D., Gervasi O., Murgante B., Pardede E., Apduhan B.O. (eds) Computational Science and Its Applications, ICCSA 2010, Lecture Notes in Computer Science, vol 6017, Springer, Berlin, Heidelberg, pp. 257–272.
- [14] Wang, Y., Singh, S., Hosking, J. and Grundy, J. (2006). An aspect-oriented UML tool for software development with early aspects. In: Proceedings of the 2006 International Workshop on Early Aspects at ICSE, EA '06, pp. 51-58.
- [15] Liu, X, Liu, S. and Zheng, X. (2009). Adapting the NFR Framework to Aspectual Use-Case Driven Approach. In: Proceedings of the Seventh ACIS International Conference on Software Engineering Research, Management and Applications, Haikou, pp. 209-214.

- [16] Mosconi, M., Charfi, A., Svacina, J. and Wloka, J. (2008). Applying and evaluating AOM for platform independent behavioral UML models. In: Proceedings of the 2008 AOSD Workshop on Aspect-Oriented Modeling, AOM '08, pp. 19–24.
- [17] Zhou, Y., Lei, G., Li, P. and Kong, L. (2008). Realizing extension use cases with AOP. In: Proceedings of the 2008 IEEE International Symposium on IT in Medicine and Education, Xiamen, 2008, pp. 1040-1044.
- [18] Jalali, A. (2015). Static Weaving in Aspect Oriented Business Process Management. In: Proceedings of the Conceptual Modeling, 34th International Conference, ER 2015, Stockholm, Springer, pp. 548–557.
- [19] De Oliveira, K.S., Franca, J.M.S. and Soares, M.S. (2013). Extensions of SysML for Modeling an Aspect Oriented Software Architecture with Multiple Views. In: Proceedings of the 10th International Conference on Information Technology: New Generations, Las Vegas, pp. 680–685.
- [20] Budwell, C.C. and Mitropoulos, F.J. (2008). The SLAI Methodology: An Aspect-Oriented Requirement Identification Process. In: Proceedings of the 2008 International Conference on Computer Science and Software Engineering, vol. 2, pp. 296–301.
- [21] Jalali, A., Maggi, F.M. and Reijers, H.A. (2015). Enhancing AO-BPM through declarative rules. In: Proceedings of the Conceptual Modeling, 34th International Conference, ER 2015, Stockholm, pp. 108–118.
- [22] Jalali, A., Ouyang, C., Wohed, P. and Johannesson, P. (2017). Supporting aspect orientation in business process management. In: Software and Systems Modeling, vol. 16, no. 3, pp. 903–925.
- [23] Mehner, K., Monga, M. and Taentzer, G. (2006). Interaction Analysis in Aspect-Oriented Models. In: Proceedings of the 14th IEEE International Requirements Engineering Conference, RE'06, pp. 69–78.
- [24] Amirat, A., Mohamed, L. and Khammaci, T. (2008). Modularization of Crosscutting Concerns in Requirements Engineering. In: The International Arab Journal of Information Technology, vol. 5, no. 2, pp. 120-125.
- [25] Amálio, N., Kelsen, P., Ma, Q. and Glodt, C. (2010). Using VCL as an Aspect-Oriented Approach to Requirements Modelling. In: Transactions on Aspect-Oriented Software Development VII: A Common Case Study for Aspect-Oriented Modeling, pp. 151–199.
- [26] Conejero, J.M., Hernández, J., Jurado, E. and van den Berg, K. (2010). Mining early aspects based on syntactical and dependency analyses. In: Science of Computer Programming, vol. 75, no. 11, pp. 1113–1141.
- [27] Singh, N. and Singh Gill, N. (2012). Towards an Integrated AORE Process Model for Handling Crosscutting Concerns. In: International Journal of Computer Applications, vol. 37, no. 3, pp. 18–24.
- [28] Keriakos, M., Hosny, H. and Aly, S.G. (2012). Context aware business process aspect modeler. In: Proceedings of the 2nd International Conference on Pervasive Embedded Computing and Communication Systems, PECCS-2012, pp. 206-213.
