

Actas

XXVI Congreso Argentino de  
Ciencias de la Computación

caaic  
2020

05 al 09 de Octubre

Congreso Virtual



Universidad Nacional  
de La Matanza

Registro del presente Libro de Actas:

**Congreso Argentino de Ciencias de la Computación**  
2020 CACIC: XXVI Congreso Argentino de Ciencias de la Computación /  
Coordinación General de Alicia Mon. - 1a ed. - San Justo: Universidad  
Nacional de La Matanza, 2021.  
Libro digital, PDF

ISBN 978-987-4417-90-9



Archivo Digital: descarga y online  
ISBN 978-987-4417-90-9

1. Computación. 2. Congreso. I. Mon, Alicia, coord. II. Título.  
CDD 004.071

## ÍNDICE

Autoridades	
UNLaM .....	1
RedUNCI .....	2
Comité Organizador .....	3
Escuela Internacional de Informática.....	3
Comité Académico .....	4
Comité Científico.....	6
<b>WORKSHOPS</b> .....	
WASI – Agentes y Sistemas Inteligentes .....	9
Modeling Human Decision Making with Defeasible Logic Programming (13337). Maximiliano Sapino, Edgardo Ferretti, Luciana Mariñelarena Dondena, Marcelo Errecalde .....	10
A Real Case of Multi-Period Water Distribution Network Design solved by a Hybrid Simulated Annealing (13346/13477). Carlos Bermudez, Hugo Alfonso, Gabriela Minetti, Carolina Salto .....	20
Image augmentation for object detection of grapevines (13354). Maximiliano Sapino <sup>1,2,3</sup> , Edgardo Ferretti, Luciana Mariñelarena Dondena, Marcelo Errecalde .....	30
Navegación autónoma mediante aprendizaje por refuerzo (13383). Ignacio Agustín Costa, Esteban José De Leo, Iris Inés Sattolo, Marisa Daniela Panizzi .....	40
Manejo de Diversidad en CHC aplicado a la optimización del costo energético en parques eólicos (13412). Martín Bilbao, Guillermo Leguizamón.....	50
Particle Swarm Optimization with Adaptive Inertia Weight using Fuzzy Logic for Large- Scale Problems (13433). Fabiola Patricia Paza, Guillermo Leguizamón, Efrén Mezura- Montes.....	60
WPDP – Procesamiento Distribuido y Paralelo.....	70
Comparación de un Algoritmo de Bidiagonalización para su Utilización en la Recuperación de Información (13380). Osvaldo Spositto, Viviana Ledesma, Gastón Procopio, Hugo Ryckeboer, Victoria Saizar y Alexis Vainberg .....	71
Comparación de Arquitecturas HPC para Computar Caminos Mínimos en Grafos. Intel Xeon Phi KNL vs NVIDIA Pascal (13405). Manuel Costanzo, Enzo Rucci, Ulises Costi, Franco Chichizola, Marcelo Naiouf .....	81
WTIAE – Tecnología Informática Aplicada en Educación.....	92
Sistema de Comunicación y Armado Multimedial de Material Pedagógico para Estudiantes Regulares y Diversos Funcionales en la Educación Superior (13338). Vanina Cecilia Chiavetta, Luis Mariano Mongelo, Marcela Dávila, María Laura Villarruel, Gisela Díaz, Néstor Ariel Pan, Elías Biscaia .....	94
Educación Virtual con Alumnos de Primer Año de Ingeniería en tiempos de aislamiento social obligatorio (13341). Artemisa Trigueros, Mabel Compagnoni, Larisa Toro.....	103
Formación de Competencias de la Industria 4.0 en Estudiantes de 1er. Año de Ingeniería (13342). Artemisa Trigueros, Mabel Compagnoni, Larisa Toro, Sabrina Gómez.....	113

Framework para configuración de batallas históricas: Generando a través del juego un espacio de aprendizaje (13345). Christian Parkinson, Roxana Martínez, Pablo Vera, Rocío Rodríguez .....	123
Juego serio de realidad virtual para acercar figuras importantes de la historia de la Informática (13381). Mariano Mazza, Cecilia Sanz, Verónica Artola .....	133
La Producción de Juegos Serios Móviles. Posibilidades y Desafíos para el Docente de Nivel Superior (13386). Edith Lovos, Iván Basciano, Evangelina Gil, Cecilia Sanz.....	144
Prácticas educativas STEAM para fomentar el desarrollo de competencias científico-tecnológicas (13390). Silvina Manganelli.....	154
Herramientas de model finding para asistir en la Construcción de Especificaciones Formales (13397). Sonia Permigiani, María Marta Novaira, Gastón Scilingo, Marcelo Arroyo.....	165
Experiencia Educativa con Tecnologías Emergentes (13436). Mariela Zuñiga, Claudia Liendo, María Rosas, Graciela Rodríguez, Nicolás Jofré, Yoselie Alvarado, Myriam Villegas, Jacqueline Fernandez, Roberto Guerrero .....	175
MEHI – Material Educativo Hipermedial Interactivo para un curso de Química General. Resultados de una experiencia áulica (13441). García Ricardo, Gorga Gladys , Bertone Rodolfo.....	186
WCGIV – Computación Gráfica, Imágenes y Visualización .....	197
Simulation of High-Visual Quality Scenes in Low-Cost Virtual Reality (13411). Matías N. Selzer, M. Luján Ganuza, Dana K. Urribarri, Martín L. Larrea, Silvia M. Castro .....	198
NATUS. A Physically-Based Rendering Engine with Real-Time Global Illumination (13415). Ignacio del Barrio, María Luján Ganuza, Silvia M. Castro .....	208
Towards defining a CAVE like system performance evaluation (13426). Nicolás Jofré, Graciela Rodríguez, Yoselie Alvarado, Jacqueline Fernández, Roberto Guerrero .....	218
ARPOT: An Augmented Reality Platform for Outdoor Teaching (13439). Matías N. Selzer, Dana K. Urribarri, Nicolás F. Gazcón Martín L. Larrea, Claudia Sereno, Antonela Volonté, Aldana Mastrandrea, Silvia M. Castro.....	229
WBDMD – Bases de Datos y Minería de Datos.....	239
Seguimiento de rodeos en establecimientos agropecuarios (13335). Pablo Rafael Palmero, Edilma Olinda Gagliardi, Maria Gisela Dorzán .....	240
Conducting a systematic review: Trends in machine learning and text mining (13372). Mariana Falco, Ignacio Berdiñas .....	250
Una Aplicación de Espacios Métricos en el Ámbito del Comercio Electrónico (13377). Libertad Speranza, María de los Ángeles de la Torre, Norma Herrera .....	260
Análisis del Desempeño de Clustering y Árboles de Decisión en la Evaluación Clínica de Microbiomas de Pacientes con Cáncer Colorrectal (13384). Laura Avila, Victoria Santa María, Luis López, Marcelo Soria, Cristóbal R. Santa María .....	270
Decisiones y lecciones aprendidas en un proceso ETL aplicado a sistemas con testimonios de delitos de lesa humanidad (13387). David Troncoso *, Agustina Buccella, Alejandra Cechich.....	279

SMOTE, Algoritmo para balanceo de clases en un estudio aplicado a la ganadería (13396). Osvaldo Sposito, Gabriel Blanco, Lorena Matteo, Marcelo Levi y Julio Bossero.....	289
Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales (13428). Luciano Marrero, Verena Olsow, Fernando Tesone, Pablo Thomas, Lisandro Delia, Patricia Pesado.....	299
Smart Grids Challenge: A competitive variant for Single Objective Numerical Optimization (13435). Fabricio Loor, M. Guillermo Leguizamón, Efrén Mezura-Montes .....	310
Representación del Trie de Sufijo: Una Evaluación Empírica (13437). Darío Ruano, Jesica Cornejo, Paola Azar, Norma Herrera .....	318
WIS – Ingeniería de Software .....	328
Transferencia de Conocimiento en la Gestión de Calidad de la Ingeniería de Software (13327). Mauricio Rozo, Inés Casanovas, Ma. Florencia Pollo Cattaneo.....	329
ScrumGame: A Serious Game to Initiate Software Trainees in Scrum (13331). Guillermo Rodríguez, Matías Glessi, Alfredo Teyseyre, Pablo Gonzalez .....	339
Revisión sistemática de la literatura sobre implementación de arquitecturas software para sistemas críticos (13332). Joaquín Acevedo, Andrea Lezcano, Emanuel Irrazábal.....	350
A Parallel Tableau Algorithm for BIG DATA Verification (13334). Fernando Asteasuain, Luciana Rodriguez Caldeira .....	360
Evaluación de un Modelo Ontológico basado en la Adecuación Funcional de la Norma ISO 25010 para la Elicitación de Requisitos de Software (13344). Rebeca Yuan, Carlos Salgado, Mario Peralta, Alberto Sánchez .....	370
Impacto de una debilidad de ciberseguridad en la arquitectura de un sistema electromédico (13353). Diego Coulombie, Agustín Reyes, Alberto Miguens .....	380
TAPIR: An Object-Oriented Programming Testing Framework based on Message Sequence Specification with Aspect-Oriented Programming (13357). Martín L. Larrea, Dana K. Urribarri.....	389
Risk Refinement in the Deployment Process of Software Systems: A Case Study (13358). Felipe Ortiz, Marisa Panizzi, Rodolfo Bertone.....	399
Aplicaciones Móviles 3D: un estudio comparativo de performance y consumo de energía (13375). Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado, Jefferson Perez Altamirano, Martín De la Canal Erbetta.....	408
Modelo de evaluación de datos utilizando el enfoque GQM (13414). Julieta Calabrese, Silvia Esponda, Ariel Pasini, Patricia Pesado .....	418
WARSO – Arquitectura, Redes y Sistemas Operativos.....	429
Parámetro H: Análisis de la Autosimilaridad de un Tráfico VoIP (13333). Sergio A. Villiani, Diego R. Rodríguez Herlein .....	430
Interfaz Humano-Máquina web amigable para dispositivo IoT (13355). Rocío Madou, Federico N. Guerrero, Enrique M. Spinelli .....	440
Extended Petri Net Processor for Embedded Systems (13367). Luis Orlando Ventre, Orlando Micolini, and Emiliano Daniele.....	450

Análisis de rendimiento y consumo para sistema embebido con requisitos de tiempo explícitos (13376). Esteban Carnuccio, Graciela De Luca, Waldo Valiente, Mariano Volker, Raúl Villca, Matías Adagio .....	460
Arquitectura Cliente-Servidor de Alto Rendimiento para servicio RTK (13389). José H. Moyano, Karina M. Cenci, Jorge R. Ardenghi.....	470
WISS – Innovación en Sistemas de Software .....	480
Motor de Reglas Desacoplado Orientado a Formato JavaScript Object Notation (13359). Marcos Maciel.....	481
Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística (13374). Carlos Barrionuevo, Jorge Ierache , Iris Sattolo .....	491
Proposal of a Multimodal Model for emotional assessment within affective computing in gastronomic settings (13379). Jorge Ierache, Facundo Nervo, Iris Sattolo, Rocio Ierache, Gabriela Chapperón .....	501
Estrategias a nivel de macro y micro-planificación en un prototipo de verbalización de ontologías (13392). Martín Bermudez, Sandra Roger .....	512
GeoPerfil Profesional: una herramienta automática de información sobre profesionales (13393). Emanuel Balcazar, Lucas Bobadilla, Waldo Fusiman, Leo Ordinez.....	522
COVID-19: un análisis comparativo de Apps (13408). Juan Fernández Sosa, Verónica Aguirre, Lisandro Delía, Pablo Thomas, Leandro Corbalán Patricia Pesado .....	532
A Language for the Specification of the Schema of Spreadsheets for the Materialization of Ontologies (13420). Sergio Alejandro Gómez, Pablo Rubén Fillottrani.....	546
Genvi: propuesta de un desarrollo tecnológico para abordar la violencia de género (13427). Juan Pablo Quiñones, Viviana Harari, Ivana Harari .....	556
Sistema para el registro de factores psicológicos que inciden en los trastornos gastrointestinales funcionales. “Digestivamente” (13442). Juan Zaragoza, Jorge Ierache, Nicolas Domínguez, Hernan Merlino, Milagros Celleri, Florencia Koutsovitits, Cristian Garay .....	569
WPSTR – Procesamiento de Señales y Sistemas de Tiempo Real .....	579
Diseño y desarrollo de un sistema embebido para una trampa pitfall con data logger (13339). Mariano Droz, Juan Ramos, Pedro Benitez, Luz Zapata, Beatriz Díaz .....	580
Dispositivo GPS para monitoreo del rendimiento deportivo (13368). Luisina Santos, Marcelo Guiguet, Pablo Luengo, Mónica Sarobe, Eduardo Alvarez, Gustavo Useglio, Federico Gómez, Matías Capelli, Carlos Di Cicco .....	590
Simplifying concurrency and monitoring on Arduino for Internet of Things (13370). Ricardo Moran, Matías Teragni, Gonzalo Zabala .....	599
Control de cámara de germinación hidropónica mediante IoT, broker Mosquitto y placas SBC en invernaderos de la provincia de Misiones (13373). Marcelo Marinelli, Myriam Kurtz, Rubén R. Urquijo, Guillermo Wurm.....	609

Análisis de una plataforma de simulación para Cloud Computing. Un caso de estudio (13402). Tomás Rosales, Julián Spinelli, Marcos Di Nardo, Román Bond, Daniel Rosatto, Diego Encinas, Fernando Romero .....	619
Administración del Tiempo Ocioso Mediante Slack Stealing en FreeRTOS (13416). Francisco E. Páez, José M. Urriza, Javier D. Orozco .....	629
Análisis Comparativo de Implementaciones HLS de Filtro Sobel en SoC (13440). Roberto Millon, Enzo Rucci, Emmanuel Frati.....	639
WIE – Innovación en Educación en Informática .....	649
Compilador para traducir ubicaciones geográficas en instrucciones atómicas en una aplicación de aprendizaje ubicuo de programación (13336). Denis Acosta, Margarita Alvarez, Elena Durán .....	650
Estrategias para trabajar competencias en una asignatura masiva de Informática. El caso de Seminario de Python (13378). Claudia Banchoff, Viviana Harari, Sofía Martin .....	660
Actividades de Investigación Científica con docentes y alumnos de grado en Informática y Ciencias de la Computación (13382). Germán Bianchini, Paola Caymes-Scutari, Patricia Ontiveros, Carina Rotella, Sandra Tagarelli, Sergio Salinas, Mariela Galdamez, Pamela Chirino, Karvin Díaz, Alejo Ponce de León, Renzo Suarez .....	670
Experiencias de articulación para formar emprendedores en los ciclos lectivos 2019 y 2020 (13406). Sonia I. Mariño y Viviana R. Bercheñi .....	680
Estrategias de evaluación formativa en la enseñanza de programación en modalidad remota de emergencia (13422). Gladys Dapozo, Cristina Greiner, Raquel Petris, María Fernanda Piragine, Ana María Company, María Cecilia Espíndola .....	690
Relevamiento de demandas formativas e innovación educativa en el nivel medio y superior en el DIIT- UNLaM (13430). Santiago Igarza, Bettina Donadello, Mariana Cipic, Nicolás Kotliar .....	700
Didáctica de las Ciencias de la Computación: Experiencias y Percepciones de Docentes de Educación Primaria (13431). Francisco Bavera, Marcela Daniele, Flavia Buffarini, Teresa Quintero, Cecilia De Dominici .....	708
WSI – Seguridad Informática.....	718
Riesgos/vulnerabilidades de una red informática en un Laboratorio EDI (13385). Fabián A. Gibellini, Roberto Muñoz, Analía L. Ruhl, Juliana Notreni, Milagros N. Zea Cárdenas, Ignacio Sánchez Balzaretto, Cecilia Sanchez .....	719
Exploración de correlaciones internas de los parámetros temporales generados en dinámicas de tecleo (13394). Nahuel González, Germán M. Concilio, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperue .....	726
Implementación de Blockchain para aseguramiento de evidencia digital en entornos Forensic Readiness (13418). Javier Díaz, Mónica D. Tugnarelli, Lucas Barboza, Mauro F. Fornaroli, Facundo N. Miño .....	736
Track de Gobierno Digital y Ciudades Inteligentes .....	745
Prototipo para la gestión de agua de lluvia para la descarga de sanitarios (13347). Leandro Marín Crespi, Cristian Schiffino, Iris Inés Sattolo, Marisa Daniela Panizzi .....	746

Marco de vinculación de datos abiertos aplicado al contexto de datos medioambientales (13388). Juan Santiago Preisegger, Alejandro Greco , Ariel Pasini, Marcos Boracchia, Patricia Pesado.....	756
Validando las dos primeras etapas de metodo_SCGE en un organismo gubernamental (13398). Marcelo Castro, Claudia Fabiana Pons, Rocío Andrea Rodríguez.....	767
Improving a Low-Cost Surveillance System (13401). Carlos Sebastián Castañeda, María José Abásolo.....	777
Desarrollo de un prototipo para estación de monitoreo de un dispositivo híbrido de termotanques solar y a gas natural instalado en la Sede Atlántica de la Universidad Nacional de Río Negro (13429). Javier Valsecchi, Guillermo Malpeli, Juan Cruz Martinez Luquez, Héctor Luis Vivas.....	787
Innovación en el Sector Público para Ciudades Inteligentes Sostenibles (13438). Rocío Muñoz, Ariel Pasini, Patricia Pesado.....	797
New technologies for innovation in public services: A review of case studies (13444). Ana Inés Dennehy, Elsa Estévez, and Pablo R. Fillotrani.....	807

# Autoridades UNLaM

Rector - Prof. Dr. Daniel E. Martínez

Vicerrector - Prof. Dr. Víctor René Nicoletti

Vicerrector Interino - Dr. Fernando L. Acosta

Decano DIIT - Mg. Jorge Eterovic

Vicedecano DIIT - Mg. Gabriel Blanco

Secretario Académico DIIT - Ing. Santiago Igarza

Secretaria de Investigaciones DIIT - Dra. Bettina Donadello

Secretaria Administrativa y de Extensión DIIT - Dra. Mariana Cipicic

Coordinador de Ing. Informática DIIT - Ing. Claudio D'Amico

Coordinadora Tecnicatura Web - Mg. Ing. Cintia V. Gioia

# AUTORIDADES REDUNCI

Coordinador Titular - Patricia Pesado (UNLP)  
Coordinador Alterno - Marcelo Estayno (UNLZ)

## JUNTA DIRECTIVA

Marcela Printista (UNSL) 2018-2020  
Fernanda Carmona (UNdeC) 2018-2020  
Mónica Tugnarelli (UNER) 2018-2020  
Hugo Padovani (UM) 2018-2020  
Jorge Eterovic (UNLaM) 2019-2021  
Claudio Aciti (UNCPBA) 2019-2021  
Marcelo Arroyo (UNRC) 2019-2021  
Marisa Panizzi (UK) 2019-2021

# COMITÉ ORGANIZADOR

Alicia Mon (Presidente)

Miriam Barone (Coordinadora)

Nora Gigante (Coordinadora Comunicación)

Leslie Salerno

Yamila Tesolin

Gladys Kaplan

Renata Guatelli

Horacio Del Giorgio

# ESCUELA INTERNACIONAL DE INFORMÁTICA

Directora: Verónica Aubín

# COMITÉ ACADÉMICO

Diego Garbervetsky (UBA - Cs. Exactas)  
Adriana Echeverria (UBA - Ingeniería)  
Patricia Pesado (UN La Plata)  
Sonia Rueda (UN Sur)  
Fabiana Piccoli (UN San Luis)  
Claudio Aciti (UNCPBA)  
Claudio Vaucheret (UN Comahue)  
Jorge Eterovic (UNLaM)  
Hugo Alfonso (UN La Pampa)  
Marcelo Estayno (UN Lomas de Zamora)  
Guillermo Feierherd (UN Tierra del Fuego)  
Gustavo Gil (UN Salta)  
Marta Lasso (UN Patagonia Austral)  
Nelson Rodriguez (UN SanJuan)  
Patricia Vivas (UADER)  
Carlos Buckle (UN Patagonia SJB)  
Mónica Tugnarelli (UN Entre Ríos)  
Gladys Dapozo (UN Nordeste)  
Kantor Raul (UN Rosario)  
Horacio Kuna (UN Misiones)  
Claudia Russo (UNNOBA)  
Fernanda Carmona (UN Chilecito)  
Diego Azcurra (UN Lanús)  
Elena Duran (UN Santiago del Estero)  
Alejandro Arroyo Arzubi (Esc. Sup. Ejército)  
Horacio Loyarte (UN Litoral)  
Marcelo Arroyo (UN Rio IV)  
Daniel Fridlender (UN Córdoba)  
Analía Herrera Cognetta (UN Jujuy)  
Luis Vivas (UN Rio Negro)  
Laura Prato (UN Villa María)

Wálter Panessi (UN Lujan)  
Maria Valeria Poliche (UN Catamarca)  
Eduardo Campazzo (UN La Rioja)  
Alejandro Oliveros (UN Tres de Febrero)  
Griselda María Luccioni (UN Tucumán)  
Martín Morales (UNAJ)  
Patricia Zachman (UN Chaco Austral)  
Antonio Foti (UN del Oeste)  
Carlos García Garino (UN de Cuyo)  
Julio Cesar Doumecq (UN de Mar del Plata)  
Hugo Padovani (UM)  
Marcelo De Vincenzi (UAI)  
Alberto Guerci (UB)  
Marisa Panizzi (U Kennedy)  
Juan Bournissen (U Adventista del Plata)  
Jorge Finocchietto (UCAECE)  
Adriana Alvarez (UP)  
Sebastián Grieco (UCA Rosario)  
Marcelo Zanitti (U Salvador)  
Rosa Giménez (U Aconcagua)  
Carlos Beyersdorf (U Gastón Dachary)  
Ariadna Guglianone (UCEMA)  
Juan Pablo Cosentino (U Austral)  
Liliana Rathmann (U Atlántida Argentina)  
Rodolfo Bertone (UCA La Plata)  
Alicia Mon (ITBA)  
Fernando Pinciroli (U Champagnat)

# COMITÉ CIENTÍFICO

Claudio Aciti (Argentina) - Presidente

Jorge Eterovic (Argentina)

María José Abásolo (Argentina)

Hugo Alfonso (Argentina)

Jorge Ardenghi (Argentina)

Marcelo Arroyo (Argentina)

Hernán Astudillo (Chile)

Sandra Baldasarri (España)

Javier Balladini (Argentina)

Luis Barbosa (Portugal)

Rodolfo Bertone (Argentina)

Oscar Bría (Argentina)

Nieves Brisaboa (España)

Carlos Buckle (Argentina)

Alberto Cañas (EE.UU)

Ana Casali (Argentina)

Silvia Castro (Argentina)

Alejandra Cechich (Argentina)

Edgar Chavez (México)

Carlos Coello Coello (México)

Uriel Cuckierman (Argentina)

Armando De Giusti (Argentina)

Laura De Giusti (Argentina)

Marcelo De Vincenzi (Argentina)

Claudia Deco (Argentina)

Beatriz Depetris (Argentina)

Javier Diaz (Argentina)

Juerguen Dix (Alemania)

Ramón Doallo (España)

Domingo Docampo (España)

Dujmovic Jozo (USA)

Marcelo Estayno (Argentina)  
Elsa Estevez (Argentina)  
Marcelo Falappa (Argentina)  
Pablo Fillottrani (Argentina)  
Jorge Finocchieto (Argentina)  
Fрати Emmanuel (Argentina)  
Daniel Fridlender (Argentina)  
Carlos García Garino (Argentina)  
Javier García Villalba (España)  
Marcela Género (España)  
Sergio Gomez (Argentina)  
Eduard Gröller (Austria)  
Roberto Guerrero (Argentina)  
Jorge Ierache (Argentina)  
Tomasz Janowski (Naciones Unidas)  
Kuna Horacio (Argentina)  
Laura Lanzarini (Argentina)  
Guillermo Leguizamón (Argentina)  
Fernando Lopez Gil (España)  
Ronald Prescott Loui (EEUU)  
Emilio Luque (España)  
Cristina Madoz (Argentina)  
Alejandra Malberti (Argentina)  
Maria Malbran (Argentina)  
Cristina Manresa Yee (España)  
Marco Javier (España)  
Mauricio Marín (Chile)  
Ramón Mas Sansó (España)  
Orlando Micolini (Argentina)  
Alicia Mon (Argentina)  
Regina Motz (Uruguay)  
Marcelo Naiouf (Argentina)  
Antonio Navarro Martín (España)

José Angel Olivas Varela (España)  
Ariel Pasini (Argentina)  
Patricia Pesado (Argentina)  
María Fabiana Piccoli (Argentina)  
Marcela Printista (Argentina)  
Álvaro Pardo (Uruguay)  
Mario Piattini (España)  
Enrico Puppo (Italia)  
Hugo Ramón (Argentina)  
Dolores Rexachs (España)  
Nora Reyes (Argentina)  
Rosabel Roig Vila (España)  
Gustavo Rossi (Argentina)  
Paolo Rosso (España)  
Sonia Rueda (Argentina)  
Francisco Ruiz (España)  
Claudia Russo (Argentina)  
Carolina Salto (Argentina)  
Cecilia Sanz (Argentina)  
Guillermo Simari (Argentina)  
Osvaldo Sposito (Argentina)  
Ralf Steinmetz (Alemania)  
Remo Suppi (España)  
Liane Tarouco (Brasil)  
Francisco Tirado (España)  
Luiz Velho (Brasil)  
Marcelo Vénere (Argentina)  
Eduardo Vendrell (España)  
Horacio Villagarcía (Argentina)  
Dante Zanarini (Argentina)

# Workshops

# Workshop Agentes y Sistemas Inteligentes

Coordinadores

Guillermo Leguizamón (UNSL)  
Carolina Salto (UNLPam)  
Daniel Fridlender (UNC)

# Modeling Human Decision Making with Defeasible Logic Programming

Maximiliano Sapino<sup>1,2,3</sup>, Edgardo Ferretti<sup>1,2</sup>, Luciana Mariñelarena Dondena<sup>1</sup>,  
and Marcelo Errecalde<sup>1,2</sup>

<sup>1</sup> Universidad Nacional de San Luis (UNSL), San Luis - Argentina

<sup>2</sup> Laboratorio de Investigación y Desarrollo en Inteligencia Computacional, UNSL

<sup>3</sup> Ph.D. Fellow at the National Council of Scientific and Technical Research  
(CONICET) - Argentina

e-mails: {ferretti,merreca}@unsl.edu.ar

**Abstract.** Decision making is an active research topic in several disciplines and it can be studied in as many ways as different research areas face the problems in this field of study. In Computer Science, decision-making problems have been mainly tackled from the research field of Artificial Intelligence; and *Argumentation* has contributed with its unique strengths. In this work, following a psychological perspective we show the adequacy of Defeasible Logic Programming to model the Dictator Game by emulating the answers contained in a survey we conducted. The Dictator Game is a well-known problem belonging to the field of experimental economic studies related to human decision making. Moreover, the obtained model is simpler than other leading approaches in the area.

**Keywords:** Human Decision Making, Defeasible Logic Programming, Experimental Economic Studies, Dictator Game

## 1 Introduction

The use of argumentative reasoning for decision making is an active research field in Artificial Intelligence (AI) [1, 4, 5, 7, 9–11, 13, 14, 17, 19, 20, 24, 25]. The state of the art is vast, and the different works related to argumentation-based decision making can be organized according to different dimensions, namely: the decision problem tackled (decision under uncertainty [1, 20], multi-criteria decision [1, 5, 11, 25], single-agent [10, 11, 13, 14], multi-agent [4, 20], etc.), how eligible alternatives are conceived (marketing approach [10, 11, 13, 14], goal-based [1, 17, 20], etc.) the underlying argumentation framework (abstract [1, 7, 9], dynamic [14], concrete [5, 10, 11, 13, 25], etc.) whether decision behavior is formalized with respect to a classical decision approach [1, 10, 11, 14] or not [4, 5, 9, 25], whether the decision process needs a particular agent architecture to be embedded in [19, 24] or not [10, 11, 13, 14], among others.

As stated in [21], it is well-known that in AI all the developed systems that concern thinking and acting, can be broadly classified depending on whether they follow a human-centered approach or a rationalist approach. The former

approach measures the systems success in terms of the fidelity to *human* performance, while the latter involves measuring the systems performance against the *ideal* one, so-called *rationality*.

Several disciplines have contributed ideas and techniques to AI, and in particular, regarding decision making, we can highlight Philosophy, Sociology, Psychology and Economics. As expected, in psychological research of individuals decision making there exist several research programs based on Simon's criticism of mainstream economic models of perfect rationality (e.g., see [23]). Moreover, as stated in [6], the decision making research program in psychology was dominated by Tversky and Kahneman's approach (e.g., see [26]) empirically testing Simon's suggestions and showing that they were correct.

In this context, [4] is an interesting paper arising from the Computer Science field, where different studies from experimental economic are modeled using an argumentative framework to reproduce the qualitative decisions that humans exhibited in the studies conducted. In particular two well-known games were studied, the Dictator game [15] and the Ultimatum game [18] and results were compared with those reported by humans with different cultural background. Argument schemes were used as the underlying argumentation approach and they were instantiated with Action-based Alternating Transition Systems (AATS) [27]. Instantiating the argument scheme used in [2] and its associated critical questions, allows to obtain the set of conflicting arguments. Once this set has been produced, in order to evaluate the arguments acceptability, they are organized in a Value-based Argumentation Framework (VAF) [3].

A VAF is an extension of the standard Argumentation Framework (AF) [8]. VAFs extend AFs in that each argument in the graph is associated with the value promoted by that argument. Whereas in an AF attacks always succeed, in a VAF they succeed only if the value associated with the attacker is ranked by the audience evaluating the VAF equal to, or higher than, the argument it attacks. Unsuccessful attacks are removed, and then the resulting framework is evaluated as a standard AF. The VAF thus accounts for elements of subjectivity in that the arguments that are acceptable are dependant upon the audience's ranking of the values involved in the scenario.

Another interesting proposal, where non-rational decision making is discussed even though it is not the primary focus of the paper, is the work presented in [14]—where the proposed argumentation-based model of decision-making generalizes the classical maximum-expected utility model using a Dynamic Argumentation Framework. Apart from this, the working methodology presented in [14] also paves the way for non-rational behavior to be taken into account in decision making by using argumentation. In fact, the methodology of [14] is similar to that of [12] in the way alternatives are compared in a pairwise manner, but where the use of decision rules is needed when decisions to be made must be rational.

Therefore, by considering the approach of [12], in the paper at hand we model human decision making by using *Defeasible Logic Programming* (DeLP) [16], a formalism that combines results of Logic Programming and Defeasible Argu-

mentation. It is worth mentioning, that in [12], the focus of the work is the combination of autonomous navigation with high-level reasoning and decision making is based on a greedy policy, by using DeLP as well. Besides this real world domain, DeLP has been used to solve other real world problems (cf. [22]). Moreover, DeLP has been successfully used to model rational decision making [10] and showing its flexibility to model human decision making as well, makes richer its application field. In particular, in this work we will tackle the Dictator game by modeling the answers contained in a survey conducted by us.

The rest of the paper is organized as follows. Section 2 introduces DeLP, the argumentation formalism used to model the Dictador game. Then, Sect. 3, presents the formulation of this game based on DeLP. Finally, Sect. 4 discuss the model proposed considering other related works and draws the conclusions.

## 2 Defeasible Logic Programming in a Nutshell

As mentioned in the introductory section, DeLP [16] is the argumentation formalism that it will be used for knowledge representation, reasoning and finally model human decision making. Besides, the literal-based criterion used to decide between conflicting arguments is also presented, since it is not the default criterion provided by the formalism.

A DeLP-program  $\mathcal{P}$  is denoted  $(\Pi, \Delta)$ , where the set  $\Pi$  containing indisputable knowleged (facts and strict rules) is distinguished from  $\Delta$  containing defeasible rules. *Facts* are ground literals representing atomic information or the negation of atomic information, by using *strong* negation (“ $\sim$ ”). *Strict rules* are denoted  $L_0 \leftarrow L_1, \dots, L_n$  and represent firm information, whereas *defeasible rules* are denoted  $L_0 \prec L_1, \dots, L_n$  and represent tentative information. In both cases, the *head*  $L_0$  is a literal and the *body*  $\{L_i\}_{i>0}$  is a set of literals. Strict and defeasible rules are ground, however, following the usual convention, some examples will use “schematic rules” with variables. It is worth noticing that strong negation is allowed in the head of program rules, and hence may be used to represent contradictory knowledge.

In DeLP, to deal with contradictory and dynamic information, *arguments* for conflicting pieces of information are built and then compared to decide which one *prevails*. An *argument* for a literal  $L$ , denoted  $\langle \mathcal{A}, L \rangle$ , is a minimal set of defeasible rules  $\mathcal{A} \subseteq \Delta$ , such that  $\mathcal{A} \cup \Pi$  is non-contradictory and there is a derivation for  $L$  from  $\mathcal{A} \cup \Pi$ . To establish if  $\langle \mathcal{A}, L \rangle$  is a non-defeated argument, *argument rebuttals* or *counter-arguments* that could be *defeaters* for  $\langle \mathcal{A}, L \rangle$  are considered, i.e. counter-arguments that by some criterion are preferred to  $\langle \mathcal{A}, L \rangle$ . Since counter-arguments are arguments, defeaters for them may exist, and defeaters for these defeaters, and so on. Thus, a sequence of arguments called *argumentation line* is constructed, where each argument defeats its predecessor in the line (for a detailed explanation of this dialectical process see [16]). The prevailing argument provides a warrant for the information it supports. A literal  $L$  is *warranted* from  $(\Pi, \Delta)$  if a non-defeated argument  $\mathcal{A}$  supporting  $L$  exists. Given a query  $Q$  there are four possible answers: YES, if  $Q$  is warranted; NO, if

the complement of  $Q$  is warranted; UNDECIDED, if neither  $Q$  nor its complement are warranted; and UNKNOWN, if  $Q$  is not in the language of the program.

In DeLP, *generalized specificity* is used as default criterion to compare conflicting arguments, but an advantageous feature of DeLP is that the comparison criterion among arguments can be replaced in a modular way. Therefore, in our proposal, we use an appropriate literal-based criterion—that was originally proposed in [10]. In a program  $\mathcal{P}$ , a subset of literals  $\text{comp-lits}(\mathcal{P}) \subseteq \Pi$  called *comparison literals* will be distinguished. This set of comparison literals will be used by the comparison criterion defined below.

**Definition 1 (L-order).** *Let  $\mathcal{P}$  be a DeLP-program and  $\text{comp-lits}(\mathcal{P})$  be the set of comparison literals in  $\mathcal{P}$ . An L-order over  $\mathcal{P}$  is a partial order over the elements of  $\text{comp-lits}(\mathcal{P})$ .*

An L-order must be provided as a set of facts within the program. These facts are written as  $L_1 > L_2$ , stating that a literal  $L_1$  is preferred to a literal  $L_2$ , and they will be used to decide when an argument is better than another. Based on a given L-order, the following argument comparison criterion can be defined.

**Definition 2 (Literal-based comparison criterion).** *Let  $\mathcal{P} = (\Pi, \Delta)$  be a DeLP-program and let “ $>$ ” be an L-order over  $\mathcal{P}$ . Given two argument structures  $\langle \mathcal{A}_1, h_1 \rangle$  and  $\langle \mathcal{A}_2, h_2 \rangle$ , the argument  $\langle \mathcal{A}_1, h_1 \rangle$  will be preferred over  $\langle \mathcal{A}_2, h_2 \rangle$  iff:*

1. *there are two literals  $L_1$  and  $L_2$  such that  $L_1 \in^* \mathcal{A}_1$ ,  $L_2 \in^* \mathcal{A}_2$ ,  $L_1 > L_2$ , and*
2. *there are no literals  $L'_1$  and  $L'_2$  such that  $L'_1 \in^* \mathcal{A}_1$ ,  $L'_2 \in^* \mathcal{A}_2$ , and  $L'_2 > L'_1$ .*

**Notation:**  $L \in^* \mathcal{A}$  iff there exists a defeasible rule  $(L_0 \prec L_1, L_2, \dots, L_n)$  in  $\mathcal{A}$  and  $L = L_i$  for some  $i$  ( $0 \leq i \leq n$ ).

### 3 Modeling the Dictator Game

In this section we show how DeLP can be used to model the Dictator Game [15]. We begin by considering the problem formulation. Following the problem statement from [4], we will consider the same limited number of options that comprise the set of alternatives (actions)  $\mathbf{A}$  and we assume 1000 units of money to be divided. The set  $\mathbf{A}$  thus comprises the following five actions corresponding to different divisions of the money, namely:  $\mathbf{a}_1 = \text{give}(70\%)$ ,  $\mathbf{a}_2 = \text{give}(100\%)$ ,  $\mathbf{a}_3 = \text{give}(50\%)$ ,  $\mathbf{a}_4 = \text{give}(0\%)$  and  $\mathbf{a}_5 = \text{give}(30\%)$ . The dictator starts having the whole money and the “motivations” to share it that we have considered are mentioned below:

**Money:** Most obvious is money’s value. This is what the economic man is supposed to maximize. Given that we need to recognize that the other player having money may be considered positively by the dictator, we need to distinguish money for the dictator himself from money for the other.

**Giving:** It can be held that giving a gift is a source of pleasure, and this is what motivates the dictator to share.

**Image:** Another consideration is the desire not to appear mean before the experimenter that motivates sharing. It could even be that one does not want to appear mean to oneself.

**Equality:** Equality, as defined by an equal distribution, characterizes a sense of fairness.

In our model, these motivations are used as the criteria to compare the alternatives among each other. Some of these motivations establish well-defined preference orderings among the alternatives; for example, the money the dictator has for himself that will be represented by the comparison literal  $ms$ , clearly produces the following ordering among the alternatives:  $a_4 \prec a_5 \prec a_3 \prec a_1 \prec a_2$ —where  $a_i \prec a_j$  denotes that action  $a_i$  is preferred to  $a_j$ . Conversely, “the money for the other” motivation ( $mo$ ) generates a mirror-like ordering of the alternatives:  $a_2 \prec a_1 \prec a_3 \prec a_5 \prec a_4$ .

The remaining motivations can be considered more subjectives and different orderings can be obtained depending on the individuals’ personality and cultural background. In this context, one possible ordering of all the alternatives according to these motivations is the one presented in Table 1, as facts that belong to set  $\Pi$  of the DeLP-program that will be used to model the game. It is worth noticing that this table contains more information than the ordering itself. The particular ordering we refer is represented as black-colored comparison literals (factual information) while the references to arguments names in other colors will be discussed later. We can see for example, in the first two columns of the table—in a pairwise comparison manner—the two above-mentioned orderings regarding  $ms$  and  $mo$  motivations.

Having as a guideline the working methodology of previous works [10, 12, 14] which have used a literal-based comparison criterion for arguments, we have built the set  $\Delta$  of our DeLP-program as shown below in Fig. 1(a). Besides, the L-order of the comparison literals is presented in Fig. 1(b). This L-order states that the preferences on the motivations used as criteria are the following:  $giv \prec mo \prec ms \prec im \prec eq$ —and it was obtained from the survey we conducted to carried out this reasearch.

Our survey is composed by 276 samples, i.e. 276 different people that played the Dictador game. They are mainly students of the National University of San Luis but also 2.9% of the people are employed in the private industry and 3.3% work in the public sector. Considering their origin, 84% were born in the capital of San Luis and the remaining 16% were born in 19 different cities from San Luis province and other provinces from Argentina. All the people live in San Luis at present. Ages ranged from 17 to 54 years old, with an approximate mean of 22 years old. Regarding educational level, 96% of the participants have completed secondary education and only 4% have completed a tertiary level. When gender is considered, 73% of the samples correspond to female individuals and 27% to male individuals. Finally, it is worth mentioning that 93% of the people do not have children, while the remaining 7% do.

The answers obtained from the survey reported that 66.67% of people chose  $\mathbf{a}_3 = give(50\%)$ , 10.87% chose  $\mathbf{a}_2 = give(100\%)$ , 9.78% chose  $\mathbf{a}_4 = give(0\%)$ ,

**Table 1.** Factual information stating how alternatives are deemed considering the comparison literals.

$ms(a_4, a_1)$	$\mathcal{A}_1$ $\mathcal{A}_2$	$mo(a_2, a_1)$	$\mathcal{A}_{21}$ $\mathcal{A}_{22}$	$giv(a_3, a_4)$	$\mathcal{A}_{41}$ $\mathcal{A}_{42}$	$im(a_1, a_4)$	$\mathcal{A}_{61}$ $\mathcal{A}_{62}$	$eq(a_3, a_1)$	$\mathcal{A}_{81}$ $\mathcal{A}_{82}$
$ms(a_4, a_2)$	$\mathcal{A}_3$ $\mathcal{A}_4$	$mo(a_2, a_3)$	$\mathcal{A}_{23}$ $\mathcal{A}_{24}$	$giv(a_3, a_5)$	$\mathcal{A}_{43}$ $\mathcal{A}_{44}$	$im(a_2, a_4)$	$\mathcal{A}_{63}$ $\mathcal{A}_{64}$	$eq(a_3, a_2)$	$\mathcal{A}_{83}$ $\mathcal{A}_{84}$
$ms(a_4, a_3)$	$\mathcal{A}_5$ $\mathcal{A}_6$	$mo(a_2, a_4)$	$\mathcal{A}_{25}$ $\mathcal{A}_{26}$	$giv(a_3, a_2)$	$\mathcal{A}_{45}$ $\mathcal{A}_{46}$	$im(a_3, a_4)$	$\mathcal{A}_{65}$ $\mathcal{A}_{66}$	$eq(a_3, a_4)$	$\mathcal{A}_{85}$ $\mathcal{A}_{86}$
$ms(a_4, a_5)$	$\mathcal{A}_7$ $\mathcal{A}_8$	$mo(a_2, a_5)$	$\mathcal{A}_{27}$ $\mathcal{A}_{28}$	$giv(a_3, a_1)$	$\mathcal{A}_{47}$ $\mathcal{A}_{48}$	$im(a_5, a_4)$	$\mathcal{A}_{67}$ $\mathcal{A}_{68}$	$eq(a_3, a_5)$	$\mathcal{A}_{87}$ $\mathcal{A}_{88}$
$ms(a_5, a_3)$	$\mathcal{A}_9$ $\mathcal{A}_{10}$	$mo(a_1, a_3)$	$\mathcal{A}_{29}$ $\mathcal{A}_{30}$	$giv(a_5, a_4)$	$\mathcal{A}_{49}$ $\mathcal{A}_{50}$	$im(a_2, a_1)$	$\mathcal{A}_{69}$ $\mathcal{A}_{70}$	$eq(a_5, a_4)$	$\mathcal{A}_{89}$ $\mathcal{A}_{90}$
$ms(a_5, a_2)$	$\mathcal{A}_{11}$ $\mathcal{A}_{12}$	$mo(a_1, a_4)$	$\mathcal{A}_{31}$ $\mathcal{A}_{32}$	$giv(a_5, a_2)$	$\mathcal{A}_{51}$ $\mathcal{A}_{52}$	$im(a_2, a_3)$	$\mathcal{A}_{71}$ $\mathcal{A}_{72}$	$eq(a_5, a_2)$	$\mathcal{A}_{91}$ $\mathcal{A}_{92}$
$ms(a_5, a_1)$	$\mathcal{A}_{13}$ $\mathcal{A}_{14}$	$mo(a_1, a_5)$	$\mathcal{A}_{33}$ $\mathcal{A}_{34}$	$giv(a_5, a_1)$	$\mathcal{A}_{53}$ $\mathcal{A}_{54}$	$im(a_2, a_5)$	$\mathcal{A}_{73}$ $\mathcal{A}_{74}$	$eq(a_1, a_2)$	$\mathcal{A}_{93}$ $\mathcal{A}_{94}$
$ms(a_3, a_1)$	$\mathcal{A}_{15}$ $\mathcal{A}_{16}$	$mo(a_3, a_4)$	$\mathcal{A}_{35}$ $\mathcal{A}_{36}$	$giv(a_4, a_1)$	$\mathcal{A}_{55}$ $\mathcal{A}_{56}$	$im(a_1, a_3)$	$\mathcal{A}_{75}$ $\mathcal{A}_{76}$		
$ms(a_3, a_2)$	$\mathcal{A}_{17}$ $\mathcal{A}_{18}$	$mo(a_3, a_5)$	$\mathcal{A}_{37}$ $\mathcal{A}_{38}$	$giv(a_4, a_2)$	$\mathcal{A}_{57}$ $\mathcal{A}_{58}$	$im(a_1, a_5)$	$\mathcal{A}_{77}$ $\mathcal{A}_{78}$		
$ms(a_1, a_2)$	$\mathcal{A}_{19}$ $\mathcal{A}_{20}$	$mo(a_5, a_4)$	$\mathcal{A}_{39}$ $\mathcal{A}_{40}$	$giv(a_1, a_2)$	$\mathcal{A}_{59}$ $\mathcal{A}_{60}$	$im(a_3, a_5)$	$\mathcal{A}_{79}$ $\mathcal{A}_{80}$		

9.42% chose  $\mathbf{a}_5 = give(30\%)$  and 3.26% chose  $\mathbf{a}_1 = give(70\%)$ . It is beyond the scope of our present study to analyze the possible reasons of these behaviors, but as mentioned above, this ordering of the alternatives resulted in the L-order presented in Fig. 1(b). This is due to the fact that having this L-order, our DeLP-program has the same choice behavior than the one observed in our survey. As expected, different answer would produce different L-orders so that the DeLP-program can perform a choice behavior conformant to the survey.

As mentioned above, Table 1 contains factual information where actions are compared in a pairwise manner considering the motivations to share the money that the dictator has. These facts were stated considering the rules of the game and our understanding of some elements of subjectivity—like for instance, if giving more than half of the money is a source of pleasure for the dictator. In our view it is not, and that is why facts like  $giv(a_2, a_1)$ ,  $giv(a_2, a_3)$ ,  $giv(a_2, a_4)$ ,  $giv(a_2, a_5)$ ,  $giv(a_1, a_3)$ ,  $giv(a_1, a_4)$ ,  $giv(a_1, a_5)$  and  $giv(a_5, a_4)$ , were not included in  $\Pi$ . In this way, taking into account the factual information present in  $\Pi$  about the comparison literal  $giv$  where alternative  $\mathbf{a}_3$  is deemed better than the other ones when compared among each other, provided us a firm reason to set this comparison literal in the first place of the L-order.

Then  $mo$  was placed second in the L-order given its direct relation with the fact that  $\mathbf{a}_2$  was the second action chosen in the survey. Following a similar pattern of reasoning  $ms$  was placed third given its relation with  $\mathbf{a}_4$ . First and second columns of Table 1 show the facts encoding the alternatives orderings according

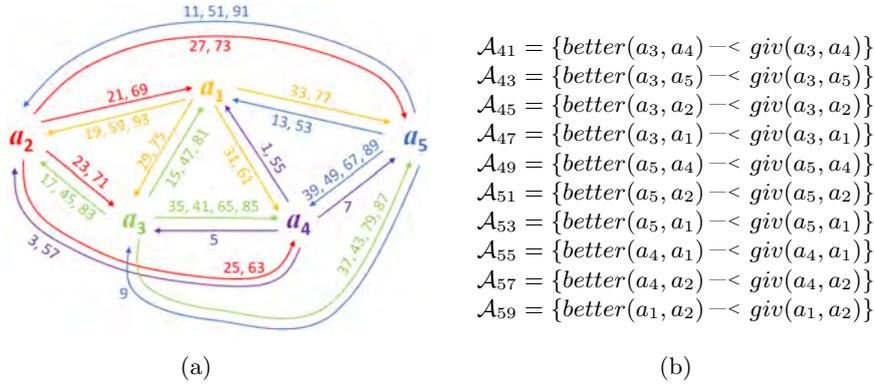
<p>(1) <math>better(X, Y) \prec ms(X, Y)</math>  (2) <math>\sim better(Y, X) \prec ms(X, Y)</math>  (3) <math>better(X, Y) \prec mo(X, Y)</math>  (4) <math>\sim better(Y, X) \prec mo(X, Y)</math>  (5) <math>better(X, Y) \prec giv(X, Y)</math>  (6) <math>\sim better(Y, X) \prec giv(X, Y)</math>  (7) <math>better(X, Y) \prec im(X, Y)</math>  (8) <math>\sim better(Y, X) \prec im(X, Y)</math>  (9) <math>better(X, Y) \prec eq(X, Y)</math>  (10) <math>\sim better(Y, X) \prec eq(X, Y)</math></p> <p>(11) <math>choose(X) \prec better(X, Y)</math>  (12) <math>\sim choose(X) \prec better(Y, X)</math></p>	<p><math>giv(Z, W) &gt; ms(W, Z)</math>  <math>giv(Z, W) &gt; mo(W, Z)</math>  <math>giv(Z, W) &gt; im(W, Z)</math>  <math>giv(Z, W) &gt; eq(W, Z)</math>  <math>mo(Z, W) &gt; ms(W, Z)</math>  <math>mo(Z, W) &gt; im(W, Z)</math>  <math>mo(Z, W) &gt; eq(W, Z)</math>  <math>ms(Z, W) &gt; im(W, Z)</math>  <math>ms(Z, W) &gt; eq(W, Z)</math>  <math>im(Z, W) &gt; eq(W, Z)</math></p>
(a)	(b)

**Fig. 1.** (a) Set  $\Delta$  containing the rules to compare the alternatives among each other and choosing the final alternative. (b) L-order over the comparison literals.

to these criteria, where as aforesaid, they are objective. Finally, comparison literals  $im$  and  $eq$  which are also subjectives were placed fourth and fifth in the L-order, respectively— $im$  was placed before than  $eq$ , given that all the actions can be compared against each other with this criterion and with  $eq$  cannot.

In this way, with an L-order defined, based on the factual information presented in Table 1 and through the application of the defeasible rules in Fig. 1(a) aiming at comparing the alternatives, 94 arguments can be built supporting whether an action  $\mathbf{a}_i$  is better or not than another action  $\mathbf{a}_j$ . For example, in Fig. 2(b), ten arguments are shown and they were built using rule (5) from Fig. 1(a) together with the facts of third column of Table 1. The names of these arguments also appear in the top-right corner of each cell of this column indicating that they can be built from the factual information included in the cell. The same notation is used for all the cells at the table, and due to space constraints it is not possible to show all these arguments in the paper at hand.

It is worth noting that the colors each argument name has (on the top-right corner of the cell), are related to the attack graph depicted in Fig. 2(a). This so-called attack graph is not the usual attack graph from abstract argumentation frameworks; it rather is a graphical representation that we found convenient to summarize how the different arguments generated support which alternative is better than another one depending on the criterion considered. As it can be observed, the nodes correspond to the alternatives and they are differently colored as well as their outcoming arcs. The labels on the arcs represent the number of the argument by which an alternative is supported better than another one. For instance, number 55 labels the arc from  $\mathbf{a}_4$  to  $\mathbf{a}_1$ , and its corresponding argument  $\mathcal{A}_{55}$  can be observed in Fig. 2(b). But also, argument  $\mathcal{A}_1$ —colored in violet in Table 1— supports that action  $\mathbf{a}_4$  is better than  $\mathbf{a}_1$  by means of rule (1) (cf. Fig. 1(a)). Besides, all the arguments names colored in gray which appear on



**Fig. 2.** (a) Attack graph relating the alternatives and the different arguments supporting their comparison. (b) Subset of arguments supporting the comparison of different alternatives based on the comparison literal *giv*.

the bottom-right corner of each cell of Table 1, correspond to the complementary arguments that can be obtained with the fact in the cell together with one of the defeasible rules having the literal  $\sim better(Y, X)$  in their heads—and that were not included in the graphical representation of Fig. 2(a).

So far, all the arguments considered aim at comparing the actions among each other, and the following literals are warranted:  $bt(a_1, a_2)$ ,  $\sim bt(a_2, a_1)$ ,  $bt(a_3, a_1)$ ,  $\sim bt(a_1, a_3)$ ,  $bt(a_3, a_2)$ ,  $\sim bt(a_2, a_3)$ ,  $bt(a_3, a_4)$ ,  $\sim bt(a_4, a_3)$ ,  $bt(a_3, a_5)$ ,  $\sim bt(a_5, a_3)$ ,  $bt(a_4, a_1)$ ,  $\sim bt(a_1, a_4)$ ,  $bt(a_4, a_2)$ ,  $\sim bt(a_2, a_4)$ ,  $bt(a_5, a_1)$ ,  $\sim bt(a_1, a_5)$ ,  $bt(a_5, a_2)$ ,  $\sim bt(a_2, a_5)$ ,  $bt(a_5, a_4)$  and  $\sim bt(a_4, a_5)$ . We have abbreviated literal *better* as *bt* to improve the sentence reading. These literals can be used together with rules (11) and (12) (cf. Fig. 1(a)) to build arguments supporting which action must be chosen.

Given that action  $a_3$  is deemed better than the other alternatives based on the most preferred comparison literal (*giv*), by using rule (12), three arguments supporting  $\sim choose(a_1)$ ,  $\sim choose(a_4)$  and  $\sim choose(a_5)$  can be built. These arguments will act as blocking defeaters for the arguments that can be built supporting  $choose(a_1)$ ,  $choose(a_4)$  and  $choose(a_5)$ , respectively. Because all of them are based on comparison literal *giv*, considering the comparison criterion used (cf. Definition 2) the answer of the DeLP interpreter is UNDECIDED. Conversely, the arguments built –by means of rule (11)– that support conclusion  $choose(a_3)$  cannot be defeated and hence  $choose(a_3)$  is warranted, and the answer of the DeLP interpreter is YES—coinciding with the most chosen action in the survey.

## 4 Conclusions

In this paper we have given an account of argumentation-based decision making in a simple scenario from experimental economics, like the Dictator game. In

particular, DeLP was used to model human choice behavior from a survey we conducted. The obtained program was simple and there was no need to use the decision device called *decision rules*; a device widely used in previous works [10, 11, 14] where rational decision making was pursued rather than human decision making. Besides, regarding the DeLP-program used in [12], the current proposal is also more concise since only two rules of the form  $choose(X) \prec better(X, Y)$  and  $\sim choose(X) \prec better(Y, X)$ , were needed to decide the final action to be chosen. Naturally, if the answers in the survey had been different, we could have had to resort to add more rules for choosing alternatives or even using decision rules despite not being modeling a rational decision behavior.

Regarding the leading work presented in [4], our proposal is simpler since all the modeling is performed within DeLP instead of performing a two-stage process to instantiate an argument scheme with an AATS to obtain the set of conflicting arguments and then building a VAF in order to evaluate the arguments acceptability. With this work, we aimed at showing the flexibility that DeLP has to model different decision making scenarios and not necessarily those strictly related to rational decisions thus making richer its application field.

**Acknowledgments.** This work was partially supported by CONICET and Universidad Nacional de San Luis (PROICO P-31816).

## References

1. Amgoud, L., Prade, H.: Using arguments for making and explaining decisions. *Artificial Intelligence* 173(3-4), 413–436 (2009)
2. Atkinson, K., Bench-Capon, T., McBurney, P.: Computational representation of practical argument. *Synthese* 152(2), 157–206 (2006)
3. Bench-Capon, T.: Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation* 13(3), 429–448 (June 2003)
4. Bench-Capon, T., Atkinson, K., McBurney, P.: Using argumentation to model agent decision making in economic experiments. *Autonomous Agents and Multi-Agent Systems* 25(1), 183–208 (2011)
5. Buron Brarda, M., Tamargo, L.H., García, A.J.: An approach to enhance argument-based multi-criteria decision systems with conditional preferences and explainable answers. *Expert Systems with Applications* 126, 171–186 (2019)
6. Campitelli, G., Gobet, F.: Herbert simon’s decision-making approach: Investigation of cognitive processes in experts. *Review of General Psychology* 14(4), 354 (2010)
7. Carstens, L., Fan, X., Gao, Y., Toni, F.: Graph Structures for Knowledge Representation and Reasoning: 4th International Workshop, GKR 2015, Revised Selected Papers, chap. An Overview of Argumentation Frameworks for Decision Support, pp. 32–49. Springer (2015)
8. Dung, P.M.: On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77(2), 321–358 (1995)
9. Fan, X., Toni, F.: Theory and Applications of Formal Argumentation: Second International Workshop, TAFA 2013, Revised Selected papers, chap. Decision Making with Assumption-Based Argumentation, pp. 127–142. Springer (2014)

10. Ferretti, E., Errecalde, M., García, A., Simari, G.: Decision rules and arguments in defeasible decision making. In: 2nd International Conference on Computational Models of Arguments (COMMA). pp. 171–182. IOS Press (2008)
11. Ferretti, E., Errecalde, M., García, A., Simari, G.: A possibilistic defeasible logic programming approach to argumentation-based decision-making. *Journal of Experimental & Theoretical Artificial Intelligence* 26(4), 519–550 (2014)
12. Ferretti, E., Kiessling, R., Silnik, A., Petrino, R., Errecalde, M.: New Trends in Electrical Engineering Automatic Control, Computing and Communication Sciences, chap. Integrating vision-based motion planning and defeasible decision making for differential-wheeled robots: A case of study with the Khepera 2 robot. LOGOS Verlag, Berlin, Germany (2010)
13. Ferretti, E., Errecalde, M.: Argumentation-Based Proofs of Endearment: Essays in Honor of Guillermo R. Simari on the Occasion of His 70th Birthday, chap. A P-DeLP Instantiation of a Dynamic Argumentation Framework for Decision Making. College Publications, England (2018)
14. Ferretti, E., Tamargo, L.H., García, A.J., Errecalde, M.L., Simari, G.R.: An approach to decision making based on dynamic argumentation systems. *Artificial Intelligence* 242, 107 – 131 (2017)
15. Forsythe, R., Horowitz, J.L., Savin, N.E., Sefton, M.: Fairness in simple bargaining experiments. *Games and Economic Behavior* 6(3), 347–369 (1994)
16. García, A., Simari, G.: Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming* 4(1-2), 95–138 (2004)
17. Matt, P., Toni, F., Vaccari, J.R.: Argumentation in Multi-Agent Systems: 6th International Workshop, ArgMAS 2009. Revised Selected and Invited Papers, chap. Dominant Decisions by Argumentation Agents, pp. 42–59. Springer (2010)
18. Nowak, M.A., Page, K.M., Sigmund, K.: Fairness versus reason in the ultimatum game. *Science* 289, 1773–1775 (2000)
19. de Oliveira Gabriel, V., Panisson, A.R., Bordini, R.H., Adamatti, D.F., Billa, C.Z.: Reasoning in bdi agents using toulmin’s argumentation model. *Theoretical Computer Science* 805, 76–91 (2020)
20. Rieke, R.D., Sillars, M.O., Peterson, T.R.: *Argumentation and Critical Decision Making*. Pearson, 8th edn. (2012)
21. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*. Prentice Hall, 3 edn. (2010)
22. Schlesinger, F., Ferretti, E., Errecalde, M., Aguirre, G.: An Argumentation-based BDI Personal Assistant. In: 23rd International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA-AIE) (2010)
23. Simon, H.A.: A behavioral model of rational choice. *The Quarterly Journal of Economics* 69(1), 99–118 (1955)
24. Sosa Toranzo, C., Ferretti, E., Errecalde, M.: Intention Reconsideration like Dichotomous Choice: an Argumentation-based Approach. In: XLIII Latin American Computer Conference (CLEI). IEEE, Córdoba, Argentina (September 2017)
25. Teze, J.C., Gottifredi, S., García, A.J., Simari, G.R.: An approach to generalizing the handling of preferences in argumentation-based decision-making systems. *Knowledge-Based Systems* 189, 105–112 (2020)
26. Tversky, A., Kahneman, D.: Rational choice and the framing of decisions. *Journal of Business* 59(4) (1986)
27. Wooldridge, M., van der Hoek, W.: On obligations and normative ability: Towards a logical analysis of the social contract. *Journal of Applied Logic* 3, 396–420 (2005)

# A Real Case of Multi-Period Water Distribution Network Design solved by a Hybrid Simulated Annealing

Carlos Bermudez<sup>1</sup>, Hugo Alfonso<sup>1</sup>, Gabriela Minetti<sup>1</sup>, and Carolina Salto<sup>1,2</sup>

<sup>1</sup> Facultad de Ingeniería, Universidad Nacional de La Pampa

<sup>2</sup> CONICET, Argentina

bermudezc,alfonsoh,minettig,saltoc@ing.unlpam.edu.ar

**Abstract.** In this work, we propose an optimization solver based on a hybrid Simulated Annealing (HSA) to optimize the water distribution network design. The problem formulation includes multi-period restrictions with time-varying demand patterns. The HSA search process is affected by the Markov Chain Length (MCL), making modifications in the network design. For that reason, we studied the HSA behaviour by considering static and dynamic methods to compute the MCL. We test the algorithms by using networks reported in the state-of-the-art and also a real and new median size network that arises from a regional necessity. The experimentation suggests the use of a dynamic method, which exhibits the balance between efficacy and efficiency.