- [29] Bošković, M., Mussbacher, G., Bagheri, E., Amyot, D., Gašević, D. and Hatala, M. (2011). Aspect-Oriented Feature Models. In: Dingel J., Solberg A. (eds) Models in Software Engineering, MODELS 2010, Lecture Notes in Computer Science, vol 6627, Springer, Berlin, Heidelberg, pp. 110–124.
- [30] Rashid A. and Moreira A. (2006) Domain Models Are NOT Aspect Free. In: Nierstrasz O., Whittle J., Harel D., Reggio G. (eds) Model Driven Engineering Languages and Systems. MODELS 2006. Lecture Notes in Computer Science, vol. 4199. Springer, Berlin, Heidelberg, pp. 155–169.
- [31] Gray, J., Bapty, T., Neema, S. and Tuck, J. (2001). Handling crosscutting constraints in domain-specific modeling. In: Communications of the ACM, vol. 44, no. 10, pp. 87–93.
- [32] Niu, N., Easterbrook, S. and Yu, Y. (2007). A Taxonomy of Asymmetric Requirements Aspects. In: Proceedings of the 10th International Conference on Early Aspects: Current Challenges and Future Directions, pp. 1–18.
- [33] Rashid, A., Cottenier, T., Greenwood, P., Chitchyan, R., Meunier, R., Coelho, R., Südholt, M. and Joosen, W. (2010). “Aspect-oriented software development in practice: Tales from AOSD-Europe”. In: Computer, California, vol. 43, no. 2, pp. 19-26.
- [34] Munoz, F., Baudry, B., Delamare, R. and Le Traon, Y. (2013). “Usage and testability of AOP: An empirical study of AspectJ”. In: Information and Software Technology, vol. 55, no. 2, pp. 252–266.
- [35] Munoz, F., Baudry, B., Delamare, R. and Le Traon, Y. (2009). “Inquiring the usage of aspect-oriented programming: An empirical study”. In: 2009 IEEE International Conference on Software Maintenance, Edmonton, pp. 137-146.
- [36] Magableh, A., Shukur, Z. and Ali, N.M. (2013). Systematic review on aspect-oriented UML modeling: A complete aspectual UML modeling framework. In: Journal of Applied Sciences, vol. 13, no. 1, pp. 1–13.
- [37] Mohite, S., Phalnikar, R., Joshi, M., Joshi, S.D. and Jadhav, S. (2014). Requirement and interaction analysis using aspect-oriented modeling. In: IEEE International Advance Computing Conference, IACC 2014, pp. 1448–1453.

- [38] Gerami, M. and Ramsin, R. (2011). A framework for extending agile methodologies with aspect-oriented features. In: Proceedings of the Fifth International Conference on Research Challenges in Information Science, Gosier, pp. 1-6.
- [39] Elrad, T., Aldawud, O. and Bader, A. (2005). Expressing Aspects using UML Behavioral and Structural Diagrams. In: Filman, R., Elrad, T., Clarke, S. and Akşit, M. (eds) Aspect-Oriented Software Development, Addison Wesley, 2005, pp. 459-478.
- [40] McConnell, S (2006). Software Estimation: Demystifying the Black Art. Microsoft Press.
- [41] Leavitt, N. (2000). Whatever Happened to Object-Oriented Databases? In: Computer, vol. 33, no. 8, Long Beach, pp. 16–19.
- [42] Jalali, A. (2014). Assessing Aspect Oriented Approaches in Business Process Management. In: Johansson B., Andersson B., Holmberg N. (eds) Perspectives in Business Informatics Research, BIR 2014, Lecture Notes in Business Information Processing, vol 194, Springer, Cham, pp. 231–245.
- [43] Sjoberg, D.I.K., Dyba, T. and Jorgensen, M. (2007). The Future of Empirical Methods in Software Engineering Research. In: Future of Software Engineering, FOSE '07, Minneapolis, pp. 358-378.
- [44] Wohlin, C., Höst, M., Henningsson, K. (2006). Empirical Research Methods in Web and Software Engineering. In: Mendes, E., Mosley, N. (eds), Web Engineering, Springer, Berlin.