**Keywords:** Water Distribution Network Design, Optimization, Simulated Annealing, Markov chain length

## 1 Introduction

The water distribution network design represents a major challenge because the distribution system has many components to be considered (pipes of various sizes for carrying water, valves for controlling the flow, service connections to the individual homes, and distribution reservoirs for storing the water to be fed into the distribution pipes). The solution concerning the layout, design, and operation of the network of pipes should result from efficient planning and management procedures. This problem is known as Water Distribution Network Design (WDND) and requires handling a big number of variables and constraints, in consequence, it is classified as NP-hard [1]. The water distribution network optimization aims to find the optimal pipe diameters in the network for a given layout and demand requirements. The optimal pipe sizes are selected in the final network, by satisfying the conservation of mass and energy and also the constraints.

Currently, the WDND formulations include the extension to a multi-period setting, i.e., time-varying demand patterns, which are more realistic and complex problem formulations. Some works expressed the design problem as a multi-objective optimization problem and applied a multi-objective evolutionary algorithm [2]. A genetic algorithm was developed to solve six small networks [3],

2 Carlos Bermudez, Hugo Alfonso, Gabriela Minetti, and Carolina Salto

which considered the velocity constraint on the water flowing through the pipes. In [4] also regarded this constraint, but the authors used mathematical programming on bigger, closer-to-reality networks. Other metaheuristics were used to tackle more complex WDND formulations [5–7]. In this line, an Iterative Local Search [8] (ILS) considered that every demand node has 24 hrs water demand pattern and included a new constraint related to the limit on the maximal velocity of water through the pipes.

The size of the cities is constantly growing, and our city, General Pico, is not an exception. Consequently, our community needs to optimize an independent WDND of a new neighborhood of 5 km<sup>2</sup>, subject to multi-period demands, hydraulic restrictions, among others. CORPICO is the organization in charge of distributing this essential element and presents a real requirement for an optimization system to determine the best engineering solution in meeting established design criteria while at the same time minimizing capital costs. The objective of our research is to develop an optimal solution based on state-of-the-art optimization techniques, which intends to support the decision making to design, plan, and management of complex water systems.

In a previous work [9], we tackled the WDND considering time-varying demand patterns and the maximum water velocity constraint as formulated in [8]. The optimizer adopted is a hybrid Simulated Annealing (HSA)[10], which optimizes the pipe diameters by using a local search technique based on GRASP. In [9], we analyzed the HSA behavior by considering the variants for two control parameters: the cooling scheme and the initial temperature. Following this research line, in this work we propose to study the third control parameter, the Markov Chain Length (MCL), and its variants, applying the obtained algorithmic approaches to a real case. The computation of its appropriate length is a difficult task to apply in practice. Consequently, we consider three different strategies to set the MCL to understand their effectiveness in terms of solution quality and execution time. We test the performance of our HSA with the HydroGen network instances [11]. Moreover, we describe and give the features of a real medium size distribution network, which is used to test the capabilities of the HSA model developed.

- The main contributions of this work are the following:
- Analysis of how the MCL influences the HSA performance when solving the WDND problem.
- Resolution of a real distribution network from our community.
- Comparison of the developed HSA solver with state-of-the-art techniques.

The remainder of this article is structured as follows. Section 2 introduces the problem definition. Section 3 explains our algorithmic proposal, HSA, to solve the WDND optimization problem and the HSA's configurations. Section 4 describes the experimental analysis and the methodology used. Then, sections 5 and 6 present the result analysis of the variants and the comparison with an ILS [8] from the literature, respectively. Finally, Section 7 summarizes our conclusions and sketches out our future work.

## 2 Multi-Period Water Distribution Network Design

The objective of the WDND problem is to minimize the Total Investment Cost (TIC) in a water distribution network design. The problem can be characterized as: simple-objective, multi-period, and gravity-fed. Two restrictions are considered: the limit of water speed in each pipe and the demand pattern that varies in time. The network can be modeled by a connected graph, which is described by a set of nodes  $N = \{n_1, n_2, \dots\}$ , a set of pipes  $P = \{p_1, p_2, \dots\}$ , a set of loops  $L = \{l_1, l_2, \dots\}$ , and a set of commercially available pipe types  $T = \{t_1, t_2, \dots\}$ . The TIC value is obtained by the formula shown in Equation 1, where  $IC_t$  is the cost of a pipe  $p$  of type  $t$ ,  $L_p$  is the length of the tube, and  $x_{p,t}$  is the binary decision variable that determines whether the tube  $p$  is of type  $t$  or not. The objective function is limited by: physical laws of mass and energy conservation, minimum pressure demand in the nodes, and the maximum speed in the pipes, for each time  $\tau \in \mathcal{T}$ .

$$\min TIC = \sum_{p \in P} \sum_{t \in T} L_p IC_t x_{p,t} \quad (1)$$

### 2.1 A Real WDND problem

The principal motivation of this research is to get involved in solving community problems, in particular the water distribution. To provide some context, the water access problem in the province of La Pampa is a priority treatment for being a scarce natural resource. CORPICO<sup>3</sup>, the supplier of this essential service in General Pico, has to design an independent drinking water distribution network for a new neighborhood of 5 km<sup>2</sup>, minimizing the network cost through the proper selection of the pipe dimensions according to consumption and the physical laws of this type of problem. The water distribution network should cover an area called “Barrio Quintas Sur” identified as Zone 2 or Z2 (shaded region of Fig 1). Initially, the network has an extension of 1.65 km<sup>2</sup> foreseeing for the next ten years an adjacent extension of 3.4 km<sup>2</sup>, identified as the zones Z1, Z3, and Z4. The network designed for Z2 is independent but requires taking into consideration the demand of the other zones to become an extra supply network or to receive water from them (bypass).

## 3 HSA for the Multi-Period WDND Problem

In this work, we continue with the study of the HSA as WDND solver, initiated in [9]. In particular, HSA consists of adapting and hybridizing the SA algorithm [12] to solve the Multi-Period WDND optimization problem. The first design issue is fixed regarding the problem to be solved. For this reason, the

<sup>3</sup> CORPICO is the Regional Cooperative for Electricity, Works and other Services in the city of General Pico, province of La Pampa, Argentina.

4 Carlos Bermudez, Hugo Alfonso, Gabriela Minetti, and Carolina Salto



Fig. 1. Map laying water networks in General Pico.

---

**Algorithm 1** HSA to solve the WDND Optimization Problem

---

```

1:  $k = 0$ ;
2:  $\text{initTemp}(T)$ ;
3:  $\text{initialize}(S_0)$ ;
4:  $TIC_0 = \text{evaluate}(S_0)$ ;
5: repeat
6:   repeat
7:      $k = k + 1$ ;
8:      $S_1 = \text{MP-GRASP\_LS}(S_0)$ ;
9:      $TIC_1 = \text{evaluate}(S_1)$ ;
10:    if  $TIC_1 < TIC_0$  then
11:       $S_0 = S_1$ ;  $TIC_0 = TIC_1$ 
12:    end if
13:     $S_2 = \text{perturbation\_operator}(S_0)$ ;
14:     $TIC_2 = \text{evaluate}(S_2)$ ;
15:    if  $(TIC_2 < TIC_0)$  or  $(\exp^{((TIC_2 - TIC_0)/T)} > \text{random}(0, 1))$  then
16:       $S_0 = S_2$ ;  $TIC_0 = TIC_2$ 
17:    end if
18:  until  $(k \bmod \text{MCL} == 0)$ 
19:   $\text{update}(T)$ ;
20: until stop criterion is met
21: return  $S_0$ ;

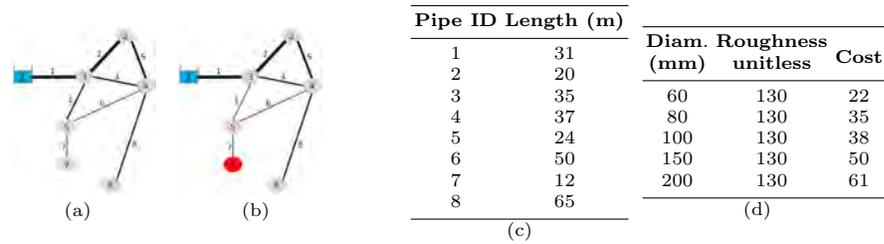
```

---

search space, solution codification, cost (evaluation) function, perturbation operator, and local search are directly related to the WDND problem. Meanwhile, the second design issue contemplates the algorithmic search strategy and includes the definition of the following control parameters: temperature, their annealing schedule, and the Markov chain length [13]. In the previous work, we analyzed different configurations of the two first HSA control parameters. In this work, we investigate the third one and the way to calculate the Markov chain length. The next sections describe the HSA algorithm and Algorithm 1 shows the HSA pseudocode.

### 3.1 HSA design issue based on problem definition

A solution for the WDND problem is a network, as shown in Figure 2 (a) and (b), which is represented by a vector. Each vector element is the diameter selected for the pipe it represents. Table 1 shows the vectors that correspond to the candidate



**Fig. 2.** Different solutions or network designs. (a) Solution 1; (b) Solution 2; (c) Pipe lengths; (d) Available pipe types with their corresponding costs.

**Table 1.** Different solutions or network designs in vector representation.

Solution	Pipe ID	1	2	3	4	5	6	7	8	Feasibility
	Length (m)	31	20	35	37	24	50	12	65	TIC
1	diam. (mm)	150	150	80	80	100	60	60	80	feasible
	cost	1550	1000	1225	1295	912	1100	264	2275	9621
2	diam. (mm)	150	150	80	60	100	60	60	80	infeasible
	cost	1550	1000	1225	814	912	1100	264	2275	9140

solutions in Figure 2(a) and (b). The TIC for each solution is calculated by the Equation 1, using the input data from tables (c) and (d) of Figure 2. The first solution is hydraulically feasible (satisfying all constraints mentioned in Section 2) and the second one is infeasible (violating the minimum pressure constraint in node 7). HSA uses the EPANET 2.0 toolkit [14] to solve the hydraulic equations, since this hydraulic solver is applied in most existing works.

Following with the HSA search components related to this problem, HSA generates a feasible initial solution  $S_0$  (line 3) applying both **HighCost** and **Lowcost** mechanisms proposed in [8]. Furthermore, the HSA is hybridized with MP-GRASP local search [8], which intensifies the search into the current region of the solution space (line 8). Then a greedy selection mechanism is performed (lines 10-12). The last search component associated with the WDND problem is the perturbation operator, used to obtain a feasible neighbor (line 13) and explore other areas of the search space. It randomly changes some pipe diameters.

### 3.2 HSA design issues specific to the search

The algorithmic search strategy compares the solution  $S_2$  generated by the perturbation operator with the current one,  $S_0$ . If  $S_2$  is worse than  $S_0$ ,  $S_2$  can be accepted under the Boltzmann probability (line 15, second condition). In this way, at high temperatures ( $T$ ) the exploration of the search space is strengthened. In contrast, at low temperatures the algorithm only exploits a promising region of the solution space, intensifying the search. In order to update  $T$ , a cooling schedule is used (line 19) and it is applied after a certain number of iterations ( $k$ ) given by the Markov chain length (line 18). Finally, HSA ends the search when the total evaluation number is reached or  $T = 0$ .

The choice of the right initial temperature has an important role in the HSA performance to find good solutions. A classical and intuitive method is used in this work that is introduced by Kirkpatrick et al. [12]. The scheme to control the cooling process is also crucial, so that the system gradually cools from a higher temperature, ultimately freezing to a global minimum state. Many attempts have been made to derive or suggest good schedules [13], being the most known cooling processes: the proportional [12], exponential [12], and logarithmic [15] schemes. A fourth cooling processes, named Random scheme, was successfully introduced and tested in order to solve this problem in [10].

The Markov chain length is the number of required transitions (moves) to reach the equilibrium state at each temperature,  $T_k$ . This number can be either static or adaptive. In the first case, the number of movements is calculated before the search starts. The static approach, named as MCLs, assumes that each temperature  $T_k$  is held constant for a sufficient and fixed number of iterations. In this work, each  $T_k$  is held constant for 30 iterations, a common number used in the scientific community.

For the adaptive cases, the Markov chain length depends on the characteristics of the search. For instance, Cardoso et al. [16] consider that the equilibrium state is not necessarily attained at each level of the temperature. Consequently, the cooling schedule is applied as soon as an improved candidate (neighbor) solution is generated. In this way, the computational effort can be drastically reduced without compromising the solution quality. This approach is referred as MCLa1. Another adaptive approach is proposed by Ali et al. [17], named as MCLa2, which uses both the worst and the best solutions found in the Markov chain (inner loop) to compute the next MCL. This strategy allows to increase the number of function evaluations at a given temperature if the difference between the worst and the best solutions increases, but if an improved solution is found the MCL remains unchanged.

## 4 Experiments with HSA's Variants

In our experiments, we use the best HSA's variant found in [9] to solve the multi-period WDND problem, named  $HSA_{Rand100}$ , which uses the random cooling scheme and 100 as initial temperature. This variant, renamed HSA in what follows, is executed under the three approaches to compute the MCL, as a consequence, three new HSA configurations arise. The stop condition of these algorithmic approaches is to reach 1,500,000 evaluations of the objective function. We performed 30 independent runs of each instance because of the stochastic nature of the algorithms, in order to gather meaningful experimental data and apply statistical confidence metrics to validate our results and conclusions. Before performing the statistical tests, we first checked whether the data followed a normal distribution by applying the Shapiro-Wilks test. Where the data was distributed normally, we later applied an ANOVA test. Otherwise, we used the Kruskal–Wallis (KW) test. This statistical study allows us to assess whether or

**Table 2.** Pipe types and their corresponding costs for the GP-Z2-2020 network.

Number	Diam. (mm)	Roughness	Cost	Number	Diam. (mm)	Roughness	Cost
1	63	110	2.85	5	315	110	69.10
2	90	110	5.90	6	400	110	110.89
3	110	110	8.79	7	450	110	140.15
4	125	110	11.00	8	630	110	273.28

not there were meaningful differences between the compared algorithms with  $\alpha = 0.01$ .

**HydroGen Networks.** The HydroGen networks [11] arise from 10 different distribution networks, named as HG-MP- $i$  with  $i \in \{1, 10\}$ . The demand nodes are divided into five categories (domestic, industrial, energy, public services, and commercial demand nodes), each one with a corresponding base load and demand pattern<sup>4</sup>. In this way, each HG-MP- $i$  network consists of five different instances, totaling 50 instances. The combinations of 16 pipe types and the number of pipes of each instance determines the network complexity. In this way, four different categories are obtained, to know HG-MP1-3, HG-MP4-6, HG-MP7-9, and HG-MP10, regarding the number of pipes of each instance.

**GP-Z2-2020: a Real Network.** The GP-Z2-2020 network, which arises from CORPICO's requirements, is composed of 222 domestic demand nodes and only one water reservoir. Moreover, this zone is connected with three other ones by means of some peripheral nodes which have different demand patterns, as explained in Section 2.1. Table 2 summarizes available pipe diameters, their corresponding roughness, and their unit costs (expressed in US dollars). The area is residential with a demand according to the current distribution of the customers in 584 plots, but considering a development pattern over a timespan of 30 year. The daily pattern demand corresponds to the summer period (based on the model demand of historical records) having a maximum resolution of one hour. The total number of possible combinations of design for a set of 8 commercial pipe types and 282 pipes is  $8^{282}$  which is difficult to test them; this shows the importance of optimization.

## 5 Analysis of the HydroGen Instance Results

In this section, we summarize and analyze the results of using the three new HSA's configurations to solve the 50 Hydrogen instances grouped by their corresponding distribution network. First, we analyze the behavior of these configurations by considering the results shown in the Table 3, taking the different MCL approaches into account. The columns 2-4 show the average of the best TIC values found by the three HSA's configurations for the 50 instances grouped by their corresponding distribution network. The minimal average TIC values found by each group are boldfaced. Last column summarizes the  $p$ -values obtained by the KW test. We follow with the best known TIC values found by

<sup>4</sup> The base loads can be found in the EPANET input files of the instances

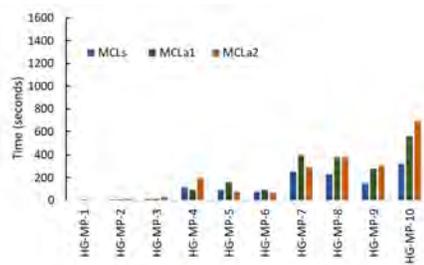
8 Carlos Bermudez, Hugo Alfonso, Gabriela Minetti, and Carolina Salto

**Table 3.** Averages of the best TIC values found by each HSA's configurations.

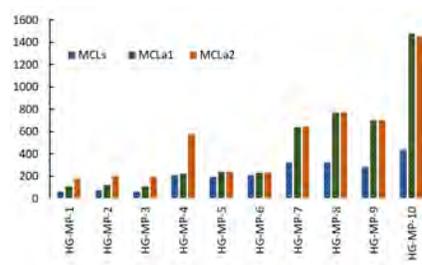
Network	MCLs	MCLa1	MCLa2	KW
HG-MP-1	<b>335723</b>	336984	337809	0,87
HG-MP-2	298430	298652	<b>297823</b>	0,85
HG-MP-3	<b>384210</b>	384665	384687	0,92
HG-MP-4	683780	<b>682361</b>	685424	0,80
HG-MP-5	719569	717268	<b>711906</b>	0,40
HG-MP-6	739923	<b>735738</b>	738108	0,93
HG-MP-7	807367	<b>792435</b>	801907	0,27
HG-MP-8	843657	841653	<b>841240</b>	0,13
HG-MP-9	823783	817325	<b>816485</b>	0,00
HG-MP-10	787399	<b>771461</b>	784385	0,08

**Table 4.** The best known TIC values found by our proposals and ILS.

Network	MCLs	MLCa1	MLCa2	ILS
HG-MP-1	<b>298000</b>	<b>298000</b>	<b>298000</b>	<b>298000</b>
HG-MP-2	245330	245330	245330	<b>245000</b>
HG-MP-3	310899	310706	<b>310493</b>	318000
HG-MP-4	592048	<b>590837</b>	592036	598000
HG-MP-5	<b>631000</b>	<b>631000</b>	<b>631000</b>	<b>631000</b>
HG-MP-6	617821	<b>609752</b>	614917	618000
HG-MP-7	648372	644568	<b>639932</b>	653000
HG-MP-8	795996	792436	<b>790037</b>	807000
HG-MP-9	716944	715863	<b>712450</b>	725000
HG-MP-10	730916	<b>712847</b>	727818	724000



**Fig. 3.** Minimum execution times consumed by each HSA's configuration.



**Fig. 4.** Total execution times consumed by each HSA's configuration.

our proposals shown in Table 4 (the minimal values are boldfaced). The performance of the HSA's configurations are compared with the ILS proposed in [8]. This metaheuristic is chosen from literature since its authors also used the HydroGen instances to test it. To ensure a fair comparison, both algorithms use the same stop criterion that is set in 1,500,000 Epanet calls. Finally, we analyze the computational effort for each HSA proposed, considering the time to find the best solution (see Figure 3) and the total execution time (see Figure 4) of the search measured in seconds, grouped by distribution network

Analyzing the results from the quality point of view for 90% of the instances, the three approaches behave statistically similar ( $p$ -value  $> \alpha$ ). The HSA with the adaptive options find the best TIC value in 80% of the cases. For the HG-MP-9 network, the HSA's behavior is significantly different, being MCLa2 the configuration for HSA that allows to find the lowest TICs values.

Focusing on computational effort analysis, for all configurations the HSA run times grow as the instance complexity increases. HSA with MCLa2 is the slowest configuration, whereas HSA with MCLs reduce significantly the total execution time (approximately 30-50%). However when the best time values to reach the best solution are analyzed, these differences are noticeably narrowed, minimizing the gap between these three approaches.

These results clearly state that the MCLa1 configuration exhibits a good tradeoff between the solution quality and the required execution time. This

**Table 5.** Results of the GP-Z2-2020 Network.

Metrics	MCLs	MCLa1	MCLa2	ILS
Minimal TIC	366684.00	358717.00	<b>347596.00</b>	355756.00
Average TIC	420949.53	408075.93	<b>401570.20</b>	429948.97
Average Best Time	<b>171.78</b>	564.26	555.44	169.55
Average Total Time	<b>186.46</b>	627.97	615.52	183.83

adaptive approach interrupts the Markov chain when the candidate solution is better than the current one, reducing the computational effort in comparison with MCLa2. Furthermore, in comparison with the literature our proposals are better than ILS in 80% of the networks and equal in the remaining ones.

## 6 Analysis of the GP-Z2-2020 Network Results

This section is devoted to the analysis of the results of the HSA when solving the real instance (GP-Z2-2020 network). Table 5 presents the results of the three HSA's configurations and the corresponding to the ILS algorithm [8] for different metrics: best and average TIC values (rows 1 and 2), the average time to find the best solution (row 3), and the average total time (row 4), expressed in seconds. The minimal values found by each metric are boldfaced.

HSA using MCLa2 finds minimal TIC values. This is an expected result since similar observations were made in the previous analysis taking into account networks of comparable complexity. An analogous situation is observed regarding the execution times. Furthermore, the three HSA solvers find better average TIC values than ILS, being HSA with MCLs the option that consumes similar execution times than ILS.

## 7 Conclusions

In this article, we analyze the influence of a relevant control parameter of the HSA solver proposed to solve the WDND optimization problem. The study includes a static (MCLs) and two adaptive (MCLa1 and MCLa2) configurations. We test the HSA's performance with three MCL configurations by using 50 Hydrogen networks. Furthermore, we introduced a new WDND instance based on a real case, which was solved successfully by this solver.

The analysis of the results obtained in the experimentation allows us to conclude the following. The adaptive configurations find better solutions than the static one by adapting the MCL to the search context. These adaptations require an extra computational effort to calculate the MCL during the search. The HSA with MCLa1 presents a good tradeoff between solution quality and effort since this adaptive method interrupts the Markov chain if a better solution than the current one is found. When our proposals are contrasted against ILS [8], the HSA's configurations outperform or equal ILS in every WDND network.

A challenging extension of this work will be to implement and test our HSA using big-data distributed frameworks to deal with larger dimension WDND networks.

10 Carlos Bermudez, Hugo Alfonso, Gabriela Minetti, and Carolina Salto

**Acknowledgments.** The authors acknowledge the support of Universidad Nacional de La Pampa (Project FI-CD-151/15 and POIR-03-2019) and the Incentive Program from MINCYT. The last author is also funded by CONICET.

## References

1. D. Yates, A. Templeman, and T. Boffey, "The computational complexity of the problem of determining least capital cost designs for water supply networks," *Engineering Optimization*, vol. 7, no. 2, pp. 143–155, 1984.
2. R. Farmani, G. A. Walters, and D. A. Savic, "Trade-off between total cost and reliability for anytown water distribution network," *Journal of Water Resources Planning and Management*, vol. 131, no. 3, pp. 161–171, 2005.
3. I. Gupta, A. Gupta, and P. Khanna, "Genetic algorithm for optimization of water distribution systems," *Environmental Modelling & Software*, vol. 14, no. 5, pp. 437–446, 1999.
4. C. Bragalli, C. D'Ambrosio, J. Lee, A. Lodi, and P. Toth, "On the optimal design of water distribution networks: a practical MINLP approach," *Optimization and Engineering*, vol. 13, no. 2, pp. 219–246, 2012.
5. R. Uma, "Optimal design of water distribution network using differential evolution," *International Journal of Science and Research (IJSR)*, vol. 5, no. 11, pp. 1515–1520, 2016.
6. R. Mansouri and M. Mohamadizadeh, "Optimal design of water distribution system using central force optimization and differential evolution," vol. 7, no. 3. [Online]. Available: <http://ijoce.iust.ac.ir/article-1-310-en.html>
7. A. Sedki and D. Ouazar, "Hybrid particle swarm optimization and differential evolution for optimal design of water distribution systems," *Adv. Eng. Informatics*, vol. 26, pp. 582–591, 2012.
8. A. De Corte and K. Sörensen, "An iterated local search algorithm for water distribution network design optimization," *Network*, vol. 67, no. 3, pp. 187–198, 2016.
9. 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.
10. 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.
11. A. De Corte and K. Sörensen, "Hydrogen," available online: <http://antor.uantwerpen.be/hydrogen> (accessed on 27 June 2018).
12. S. Kirkpatrick, C. G. Jr, and M. Vecchi, "Optimization by simulated annealing," *Science*, no. 220, pp. 671–680, 1983.
13. E.-G. Talbi, *Metaheuristics: From Design to Implementation*. Wiley, 2009.
14. L. Rossman, *The EPANET Programmer's Toolkit for Analysis of Water Distribution Systems*, 1999.
15. B. Hajek, "Cooling schedules for optimal annealing," *Mathematics of Operations Research*, vol. 13, no. 2, pp. 311–329, 1988.
16. M. Cardoso, R. Salcedo, and S. de Azevedo, "Nonequilibrium simulated annealing: A faster approach to combinatorial minimization," *Ind. Eng. Chem. Res.*, vol. 33, pp. 1908–1918, 1994.
17. M. Ali, A. Törn, and S. Viitanen, "A direct search variant of the simulated annealing algorithm for optimization involving continuous variables," *Computers & Operations Research*, vol. 29, no. 1, pp. 87 – 102, 2002.

# Image augmentation for object detection of grapevines

Tatiana Sofía Parlanti<sup>1</sup>, Alejandro Martín Lobos<sup>1,2</sup>, Luis G. Moyano<sup>2,3</sup>, and Emmanuel N. Millán<sup>2,4</sup>

<sup>1</sup> Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, Mendoza, Argentina

<sup>2</sup> CONICET, Mendoza, Argentina

<sup>3</sup> Grupo de Física Estadística e Interdisciplinaria, Centro Atómico Bariloche, Bariloche, Argentina

<sup>4</sup> ITIC and ITU, Universidad Nacional de Cuyo, Mendoza, Argentina.  
emmanueln@gmail.com

**Abstract.** Machine Learning methods are widely used for data analysis in various areas. In this work we use Neural Networks for image analysis in order to detect grape fruit clusters. A set of manually tagged images is built and a comparison is made between different data augmentation techniques in order to analyse the best way to expand the image set. The technique presented here obtained up to 13% better detection performance starting with only 100 images for training. The types of transformations and filters that worked the best for these images are discussed. In addition, training and detection times in five different hardware infrastructures, both CPU and GPUs, are briefly discussed.

**Keywords:** machine learning, neural networks, deep learning, object detection

## 1 Introduction

Machine Learning (ML) is the field of study that allows computers to have the ability to learn without having been explicitly programmed to do so [10]. ML is often associated with automatic detection of patterns or useful information in data. The use of Machine Learning for data analysis is widely used in various areas, from social to health and engineering [4,7]. This is particularly the case in the area of automatic object recognition, which has direct applications in agricultural processes [11].

A general need for agricultural producers is to have an adequate process for estimating harvest; in particular, in our local region, grape harvest. Traditionally, the most accurate and fastest methodology for estimating the harvest of grapes is given by the count of bunches per plant and weight estimation. The extrapolation of the individual result to all plants determines the projected production value. Direct visual estimation is also used, but is subjective. To automate the analysis of grapevine images with the objective of improving the

estimation of the grape harvest forecast is seen as a problem that can be addressed by ML, especially given the advances seen in recent years. For example, [1] presented an algorithm based on mathematical morphology and pixel classification for grapevine berry counting, however images were taken placing a dark cardboard behind the cluster, which is difficult for an operator to implement. Then, yield prediction based on RGB images acquired “on-the-go” from an all-terrain vehicle was presented in [2], however the capture of images was during the night, which means that operators have to work in an unusual time window. Also, ML has been used to identify apples at different times of growth using YOLOv3 [14], and [3] presented an adaptation of crowd-counting algorithms for the estimation of harvest yield for grape berries.

The relevant steps to automate the forecast process are: first, image acquisition, then automatic cluster identification, and finally, cluster analysis for harvest forecast. In this work we focus on the first stages of this process: image acquisition, labeling, training and detection with Neural Networks.

A common limitation of Neural Networks (NN) is that they typically require hundreds to thousands of training images for the correct detection of objects. To address this, we first capture images on various vineyard farms, which were then hand-tagged. A NN was trained using the images obtained from vineyards, and was later used to generate detection models. A comparison of the resulting detections, obtained from the same set of images made by trained models with 100 and 500 photographs, was made. Afterwards, a series of transformations were applied to 100 images in the first set in order to increase the number of data, which were used to train again and obtain new, improved models. Finally, the detection percentages of these models were compared against the detections of the two initial sets of images.

The work is organized as follows. Section 2 describes the computational tools we use to carry out the analysis (2.1), our images (2.2), and the filters and transformations applied to them (2.3). We then show and discuss our results in Sec. 3. Finally, we elaborate conclusions and state future lines of research in Sec. 4.

## 2 Materials and Methods

The Python ImageAI library [6] was used to carry out the different analyses, since it allows training custom object detection models on image datasets that are in Pascal VOC annotation format. ImageAI uses the YOLOv3 architecture [8,9], and only requires a few lines of code.

For the training stage, a pre-trained model provided by the same ImageAI library was used, as faster progress is obtained using transfer learning compared to training from scratch. The number of epochs used in the training was 100. Also, batch size and learning rate were set at 2 and 0.0001 respectively. Different models were generated, which were then evaluated by calculating the mAP score (mean Average Precision) of each one in order to pick the most accurate, com-

puted over the validation image set. Source code is available from the authors via email request.

## 2.1 Machine Learning Algorithms

You Only Look Once (YOLOv3) algorithm [8,9] is a real-time object detection system, which resizes the input images to  $416 \times 416$ , has 53 convolutional layers in its NN architecture and thresholds the resulting detections by the model's confidence.

For its operation the NN divides the image into a grid, and predicts bounding boxes and the level of certainty of each box for each region. The confidence scores quantify the likeliness that a given region contains an object, as well as its accuracy.

## 2.2 Image dataset

The set of images used was assembled from photographs taken of espaliers from different vineyards, two of which are in the Valle de Uco. Additionally, bunches from espaliers belonging to the Instituto Nacional de Tecnología Agropecuaria (INTA), in Lujan de Cuyo were photographed. Two smartphones were used to capture the images (with 25- and 13-megapixel sensors). The varieties photographed included Malbec and Cabernet Sauvignon, both dark-colored, used to make red wine. Close-ups of the bunches and general plans of the espaliers were made.

Each image was then annotated in Pascal VOC format, with the `labelImg` tool [15], which is straightforward to apply, and generates files with `.xml` extension that store the coordinates of the upper left and lower right vertices of each manually generated annotation. In this way, more than 800 images were tagged, corresponding to more than 3000 grape bunches.

These images, and their respective annotations, were separated into three sets, which we will call *bunch600*, *bunch120* and *detect*, detailed below:

- **bunch600**: consists of 500 images used as a training set (2253 annotated bunches), and 100 images used for the validating set (472 annotated bunches). These sets include photographs taken with bunches in the foreground and in the background;
- **bunch120**: consists of 100 images used as a training set (593 annotated bunches), and 20 images used for the validation stage (102 annotated bunches). In this case all the images used were in the foreground;
- **detect**: consists of 100 images (641 annotated bunches), which were used after the training to test the chosen model, generated in the first stage.

## 2.3 Image filters and transformations

Increasing the data present in the dataset (known as *image augmentation*) is a common technique that seeks to increase the number of images in a sparse training set, based on transformations such as rotation, translation, noise addition,

etc. [5,12]. Image augmentation relies on how NNs interpret image characteristics: small modifications are interpreted as completely different images by the NN.

In this sense, a series of filters and transformations were applied to the images of *bunch120* using the ImageMagick software [13]. Two types of tests were performed. On the one hand, tests the training and validation set was duplicated with the reflection transformation, the solarize filter and the contrast limited adaptive histogram equalization were used, as detailed below.

- **mirror**: horizontal mirroring;
- **solarize**: when solarizing, the color of the pixels whose luminosity exceeds a certain limit is inverted, in this case the limit was set at 60%;
- **clahe**: adaptive histogram equalization improves local contrast and enhances edge definitions in each region of an image. To apply clahe, a grid of defined pixel width and height tiles is used, where each box contains a fixed number of classes or bins, and a limit is established for localized contrast changes. In this case the boxes were 5% of the width and height of the image, and contained 128 bins each. In addition the contrast limit was set at 4.

On the other hand, the training and validation set was quintupled by applying four different values of the same type of filter:

- **laplacian noise**: with attenuation values 1.5, 3.0, 4.5 y 6.0;
- **paint**: blurred from “drops” of paint that blends the colors present in an environment and turns them into a single-color area. Used radius values: 3, 5, 7 y 9.

The *noise*, *paint*, *solarize* and *clahe* filters modify the images, but not the labels of the grape clusters. On the other hand, the *mirror* filter requires a transformation to be applied to *.xml* files, so labels are also mirrored. Thus, from the 100 images of *bunch120*, five new training sets were obtained, three around 200 images each, and two around 500 images each (matching, in number, the *bunch600* dataset). Similarly, the same transformations were applied to the validation images, to maintain the proportions between both sets. Fig. 1 shows the types of filters applied to the original image (a).

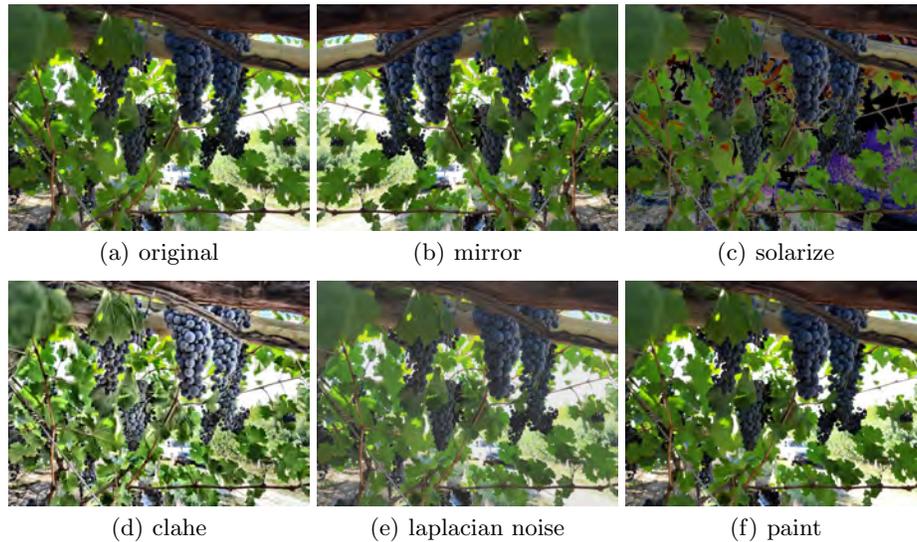
### 3 Results and Discussion

Results of the detections from chosen models obtained by the NN are presented.

#### 3.1 Hardware and Software Description

The analysis of the images is carried out mainly using GPUs. The following equipment belonging to the Toko cluster of the FCEN UNCuyo was used:

- Toko computing node with four AMD **Opteron** 6376 CPU, with 16 CPU cores at 2.3GHz each (64 cores total), 128 GB of RAM.



**Fig. 1.** Example of an original image and its transformations.

- Toko computing node with two AMD **EPYC** 7281 CPU, with 16 CPU cores at 2.1GHz each (32 cores total), 128 GB of RAM.
- Toko computing node with one AMD **Ryzen** 7 2700 with 8 cores and 16 threads running at 3.2 GHz with 64 GB of DDR4 RAM memory.
- Two AMD FX-8350 nodes with 8 cores running at 4 GHz and 32GB of DDR3 RAM. One node with an NVIDIA GeForce GTX **Titan X** (Maxwell GM200 architecture) with 12 GB of memory, and another node with an NVIDIA GeForce GTX **Titan Xp** (Pascal GP102) with 12 GB of memory.

The Toko cluster is running Slackware Linux current 2020 64bit with kernel 5.4.28, gcc 9.2, NVIDIA driver 410.73 and CUDA 10. The software used to perform the analysis was developed around the imageAI 2.1.5 library with tensorflow-gpu 1.13.1, keras 2.2.4, and python 3.7.2. In the section 3.3 we discuss the performance of the NN running in this hardware.

### 3.2 Image detection

Frequently, the same cluster is divided at detection and reported as two or three different objects, depending on whether there are leaves or branches partially obstructing it. In this situation, it was considered that each one of these detections corresponded to the real cluster if the intersection between both rectangles (the detection and the hand-labeled rectangles) was greater than 25%.

Also, a manually tagged cluster is considered “well detected” if the rectangle detected by the NN and the tagged rectangle match 60% or more. If a labeled cluster corresponds to several smaller detected rectangles, it is considered “well

detected” if the sum of the surfaces in the small rectangles represents at most 60% of the real surface of the labeled rectangle.

As explained in Sec. 2.2, *detect* is the set of images that was used to test the models. It consists of 100 images, containing 641 labeled clusters in total. These images were tagged in order to compare them against NN detections, but the NN did not had access to the tags at any time during training. The Table 1 shows the difference in detections between *bunch120* and *bunch600*, which have 100 and 500 training images, respectively. The name “originals” means that the images do not have any type of filter or transformation applied.

Table 2 compares the detections achieved by the combination of *bunch120* (original images) and the application of a transformation type, which increases the number of images to 200 of training. On the other hand, the Table 3 compares the combinations of four filters, reaching 500 training images per test.

**Table 1.** *bunch120* vs. *bunch600*

	originals <i>bunch120</i>	originals <i>bunch600</i>
Total clusters detected	455	593
Total clusters not detected	186	48
Total other detections	125	211
<b>Percentage of correct detections</b>	<b>70.98%</b>	<b>92.51%</b>
Average certainties	50.3	52.45
Standard deviation certainties	9.73	9.8

**Table 2.** Dataset detections that duplicate the training set

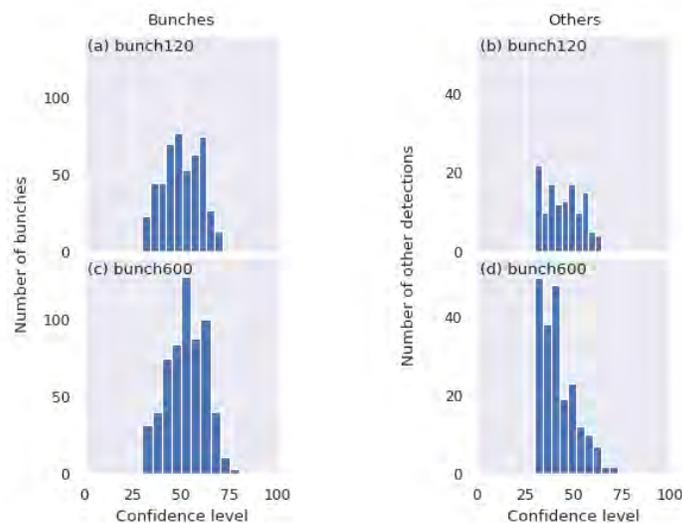
	originals + clahe	originals + mirror	originals + solarize
Total clusters detected	458	480	536
Total clusters not detected	183	161	105
Total other detections	80	123	217
<b>Percentage of correct detections</b>	<b>71.45%</b>	<b>74.88%</b>	<b>83.61%</b>
Average certainties	51.38	55.92	54.45
Standard deviation certainties	11.38	11.25	13.82

**Table 3.** Dataset detections that quintuple the training set

	originals + laplacian	originals + paint
Total clusters detected	454	467
Total clusters not detected	187	174
Total other detections	108	86
<b>Percentage of correct detections</b>	<b>70.82%</b>	<b>72.85%</b>
Average certainties	46.87	48.07
Standard deviation certainties	8.57	13.29

Regarding the level of certainty, a distinction is made between detections that correspond to clusters, from detections of other objects that could —or not— be partial-view clusters, not considered in the labeling. The different tables, in addition, report the average and standard deviation of the certainties of the detections that correspond to real clusters. Figure 2 shows the percentages of certainty of the detections obtained after training with images of *bunch120* and *bunch600*, without the application of transformations or filters. On the left,

graphs (a) and (c) take into account the certainties for labeled clusters, while graphs (b) and (d) on the right take into account detections from other objects.



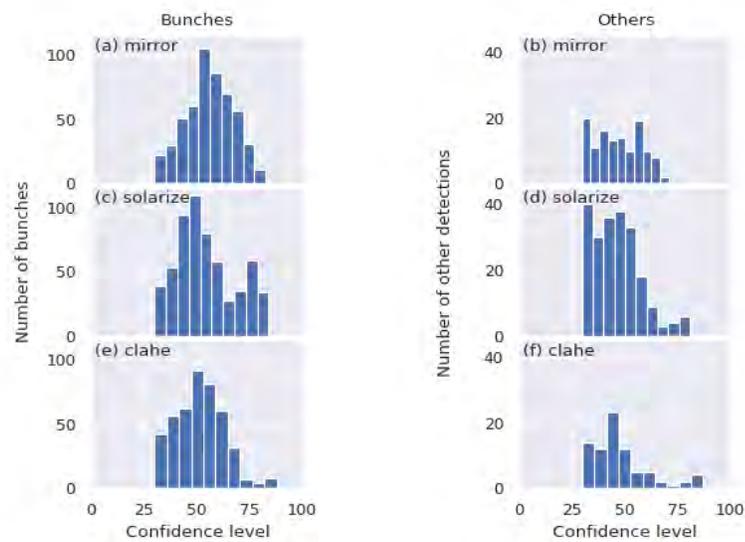
**Fig. 2.** Percentages of certainty of the detections obtained after training with the original images of *bunch120* (above) and *bunch600* (below).

Similarly, Fig. 3 has on its left the cluster detections, obtained from the training carried out with the images of *bunch120* together with the reflection transformation (a), next to the solarize filter (c), and next to the clahe equalization (e). On the right, instead, are represented the detections of other objects. Likewise, Fig. 4 shows the graphics corresponding to the training with images with Laplacian filter and paint.

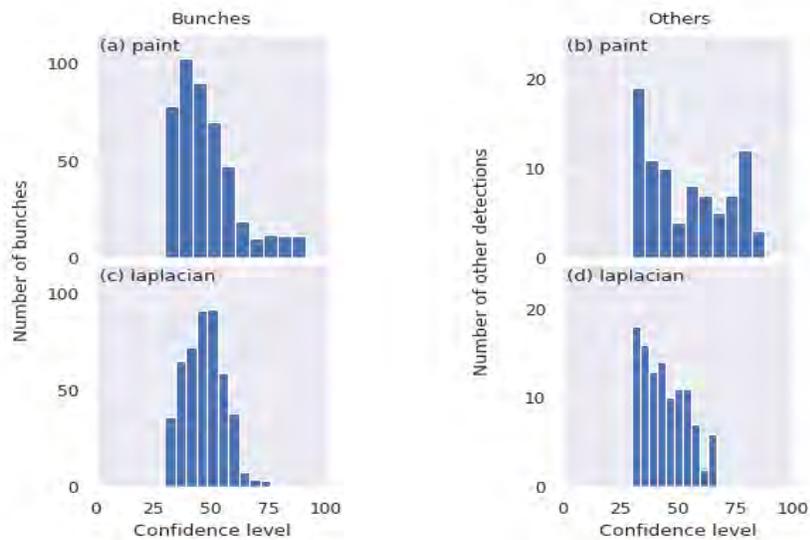
As expected, the NN reaches a higher detection of 92% from the dataset *bunch600*, while detection from *bunch120* only reaches 70%. However, the application of some *image augmentation* techniques leads to an increase in the number of detections. Duplicating the original dataset using filters such as mirroring or solarizing achieves an increase from 4% to 13%. However, quintupling the original dataset only accounts for 2% more detection when using the paint filter.

It is remarkable that a higher detection rate was obtained among the original images plus their mirrors (200 images), than the original ones plus four paint filter applications (500 images). Simply duplicating the data using this particular transformation can achieve better results than quintupling the data, with the addition of less computer time.

On the other hand, the tests of solarize filter, mirrored images, clahe equalization and application of four values of paint are the ones that present higher percentages of certainty, even higher than those of *bunch600*. However, the use of solarize is also the one that leads to additional “other detections”.



**Fig. 3.** Percentages of certainty of the detections obtained after training with the original images of *bunch120* together with their reflections (above), together with the images with solarization (middle), and together with the clahe equalization (below).



**Fig. 4.** Percentages of certainty of the detections obtained after training with the original images of *bunch120* together with four values of paint (above) and four values of Laplacian noise (below).

### 3.3 Computational performance

The estimation of performance was a secondary objective of this work. The analysis was performed in five different hardware infrastructures, both CPUs and GPUs. For this analysis the data set *bunch120* was used and the time is equivalent to one experiment. Table 4 shows that the Titan Xp GPU, as expected, has the best performance being  $\sim 3$  times faster than 64 cores of the Opteron 6376 node and  $\sim 1.4$  times faster than the previous generation Titan X GPU. The two GPUs have superior performance to the fastest CPU, the AMD Epyc 7281.

The training time for 100 experiments (section results 3.2) on the Titan Xp GPU, from the original *bunch120* image set and mirror (200 images) took  $\sim 11$  hours and original *bunch120* and 4 filters paint (500 images) took  $\sim 31$  hours. In the case of *bunch600* (500 images) the training took  $\sim 30$  hours. Detection time is  $\sim 1.5$  seconds per image using the model generated by *bunch120* + mirror on the same Titan Xp GPU. Due to these training times the number of filters that could be tested were limited as the use of the Toko cluster is shared amongst other users.

**Table 4.** Mean time in seconds for five separate training executions in five different CPUs and GPUs infrastructures.

Hardware	Mean time (s)	Standard deviation
Opteron	1652	35
Ryzen	1080	29
Epyc	962	21
Titan X	787	7
Titan Xp	563	12

## 4 Conclusions and Future Work

Cluster detection models were trained from new constructed datasets of 100 and 500 images each. Their detection percentages were compared, as well as the models obtained after using data augmentation techniques. With the methodology presented in this work, it was possible to increase the number of detections by 13%, although it was not possible to match the detection percentage obtained from the training with 500 images. Despite expectations, not all possible transformations lead to better results, as in the case of Laplacian noise application. Although a greater number of detections is achieved from a large training set, the difference that can be obtained from the application of some specific modifications to a smaller dataset is remarkable.

As future work, we expect to generate new models from the application of different transformations in the images, such as new translations or rotations, and apply adversarial learning techniques to re-train the NN. The next step is to increase the number and variety of images in the *bunch600* dataset to increase the number of images present in the dataset; we expect that such a step will decrease the number of “other detections”.

## Acknowledgments

We acknowledge Ing. Marcos Montoya and Ing. Julieta Dalmasso from INTA, Scattareggia family and Coletto Vineyard for giving us access to their grape fields. We acknowledge support from CONICET and SIIP Type 4 UNCuyo grants. This work was performed in the Toko Cluster from FCEN-UNCuyo, that is part of SNCAD, MinCyT, Argentina.

## References

1. Aquino, A., Diago, M.P., Millán, B., Tardáguila, J.: A new methodology for estimating the grapevine-berry number per cluster using image analysis. *Biosystems Engineering* 156, 80–95 (apr 2017)
2. Aquino, A., Millan, B., Diago, M.P., Tardaguila, J.: Automated early yield prediction in vineyards from on-the-go image acquisition. *Computers and Electronics in Agriculture* 144, 26–36 (jan 2018)
3. Coviello, L., Cristoforetti, M., Jurman, G., Furlanello, C.: In-field grape berries counting for yield estimation using dilated cnns. arXiv preprint arXiv:1909.12083 (2019)
4. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* 521(7553), 436–444 (may 2015)
5. Mikolajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary PhD Workshop (IIPhDW). IEEE (may 2018)
6. Moses, Olafenwa, J.: Imageai, an open source python library built to empower developers to build applications and systems with self-contained computer vision capabilities (mar 2018–), <https://github.com/OlafenwaMoses/ImageAI>
7. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Édouard Duchesnay: Scikit-learn: Machine learning in python. *Journal of Machine Learning Research* (2012)
8. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (jun 2016)
9. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement (2018), <http://arxiv.org/pdf/1804.02767v1:PDF>
10. Samuel, A.L.: Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 3(3), 210–229 (jul 1959)
11. Shai Shalev-Shwartz, S.B.D.: Understanding Machine Learning. Cambridge University Pr. (2014), [https://www.ebook.de/de/product/22370006/shai\\_shalev\\_shwartz\\_shai\\_ben\\_david\\_understanding\\_machine\\_learning.html](https://www.ebook.de/de/product/22370006/shai_shalev_shwartz_shai_ben_david_understanding_machine_learning.html)
12. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *Journal of Big Data* 6(1) (jul 2019)
13. The ImageMagick Development Team: Imagemagick, <https://imagemagick.org>
14. Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., Liang, Z.: Apple detection during different growth stages in orchards using the improved YOLO-v3 model. *Computers and Electronics in Agriculture* 157, 417–426 (feb 2019)
15. Tzotalin: LabelImg. Git code (2015), <https://github.com/tzotalin/labelImg>

## Navegación autónoma mediante aprendizaje por refuerzo

Ignacio Agustín Costa<sup>1</sup>, Esteban José De Leo<sup>1</sup>, Iris Inés Sattolo<sup>1</sup>, Marisa Daniela Panizzi<sup>1</sup>

Escuela Superior de Ingeniería, Informática y Ciencias Agroalimentarias. Universidad de Morón. Cabildo 134 (B1708JPD). Partido de Morón, Argentina.

[costa.nacho@gmail.com](mailto:costa.nacho@gmail.com); [deleoesteban@gmail.com](mailto:deleoesteban@gmail.com); [iris.sattolo@gmail.com](mailto:iris.sattolo@gmail.com); [marisapanizzi@outlook.com](mailto:marisapanizzi@outlook.com)

**Resumen.** Existen diversas situaciones, como derrumbes edilicias o catástrofes, en donde seres vivos quedan atrapados y su supervivencia depende de encontrar el camino de salida del lugar. Este trabajo presenta el diseño e implementación de un sistema de aprendizaje mediante la utilización de Redes Neuronales Artificiales (RNA) con aprendizaje por refuerzo, con el propósito de resolver laberintos de estructura desconocida con un robot autónomo (RA). Se definieron un conjunto de políticas y puntajes para llevar a cabo el aprendizaje por parte del agente, y posteriormente encontrar la solución buscada. La RNA fue implementada sobre las librerías Tensorflow y Keras. El RA es comandado por Arduino, el cual se comunica con una aplicación de escritorio. Finalmente, se diseñó una maqueta adaptable a cualquier laberinto para poder ejecutar la resolución mediante el RA. Los resultados indican que el aprendizaje por refuerzo es apropiado para resolver esta problemática, dado que es adaptable a problemas complejos.

**Palabras clave:** aprendizaje por refuerzo, red neuronal artificial, RA, q-learning, resolución de laberintos, exploración de estructuras desconocidas.

### 1 Introducción

Actualmente, existen situaciones y escenarios reales donde el ambiente es desconocido para los individuos que se encuentran en él, y es imperativo encontrar una salida. Algunas situaciones pueden ser catástrofes naturales o derrumbes edilicios, en donde seres vivos quedan atrapados entre restos o escombros, y su supervivencia depende de encontrar el camino de salida del lugar.

Existen diversos algoritmos que proporcionan respuestas acordes al problema de encontrar el trayecto óptimo dentro de un ambiente desconocido, algunos de ellos son: el algoritmo estrella [1] y doblar siempre a la izquierda/derecha. Los mismos, aplican procedimientos para encontrar un camino válido, dentro de todos los posibles. Es decir, están restringidos a estructuras simples en donde hay espacios físicos ocupados, y otros libres. La resolución al problema sólo recae en recorrer los espacios libres hasta encontrar una salida. El ejemplo más sencillo sería imaginar a una persona rodeada de

escombros resultantes de un derrumbe, en donde se desea encontrar un camino seguro de salida de la zona.

Pero, hay escenarios en donde aplican reglas más complejas. Por ejemplo, si la persona en medio del derrumbe debe rescatar a otros antes de encontrar la salida, estos algoritmos no nos servirían.

Otro caso sería, si un drone debe recorrer la zona de derrumbe para identificar los daños ocasionados. Según las características técnicas del dispositivo, podría elevarse por encima de obstáculos de cierta altura, pero no estructuras demasiado altas.

De acuerdo a los problemas mencionados, en esta investigación se propuso como solución crear un sistema de aprendizaje para la resolución de caminos en ambientes complejos de estructura desconocida, y su ejecución con un agente autónomo.

Se desarrolló una aplicación que permita ingresar los datos de entrada del entorno, realice el entrenamiento mediante una RNA y envíe la solución al RA mencionado.

Este artículo se organiza de la siguiente manera: se presentan los Trabajos relacionados (en la sección 2), el desarrollo de la solución (en la sección 3), su validación (en la sección 4) y, por último, las conclusiones y las futuras líneas de trabajo.

## 2 Trabajos relacionados

Se realizó un Estudio de Mapeo Sistemático (en inglés, Systematic Mapping Study o SMS) de acuerdo con el proceso propuesto en [2] para analizar el estado del arte sobre el aprendizaje automático para resolución de caminos de estructuras desconocidas. Por razones de espacio, en [3] se encuentra información sobre el SMS.

En este estudio se seleccionaron 28 estudios primarios de un conjunto inicial de 2450 artículos que arrojó la búsqueda. Estos se obtuvieron realizando una búsqueda automática en las librerías digitales Google Scholar, IEEE Xplore y Scielo, en el período comprendido entre el año 2008 y 2018.

Una vez analizados los estudios primarios, se concluyó que:

- El 50% de las publicaciones seleccionadas (14 artículos), hacen uso de algoritmos de resolución de caminos los cuáles no realizan un procedimiento de aprendizaje. En cambio, sólo el 14% de las publicaciones utilizan modelos de aprendizaje.
- Dentro de los modelos de inteligencia artificial, el algoritmo genético, las RNA tipo SOM y convolucionales fueron los predilectos, ocupando un 8%, 4% y 4%, respectivamente.
- Un 40% de las publicaciones que construyeron RA (20 publicaciones) optaron por hacer un diseño propio de la electrónica del RA. El restante 60% se volcó a marcas reconocidas de microcontroladores. Las más populares fueron Arduino con un total de 4 artículos (20%) y en segundo lugar Pioneer con 3 artículos (15%).
- Como lenguajes predilectos se aprecia una amplia proporción de uso de Matlab y C++, siendo un 33% cada uno. Ambos lenguajes, de bajo nivel, permiten realizar cálculos específicos para realizar estas operaciones. Relacionado con los resultados de la PI3, los RA desarrollados con la placa Raspberry Pi, utilizaron Python. En contraparte, los RA desarrollados con la placa Arduino, utilizan C++ como lenguaje.

- Ninguno de los RA sigue un recorrido preestablecido, se encontró que todos los modelos diseñados avanzan hasta encontrarse con un obstáculo, para luego moverse en una dirección sin destino determinado. Esto se hace hasta toparse con la solución.
- La mayor cantidad de fuentes de congresos fue obtenida de Argentina y España, con un resultado de 4 cada uno. Con respecto a las publicaciones de revistas, Colombia y Ecuador fueron mayoría, con un total de 4 y 3 respectivamente.

### 3 Desarrollo de la solución

Para la construcción de la solución se definieron los siguientes componentes:

- Maqueta.
- Aplicación de escritorio.
- Robot Autónomo.

En [3] se describe con un mayor nivel de detalle, la maqueta del laberinto y la aplicación de escritorio.

El ambiente a recorrer será representado por una maqueta a escala. La misma consta de un papel subdividido en cuadrados que representa una sección del terreno. Sobre el papel se ubicarán paredes de madera que delimitarán los caminos a recorrer.

El usuario debe plasmar el escenario real en la maqueta y en la aplicación de escritorio. La misma recibirá como input la estructura a recorrer. Una vez ingresada, el usuario inicia el proceso del entrenamiento. Durante el mismo, la aplicación muestra en pantalla los datos del entrenamiento. Una vez finalizado, se habilita la ejecución del RA para recorrer camino óptimo hacia la salida de la estructura.

El RA es el encargado de recorrer la maqueta. El mismo está diseñado con un microcontrolador Arduino y se comunica a través del protocolo Bluetooth con la aplicación de escritorio. El usuario inicia la ejecución luego del entrenamiento, el RA recibe la información y recorre la estructura hasta la salida.

A continuación, se presenta el modelo de aprendizaje (sección 3.1.) y la implementación de la solución (sección 3.2.).

#### 3.1 Modelo de aprendizaje

El modelo de aprendizaje utilizado es el aprendizaje por refuerzo. El mismo propone obtener recompensas por cada acción realizada pero la retroalimentación no es constante, dado que un sólo movimiento no resolverá el problema. Será un conjunto de los mismos, el que recompense al agente eventualmente. Basado sólo en esas recompensas, el agente tiene que aprender a comportarse en el medio ambiente. No se le dice qué acciones tomar, si no que él debe experimentar para encontrar qué acciones lo llevan a una mayor recompensa. Los casos más desafiantes son los que no llevan a una recompensa inmediata, si no en las siguientes situaciones.

Se considera modelar la resolución del laberinto, como un Proceso de Decisión de Markov [4]. En donde cada estado se verá constituido por la estructura a recorrer y la posición del agente. La transición entre estados se dará a partir de un movimiento del

agente en el entorno. Esto resultará en una recompensa positiva o negativa, cuantificada por un número entero.

Q-Learning es la técnica de aprendizaje por refuerzo que será utilizada. La misma establece la necesidad de diseñar una política que defina el modo en que el agente se comporta en un momento dado. El objetivo es que el agente aprenda de dicha política. Cada par estado-acción resultará en un valor “Q” (derivado de la palabra inglesa “quality”), que representa la recompensa futura máxima [5].

El objetivo de Q-Learning es desarrollar una política que maximice las recompensas que se obtienen al ejecutar una acción, esto es, tomar la mejor decisión en un momento y contexto dado. Se denomina Explotación, a tomar decisiones en base a lo aprendido hasta el momento por el algoritmo. De esta manera se conoce el resultado esperado, y se refuerzan estas decisiones [6].

En contraposición, la Exploración se refiere a tomar decisiones que normalmente no tomaríamos, con la posibilidad de ganar nuevo conocimiento [7].

Como motor de aprendizaje y almacenamiento del conocimiento, se utilizará una RNA Prealimentada (feed-forward en inglés). La misma, recibirá como entrada la estructura del laberinto y la posición del RA. La salida serán los 4 movimientos posibles (arriba, abajo, izquierda y derecha) con sus probabilidades, según la predicción de la próxima acción.

### 3.2 Implementación de la solución

El laberinto será modelado como una matriz de números enteros. Cada celda correspondiente a la cuadrilla tendrá un valor asociado representando su estado actual:

- 0 = Celda libre: Es aquella celda a la cual el RA puede ingresar.
- 1 = Celda ocupada: Es aquella celda a la cual el RA no puede ingresar.
- 99 = Celda del robot: Es la celda en la cual se encuentra el RA ubicado.