- [45] Pinciroli, F., Barros Justo, J.L., Zeligueta, L. and Palma, M. (2018) Systematic Mapping Protocol. Coverage of Aspect-Oriented Methodologies for the Early Phases of the Software Development Life Cycle. In: Proceedings of the II International Congress on Computer Sciences and Information Systems, CICCSI 2018, Mendoza.
- [46] Pinciroli, F. (2011). Consideraciones para un proceso de desarrollo de software de calidad orientado a aspectos. In: Proceedings of the Sexto Encuentro de Investigadores y Docentes de Ingeniería, EnIDI 2011, San Rafael.
- [47] Pinciroli, F., Barros Justo, J.L. and Forradellas, R. (2017). Aspect-Oriented Business Process Modeling Approaches: An assessment of AOP4ST. In: Proceedings of the 46 Jornadas Argentinas de Informática, JAIIO 2017, pp. 40-47.
- [48] Pinciroli, F. and Zeligueta, L. (2016). El modelo de negocio en AOP4ST. In: Proceedings of the XVIII Workshop de Investigadores en Ciencias de la Computación, WICC 2016, Entre Ríos, pp. 423-426.
- [49] Pinciroli, F. (2016). Aspect-oriented business process composition rules in AOP4ST. In: Proceedings of the 35th International Conference of the Chilean Computer Science Society, Valparaíso, pp. 1–6.
- [50] Pinciroli, F. and Albino, G. (2020). Reglas de composición para modelos de procesos orientados a aspectos con BPMN 2.0. In: Proceedings of the 8° Congreso Nacional de Ingeniería en Informática e Ingeniería en Sistemas de Información, CONAIIISI 2020, San Francisco.
- [51] Pinciroli, F. and Barros Justo, J.L. (2017). Early aspects in ‘Aspect-Oriented Process for a Smooth Transition’. In: Proceedings of the XXIII Argentine Congress of Computer Science, La Plata, pp. 692-701.
- [52] Pinciroli, F. (2017). Concern detection along the requirement development. In: Entre Ciencia e Ingeniería, vol. 12, no. 23, Pereira, pp. 117-122.
- [53] Pinciroli, F. (2018). Requisitos de usuario y gestión de la demanda en AOP4ST. In: Proceedings of the XX Workshop de Investigadores en Ciencias de la Computación, WICC 2018, Buenos Aires, pp. 511-515.
- [54] Pinciroli, F. and Palma M. (2020). Desarrollo de requisitos en “Aspect-Oriented Process for a Smooth Transition”. In: Proceedings of the XXII Workshop de Investigadores en Ciencias de la Computación, WICC 2020, El Calafate.
- [55] Pinciroli, F. (2020). Explicit and implicit join point designation in aspect-oriented business modeling. In: Proceedings of the XLVI Conferencia Latinoamericana de Informática – CLEI 2020, Loja.
- [56] Pinciroli, F. (2016). Improving software applications quality by considering the contribution relationship among quality attributes. In: Proceedings of the 3rd International Workshop of Computer Antifragility and Antifragile Engineering, ANTIFRAGILE 2016, Procedia Computer Science, vol. 83, pp. 970–975.
- [57] Pinciroli, F. (2016). An HCI quality attributes taxonomy for an impact analysis to interactive systems design and improvement. In: Pesado, P.M. et al. (eds) Computer Science and Technology Series, Proceedings of the XXII Argentine Congress of Computer Science, Selected Papers, La Plata, pp. 133-143.
- [58] Pinciroli, F. (2019). Modeling the Static View in Aspect-Oriented Software Development. In: Proceedings of the III International Congress on Computer Sciences and Information Systems, CICCSI 2019, Mendoza.
- [59] Pinciroli, F. (2016). Aspect-Oriented Process for a Smooth Transition. In: Proceedings of the Ph.D. Symposium, IEEE 11 Congreso Colombiano de Computacion, Popayán.
- [60] Pinciroli, F. (2020). A maturity model for the Integrated Agile Transformation Model™. In: Proceedings of the IV International Congress on Computer Sciences and Information Systems, CICCSI 2020, Mendoza.