Para la representación del laberinto se creó una maqueta adaptable a cualquier diseño. Para el terreno se imprimió un plotter con la cuadrícula de 6 filas y 6 columnas (36 celdas en total).

Para representar las celdas ocupadas, se fabricaron paredes de madera de 3mm cortadas con láser para una mayor precisión.

Como interfaz con el usuario, se creó una aplicación de escritorio en Python 3.7 utilizando la librería gráfica Tkinter. Este programa tomará como input el laberinto provisto por el usuario y permitirá realizar el aprendizaje por parte del algoritmo. Durante el proceso de entrenamiento, se visualizarán datos estadísticos correspondientes. Luego de finalizado el entrenamiento, se permitirá acceder a reportes con métricas de los resultados del mismo. Asimismo, se mostrará una animación con la solución encontrada. Por último, el programa se comunicará a través de Bluetooth, utilizando la librería X, con el RA enviando la información de la resolución.

Para el entrenamiento, como parámetro de entrada necesitaremos la matriz del laberinto. Con las dimensiones de la estructura se crea la RNA y se iniciará con valores de pesos aleatorios.

Se define a un “estado” como una foto del laberinto en un momento determinado. La misma representa al entorno y la posición del RA. Llamaremos “transición” al cambio de un estado a otro, ante una acción (movimiento de celda). Utilizaremos el modelo de Proceso de Decisión de Markov.

Se ejecutarán igual cantidad de corridas como epochs definidos, por defecto 1000. Llamamos una “corrida” a la serie completa de movimientos del RA desde la posición inicial hasta la salida o la penalización máxima. Para cada corrida, la posición inicial del RA será determinada de forma aleatoria por el algoritmo.

En cada corrida, se ejecutarán una cantidad de movimientos (acciones) a los cuales consideraremos episodios. En cada episodio, se evalúa si realizar explotación o exploración. Establecemos el coeficiente de exploración inicial al 10%. Al superar el 90% de corridas exitosas, se disminuye al 5%. Para explotar, se envía el estado a la RNA y se consulta el movimiento a realizar.

Cada episodio se almacenará en una memoria interna. El registro del episodio estará compuesto por cinco partes: Estado Inicial, Acción, Recompensa, Estado Final e indicador de finalización de corrida.

Para poder definir las reglas del problema y encontrar la solución requerida, se definió una política. La misma buscará maximizar la recompensa obtenida ante cada acción.

A través de un sistema de puntajes, la política le permitirá al RA aprender y así hallar un camino óptimo. Cada movimiento, obtendrá una recompensa positiva o negativa. La corrida comprenderá la suma de puntajes de todos los movimientos.

En la Tabla 1 se detallan las acciones junto con sus respectivos puntajes.

**Tabla 1.** Puntajes por acción.

Acción	Puntaje
Movimiento a una celda libre	-0.04
Movimiento a una celda ya visitada	-0.25
Encontrar la salida	1
Movimiento a una celda ocupada	-10
Comenzar la corrida en una celda bloqueada	Puntaje negativo límite
Movimiento fuera de los límites del laberinto	Puntaje negativo límite

Para alentar a realizar la mínima cantidad de movimientos posible, se penaliza con -0.04 cada ingreso a una celda libre.

Para alentar a no volver sobre sus propios pasos, se penaliza con -0.25 al volver a ingresar a una celda visitada previamente.

Al encontrar la salida, se obtiene +1, la cual es la máxima recompensa.

Si durante la corrida se llega al puntaje negativo límite, se considera como perdida la corrida, dado que el RA cometió demasiados errores. El cálculo del límite es:  $-0.5 * \text{tamaño laberinto}$ .

Para desalentar a ingresar a una celda que se encuentra ocupada, daremos la penalidad máxima igual al puntaje negativo límite. Este movimiento sólo se permite durante el entrenamiento.

Para desalentar a ir por fuera de los límites del laberinto, daremos la penalidad máxima igual al puntaje negativo límite. Este movimiento sólo se permite durante el entrenamiento. Al comienzo, el RA cometerá errores, de los cuales irá aprendiendo en pos de obtener mayores recompensas.

Luego de cada episodio, la RNA aprende en base a la recompensa obtenida. El entrenamiento continuará durante todas las corridas o hasta llegar a un nivel de confianza aceptable. Se determina que un nivel de aprendizaje es aceptable, si encuentra la salida del laberinto partiendo de cualquier celda libre.

Se utilizará una RNA prealimentada multicapa. Poseerá una capa oculta, siendo 3 en total. La entrada de la red será la matriz que representa al laberinto convertida en un vector. Esto quiere decir, que la cantidad de nodos de la capa de entrada será igual a la cantidad de celdas. La capa oculta, tendrá el mismo tamaño que la capa de entrada. La capa de salida poseerá 4 nodos: uno por cada movimiento posible (arriba, abajo, izquierda y derecha). El resultado de cada nodo será el valor Q estimado para ese contexto y movimiento.

Como función de activación, utilizaremos la Función Sigmoide (S-Shaped function) [8]. De esta forma, podremos permitir a la red activar binariamente cada nodo.

Como optimizador de gradientes estocásticos, utilizaremos el método Adam. [8]

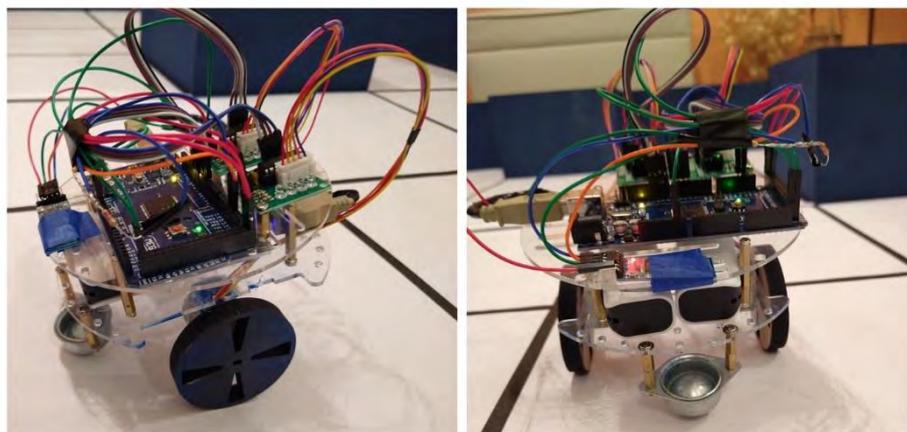
Como función de pérdida utilizaremos el Error Cuadrático Medio (Mean Squared Error) [8]. El mismo, es un estimador que mide el promedio de los errores al cuadrado.

La red será programada en el lenguaje Python 3.7.0. Se utilizaron las bibliotecas Keras 2.2.4 y Tensorflow 1.15.0, se construyó la red y se ejecutaron todas las operaciones matemáticas necesarias para predecir, aprender y calcular métricas de estos procesos.

En la Tabla 2 se presentan las características y componentes del RA, y en la Figura 1 el RA construido.

**Tabla 2.** Características y componentes del RA.

Componente	Cantidad	Detalle
Chasis	1	Estructura circular con dos niveles
Ruedas	3	Dos ruedas laterales de madera y una rueda tipo bola giratoria
Microcontrolador	1	Arduino Mega 2560
Motor	2	Motor “paso a paso” modelo 28BYJ
Controlador de Motor	2	Driver modelo ULN2003
Alimentación	2	Batería externa de 5V
Comunicación	1	Módulo Bluetooth HC-05



**Fig. 1.** Robot autónomo.

El sistema de navegación recibe la información de los movimientos a realizar y los ejecuta en el orden indicado.

El RA es capaz de realizar tres tipos de movimiento: Avanzar, Doblar a la derecha y Doblar a la izquierda.

Los motores paso a paso, como su nombre lo indica, dividen un giro completo del eje en 4096 pasos. Dicho esto, para ponerlos en funcionamiento se decide una dirección y una cantidad de pasos a ejecutar.

El movimiento de avanzar acciona los dos motores hacia delante y ejecuta 6400 pasos. Este movimiento contempla que el RA este posicionado en el medio de un cuadrado y finalice en el medio del cuadrado siguiente.

El movimiento de doblar a la derecha es un giro de 90° grados en sentido horario. Acciona el motor de la izquierda hacia delante y el de la derecha hacia atrás, ambos motores ejecutan 1800 pasos cada uno.

El movimiento de doblar a la izquierda es un giro de 90° grados en sentido antihorario. Acciona el motor de la derecha hacia delante y el de la izquierda hacia atrás, ambos motores ejecutan 1800 pasos cada uno.

#### **4 Validación de la solución**

Luego del diseño e implementación de la solución propuesta se ejecutaron diversas pruebas integradoras. Las mismas contemplaron distintos niveles de complejidad de aprendizaje, por medio de laberintos de diferentes diseños. A continuación, se presenta un caso de los tantos realizados.

#### 4.1 Caso de prueba: Laberinto simple con diversos caminos posibles

Para este caso de prueba se diseñó un laberinto con algunas celdas ocupadas dispersas, pero con la mayoría de celdas libres. Esto genera múltiples caminos incluso combinables.

El objetivo de esta prueba es que el algoritmo entrene, buscando una solución para este laberinto. Como resultado esperado, deberá elegir el camino más óptimo, de entre los tantos caminos posibles.

Luego de la ejecución y posterior análisis de los resultados obtenidos, podemos concluir que:

- El entrenamiento finalizó en el epoch 65. Este desempeño es aceptable, dado que encontró la solución óptima a los 9 minutos.
- El algoritmo pudo elegir el camino más óptimo entre tantos posibles.
- El camino elegido contiene trayectos rectos, para evitar movimientos innecesarios.
- A partir del epoch 18 del entrenamiento, el algoritmo comenzó a encontrar posibles soluciones al laberinto.
- En el epoch número 30 del entrenamiento, se alcanzó un win rate del 50%, habiendo encontrado ya 10 soluciones posibles.
- Pasados 5 minutos de entrenamiento, comienza una crecida constante de nuevas soluciones encontradas.
- Al principio del entrenamiento, la función de pérdida devuelve valores altos. Con el transcurso del tiempo, el aprendizaje se ve reflejado en la notoria disminución de la función a partir del epoch 23.
- El RA demoró 40 segundos en recorrer el laberinto.

El diseño y la solución obtenida se pueden visualizar en la Figura 2.

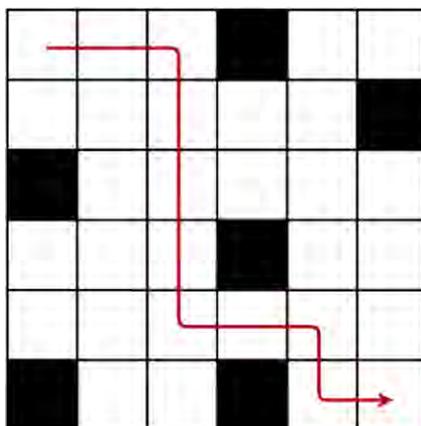


Fig. 2. Diseño y resolución del laberinto del primer caso de prueba.

## 5 Conclusiones y futuras líneas de trabajo

En este trabajo se propuso implementar un algoritmo que aprenda por refuerzo a resolver la problemática planteada, haciendo uso de un sistema de políticas y puntajes. Para modelar el problema se consideró la estructura como un laberinto y el agente junto con sus movimientos y estados como un proceso de decisión de Markov. Esto nos permitió considerar cada estado como la ubicación del agente en el escenario, y cada movimiento como una transición de estados.

Q-Learning fue la técnica de aprendizaje por refuerzo que se utilizó. La misma permitió definir una política que constituyó un conjunto de reglas que determinaron el accionar del agente durante el aprendizaje. Este aprendizaje es obtenido mediante un proceso de entrenamiento, el cual consiste en la búsqueda de una salida desde diferentes puntos de partida. Esto es ejecutado durante varias iteraciones y finalmente evaluado para comparar soluciones. Dicha evaluación está dada por la comparación de puntajes obtenidos en cada solución. El sistema de puntaje diseñado busca incentivar al agente a realizar movimientos válidos en búsqueda de la solución del camino óptimo, y desalentar movimientos inválidos como así también trayectorias innecesarias.

La implementación presentada es aplicable a cualquier escenario y adaptable a cualquier regla del entorno. Esto se puede llevar a cabo mediante la redefinición de políticas y puntajes, estableciendo una lógica de comportamiento.

Por otra parte, este trabajo de investigación no sólo estuvo enfocado a la resolución de estructuras desconocidas, sino que también a la ejecución de la solución por medio de un RA. La combinación del algoritmo de aprendizaje y su ejecución por medio del RA, demostró proveer una solución integral a la problemática con un alto grado de satisfacción. Mediante las pruebas de distintos tipos de motores y estructuras de autos, se pudo concluir que, teniendo un alto grado de precisión en los mismos, son prescindibles los diversos sensores que dan soporte adicional a la navegación.

A pesar de los avances realizados en este trabajo, en cuanto a la navegación autónoma en estructuras desconocidas, aún quedan diversos desafíos por resolver. Estos desafíos contemplan desde etapas de reconocimiento del escenario hasta el diseño del RA adaptable a las necesidades del problema. A continuación, se enumeran las diferentes líneas de trabajo futuro surgidas de estos desafíos:

- Reconocimiento del escenario por parte del RA, creando un mapa del entorno que servirá como entrada al proceso de aprendizaje.
- Evaluación de las condiciones físicas del entorno para detectar posibles riesgos. Esto se puede llevar a cabo, incorporando diversos sensores como por ejemplo, detector de llama, sensor de temperatura, sensor de humo, etc.
- Transmisión de video y audio en tiempo real del escenario.
- Implementar un sistema de visión artificial con el fin de detectar posibles víctimas.
- Desarrollar un sistema escalable de coordinación de trabajo en conjunto entre varios RA para dividir tareas.
- Ampliar el rango de comunicación entre la aplicación y el RA para operar a mayor distancia.

## 6 Referencias

1. Cui, X., & Shi, H. (2011). *A\*-based Pathfinding in Modern Computer Games*. IJCSNS International Journal of Computer Science and Network Security, 11(1), 125-130.
2. Kitchenham, B., Brereton, P., & Budgen, D. (2015). *Evidence-Based Software Engineering and Systematic Reviews*. USA: Chapman and Hall 1 st. Editon. Chapman and Hall/CRC.
3. Costa I., De Leo E., Panizzi M., Sattolo I. Anexo - Navegación autónoma mediante aprendizaje por refuerzo. <https://doi.org/10.6084/m9.figshare.12676685.v1>
4. Hamdy, T. (2004). *Investigacion de operaciones*. México: Pearson Educación.
5. Nazia, H. (2019). Hands-On Q-Learning with Python: *Practical Q-learning with OpenAI Gym, Keras and Tensorflow*. Reino Unido: Packt.
6. Sean, S., Yang, W., & Rajalingappaa, S. (2018). *Python Reinforcement Learning Projects: Eight hands-on projects exploring Reinforcement learning algorithms using TensorFlow*. Reino Unido: Packt.
7. Ameet, J. (2020). *Machine Learning and Artificial Intelligence*. EEUU: Springer.
8. Sharma, A., Ravi Vishwesh, S., & Beyeler, M. (2019). *Machine Learning for OpenCV 4: Intelligent algorithms for building image processing apps using OpenCV 4, Python, and scikit-learn*. Reino Unido: Packt.

# Manejo de Diversidad en CHC aplicado a la optimización del costo energético en parques eólicos

Martin Bilbao<sup>1</sup>, Guillermo Leguizamón<sup>2</sup>

<sup>1</sup> Universidad Nacional de la Patagonia Austral, Caleta Olivia, Argentina  
mbilbao@uaco.unpa.edu.ar

<sup>2</sup> Universidad Nacional de San Luis, Argentina  
legui@unsl.edu.ar

**Resumen** En este trabajo analizaremos diferentes mecanismos de manejo de diversidad para el algoritmo CHC (Crossover elitism population, Half uniform crossover combination, Cataclysm mutation) para resolver problemas de optimización en parques eólicos de energía. El algoritmo CHC convencional contiene un mecanismo de reinicio poblacional aleatorio, esto conlleva a la posibilidad de perder cierto conocimiento adquirido si no se maneja adecuadamente. Es por ello que estudiaremos otros mecanismos de reinicio poblacional que tengan en cuenta el conocimiento adquirido durante su evolución para intentar lograr mejor convergencia en los resultados. El objetivo final es minimizar el costo del KWh analizando tres variantes de reinicio poblacional y cómo impactan en los resultados finales con respecto a la versión de CHC convencional.

**Keywords:** Optimización, Energía eólica, Metaheurísticas, Inteligencia Artificial

## 1. Introducción

El viento es un recurso de la naturaleza que ha acompañado a la humanidad desde sus orígenes. Si bien el interés de utilizar el recurso del viento para producir energía se dio debido a la crisis de petróleo vivida en el 1973 [1], tiene grandes soportes para ser desarrollado e implementado. Uno de ellos es el aspecto económico ya que utilizamos un recurso natural para producir energía y el otro es una cuestión social y ambiental que están descritas en el protocolo de Kioto impulsado por las naciones unidas [2] y el libro blanco de energía impulsado por la unión europea [3].

Uno de los aspectos más importantes para la diseño de un parque eólico, es la forma de distribuir los molinos en diferentes tipos terreno, ya que deben ser ubicados de tal manera que no produzca interferencia con los demás y de forma que se obtenga la mayor eficiencia del mismo. Las cuestiones técnicas de distribución de molinos eólicos se pueden estudiar en los siguientes trabajos [4].

En el estado del arte se han desarrollado muchos trabajos relacionados a la temática en cuestión. Diversas propuestas han sido consideradas según las características de la función objetivo y la existencia de restricciones [6-7]. Por otro lado los enfoques metaheurísticos se presentan como una alternativa altamente viable, no sólo por su robustez y flexibilidad, sino por la calidad de los resultados encontrados en tiempos razonables [8-10] en donde aplicaron Algoritmos Genéticos para diferentes instancias de viento en un parque eólico.

Otros trabajos más recientes son los aportes de Kusiak et al. [12], Wanger et al. [13-14] y Wilson et al. [15-16] donde aplican enfoques multiobjetivos, estrategias evolutivas y modelos distribuidos de algoritmos genéticos para la resolución de instancias más complejas con restricciones. Para una reseña más completa de publicaciones relacionadas al diseño de parques eólicos de energía se recomienda el siguiente trabajo [17].

El resto del artículo se divide de la siguiente manera: en la sección 2 se explicará el modelo energético utilizado, así como los modelos de potencia, costos y función objetivo del problema, en la sección 3 se explicará el algoritmo CHC y las variantes de reinicio poblacional, en la sección 4 se mostrarán los experimentos realizados para los casos de estudios propuestos y en la sección 5 las conclusiones y trabajos futuros.

## 2. Modelado del problema de optimización de parques eólicos

El Wake Free Ratio es un parámetro determinante en el diseño de un parque eólico ya que nos permite saber la utilidad máxima obtenida del recurso eólico por  $n$  molinos en función al máximo teórico posible. Se sabe que, por el efecto estela [5] que produce un molino al estar próximo a otro, se genera automáticamente una reducción de la velocidad del viento, por ende una reducción de la energía producida (Wake Loss Energy). Para calcular el Wake Free Ratio se utiliza la Ecuación 1:

$$wfr = \frac{E_{tot}}{E_{wf} * n}, \quad (1)$$

siendo  $E_{tot}$  la energía total generada en el parque eólico teniendo en cuenta pérdidas por obstáculos e interferencias (Ecuación 5),  $E_{wf}$  la energía teórica posible que se puede obtener del parque eólico sin tener en cuenta pérdidas de ningún tipo y  $n$  es la cantidad de aerogeneradores que tiene instalado el parque eólico en el momento del cálculo.

El costo de producir un KWh es un factor importante para el mercado económico y la viabilidad del proyecto de construcción del parque eólico. El cálculo del KWh está asociado al Wake Free Ratio obtenido anteriormente y a la cantidad de molinos instalados, como se puede observar en la Ecuación 2:

$$costo_{kwh} = \frac{(c_t \cdot n + c_s \cdot \lfloor (\frac{n}{m}) \rfloor) (\frac{2}{3} + \frac{1}{3} \cdot e^{-0.00174n^2}) + c_{OM} \cdot n}{\frac{(1-(1+r)^y)}{r}} \cdot \frac{1}{8760 * E_{tot}} + \frac{0,1}{n} \quad (2)$$

En este caso,  $n$  y  $E_{tot}$  son las variables de la instancia, donde  $n$  es la cantidad de aerogeneradores de la configuración y  $E_{tot}$  es la cantidad de energía producida por el parque eólico, que despejada de la ecuación 3 se obtiene lo siguiente:

$$E_{tot} = wfr * E_{wf} * n. \quad (3)$$

Los demás parámetros son  $c_t$ , el costo de una turbina, desde la creación hasta su instalación,  $c_s$ , el costo de subestación,  $m$ , la cantidad de molinos que entran por subestación,  $r$ , la tasa de interés anual,  $c_{OM}$ , el costo de operación y mantenimiento anual por molino e  $y$ , el tiempo de vida del parque eólico en años.

## 2.1. Modelos eólicos utilizados

Los siguientes modelos eólicos reales serán utilizados para evaluar el comportamiento de las técnicas evolutivas. Dichos elementos se describen a continuación.

La distribución de Weibull o rosa de los vientos utilizada para el escenario se describe mediante la tabla 1 donde se encuentran los parámetros de forma ( $c$ ) y escala ( $k$ ), el factor de forma  $\omega$  (probabilidad de que ocurra dicha velocidad de viento), como así también la dirección del viento en grados para todo el terreno.

$i-1$	$\theta_{i-1}$	$\theta_i$	$k$	$c$	$\omega_{i-1}$	$i-1$	$\theta_{i-1}$	$\theta_i$	$k$	$c$	$\omega_{i-1}$
0	0	15	2.0	7.0	0.0002	12	180	195	2.0	10.0	0.1839
1	15	30	2.0	5.0	0.0080	13	195	210	2.0	8.5	0.1115
2	30	45	2.0	5.0	0.0227	14	210	225	2.0	8.5	0.0765
3	45	60	2.0	5.0	0.0242	15	225	240	2.0	6.5	0.0080
4	60	75	2.0	5.0	0.0225	16	240	255	2.0	4.6	0.0051
5	75	90	2.0	4.0	0.0339	17	255	270	2.0	2.6	0.0019
6	90	105	2.0	5.0	0.0423	18	270	285	2.0	8.0	0.0012
7	105	120	2.0	6.0	0.0290	19	285	300	2.0	5.0	0.0010
8	120	135	2.0	7.0	0.0617	20	300	315	2.0	6.4	0.0017
9	135	150	2.0	7.0	0.0813	21	315	330	2.0	5.2	0.0031
10	150	165	2.0	8.0	0.0994	22	330	345	2.0	4.5	0.0097
11	165	180	2.0	9.5	0.1394	23	345	360	2.0	3.9	0.0317

**Tabla 1.** Distribución de vientos para el escenario de estudio

El modelo de obtención de energía está basado en la generación de energía de cada aerogenerador, la distribución de Weibull anteriormente descrita, el

modelo de estela y la interferencia con otros aerogeneradores y la orografía del terreno. Para un aerogenerador simple, la energía producida se calcula como muestra la ecuación 4:

$$E_{tot} = \int_{\theta} P(\theta) \int_v p_v(v(\theta), c(\theta), k(\theta)) \beta^i(v) \quad (4)$$

Donde  $\beta^i(v)$  es la curva de potencia del aerogenerador utilizado en función a una velocidad de viento  $v$  dado,  $p_v(v(\theta), c(\theta), k(\theta))$  es la distribución de probabilidad de Weibull (Tabla 1) en función a la velocidad del viento, el factor de forma y de escala, y  $\int_{\theta} P(\theta)$  es la función de dirección del viento, el cual varía entre  $0^\circ$  y  $360^\circ$ .

Para calcular la interferencia entre molinos utilizaremos el modelo de estela propuesto por Park [11] que calcula las velocidades de viento en los puntos interiores a la estela y los nuevos parámetros de forma y escala, ya que ellos también se ven afectados, es por ello que la fórmula para la obtención de energía debe adaptarse al modelo de estela como muestra la siguiente ecuación 5:

$$E^{farm} = \left( \sum_i \int_{\theta} P(\theta) \int_v p_v^\theta(v, c_i(\theta), k_i(\theta), x_i, y_i, X, Y) \beta^i(v) \right) \quad (5)$$

El procedimiento para calcular el efecto del modelo estela entre los aerogeneradores utilizando el modelo Park se muestra en el Algoritmo 1

---

#### Algoritmo 1 Modelo de estela Park

---

**Require:** dado  $X, Y$  la posición de un aerogenerador,  $C_t$  thrust coefficient y  $k$  el factor de escala;

$a \leftarrow 1 - \sqrt{1 - C_t}$ ;

$b \leftarrow \frac{k}{R}$ ; /\*  $R$  es el rotor del aerogenerador \*/

$u \leftarrow \text{unitstepfunction}$ ;

$o \leftarrow (x_i - x_j) \cos(\theta) + (y_i - y_j) \sin \theta$ ;

$d_{i,j} \leftarrow \|o\|$ ;

$\alpha \leftarrow \tan^{-1} k$ ;

**para**  $i = 1$  hasta  $n$  **hacer**

**para**  $\theta = 0^\circ$  hasta  $360^\circ$  **hacer**

**para**  $j = 1$  hasta  $n - 1$  and  $j \neq i$  **hacer**

$\delta_{i,j} \leftarrow \cos^{-1} \left( \frac{o + \frac{R}{k}}{\sqrt{(x_i - x_j + \frac{R}{k} \cos(\theta))^2 + (y_i - y_j + \frac{R}{k} \sin(\theta))^2}} \right)$ ;

$Vdef_{i,j} \leftarrow u(\delta_{i,j} - \alpha) \frac{a}{(1 + b d_{i,j})^2}$ ;

**fin para**

$Vdef_i^\theta \leftarrow \sqrt{\sum_j (Vdef_{i,j})^2}$ ;

$c_i(\theta) \leftarrow c_i(\theta) \times (1 - Vdef_i)$ ;

**fin para**

**fin para**

---

La curva de potencia del aerogenerador es una función que muestra la salida de potencia obtenida en función a la velocidad del viento incidente en el aerogenerador. La misma consta de tres parámetros importantes como la velocidad inicial  $v_{in}$ , la velocidad nominal  $v_{rated}$  y la velocidad de corte  $v_{cut}$ . La fórmula que determina la obtención de energía del aerogenerador se puede observar en la Ecuación 6:

$$P_i(kW) = \begin{cases} 0 & \text{si } v_x < v_{in}, \\ \lambda \cdot v + \eta & \text{si } v_{in} \leq v_x \leq v_{rated}, \\ P_{rated} & \text{si } v_{rated} \leq v_x \leq v_{cut}, \\ 0 & \text{si } v_{cut} < v_x \end{cases} \quad (6)$$

Para esta instancia se ha dispuesto resolver el problema de optimización, que tiene como característica a considerar, la cantidad de aerogeneradores máximos posibles a utilizar (100 en este caso), la energía máxima posible que se puede obtener (7315.27 kw), las dimensiones ancho y largo en metros del terreno (3000 x 3000 metros) sin obstáculos considerados y la distribución de vientos considerada en la Tabla 1.

- Objetivo: Minimizar el costo de producción del Kw
  1. Maximizar el número de aerogeneradores dados por la instancia.
  2. Minimizar el costo de producción del Kw (fitness).

### 3. Algoritmo CHC y Variantes Propuestas

El algoritmo CHC [18] se ha comportado de manera eficiente en trabajos presentados anteriormente [19-21], es por ello que resulta interesante observar su desempeño en instancias reales obtenidas de simuladores de parques eólicos y analizar en terrenos complejos con funciones objetivos realísticas.

CHC es un algoritmo elitista que al seleccionar los padres mediante algún mecanismo, genera el cruzamiento siempre y cuando los individuos sean lo suficientemente diferentes. Para ello se mide la distancia de Hamming entre ellos y si es mayor a un umbral definido, se efectúa el cruce mediante un mecanismo llamado HUX (Heterogeneous Uniform Crossover), que consiste en mantener en sus descendientes las posiciones comunes entre los padres y a las posiciones diferentes, asignársela de manera equitativa, intentando mantener la diferencia entre ellos (diversidad). Esto es importante ya que CHC no tiene mecanismo de mutación directo, sino que mantiene en parte la diversidad generando hijos diferentes en la cruce.

En caso de que la población no logre mejorar porque se produce un estancamiento ante la falta de diversidad, dada una serie de iteraciones consecutivas, se produce un cataclismo en la población con el objetivo de introducir nuevos individuos y así lograr la diversidad necesaria para salir del estancamiento. En

su versión canónica, dicho reinicio poblacional consta de mantener el mejor individuo y formar la población restante tomando como plantilla la mejor solución encontrada y perturbarla bit a bit con una probabilidad de 35 %.

En nuestro caso hemos propuesto tres variantes al mecanismo de reinicio poblacional dado a que en estudios previos se ha observado que este mecanismo es muy importante para la convergencia del algoritmo ya que controla y mantiene el conocimiento adquirido por el problema durante su etapa de exploración.

- **MACHC (Modelo Aleatorio CHC)**: Reinicio Poblacional preservando el mejor individuo y los demás se generan en forma aleatoria.
- **MSCHC (Modelo Sesgado CHC)**: Reinicio Poblacional con probabilidad de mutación estática equivalente al sesgo de la distribución de energía en las posiciones factibles del parque eólico.
- **MECHC (Modelo Enriquecido CHC)**: Reinicio Poblacional con probabilidad dinámica equivalente a la diversidad genética de cada bit en función al mejor individuo, es decir que cada bit del mejor individuo se mantiene en los siguientes individuos con una probabilidad equivalente a la diversidad genética de ese bit a nivel poblacional.

La propuesta es resolver el escenario con cada una de las variantes mencionadas anteriormente y luego analizar la influencia del manejo de diversidad en los resultados obtenidos.

#### 4. Experimentos realizados

En la siguiente sección, se resolverá una instancia real diseñada con un simulador comercial de diseño de parques eólicos, donde se estudian varios mecanismos de reinicio poblacional con el objetivo de minimizar el costo de kWh. Para comparar los resultados obtenidos con el algoritmo CHC se ha utilizado un algoritmo genético canónico, que resolverá las mismas instancias con los mismos parámetros para que sean totalmente comparables y poder realizar luego un análisis estadístico de los resultados. Para ello se han definido los siguientes parámetros de configuración para el aerogenerador: Potencia nominal = 1500 Kw, R (diámetro Rotor) = 38,5 metros,  $\lambda$  (pendiente) = 140,86,  $\eta = -500$ ,  $v_{in} = 3,5$  m/s,  $v_{out} = 20$  m/s y  $v_{rated} = 14$  m/s.

Para este escenario, se han realizado 2.000 evaluaciones que corresponden a 100 iteraciones sobre el algoritmo, y se ejecutaron 30 corridas independientes para luego comparar cada una de las instancias mediante un análisis estadístico.

Como objetivo principal, lo que se intenta optimizar es el costo de producción del Kw. Dicha función depende del Wake Free Ratio calculado anteriormente y de ciertos parámetros económicos que fueron tenidos en cuenta en la viabilidad del proyecto. La función tiene dos variables confrontadas: cantidad de turbinas

y costo del kWh. La ecuación 2 se utilizará como función de fitness a minimizar. En este caso se utilizarán los tres modelos de reinicio poblacional al momento de detectar un estancamiento en la población. Todas las propuestas serán comparadas con un algoritmo genético.

Los demás parámetros son constantes y pertenecen a diferentes costos en dólares del modelado del parque eólico,  $c_t$  es el costo de una turbina = 750.000,  $c_s$  el costo de la subestación = 8.000.000,  $m$  la cantidad de turbinas por subestación = 30,  $r$  la tasa de interés anual = 0,03,  $c_{OM}$  el costo de operación y mantenimiento por turbina = 20.000 e  $y$  el tiempo de vida útil del campo (años) = 20.

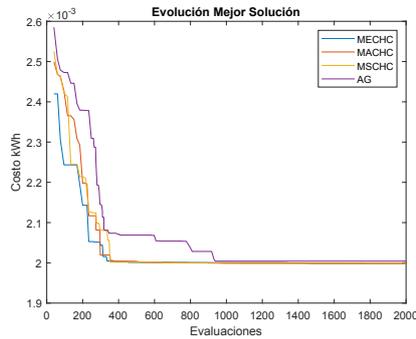
Los resultados más relevantes se muestran en la tabla 2 donde podemos observar los valores promedios de las 30 ejecuciones del escenario analizado. Como se puede ver MECHC obtuvo un costo promedio de 0,0020 dólares con un esfuerzo computacional de aproximadamente un cuarto del total de evaluaciones. La cantidad de molinos de la configuración encontrada es de 89 molinos.

Algoritmo	Costo	DvSt.	Evaluaciones	Generaciones	Cant.Turbinas
AG	0,0024122	+/- 0,001	1.968	98,50	79,15
MACHC	0,0022446	+/- 1,67E-05	1.468	67,90	80,10
MSCHC	0,0021946	+/- 0,0002	1.839	94,46	85,93
MECHC	0,0020505	+/- 1,65E-05	487	10,66	89

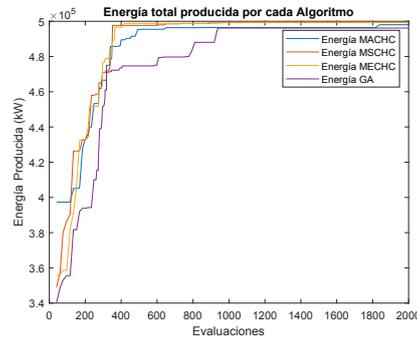
**Tabla 2.** Resultados Escenario Propuesto

En la figura 1 se puede ver cómo varía el costo energético de la mejor solución encontrada para ambos algoritmos. Se visualizan grandes mejoras en la función al inicio de la ejecución, luego de forma escalonada se logran avances importantes luego de la mitad de evaluaciones. Para el caso de ME-CHC la mejora fue continua inclusive al final de las evaluaciones. AG logró una mejora importante hasta la mitad de evaluaciones y luego no logró converger a una solución mejor. Tanto MACHC como MSCHC tuvieron un comportamiento similar, un poco inferior a MECHC.

En la figura 2 se puede observar la variación de energía a medida que avanzan las evaluaciones de todos los algoritmos. El análisis realizado es interesante ya que se puede ver que hay una variación importante en la energía producida manteniendo la misma cantidad de molinos, esto es debido a que existen muchas maneras de distribuir la misma cantidad de molinos en el parque y cada una de ellas tendrá un efecto importante en el Wake Free Ratio, por ende variará la energía real obtenida. En el caso de MECHC y MSCHC se ha obtenido un valor de energía de 499.560, 35kW, para MACHC se obtuvo 498.186, 14kW y GA 496.128, 57kW. Se puede observar la variación energética a medida que avanzan



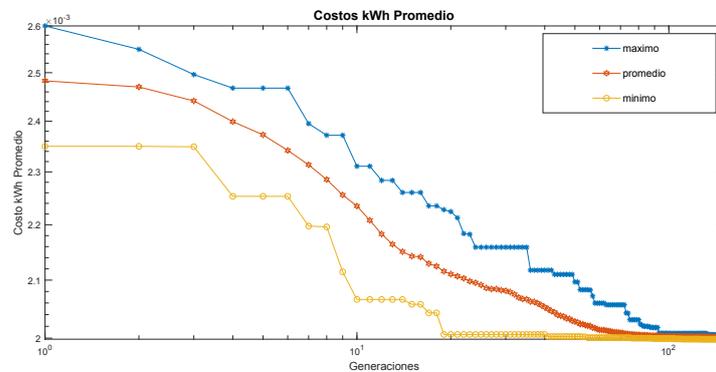
**Figura 1.** Evolución del costo energético Kwh para los 4 algoritmos



**Figura 2.** Energía Producida para los 4 algoritmos

las evaluaciones sin modificar la cantidad de molinos. Es muy importante tener en cuenta la relación que existe entre el Wake Free Ratio y la cantidad de energía producida.

MECHC ha tenido el mejor comportamiento en esta instancia. La forma de manipular el reinicio poblacional generando diversidad con conocimiento del problema fue clave para una mejor convergencia. En la figura 3 se observan los promedios del mejor fitness, los valores máximos y mínimos de las 30 ejecuciones.



**Figura 3.** Valores promedios para las 30 ejecuciones de MECHC

Para poder comprobar si existen diferencias significativas entre cada una de las instancias evaluadas, realizaremos el análisis estadístico *Kruskal Wallis*, con un  $\alpha = 0,05$ , para determinar si existen diferencias entre todas las instancias. Luego utilizaremos *Wilcoxon de a pares* para analizar en grupos donde existe

tal diferencia. De esta manera podremos realizar un ranking para determinar un orden de mayor a menor en cuanto al desempeño del operador de reinicio poblacional. Para este análisis se utilizará la función objetivo costo como hipótesis. Al realizar el test de kruskal wallis, el p-value da un valor igual a 0,0028, como es menor al valor de significancia  $\alpha$ , existen diferencias significativas en el desempeño de los tres casos analizados. Aplicando wilcoxon de a pares se obtiene lo siguiente:

- MACHC-MSCHC: p-value = 0,10 por lo tanto no existe diferencia significativa en sus desempeños
- MSCHC-MECHC: p-value = 0,021 por lo tanto existe diferencia significativa en sus desempeños, en este caso MECHC tuvo mejor desempeño
- MECHC-MACHC: p-value = 0,018 por lo tanto existe diferencia significativa en sus desempeños, en este caso MECHC tuvo mejor desempeño

Por lo tanto, se concluye que, en función al análisis estadístico, el mecanismo de reinicio poblacional MECHC ha mejorado el desempeño del algoritmo CHC, ya que los nuevos individuos, luego de un reinicio poblacional, se crean en función al conocimiento que tiene de cada una de las posiciones en particular dentro del campo eólico. Si es una posición interesante, tendrá una diversidad mayor en esa posición a nivel poblacional.

## 5. Conclusiones

Se puede observar que CHC se ha comportado correctamente ante instancias más realistas, pudiendo inclusive superar al algoritmo genético propuesto. Se observó que al manipular de forma eficiente la diversidad utilizando el conocimiento adquirido del problema mejoró notablemente los resultados obtenidos. MECHC utiliza la matriz energética del modelo teórico como base para la adquisición del conocimiento en las posiciones de mayor potencial eólico, dado que existen mejores vientos para aprovechar lo que se traduce en mayor concentración de energía capturable en la zona. Como trabajo futuro se propone trabajar con un dominio continuo en el espacio de búsqueda donde los molinos puedan ser ubicados en cualquier posición factible del campo eólico.

## Referencias

1. D. Andersen, *Review of historical and modern utilization of wind power*, Risø-R-1205, 2007, disponible online en <http://www.risoe.dk/rispubl/VEA/dannemand.htm>
2. Naciones Unidas, *Protocolo de Kioto*, disponible online en <http://unfccc.int/resource/docs/convkp/kpspan.pdf>.
3. Comisión de la Unión Europea, *Libro blanco de la energía*, disponible online en <http://europa.eu.int/coo/of/white/>

4. J.F. Walker, N. Jenkins, *Wind energy technology*, John Wiley and Sons, Ltd./Inc., New York, 1997.
5. I. Katic, J. Hojstrup and N. O. Jensen, *A Simple Model for Cluster Efficiency*, European Wind Energy Association Conference and Exhibition, Rome-Italy, pp. 407-410, 7-9 October 1986.
6. S. Donovan *An improved mixed integer programming model for wind farm layout optimization*. YPP entry, 2008.
7. C.N. Elkinton J.F. Manwell & J.G.McGowan, *Algorithms for Offshore Wind Farm-Layout Optimization*. Wind Engineering, Volume32, Number 1, pp. 67-84(18). MultiScience Publishing Co Ltd. 2008.
8. S. A. Grady, M. Y. Hussaini, M.M. Abdullah *Placement of wind turbines using genetic algorithms Renewable Energy*, Volume 30, Issue 2, pp. 259-270. 2005.
9. J.C. Mora, J.M. Calero Barón, J.M. Riquelme Santos, M. Burgos Payán *An evolutive algorithm for wind farm optimal design*. Neurocomputing, Volume 70 , Issue 16-18, pp. 2651-2658, October 2007.
10. H.S. Huang, *Distributed Genetic Algorithm for Optimization of Wind Farm Annual Profits*. Intelligent Systems Applications to Power Systems, International Conference on Volume , Issue , 5-8 ,Page(s):1 - 6. 2007.
11. S. Frandsen, *On the wind speed reduction in the center of large clusters of wind turbines*, Proceedings of the European Wind Energy Association Conference and Exhibition EWEC 91 Amsterdam, Netherlands. pp. 375-80. 1991.
12. A.Kusiak and Z.Song. *Design of wind farm layout for maximum wind energy capture*. Renewable Energy,35(3):685-694, 2010.
13. M.Wagner, K.Veeramachaneni, F.Neumann, and Una-May O'Reilly. *Optimization the Layout of 1000 Wind Turbines*. In European Wind Energy Association Annual Event, 2011.
14. M.Wagner, J.Day, F.Neumann. *A Fast and Effective Local Search Algorithm form Optimization the Placement of Wind Turbines*. Renewable Energy, 51(0):64-70,2013
15. D.Wilson, S.Cussat-Blanc, K.Veeramachaneni, Una-May O'Reilly and E.Awa. *On Learning to generate wind farm layouts*. In Proceedings of the fifteenth Annual Conference on Genetic and Evolutionary Computation Conference, ACM,2013
16. D.Wilson, S.Cussat-Blanc, K.Veeramachaneni, Una-May O'Reilly and Hervé Luga. *A Continuous developmental model for wind farm layout optimization*. In Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, pages 745-752, ACM,2014.
17. J. Serrano González, M. Burgos Payán, J. M. Riquelme Santos, F. González-Longattb *A review and recent developments in the optimal wind-turbine micro-siting problem*, Renewable and Sustainable Energy Reviews, <https://doi.org/10.1016/j.rser.2013.09.027>, 2014.
18. L.J. Eshelman. *The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination*. In Foundations of Genetic Algorithms,1, pages 265-283. Morgan Kaufmann, 1991.
19. M. Bilbao, E. Alba, *CHC and SA applied to wind energy optimization using real data*, CEC 2010, Barcelona-España, DOI: 10.1109/CEC.2010.5586395. 2010.
20. M. Bilbao, E. Alba, D. Pandolfi, *Metaheurísticas aplicadas al diseño de parques eólicos de energía*, SELASI,Venezuela, p. 35-42, ISBN: 978-980-7630-00-9. 2013.
21. M. Bilbao, G. Leguizamón , *Multicore Parallelization of CHC for Optimal Aerogenerator Placement in Wind Farms*, CACIC, Córdoba, Argentina, p. 85-94.ISBN: 978-987-688-377-1. 2019.

# Particle Swarm Optimization with Adaptive Inertia Weight using Fuzzy Logic for Large-Scale Problems

Fabiola-Patricia Paz<sup>a,1</sup>, Guillermo Leguizamón<sup>a,2</sup>, Efrén Mezura-Montes<sup>b,3</sup>

<sup>a</sup>LIDIC, Universidad Nacional de San Luis, Ejército de Los Andes 950, 5700, San Luis, ARGENTINA - CONICET

<sup>b</sup>Artificial Intelligence Research Center, University of Veracruz, Sebastian Camacho 5, Centro, Xalapa, Veracruz, 91000, MEXICO

<sup>1</sup>fabypaz@gmail.com.ar, <sup>2</sup>legui@unsl.edu.ar, <sup>3</sup>emezura@uv.mx

**Abstract.** In this paper an alternative approach is proposed to improve the convergence of Particle Swarm Optimization (PSO) algorithm by adapting the inertial weight parameter with a fuzzy logic system to solve large-scale optimization problems. The PSO algorithm is a population-based metaheuristic inspired by the social behavior of birds, and it has been applied to numerous optimization problems successfully. However, one of its main disadvantages is the decaying performance when applied to complex and large-scale problems. The proposed algorithm uses the fuzzy system to dynamically calculate a value of the Inertia Weight parameter during the search process to find better solutions. After carrying out experiments on a well-known benchmark for large-scale optimization, the proposed approach provides a competitive performance.

**Keywords:** Fuzzy logic, Particle Swarm Optimization, convergence control, adaptive inertia weight.

## 1 Introduction

Particle Swarm Optimization (PSO) is an algorithm based on Swarm Intelligence theory, and inspired by the social behavior of certain animals when they interact with another of their same species to achieve a common goal. It was proposed by Kennedy and Eberhart in 1995 and was developed to emulate the cooperative work in bird flocks [1] [2]. Since its original version, PSO has attracted the researchers interest mainly due to its competitive performance in complex search spaces and also its simplicity to code it. However, as other swarm intelligence algorithms, PSO may suffer from premature convergence, especially in complex or large-scale problems [20] [15]. One of the main reasons for this undesired behavior is due to the rapid movement of particles when exchanging information, leading to a low diversity of the swarm in those initial iterations. In recent years, several techniques have been proposed to avoid PSO's premature convergence [3]. In the literature, fuzzy logic has been successfully used to improve the performance of swarm intelligence algorithms. Some published works that were of interest for this work are the following: Olivas et al. [7] [8] [14] implemented PSO and ant colony optimizer (ACO), Pérez et al. [6] the Bat algorithm, Sombra et al. [9] the gravitational search algorithm, Valdez et al. [10] and [11] a set of algorithms such as PSO, genetic algorithm (GA), and ant colony optimization (ACO), Norouzzadeh et al. [12] used PSO, Ochoa et al. [13] adopted differential evolution (DE) and Kumar et al. [16] PSO and other meta-heuristics. However, most of the before mentioned references solved small-scale problems (less than 500 dimensions). On the other hand, large-scale global optimization (LSGO) problems are challenging because the search space grows exponentially as the number of dimensions increases.

Motivated by the above mentioned, in this paper we propose two variants of the PSO algorithm using a fuzzy logic system to solve LSGO problems seeking high-quality solutions with a low computational cost. A set of popular benchmark problems widely used in the literature is solved in this work as they represent different characteristics of real-world instances [17] [20]. Another contribution of the proposal is that two

blurred systems are designed to control the inertia weight with a different approach. The iteration number and the swarm diversity are defined as input variables, while the inertia weight and a variable for resetting a part of the swarm are the output variables. Both, input and output variables are combined into two fuzzy systems designed exclusively to handle LSGO. A comparative study between the systems is then carried out to analyze the quality of the results.

This work is organized as follows: Section 2 presents the Inertia weight PSO algorithm and the Constriction Factor approach. In Section 3, the fuzzy system and its fundamental elements are defined. Section 4 details the parameter control based on fuzzy logic for the PSO. In Section 5 the experiments are conducted. Finally, Section 6 presents the conclusions obtained.

## 2 Particle Swarm Optimization

The PSO algorithm is initialized with a set of randomly generated candidate solutions, called particles, within a D-dimensional search space. Each particle  $x_i(t)$  represents a position within the search space, and has its assigned velocity vector  $v_i(t)$  and its best position  $p_{best_i}(t)$  found so far. The best particle of the whole swarm, based on the fitness function value, is called ( $g_{best}(t)$ ). A particle flies to a new position by using Equation (1), called velocity update, and Equation (2), known as flight operator:

$$v_i(t+1) = w * v_i(t) + c_1 * r_1 * (p_{best_i}(t) - x_i(t)) + c_2 * r_2 * (g_{best}(t) - x_i(t)) \quad (1)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (2);$$

where  $v_i$  is the velocity of particle  $i$ ,  $x_i$  is the current position of particle  $i$ ;  $r_1$  and  $r_2$  are random numbers with uniform distribution between  $[0,1]$ ;  $p_{best_i}$  is the best position found by particle  $i$ ; and  $g_{best_i}$  is the best particle of the swarm. The cognitive and social learning coefficients are represented by  $c_1$  and  $c_2$ , respectively. These values are generally constant, but can also be dynamic.  $w$  is a static or dynamic inertia weight [1] [2] [15]. The incorporation of  $w$  in the original PSO algorithm allows a better control of the current particle velocities. With this inertia weight, the exploration and exploitation capabilities can be influenced. A higher inertia weight facilitates exploration, while a lower inertia weight favors exploitation. An appropriate  $w$ -value provides a balance in the search for the best solutions and helps to reduce the computational cost measured by the number of solution evaluations [15].

Another approach used in PSO variants is the constriction factor proposed by Clerc [19]. The velocity equation is defined in Equation (3):

$$v_i(t+1) = X * v_i(t) + c_1 * r_1 * (p_{best}(t) - x_i(t)) + c_2 * r_2 * (g_{best}(t) - x_i(t)) \quad (3)$$

and

$$X = \frac{2}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|},$$

where  $\varphi = c_1 + c_2$  y  $\varphi \geq 4$ . The parameter  $X$  controls the extent of exploration and exploitation. A  $X$  value close to 0 indicates a high degree of exploitation, whereas a value close to 1, a higher degree of exploration is obtained. This value can be constant or can change through the search process, usually with high values in the first iterations and low values late in the search.

### 3 Fuzzy Logic System

Fuzzy Logic (FL) is based on the fuzzy set theory proposed by Lotfi Zadeh in 1965. This approach forms the basis for approximate human reasoning using fuzzy if-then rules as a tool that allows each object to be assigned a degree of membership in one class or another in a smooth and flexible way [4] [5]. Fuzzy inference is the process of obtaining an output value for an input value using fuzzy set theory. There are two types: the Mamdani model and the TSK model (Takagi, Sugeno and Kang). The one proposed by Mamdani in 1975, could be described in four main steps. The first step is the *Fuzzification* of Input Variables, which consists of taking the values of the variables and determining the degree of membership in the fuzzy sets. The next step is the rule *evaluation*. The usual knowledge representation in fuzzy terms is done by **IF A THEN B** rules, where A is the antecedent and B the consequent. If a rule has multiple antecedents, the AND or OR operator is used to obtain a single number that represents the result of the evaluation. This result is applied to the consequent by means of an implication operator. The output *aggregation* consists in the unification of all the output membership functions, combining them to obtain a single fuzzy set. Finally, *defuzzification* is the final result of the system [21].

### 4 Control of Parameters through Fuzzy Logic

A fuzzy system needs to know the condition of the swarm in order to detect swarm stagnation. Inspired by [6] [7] [8] [10] [11] [14], the first input variable is *Iteration*, defined in Equation (4) and its value varies between 0 and 1. It can be understood as the progress of the optimization process.

$$Iteration = \frac{iteration_{current}}{Maximum\ of\ iterations} \quad (4)$$

The second input variable is Diversity. Diversity measures the distances between the particles ( $x_i$ ) with respect to the best particle ( $G_{best}$ ) of the swarm at each generation. It is calculated using the Euclidean Distance defined in Equation (5), where  $ns$  is the number of particles and  $nx$  is a dimension of particle  $i$ . The variables  $minDiv$  and  $maxDiv$  represent the minimum and maximum diversity respectively and are used to normalize diversity at each generation. That is, when  $DivNorm$  has a value close to 0 the diversity will be low, otherwise, the diversity will be high [7] [8].

$$Diver = \frac{1}{ns} \sum_{i=1}^{ns} \sqrt{\sum_{d=1}^{nx} (x_{i,d} - G_{best,d})^2}$$

$$DivNorm = \begin{cases} 0, & \text{when } minDiv = maxDiv \\ \frac{Diver - minDiv}{maxDiv - mindiv}, & \text{otherwise} \end{cases} \quad (5)$$

Both, *Iteration* and *Diversity* variables have been used in different approaches [6] [7] [8] [10] [11] [14], and have showed to be suitable indicators for examining swarm conditions during the search process. Therefore, they are adopted in this work.

The objective of the fuzzy system is to control the Inertia Weight to improve particle velocities. Shi et al. [15] use an inertia weight ( $w$ ) that decreases linearly during the iterations and achieved an improved performance in several applications. Eberhart et al. [18] show that  $w$  can be configured to be equivalent to the constriction factor equations, then defining the constriction factor as a special case of the inertia weight. Based on these works, an adaptive inertia weight with a constriction factor approach is proposed as an output variable, where  $c_1$  and  $c_2$  will be constant and  $c_1 + c_2 \geq 4$ . The velocity would then be defined by Equation (6):

$$v_i(t+1) = \mathbf{w} * [v_i(t) + c_1 * r_1 * (p_{best}(t) - x_i(t)) + c_2 * r_2 * (g_{best}(t) - x_i(t))] \quad (6)$$

the velocity will be limited by the value of  $w$  and its value will be defined between  $[0,1]$ .

Finally, there is a second output variable for the fuzzy system, called  $\sigma$ . This variable is the swarm ratio for the reset of randomly selected particles. This is important because it can potentiate  $w$  if, for example, the population starts with a low diversity. Therefore, *Diversity* as well as *Iteration* influence the value that  $\sigma$  can take. Its maximum proportion value must be chosen carefully. Otherwise the algorithm may never converge. A  $\sigma$  value is empirically defined between  $[0, 0.2]$ . Fig. 1 shows the pseudocode of the implementation of the  $\sigma$  variable.

```

particles=permute(numSwarm)
Numpartic=Sigma*numSwarm
For i=1:Numpartic
    Pos(particles(i))
    Reset Pos
End

```

Fig. 1. Implementation of the  $\sigma$  variable

The membership functions of the fuzzy input and output variables are determined according to previous experiments. In Fig. 2-5 the membership functions for each one of the fuzzy variables are shown.

Both fuzzy systems are of Mamdani type and are ideal for this type of control [6] [7] [8] [10] [11] [14]. The first fuzzy system is called FPSO1 and the second one as FPSO2. Each one of the systems is detailed below.

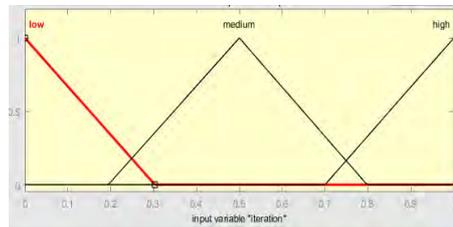


Fig. 2. Input variable: *Iteration*

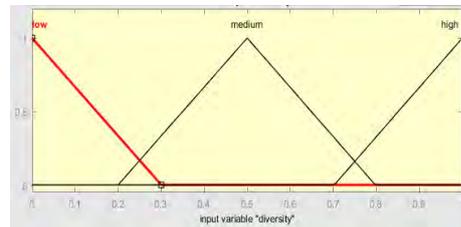


Fig. 3. Input variable: *Diversity*

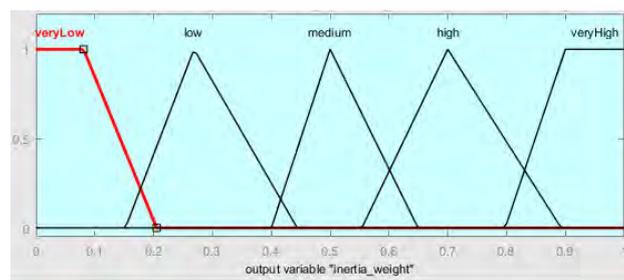


Fig. 4. Output variable: *inertia weight*

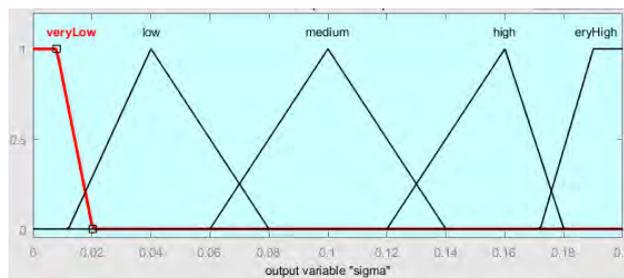


Fig. 5. Output variable: *Sigma*

### 3.1 FPSO1

The first variant proposed, called FPSO1, will have two input variables and one output variable, see Fig. 6. The input variables are *Iteration* and *Diversity* according to Figs. 2 and 3.

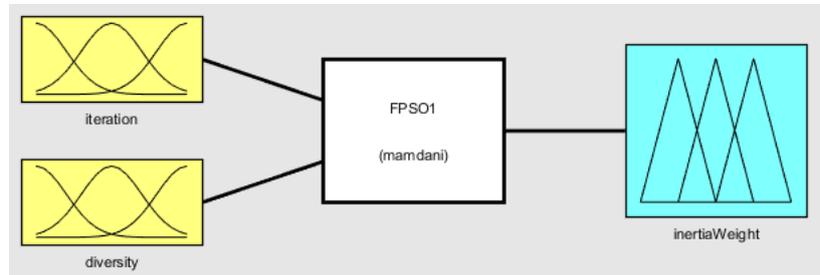


Fig. 6. First Fuzzy System FPSO1

The rules of FPSO1 can be seen in Fig. 7. In general, it is observed that when diversity is low in the initial iterations, the value of the inertia weight must be high so as to help exploring more promising areas. On the other hand, if the diversity is low in the final iterations, the inertia weight must have a very low value to favor exploitation of the area already found. The value assigned for the inertia weight is obtained by the centroid method.

1. If (iteration is low) and (diversity is low) then (inertiaWeight is veryHigh)
2. If (iteration is medium) and (diversity is low) then (inertiaWeight is medium)
3. If (iteration is high) and (diversity is low) then (inertiaWeight is veryLow)
4. If (iteration is low) and (diversity is medium) then (inertiaWeight is high)
5. If (iteration is medium) and (diversity is medium) then (inertiaWeight is medium)
6. If (iteration is high) and (diversity is medium) then (inertiaWeight is veryLow)
7. If (iteration is low) and (diversity is high) then (inertiaWeight is veryHigh)
8. If (iteration is medium) and (diversity is high) then (inertiaWeight is low)
9. If (iteration is high) and (diversity is high) then (inertiaWeight is veryLow)

Fig. 7. Rule for fuzzy system FPSO1

### 3.2 FPSO2

The second proposal, called FPSO2, differs from the previous one in the fact that it has an additional output variable. This variable is *Sigma* as can be seen in Fig. 5. The rest of the variables are the same defined in Fig. 2, 3, and 4. Therefore, FPSO2 is defined as shown in Fig. 8 and the rules are shown in Fig. 9. The value assigned for the output variables (inertia weight and sigma) are obtained by the centroid method.

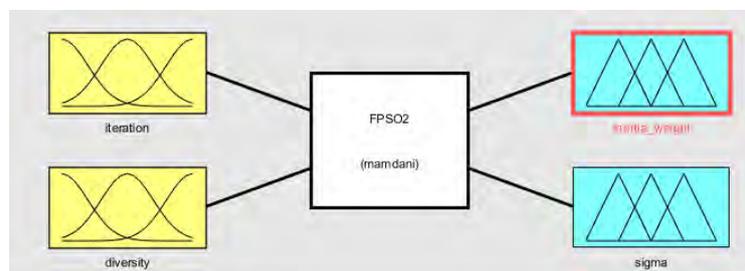


Fig. 8. Second fuzzy system FPSO2

1. If (iteration is low) and (diversity is low) then (inertiaWeight is veryHigh) (sigma is veryHigh)
2. If (iteration is medium) and (diversity is low) then (inertiaWeight is medium) (sigma is low)
3. If (iteration is high) and (diversity is low) then (inertiaWeight is veryLow) (sigma is veryLow)
4. If (iteration is low) and (diversity is medium) then (inertiaWeight is high) (sigma is high)
5. If (iteration is medium) and (diversity is medium) then (inertiaWeight is medium) (sigma is medium)

6. If (iteration is high) and (diversity is medium) then (inertiaWeight is veryLow) (sigma is veryLow)
7. If (iteration is low) and (diversity is high) then (inertiaWeight is veryHigh) (sigma is high)
8. If (iteration is medium) and (diversity is high) then (inertiaWeight is medium) (sigma is low)
9. If (iteration is high) and (diversity is high) then (inertiaWeight is veryLow) (sigma is veryLow)

**Fig. 9.** Rule for fuzzy system FPSO2

## 5 PSO variants using FPSO1 and FPSO2

The simple PSO algorithm is combined with the fuzzy systems to generate the following variants:

- *FuzzyPSO1*: This variant is built with the simple PSO algorithm and the fuzzy system FPSO1 and its pseudocode is shown in Fig. 10.

```

Swarm Initialization
Repeat
  calculate_Iteration Eq. (4)
  calculate_Diversity Eq. (5)
  Fis=FPSO1(Iteration,Diversity)
  Update velocity Eq. (6)
  Update position Eq. (2)
  Update pbest
  Update gbest
Until Maximum_iteration

```

**Fig. 10.** Pseudocode of the *FuzzyPSO1* algorithm

- *FuzzyPSO2*: the simple PSO algorithm is combined with the fuzzy system FPSO2 and the pseudocode of this variant can be seen in Fig. 11.

```

Swarm Initialization
Repeat
  calculate_Iteration Eq. (4)
  calculate_Diversity Eq. (5)
  Fis=FPSO2(Iteration,Diversity)
  Reinitialize particles with sigma
  Update velocity Eq. (6)
  Update position Eq. (2)
  Update pbest
  Update gbest
Until Maximum_iteration

```

**Fig. 11.** Pseudocode of the *FuzzyPSO2* algorithm

## 6 Experiments

Table 1 presents seven functions for large-scale global optimization, all of them can scale to different dimensions.

**Table 1.** Large-scale benchmark problems

	<b>Function</b>	<b>Limits</b>	<b>Optimum</b>	
F1	unimodal	Shifted Sphere Function	$x \in [-100,100]$	0
F2		Shifted Schwefel's Problem 2.21	$x \in [-100,100]$	0
F3		Shifted Rosenbrock's Function	$x \in [-100,100]$	0
F4		Shifted Rastrigin's Function	$x \in [-5,5]$	0
F5	multimodal	Shifted Griewank's Function	$x \in [-600,600]$	0
F6		Shifted Ackley's Function	$x \in [-32,32]$	0
F7		FastFractal "DoubleDip" Function	$x \in [-1,1]$	unknown

## 6.1 Experimental Results

The comparisons of the simple PSO algorithm, and the *FuzzyPSO1* and *FuzzyPSO2* proposals were performed with the functions defined in Table 1 for 500 (500 -D) and 1000 (1000-D) dimensions. The following configurations were considered: population size 50 for 500-D and 100 for 1000-D, based on [19] the cognitive learning coefficient  $c_1=2$ , social learning coefficient  $c_2=3$  ( $c_1+c_2>4$ ), maximum velocity equal to the maximum value of the variables. There were  $5E+03*D$  fitness evaluations (FEs) and 25 runs for each function as suggested in [20] and the best and mean statistical values were obtained and the better one is highlighted with gray background. The results of the seven test functions in 500-D and 1000-D are shown in Table 2 and representative convergence plots based on the median run of the three algorithms are presented in Fig. 12.

**Table 2.** Experimental results of F1-F7 in 500-D and 1000-D (mean (best))

Fun	D	PSO	<i>FuzzyPSO1</i>	<i>FuzzyPSO2</i>
F1	500	8,930E+05(7,981E+05)	3,132E+05(2,6E+05)	<b>2,901E+04 (2,1E+04)</b>
	1000	3,276E+06(2,4E+06)	1,155E+06(1,0E+06)	<b>9,227E+04 (7,6E+04)</b>
F2	500	1,535E+02 (1,5E+02)	1,350E+02(1,3E+02)	<b>9,491E+01(9,5E+01)</b>
	1000	1,724E+02 (1,7E+02)	1,522E+02(1, 5E+02)	<b>9,535E+01(9,5E+01)</b>
F3	500	5,318E+11(4,6E+11)	6,483E+10(3,6E+10)	<b>1,198E+03(9,0E+02)</b>
	1000	4,188E+12(2,3E+12)	5,726E+11(4,6E+11)	<b>3,379E+03(2,7E+03)</b>
F4	500	7,832E+03(7,4E+03)	5,338E+03(4,6E+03)	<b>3,467E+03(2,6E+03)</b>
	1000	1,949E+04(1,7E+04)	1,427E+04(1,3E+04)	<b>9,231E+03(8,1E+03)</b>
F5	500	8,224E+03(7,3E+03)	2,723E+03(2,2E+03)	<b>1,989E+02(1,5E+02)</b>
	1000	2,675E+04(2,3E+04)	1,015E+04(8,7E+03)	<b>7,828E+02(7,0E+02)</b>
F6	500	2,080E+01(2,0E+01)	2,041E+01(2,0E+01)	<b>1,775E+01(1,7E+01)</b>
	1000	2,144E+01(2,1E+01)	2,070E+01(2,1E+01)	<b>1,958E+01(1,9E+01)</b>
F7	500	-3,514E+03(-3,6E+03)	-4,808E+03(-5,0E+03)	<b>-4,943E+03(-5,6E+03)</b>
	1000	-6,646E+03(-6,8E+03)	-8,929E+03(-9,3E+03)	<b>-9,316E+03(-1,0E+04)</b>

The Wilcoxon Signed Rank test [22] was computed to determine if there are significant differences among the proposed algorithms. The results of this test are shown in Table 3 and were calculated using a significance level ( $\alpha$ ) of 0.05. The null hypothesis ( $H_0$ ) indicates that there is no significant difference between the "mean" values of the compared algorithms. While the alternative hypothesis is ( $H_1$ ) indicates that if there are significant differences between the "mean" values of the compared algorithms. Three comparisons were made: (1) *FuzzyPSO1* with *PSO*, (2) *FuzzyPSO2* with *PSO*, and (3) *FuzzyPSO1* with *FuzzyPSO2*.

**Table 3.** Wilcoxon test results with  $\alpha=0.05$

Algorithms	D	p-value	Decision
FuzzyPSO1 / PSO	500	0.01563	H0 is rejected
	1000	0.01563	H0 is rejected
FuzzyPSO2 / PSO	500	0.01563	H0 is rejected
	1000	0.01563	H0 is rejected
FuzzyPSO2 /FuzzyPSO1	500	0.01563	H0 is rejected
	1000	0.01563	H0 is rejected

*PSO with FuzzyPSO1*: *FuzzyPSO1* performed significantly better against *PSO*. Table 2 shows that the results of *FuzzyPSO1* are superior in all functions. However, both algorithms were affected by dimensionality. The graphs in Fig. 12 show that the convergence of *FuzzyPSO1* is better and faster than *PSO*.

*PSO with FuzzyPSO2*: The *FuzzyPSO2* variant outperforms simple *PSO* in the different types of high dimensionality problems. In terms of scalability from 500-D to 1000-D, *FuzzyPSO2* is much less affected than *PSO*. The convergence of *FuzzyPSO2* is better in both unimodal (F1-F2) and multimodal (F3-F7) functions (see Fig. 12).

*FuzzyPSO1 with FuzzyPSO2*: The performance of *FuzzyPSO2* is statistically better than that of *FuzzyPSO1* in 500-D and 1000-D in all 7 test functions. *FuzzyPSO2* converges faster and better on test problems F1, F2, F3, F5, F6. However, in F7 both,

500-D and 1000-D, *FuzzyPSO2* seems to get trapped in local optima in the first cycles, following a convergence to better solutions late in the search. F7 is a complex multimodal function that represents many real-world problems, and both variants had a similar convergence, as seen in Fig.12.

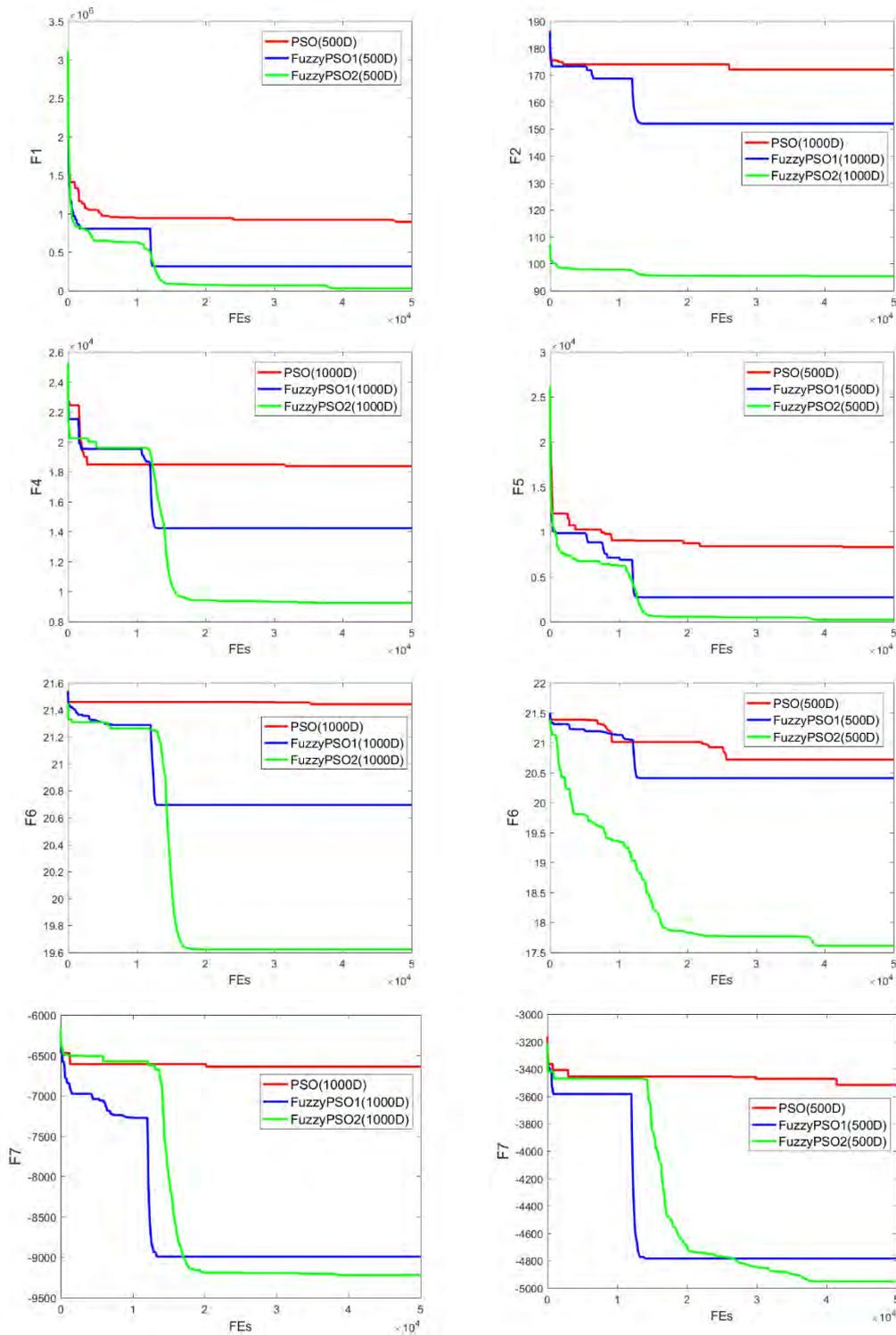


Fig. 12. F1-F7 representative convergence plots at 1000-D and 500-D based on median run.

## 7. Conclusions

In this paper two variants named *Fuzzy*PSO1 and *Fuzzy*PSO2 were presented to solve large-scale global optimization problems. The original PSO was combined with the FPSO1 fuzzy system, while the original PSO was merged with the FPSO2 fuzzy system. These systems dynamically manage the inertial weight according to the number of iterations and population diversity. They differ in the fact that FPSO2 restarts a proportion of the population if the diversity is low in the first iterations and that proportion decreases as the iterations increase. It was shown that the *Fuzzy*PSO2 variant had a better performance in all test functions compared to *Fuzzy*PSO1 and a simple PSO. However, in the case of F1 and F3 functions their performance is not yet the expected, and more work is needed to improve the results. Large-scale global optimization problems are challenging and finding the optimal one is not an easy task. However, the results obtained in this initial work are encouraging and suggest that fuzzy concepts can improve the scalability of PSO. As future work we will focus on the performance in unimodal functions and we will also test our approach in large-scale constrained optimization problems.

## References

1. Kennedy, J., Eberhart, R.C. (2001) Swarm Intelligence. Morgan Kaufmann, San Francisco (2001)
2. Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In Proceedings of ICNN'95-International Conference on Neural Networks (Vol. 4, pp. 1942-1948). IEEE.
3. Jie, J., Zeng, J., & Han, C. (2006). Adaptive particle swarm optimization with feedback control of diversity. In International Conference on Intelligent Computing (pp. 81-92). Springer, Berlin, Heidelberg.
4. Zadeh, L.: Fuzzy sets. Information & Control 8, 338–353 (1965). DOI: 10.1016/S0019-9958(65)90241-X.
5. Olivas, J. A. (2001). La lógica borrosa y sus aplicaciones. Pag. Web arantxa. ii. uam.es/dcamacho/lógica/recursos/fuzzy-into-esp. pdf.
6. Pérez, J., Valdez, F., Castillo, O. et al. (2017) Interval type-2 fuzzy logic for dynamic parameter adaptation in the bat algorithm, Soft Computing 21: 667. DOI: 10.1007/s00500-016-2469-3.
7. Olivas, F., Valdez, F., & Castillo, O. (2013). Particle swarm optimization with dynamic parameter adaptation using interval type-2 fuzzy logic for benchmark mathematical functions. In 2013 World Congress on Nature and Biologically Inspired Computing (pp. 36-40). IEEE.
8. Olivas, F., & Castillo, O. (2013). Particle swarm optimization with dynamic parameter adaptation using fuzzy logic for benchmark mathematical functions. In Recent Advances on Hybrid Intelligent Systems (pp. 247-258). Springer, Berlin, Heidelberg.
9. Sombra, A., Valdez, F., Melín, P., & Castillo, O. (2013). A new gravitational search algorithm using fuzzy logic to parameter adaptation. In 2013 IEEE Congress on Evolutionary Computation (pp. 1068-1074). IEEE.
10. Valdez, F., Melín, P., & Castillo, O. (2011). An improved evolutionary method with fuzzy logic for combining particle swarm optimization and genetic algorithms. Applied Soft Computing, 11(2), 2625-2632.
11. Valdez, F., Melín, P., & Castillo, O. (2014). A survey on nature-inspired optimization algorithms with fuzzy logic for dynamic parameter adaptation. Expert systems with applications, 41(14), 6459-6466.
12. Norouzzadeh, M. S., Ahmadzadeh, M. R., & Palhang, M. (2012). LADPSO: using fuzzy logic to conduct PSO algorithm. Applied Intelligence, 37(2), 290-304.
13. Ochoa, P., Castillo, O., & Soria, J. (2017). Differential evolution using fuzzy logic and a comparative study with other metaheuristics. In Nature-Inspired Design of Hybrid Intelligent Systems (pp. 257-268). Springer, Cham.
14. Olivas, F., Valdez, F., Castillo, O., Gonzalez, C. I., Martínez, G., & Melín, P. (2017). Ant colony optimization with dynamic parameter adaptation based on interval type-2 fuzzy logic systems. Applied Soft Computing, 53, 74-87.
15. Shi, Y., & Eberhart, R. C. (1998). Parameter selection in particle swarm optimization. In International conference on evolutionary programming (pp. 591-600). Springer, Berlin, Heidelberg.

16. Kumar, S., & Chaturvedi, D. K. (2011). Tuning of particle swarm optimization parameter using fuzzy logic. In 2011 International Conference on Communication Systems and Network Technologies (pp. 174-179). IEEE.
17. Awad, N. H., Ali, M. Z., Liang, J. J., Qu, B. Y., Suganthan, P. N., & Definitions, P. (2016). Evaluation Criteria for the CEC 2017 Special Session and Competition on Single Objective Real-Parameter Numerical Optimization, Nanyang Technological University, Jordan University of Science and Technology and Zhengzhou University. Tech. Rep.
18. Eberhart, R. C., & Shi, Y. (2000). Comparing inertia weights and constriction factors in particle swarm optimization. In Proceedings of the 2000 congress on evolutionary computation. CEC00 (Cat. No. 00TH8512) (Vol. 1, pp. 84-88). IEEE.
19. Clerc, M. (1999) The swarm and the queen: towards a deterministic and adaptive particle swarm optimization, Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406), Washington, DC, USA, 1999, pp. 1951-1957 Vol. 3, doi: 10.1109/CEC.1999.785513.
20. Tang, K., Yao, X., Suganthan, P. N., MacNish, C., Chen, Y. P., Chen, C. M., & Yang, Z. (2007). Benchmark functions for the CEC'2008 special session and competition on large scale global optimization. Nature Inspired Computation and Applications Laboratory, USTC, China, 24.
21. González Morcillo C. (2005). Lógica Difusa. Una introducción práctica. Técnicas de Soft Computing.
22. Derrac, J., García, S., Molina, D., & Herrera, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. Swarm and Evolutionary Computation, 1(1), 3-18.
23. Li, X., & Yao, X. (2011). Cooperatively coevolving particle swarms for large scale optimization. IEEE Transactions on Evolutionary Computation, 16(2), 210-224.

# Workshop Procesamiento Distribuido y Paralelo

## Coordinadores

Fabiana Piccoli (UNSL)

Laura De Giusti (UNLP)

Carlos García Garino (UNCu)

## Comparación de un Algoritmo de Bidiagonalización para su Utilización en la Recuperación de Información

Osvaldo Sposito<sup>1</sup>, Viviana Ledesma<sup>1</sup>, Gastón Procopio<sup>1</sup>, Hugo Ryckeboer<sup>1</sup>,  
Victoria Saizar<sup>1</sup> y Alexis Vainberg<sup>1</sup>

<sup>1</sup> Departamento de Ingeniería e Investigaciones Tecnológicas,  
Universidad Nacional de La Matanza, Buenos Aires, Argentina.  
{sposito; vledesma, gprocopio, hugor, vsaizar, avainberg  
}@unlam.edu.ar

**Abstract.** Este artículo presenta parte del trabajo realizado en el marco de una investigación que pretende optimizar un Sistema de Recuperación de Información, mediante la implementación y evaluación de distintos algoritmos secuenciales y paralelos para resolver eficientemente la Descomposición en Valores Singulares. Tal proceso comienza con llevar la matriz inicial a la forma bidiagonal, lo que puede consumir más del 70% del tiempo total del proceso. Por ello, como trabajo preliminar se han estudiado distintos métodos de bidiagonalización. Este trabajo se relaciona al desarrollo e implementación de un algoritmo de bidiagonalización alternativo para comparar posteriormente su comportamiento en distintas arquitecturas, en particular, las basadas en unidades de procesamiento gráfico, monoprocesadores y multiprocesadores. La experiencia de este estudio concreto ha permitido un análisis de rendimiento al ejecutar el algoritmo en cada implementación, cuando se varía el tamaño de las matrices, identificando problemas mínimos en GPU en cuanto a diferencias en la precisión de datos.

**Keywords:** Descomposición de Valores Singulares, Bidiagonalización, Sistema de Recuperación de Información.

### 1 Introducción

La Indexación Semántica Latente (ISL) es un método para la búsqueda de información en documentos a través de la indexación de términos [1], lo cual involucra la aplicación de algoritmos matemáticos especializados a fin de simular el análisis que realizaría una persona. Una técnica ampliamente utilizada a tal fin es la Descomposición en Valores Singulares (DVS), luego la recuperación se realiza a partir de los valores y vectores singulares obtenidos al aplicar dicha técnica [2].

Este trabajo se realiza en el contexto de un proyecto de investigación que tiene por objetivo optimizar la resolución de la DVS, para su posterior inserción en un Sistema de Recuperación de Información (SRI) desarrollado por el mismo equipo. En particular, la mejora se enfoca en implementar algoritmos que permitan resolver la primera fase del proceso de la DVS, conocido como bidiagonalización [3]. Se ha

comprobado que esta fase es la que más tiempo insume, estudios realizados muestran que puede consumir más del 70% o 90% del tiempo total para obtener todos los vectores singulares o solo los valores singulares, respectivamente [4]. Esto hace que el método de bidiagonalización, con un alto nivel de paralelismo de sus operaciones, sea un excelente candidato para la utilización de unidades de procesamiento gráfico (GPU, por sus siglas en inglés) ya que estas brindan la potencia de procesamiento requerida. Son varios los autores que han presentado trabajos en los que proponen la utilización de GPU para la bidiagonalización [5], [6].

En principio se han evaluado distintos métodos de bidiagonalización y, a modo inicial, se implementó un algoritmo genérico, secuencial, que sirvió para evidenciar el funcionamiento interno del proceso. Posteriormente, se ha orientado el estudio hacia algoritmos que puedan ser implementados en plataformas paralelas.

Luego de estudiar distintas variantes, se ha decidido adaptar y desarrollar un algoritmo alternativo propuesto por Barlow, Bosner y Drmač [7] (en adelante, Barlow), para evaluar su implementación en tres arquitecturas diferentes, basadas en CPU monoprocesador, multiprocesador y en GPU. Vale aclarar que se han encontrado trabajos en los que el algoritmo en estudio es probado en MatLab, en CPU secuencial y paralelo, pero nunca ha sido probado en GPU. Para la implementación se utilizó el framework de CUDA con el fin de comparar los tiempos de respuesta resultantes cuando se aplica el algoritmo en matrices de distintos tamaños.

El resto del artículo se organiza de la siguiente manera: la Sección 2 repasa los algoritmos de bidiagonalización en el contexto de los métodos utilizados para la recuperación de información; la Sección 3 describe los algoritmos implementados para este trabajo; la Sección 4 muestra los resultados experimentales obtenidos; y finalmente, la Sección 5 presenta las principales conclusiones e ideas para avanzar en esta investigación.

## 2 Métodos para la Recuperación de Información

Un SRI necesita componerse, por una parte, de un formalismo que permita representar documentos y consultas y, por otra parte, de una medida de similitud entre un documento y una consulta. En la actualidad conviven una variedad de modelos basados en distintos paradigmas para representar tanto documentos como consultas en SRI y comparar la semejanza de tales representaciones [8]. Entre estos, se destacan los modelos clásicos: el modelo booleano, el modelo vectorial y, el modelo probabilístico. En el modelo vectorial se seleccionan las palabras útiles, que por lo general son todos los términos del documento a excepción de las palabras semánticamente vacías, este proceso se enriquece utilizando técnicas de lematización y etiquetado [9]. El trabajo presentado en este artículo está circunscripto en una variante del método de recuperación vectorial, la ISL.

El método de ISL permite la búsqueda de información en documentos mediante la indexación de sus términos [1]. Involucra la definición de un espacio semántico donde los términos y los documentos altamente relacionados son colocados unos cerca de otros, reflejando los patrones de asociación entre los datos más importantes e

ignorando los menos importantes, es decir los que tienen menor influencia al momento de la recuperación.

La aplicación de la ISL, como se dijo antes, implica la utilización de algoritmos matemáticos especializados, que como resultado simulan el análisis que realizaría una persona. Por otro lado, con el método de ISL se pretende resolver dificultades durante la recuperación causadas por problemas de sinonimia y polisemia (o equivocidad del habla corriente). Por ejemplo, si la búsqueda se realiza a partir de la palabra “estación”, la cual tiene múltiples significados (polisemia) una búsqueda literal de la palabra produciría muchos resultados posibles (estación de tren, estación del año, etc.). Si lo que se desea buscar es “estación del año”, resultaría de interés que los resultados incluyan palabras distintas, pero con un significado igual o parecido, por ejemplo “temporada”, “época” y así por el estilo (sinonimia). Por lo tanto, aplicando la ISL es posible buscar por conceptos o definiciones en contraste a lo que sería una búsqueda literal. Para ello, un primer recurso es trabajar con lexemas y no con palabras, ya que palabras derivadas de una misma raíz comparten buena parte de la carga semántica.

En general, al indexar los términos de los documentos, la matriz de documentos resultante se vuelve muy grande, por tal razón, a fin de acelerar el proceso de recuperación de información, suelen aplicarse técnicas de reducción de la dimensionalidad con el fin de transformar dicha matriz en una de menores dimensiones, pero capaz de reflejar las características de la matriz original al momento de llevar adelante las búsquedas. Para tal propósito se aplica la DVS, una técnica de factorización de matrices que permite descomponer una matriz en varias matrices que presentan las propiedades más significativas de la matriz original [1], [10], [11]. Así, una matriz  $A$  de tamaño  $t \times d$  descompuesta con DVS produce tres matrices, tal como se puede observar en la figura 1.

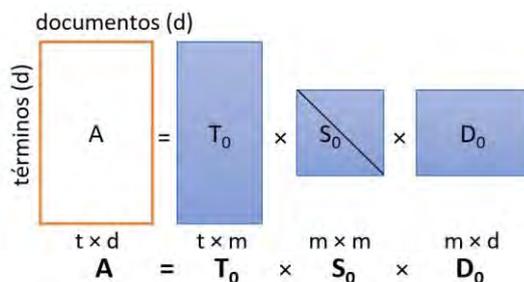


Fig. 1. Reducción de dimensiones en DVS. Fuente: [1].

Las columnas de  $T_0$  y  $D_0$  son ortonormales (ortogonales y de tamaño uno) y son las matrices izquierda y derecha respectivamente, de vectores singulares y,  $S_0$  es una matriz diagonal compuesta de los valores singulares de  $A$ . El triple producto indicado da una matriz de  $t \times d$  de rango  $m$ . De todas las matrices de  $t \times d$  de rango  $m$  que aproximen a  $A$ , la de menor error, es decir distancia, es una que comparte los mayores  $m$  autovalores de  $A$ , obtenidos en una descomposición DVS y anula los restantes, comparte sus autovectores. Habiendo autovalores nulos sus correspondientes

componentes en los autovalores no tienen influencia y por lo tanto son recortados a tamaño  $m$ .

La ventaja de utilizar estos modelos de orden reducido es que simplifica la comprensión del sistema, reduce el coste computacional en los problemas de simulación, lo cual a su vez implica menor esfuerzo computacional en el diseño de controladores numéricamente más eficientes y se obtienen leyes de control más simples [12]. Esto justifica la importancia y necesidad de buscar modelos matemáticos simplificados que aproximen al máximo el comportamiento del sistema original. El modelo resultante, que tendrá un número menor de estados que el sistema original, se denomina modelo reducido o modelo de orden reducido, mientras que se conoce como reducción del modelo al procedimiento utilizado para conseguirlo.

Existen dos tipos principales de algoritmos que se aplican al cálculo computacional de la DVS de una matriz real, el método unilateral de Jacobi y aquellos que se basan en la bidiagonalización [10]. El número de operaciones para los distintos algoritmos se encuentra en el orden de  $O(n^3)$ , las diversas propuestas y mejoras que han surgido buscan disminuir operaciones costosas en tiempo. El trabajo de este equipo de investigación se enfoca en los algoritmos basados en bidiagonalización, los cuales aplican transformaciones ortogonales con el fin de obtener una forma bidiagonal para luego conseguir la DVS de la matriz bidiagonal.

## 2.1 Algoritmos Aplicados para la Bidiagonalización

Tal como se explicó anteriormente, la reducción bidiagonal de una matriz densa general se usa muy frecuentemente como un paso preliminar para el cálculo de la DVS [13]. A partir de la revisión en la literatura se descubrió que existen distintos métodos para la bidiagonalización de una matriz, los enfoques más tradicionales utilizan las transformaciones de Householder por la izquierda y por la derecha de la matriz [10], [14], [15]. Algunos estudios demuestran que dichos métodos presentan dos desventajas: cuando las matrices son de grandes dimensiones requieren tiempos de computación elevados y además repercuten negativamente en los costos de comunicación de una implementación paralela del algoritmo en sistemas de memoria distribuida [16], [17]. De hecho, según Ltaief [13], el número total de operaciones para dicho algoritmo sea  $8/3(n^3)$ , pudiendo ser  $n$  previsible de varios miles.

Pretendiendo dar una solución a tales problemas han surgido diversos trabajos, entre estos se encuentran la propuesta de Ralha [18], mejorada más adelante por Barlow [7], orientada a conseguir un método más sencillo de paralelizar que los métodos tradicionales. En esta propuesta la bidiagonalización es unilateral, es decir, las transformaciones de Householder son aplicadas solamente por el lado derecho de la matriz. Posteriormente, Da Silva Sanches de Campos [17] presenta una mejora al método de Barlow con el objetivo de reducir el número de comunicaciones necesarias para una implementación paralela destinada a sistemas de memoria distribuida.

Las operaciones utilizadas en el proceso hacen que el método de bidiagonalización sea altamente paralelizable [19]. De más está decir, que la correcta ejecución de algoritmos paralelos depende fuertemente de que los tamaños de las matrices se adapten a las capacidades de la máquina donde estos se ejecutan, por lo que, en

matrices de alta dimensionalidad, aparecen problemas como el espacio en la memoria, la correctitud del algoritmo y el incremento en los tiempos de ejecución.

Con lo anterior presente, se han realizado numerosos trabajos que incluyen estudios comparativos en cuanto al rendimiento al bidiagonalizar matrices de distintos tamaños cuando se utilizan distintas implementaciones variando la arquitectura. Entre estos, se han contrastado implementaciones secuenciales y paralelas sobre una arquitectura homogénea basada en CPU [17], se han experimentado algoritmos en mosaico con distinta cantidad de nodos multinúcleo de un sistema de memoria compartida distribuida en paralelo [4], [20]. Otros han buscado aprovechar la capacidad que ofrecen las GPU y experimentaron su uso aplicando algoritmos en arquitecturas tanto homogéneas [5], [3] como también heterogéneas en las que se combinan el uso de CPU con GPU [21].

En este trabajo se decide poner especial interés en uno de los algoritmos alternativos de bidiagonalización, el propuesto por Barlow en [7], dado que está pensado para soportar el paralelismo, lo cual está en consonancia con el objetivo de la investigación en curso, lograr una implementación a partir de una arquitectura basada en GPU.

### 3 Implementación de Algoritmos de Bidiagonalización

En principio se evaluaron distintos métodos para el cálculo de la bidiagonalización y, a modo inicial, se implementó un algoritmo genérico, secuencial, que sirvió para evidenciar el funcionamiento interno del proceso. Dicho algoritmo, basado en las transformaciones de Householder [4], expresado a continuación en la figura 2 como algoritmo, ha sido desarrollado en el lenguaje C#, parte de este trabajo ha sido presentado en [22]. La idea era que este, aunque secuencial, sirviera de base tanto para comprender el proceso en sí mismo, como también, para tomarlo como referencia en el diseño e implementación del algoritmo de Barlow.

---

**Algoritmo 1: Reducción Bidiagonal vía Reflectores de Householder**

---

```

1 for  $j = 1$  to  $n$  do
2    $x = A_{j:n,j}$ 
3    $u_j = \text{sign}(x_1) \|x\|_2 e_1 + x$ 
4    $\hat{u}_j = u_j / \|u_j\|_2$ 
5    $A_{j:n,j:n} = A_{j:n,j:n} - 2 u_j (\hat{u}_j^* A_{j:n,j:n})$ 
6   if  $j < n$  then
7      $x = A_{j,j+1:n}$ 
8      $v_j = \text{sign}(x_1) \|x\|_2 e_1 + x$ 
9      $\hat{v}_j = v_j / \|v_j\|_2$ 
10     $A_{j:n,j+1:n} = A_{j:n,j+1:n} - 2 (A_{j:n,j+1:n} \hat{v}_j) \hat{v}_j^*$ 

```

---

**Fig. 2.** Algoritmo de bidiagonalización basado en Householder. Fuente: [4].

Aunque existía la posibilidad de utilizar la biblioteca LAPACK<sup>1</sup>, disponer del código en C# ofrece como ventaja, por una parte, permitir la comprensión de cada etapa interna del proceso, y por otra sienta las bases para que este código posteriormente pueda ser adaptado a diferentes algoritmos de bidiagonalización, e implementarlos en otras arquitecturas paralelas, en particular, aquellas basadas en GPU, a fin de analizar su eficiencia.

El algoritmo de Barlow, objeto de este estudio, y explicado en detalle en [7], consiste en un método para la bidiagonalización de matrices densas en el que las transformaciones de Householder se aplican únicamente por el lado derecho de la matriz. Con esto es posible definir todas las operaciones en términos de las columnas de la matriz a transformar, lo cual permite el desarrollo de implementaciones paralelas de un modo más simplificado en comparación con los métodos tradicionales, por otra parte, se logra reducir las comunicaciones que se necesitan.

El método de Barlow se puede así expresar, en forma de algoritmo, como se muestra en la figura 3, de la siguiente manera:

---

Algoritmo 2: BarlowBidiagonalización ( $A, \alpha, \beta, Q$ )

---

```

1 for  $r = 1, 2, \dots, n - 2$  do
2    $\alpha_r = \|A(:, r)\|_2$ 
3    $q_r = \frac{A(:, r)}{\alpha_r}$ 
4    $x_r = A(:, r + 1 : n)^t q_r$ 
5    $H_r$  tal que  $H_r^t x_r = \beta_r e_1$ 
6    $A(:, r + 1 : n) = A(:, r + 1 : n) H_r$ 
7    $A(:, r + 1) = A(:, r + 1) - \beta_r q_r$ 
8 end
9  $\alpha_{n-1} = \|A(:, n - 1)\|_2$ 
10  $q_{n-1} = \frac{A(:, n - 1)}{\alpha_{n-1}}$ 
11  $\beta_{n-1} = q_{n-1}^t A(:, n)$ 
12  $A(:, n) = A(:, n) - \beta_{n-1} q_{n-1}$ 
13  $\alpha_n = \|A(:, n)\|_2$ 
14  $q_n = \frac{A(:, n)}{\alpha_n}$ 

```

---

**Fig. 3.** Algoritmo de bidiagonalización unilateral de Barlow. Fuente: [17].

Como se puede observar, hay un ciclo principal en el que se trabaja la matriz principal por columnas, va desde la columna 0 hasta la antepenúltima columna ( $n-2$ ). Además de la matriz inicial, hay 3 variables principales que se utilizan a lo largo de todo el algoritmo:

- $\alpha$ : es un vector de  $n$  elementos que contiene los valores de la diagonal principal.
- $\beta$ : es un vector de  $n-1$  elementos que contiene los valores de la diagonal superior.

<sup>1</sup> <http://www.netlib.org/lapack/>

- $q$ : es una matriz con idénticas dimensiones que la matriz principal, es una matriz de trabajo interno.

En cada iteración se va completando: una posición en el vector  $\alpha$  de elementos de la diagonal principal; una posición en el vector  $\beta$  de elementos de la diagonal superior, esto se representa en las líneas 5 y 6 del algoritmo en las que se aplican las reflexiones de Householder; una columna de la matriz  $q$ ; y además, se modifica la matriz inicial que luego se lee en las iteraciones subsiguientes. Al terminar el ciclo se completan las posiciones restantes de  $\alpha$ ,  $\beta$  y las columnas que quedan de la matriz  $q$ .

Este algoritmo ha sido paralelizado y desarrollado para ser implementado sobre las tres arquitecturas mencionadas previamente, monoprocesador, multiprocesador y GPU, con el fin de comparar el rendimiento en cada una de estas. A continuación se resumen algunos resultados obtenidos.

## 4 Resultados Experimentales

Las implementaciones han sido desarrolladas utilizando el lenguaje *C#*, en conjunto con el framework CUDA, versión 6.5.

Las características del equipo utilizado para las pruebas de este experimento son:

- CPU: AMD Ryzen 5 2600 6 núcleos 12 threads a 3.6 GHz
- Memoria: 2 x 8GB DDR4 Crucial Ballistix 2400 Mhz
- GPU: NVIDIA GEFORCE GTX 1050 2GB

En la tabla 1 se presentan los tiempos de ejecución en milisegundos de cada una de las implementaciones realizadas para este estudio. Para las pruebas se utilizaron matrices cuadradas de distintas dimensiones, y tomando como fundamento las conclusiones obtenidas por Da Silva Sanches de Campos en [17], estas contienen valores aleatorios, dado que estos no tienen incidencia en los resultados esperados.

**Tabla 1.** Tiempos de ejecución del algoritmo en milisegundos para cada arquitectura

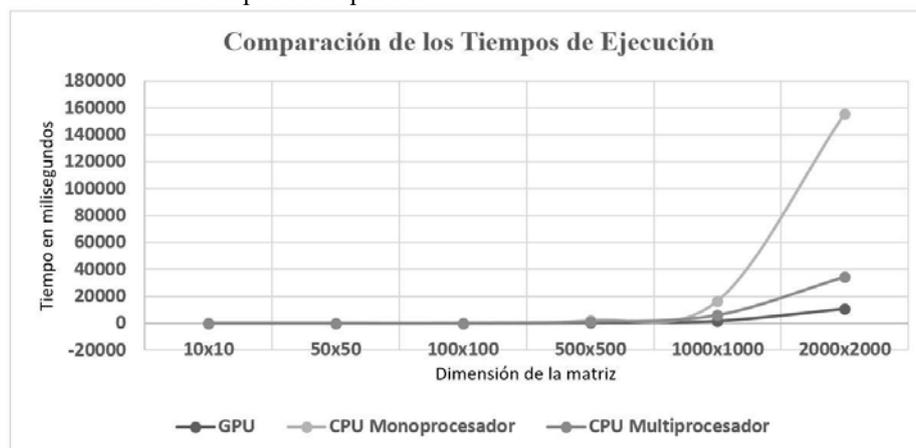
Dimensión Matriz	GPU	CPU	
		Monoprocesador	Multiprocesador
10x10	25	1	31
50x50	33	4	47
100x100	50	18	65
500x500	394	2005	1060
1000x1000	1691	16309	6008
2000x2000	10730	155418	34763

De los tiempos obtenidos, como resultado de las pruebas, se observa que cuando las matrices son de menor dimensión es conveniente ejecutar este tipo de algoritmos en CPU monoprocesador. Cuando la matriz comienza a superar las dimensiones, aproximadamente a partir de 200\*200 o 300\*300, la GPU mejora notoriamente el tiempo de ejecución con respecto a CPU monoprocesador y CPU multiprocesador. Esto se pone en evidencia, por ejemplo, observando los tiempos insumidos para la

matriz de tamaño 2000\*2000, donde la GPU logró reducir los tiempos de ejecución en un 93% y 69%, con relación a CPU monoprocesador y multiprocesador respectivamente. En cuanto a la CPU multiprocesador, se puede observar que es constante el tiempo de resolución y este se incrementa lentamente hasta las dimensiones aproximadas entre 200\*200 y 300\*300.

Debe considerarse que en el proceso monoprocesador se realiza el proceso en serie, tomándose en un principio una columna, se aplica las operaciones matemáticas correspondientes y luego se modifica la matriz general en base al resultado de la columna resultante de dichas operaciones. En cambio, en el proceso multiprocesador las operaciones son distribuidas entre los núcleos disponibles, de esta manera cada thread resuelve una porción de las operaciones correspondientes acelerando de este modo la resolución del algoritmo.

Se ha recurrido a la presentación de un gráfico, que se muestra en la figura 4, en donde pueden apreciarse las evoluciones y las respectivas variaciones en las mediciones de los tiempos de ejecución en la medida que el tamaño de la matriz se incrementa. A medida que la dimensión de la matriz va aumentando, los tiempos entre CPU monoprocesador y GPU se asemejan. A modo resumido, se puede acotar que es ventajoso ejecutar este algoritmo en CPU monoprocesador para matrices de dimensiones inferiores a 300\*300, en cambio, para aquellas de mayor dimensión será conveniente una arquitectura basada GPU la cual, como se puede visualizar, mejora notablemente los tiempos de respuesta.



**Fig. 4.** Gráfico comparativo de los tiempos de ejecución del algoritmo de Barlow en las distintas arquitecturas aplicado a matrices de diferentes dimensiones.

Respecto a la precisión de los datos, al comenzar las mediciones a partir de las matrices de menor dimensión se observó, en algunos resultados, que cuando el algoritmo fue ejecutado en GPU existía una diferencia en el 13° decimal con respecto a las ejecuciones en CPU, tanto en monoprocesador como en multiprocesador, para las cuales los valores obtenidos fueron coincidentes. Sin embargo, a medida que la dimensión de la matriz crece, se detecta que esta diferencia comienza a incrementarse, por ejemplo, en matrices de dimensión 1000\*1000 se encontraron diferencias en el 11° decimal. A partir de lo anterior, aunque las variaciones encontradas entre las

distintas implementaciones son ínfimas, teniendo en cuenta que las operaciones son las mismas, habría que investigar si los productos y las sumas en ambos tipos de procesadores son coincidentes, para de este modo evaluar la eficiencia del algoritmo en GPU.

## 5 Conclusiones

En el presente trabajo se presentó el desarrollo de un algoritmo alternativo que permite resolver el problema de la bidiagonalización de matrices densas y una comparación al implementarlo variando la arquitectura. El algoritmo fue probado adaptándolo a tres arquitecturas distintas: basada en GPU, CPU monoprocesador y multiprocesador. Se realizó un análisis de los tiempos resultantes para cada una de las implementaciones, observando la mejora en el rendimiento al paralelizar el algoritmo en una arquitectura basada en GPU para matrices de dimensiones mayores, cuando las matrices son pequeñas, en el orden de hasta  $300 \times 300$ , es preferible una arquitectura CPU monoprocesador. Es posible afirmar que las discrepancias en la precisión de los datos detectadas durante la ejecución sobre GPU son ínfimas, de todas maneras, sería necesario estudiar con mayor detalle los cálculos de cada procesador, incluso considerando matrices de mayor tamaño, para obtener conclusiones que ayuden a determinar si el comportamiento del algoritmo en dicha implementación se puede considerar exitosa.

En una siguiente etapa se pretende explorar si soluciones híbridas logran una aceleración en los cómputos, asignando en cada parte, CPU o GPU, aquellas tareas en las cuales mejor se desempeñan. Por otro lado, se espera poner a prueba la optimización conseguida en el SRI desarrollado por el equipo con el fin de comprobar el nivel de impacto alcanzado en la productividad del proceso.

**Agradecimientos.** Se agradece al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, el presente trabajo se financia en el marco del proyecto PROINCE C225.

## Referencias

1. Deerwester, S., Dumais, S., Furnas, G., Landauer, T. & Harshman, R.: Indexing by latent semantic analysis. *Journal of the American Society for Information Science*. vol. 41(6), pp.391–407. 1990.
2. Mamani Roque, M.: Descomposición en Valores Singulares y Análisis Semántico Latente. Tesis de Maestría. Universidad Politécnica de Valencia, España, 2018.
3. Dong, T., Haidar, A., Tomov, S. & Dongarra, J.: Optimizing the SVD Bidiagonalization Process for a Batch of Small Matrices. *Linear Algebra and Its Applications*, ELSEVIER, vol. 108, pp. 1008-1018, 2017.
4. Ltaief, H., Luszczek, P., & Dongarra, J.: High performance bidiagonal reduction using tile algorithms on homogeneous multicore architectures. *ACM Transactions on Mathematical Software*, vol. 39(3), 2013.

5. Lahabar, S. & Narayanan, P.: Singular Value Decomposition on GPU using CUDA. IEEE International Symposium on Parallel & Distributed Processing, pp. 1-10, 2009.
6. Liu, F., Seinstra, F.: GPU-based parallel householder bidiagonalization. HPDC '10 Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, pp. 288-291, 2010.
7. Barlow, J., Bosner, N., Drmač, Z.: A new stable bidiagonal reduction algorithm. Linear Algebra and Its Applications, ELSEVIER, vol. 397, pp. 35-84, 2005.
8. Tolosa, G. & Bordignon, F.: Introducción a la Recuperación de Información: Conceptos, modelos y algoritmos básicos. Universidad Nacional de Luján, Argentina, 2008. Recuperado el 29/06/2020 de: <http://eprints.rclis.org/12243/1/Introduccion-RI-v9f.pdf>
9. Jaimes, L. & Riveros, F.: Modelos clásicos de recuperación de la información. Revista Integración. Escuela de Matemáticas. Universidad de Santander, vol. 23(1), pp. 17-26, 2005.
10. J. Demmel, M., Gu, S. Eisenstat, et al.: Computing the Singular Value Decomposition with High Relative Accuracy. Linear Algebra and its Application, vol. 299, pp. 21-80, 1999.
11. Berry, M., Dumais, S. & O'Brien, G.: Using Linear Algebra For Intelligent Information Retrieval. Society for Industrial and Applied Mathematics, Review, vol. 37(4), pp. 573-595. Philadelphia, USA, 1995.
12. Fortuna, L., Nunnari, G. & Gallo, A.: Model order reduction techniques with applications in electrical engineering. Springer-Verlag, 1992.
13. Ltaief, H., Kurzak, J. & Dongarra, J.: Parallel Two-Sided Matrix Reduction to Band Bidiagonal Form on Multicore Architectures. IEEE Transactions on Parallel and Distributed Systems, vol. 21(4), pp. 417 – 423, 2010.
14. Golub, G. & Reinsch, C.: Singular Value Decomposition and Least Squares Solutions, Handbook Series Linear Algebra, vol. 14, pp. 403-420, 1970.
15. Chan, T.: An Improved Algorithm for Computing the Singular Value Decomposition. ACM Transactions on Mathematical Software, vol. 8(1), pp. 72-83, 1982.
16. Sangwine, S. & Le Bihan, N.: Quaternion Singular Value Decomposition based on Bidiagonalization to a Real Matrix using Quaternion Householder Transformations. Applied Mathematics and Computation, ELSEVIER, 182(1): 727-738, 2006.
17. Da Silva Sanches de Campos, C.: Algoritmos de Altas Prestaciones para el Cálculo de la Descomposición en Valores Singulares y su Aplicación a la Reducción de Modelos de Sistemas Lineales de Control. Tesis Doctoral. Departamento de Sistemas Informáticos y Computación, Universidad Politécnica de Valencia, España, 2014.
18. Ralha, R.: One-sided reduction to bidiagonal form. Linear Algebra and Its Applications, ELSEVIER, 358(1-3): 219-238, 2003.
19. Guerrero López, D.: Algoritmos Paralelos para la Reducción de Sistemas Lineales de Control Estables. Tesis doctoral. Departamento de Sistemas Informáticos y Computación, Universidad Politécnica de Valencia, España, 2015.
20. Faverge M., Langou, J., Robert, Y. & Dongarra, J.: Bidiagonalization and R-Bidiagonalization: Parallel Tiled Algorithms, Critical Paths and Distributed-Memory Implementation. IEEE Transactions on Parallel and Distributed Processing Symposium, 668 - 677, 2017.
21. Hernández Cortés, J.: Implementación paralela y heterogénea de la transformación de Householder y sus aplicaciones. Tesis de Maestría. Departamento de Computación, Unidad Zacatenco, México, 2017.
22. Spositto, O., Ledesma, V. & Procopio, G.: Aplicación de la Descomposición de Valores Singulares a un Sistema de Recuperación de Información. Revista Digital del Departamento de Ingeniería (ReDDI). Universidad Nacional de La Matanza, vol. 4(2), 2019.

# Comparación de Arquitecturas HPC para Computar Caminos Mínimos en Grafos. Intel Xeon Phi KNL vs NVIDIA Pascal

Manuel Costanzo<sup>1</sup> , Enzo Rucci<sup>1</sup> \*, Ulises Costi<sup>2</sup>, Franco Chichizola<sup>1</sup> ,  
and Marcelo Naiouf<sup>1</sup> 

<sup>1</sup> III-LIDI, Facultad de Informática, UNLP – CEA CICPBA.  
La Plata (1900), Bs As, Argentina

{mcostanzo,erucci,francoch,maniouf}@lidi.info.unlp.edu.ar

<sup>2</sup> Facultad de Informática, UNLP.  
La Plata (1900), Bs As, Argentina  
ulises\_92\_lp@hotmail.com

**Resumen** En la actualidad, uno de los principales desafíos de los sistemas de cómputo de alto rendimiento consiste en mejorar su rendimiento manteniendo el consumo energético en niveles aceptables. En ese sentido, una estrategia consolidada consiste en emplear aceleradores como las GPUs o los procesadores many-core Intel Xeon Phi. En este trabajo, se describen y comparan dispositivos de las arquitecturas NVIDIA Pascal e Intel Xeon Phi Knights Landing. Seleccionando el algoritmo de Floyd-Warshall como caso representativo de aplicaciones de grafos y limitadas por memoria, se desarrollaron implementaciones optimizadas con el fin de analizar y comparar el rendimiento y la eficiencia energética en ambos dispositivos. Contrariamente a lo considerado en el análisis preliminar, se encontró que el rendimiento de ambos dispositivos resultó comparable mientras que, respecto a la eficiencia energética, el Xeon Phi se mostró superior.

**Keywords:** Caminos mínimos · Floyd-Warshall · Xeon Phi · Knights Landing · NVIDIA Pascal · Titan X

## 1. Introducción

En la última década, la búsqueda por mejorar la eficiencia energética de los sistemas de cómputo de altas prestaciones (HPC, por sus siglas en inglés) ha incentivado la tendencia hacia la computación heterogénea y las arquitecturas masivamente paralelas [6]. Los sistemas heterogéneos combinan CPUs con aceleradores (como pueden ser las GPUs de NVIDIA y AMD o los coprocesadores many-core Xeon Phi de Intel), delegando en estos las secciones de código con alta demanda computacional. De acuerdo con el ranking Green500<sup>3</sup>, la edición de

\* Corresponding author.

<sup>3</sup> <https://www.top500.org/green500/>

Noviembre de 2010 contó con 17 sistemas que integraban aceleradores. Sin embargo, 10 años más tarde, este número se ha incrementado a 145, evidenciando la gran popularidad de esta estrategia.

En la actualidad, se puede considerar a las GPUs como la clase de acelerador dominante debido a su alta potencia de cómputo y eficiencia energética, además de su bajo costo de adquisición. En sentido opuesto, uno de sus puntos débiles es la necesidad de aprender un lenguaje específico para poder aprovecharlas al máximo, como pueden ser CUDA y OpenCL. Entre las empresas fabricantes, NVIDIA se destaca como la mayor proveedora para el segmento de alto rendimiento.

Por su parte, Intel ha presentado recientemente la segunda generación de sus procesadores Xeon Phi, con nombre clave Knights Landing (KNL). A diferencia de su antecesor Knights Corner (KNC), KNL puede operar como un procesador autónomo. Entre sus principales características, se pueden mencionar la gran cantidad de núcleos con soporte para hyper-threading, la incorporación de las instrucciones vectoriales AVX-512 de 512 bits y la integración de una memoria de alto ancho de banda (HBM, por sus siglas en inglés), entre otras [10]. Más allá de la generación, lo que distingue a esta familia es la compatibilidad con arquitecturas x86, lo que les permite a los programadores emplear modelos tradicionales en HPC, como OpenMP y MPI.

Debido a que cada acelerador tiene sus ventajas y desventajas para determinadas clases de problemas [16,2,13], la selección de la mejor opción para una aplicación dada se transforma en una cuestión fundamental en la búsqueda de lograr el máximo rendimiento. De manera de poder proveer algunas reglas generales para dicha selección, este artículo presenta un análisis de rendimiento entre dos arquitecturas HPC diferentes (Intel Xeon Phi KNL vs NVIDIA Pascal). Como caso de estudio, se seleccionó el algoritmo de Floyd-Warshall (FW) para el cómputo de caminos mínimos en un grafo, como caso representativo de aplicaciones de grafos y que se encuentren limitadas por memoria [12]. Se espera que el análisis presentado sea de ayuda para equipos de desarrollo al momento de elegir la arquitectura más conveniente para sus aplicaciones.

El resto del trabajo se organiza de la siguiente forma. En la Sección 2 se presentan los antecedentes y trabajos relacionados a la presente investigación. A continuación, en la Sección 3 se describen las implementaciones usadas mientras que en la Sección 4 se detalla el trabajo experimental realizado y se analizan los resultados obtenidos. Finalmente, en la Sección 5, se presentan las conclusiones y posibles líneas de trabajo futuro.

## 2. Antecedentes y Trabajos Relacionados

En primer lugar, se describen y comparan brevemente las arquitecturas Intel Xeon Phi KNL y NVIDIA Pascal. Luego, se describe el algoritmo FW para cómputo de caminos mínimos en un grafo. Finalmente, se detallan algunos trabajos relacionados al presente artículo.

## 2.1. Intel Xeon Phi KNL vs NVIDIA Pascal

**Intel Xeon Phi KNL.** A diferencia de una GPU o un co-procesador como el KNC, el KNL se trata de un procesador autónomo, capaz de bootear sistemas operativos y acceder directamente a la memoria DDR. La unidad de replicación escalable de la arquitectura KNL es el *tile*. Cada tile alberga 2 núcleos, una caché L2 compartida entre ambos núcleos, y una porción del directorio distribuido. Los núcleos de un tile implementan *simultaneous multithreading* (4 hilos hw por núcleo) y ejecución fuera de orden en su pipeline, además de contar con 2 unidades vectoriales que soportan las nuevas instrucciones AVX-512 de 512 bits [10].

Un chip de KNL esta conformado por entre 32 y 36 tiles activos (entre 64 y 72 núcleos), según el modelo en particular. Los tiles están interconectados por una malla 2D, con coherencia de cache basada en directorio distribuido y protocolo MESIF. Esta malla puede ser configurada en cinco modos de ejecución distintos (modos cluster). Según el modo de ejecución seleccionado, se definirá cómo el directorio distribuido se repartirá entre los tiles a lo largo del chip, lo que impacta en la latencia y el ancho de banda en el acceso a memoria.

El KNL trae una HBM de 16GB llamada MCDRAM, la cual está integrada en el mismo paquete del procesador. Esta memoria puede ser configurada en 3 modos distintos. En el modo *flat*, el espacio de direcciones es mapeado entre las dos memorias (MCDRAM y DDR), de modo que es el programador quien tiene la responsabilidad de definir cómo usarlas. Por su parte, el modo *caché* deja que el sistema se encargue de administrar la MCDRAM como una caché de la DDR. Por último, el modo *hybrid* asigna una parte de la MCDRAM como flat, y otra como cache [1].

**NVIDIA Pascal.** Pascal es la penúltima microarquitectura presentada por NVIDIA para el segmento de alto rendimiento, sucesora de Maxwell. En este segmento, una GPU de esta familia puede contar con hasta 3840 núcleos CUDA distribuidos entre (a lo sumo) 60 Streaming Multiprocessors (SM). A comparación de su predecesora, Pascal duplica la cantidad de registros por núcleo y aumenta el tamaño disponible de memoria compartida.

Esta microarquitectura presenta varias mejoras significativas respecto a Maxwell. Entre ellas, se puede destacar la incorporación en algunos de sus chips de una HBM de hasta 16GiB, que les permite alcanzar un ancho de banda de hasta 720 GB/s. También el reemplazo del tradicional bus PCIe para la comunicación CPU-GPU por un bus de alta velocidad (NVLink), que mejora significativamente la velocidad de comunicación. Por último, la provisión de una *memoria unificada*, consistente de un espacio de direccionamiento virtual entre las memorias de la CPU y la GPU, con el propósito de simplificar la programación [9].

Respecto al pico de rendimiento, algunos de los chips Pascal pueden alcanzar tasas de rendimiento en precisión doble (DP) de la mitad de precisión simple (SP). En sentido contrario, pueden duplicar el rendimiento de SP si computan en precisión media (PM) [4].

**Tabla 1.** Comparación entre Intel Xeon Phi KNL 7230 y NVIDIA Titan X

Dispositivo	Intel Xeon Phi KNL	NVIDIA Titan X
Chip	7230	GP102
Frecuencia	1.3-1.5 GHz	1.42-1.53 GHz
Cores	64 (256 hilos hw)	28 SMs (3584 cores CUDA)
Cache	1 MB L2	3 MB L2
SIMD	512-bit	-
HBM	16GB MCDRAM (450 GB/s)	-
Memoria RAM	192GB DDR4 (115.2 GB/s)	12GB GDDR5X (480.4 GB/s)
Bus	-	PCI-Express 3.0 x16
Rend. pico SP (DP)	6 (3) TFLOPS	10.97 (0.342) TFLOPS
TDP	215W	250W
TFLOPS/W (SP/DP)	0.028 / 0.014	0.051 / 0.001
Fecha de lanzamiento	Junio de 2016	Agosto de 2016

**Breve comparación.** La Tabla 1 presenta una comparación de las arquitecturas estudiadas y, en particular, considera los modelos utilizados en el trabajo experimental. Desde el punto de vista del pico de rendimiento teórico en SP, la Titan X resulta bastante superior al KNL 7230 (10.97 TFLOPS vs 6 TFLOPS). Sin embargo, debido al débil soporte de la primera para DP, es el KNL en este caso el que tiene una notable superioridad (3 TFLOPS vs 0.342 TFLOPS).

Respecto a la memoria principal, el KNL cuenta con una HBM que le permite estar casi a la altura de la Titan X en cuanto a ancho de banda, ya que el primero cuenta con tecnología DDR4 mientras que la segunda con GDDR5X. Sin embargo, por ser un coprocesador, la Titan X cuenta con un tamaño mucho menor de memoria respecto al KNL, que es un *host* en sí mismo.

Por último, al considerar la eficiencia energética (teórica), si bien la Titan X tiene una mayor potencia de diseño término (TDP), su ratio TFLOPS/W en SP casi duplica al del KNL por su mayor pico de rendimiento teórico. Por el contrario, el pobre desempeño de la Titan X para DP, lleva a que el KNL sea ampliamente superior en este caso.

A pesar de que ya han pasado algunos años de su lanzamiento, ambas arquitecturas se mantienen vigentes, como lo muestra la última edición del ranking Top500<sup>4</sup> donde 17 sistemas están equipados con Xeon Phi KNL y otros 30 con GPUs de la familia Pascal.

## 2.2. Cómputo de Caminos Mínimos en un Grafo

**Algoritmo FW.** El pseudo-código de FW se muestra en la Fig. 1. Dado un grafo  $G$  de  $N$  vértices, FW recibe como entrada una matriz densa  $D$  de  $N \times N$  que contiene las distancias entre todos los pares de vértices  $G$ , donde  $D_{i,j}$  representa la distancia del nodo  $i$  al nodo  $j$ <sup>5</sup>. FW computa  $N$  iteraciones, evaluando en la

<sup>4</sup> Top500 [www.top500.org](http://www.top500.org)

<sup>5</sup> Si no existe un camino entre los nodos  $i$  y  $j$ , se asigna *infinito* a su distancia (usualmente representado como el valor positivo más grande)

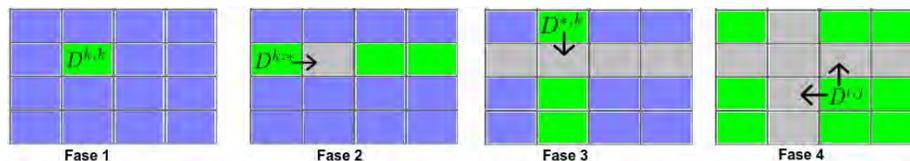
$k$ -ésima iteración todos los posibles caminos entre los vértices  $i$  y  $j$  que tienen a  $k$  como vértice intermedio. Como resultado, produce una matriz actualizada  $D$ , donde  $D_{i,j}$  contiene ahora la distancia mínima entre los nodos  $i$  y  $j$  hasta ese paso. Complementariamente, FW va construyendo una matriz adicional  $P$  que registra los caminos asociados a las distancias mínimas.

**Algoritmo FW por bloques.** A primera vista, la estructura de triple bucle anidado de este algoritmo resulta similar al de la multiplicación de matrices densas (MM). Sin embargo, como las lecturas y escrituras se realizan sobre la misma matriz, los tres bucles no pueden ser intercambiados libremente, como es el caso de MM. A pesar de esto, el algoritmo FW puede ser computado por bloques bajo ciertas condiciones [15].

El algoritmo FW por bloques (FWB) divide a la matriz  $D$  en bloques de tamaño  $TB \times TB$ , totalizando  $(N/TB)^2$  bloques. El cómputo se organiza en  $R = N/TB$  rondas, donde cada ronda consta de 4 fases ordenadas de acuerdo a las dependencias de datos entre los bloques:

1. Fase 1: actualizar el bloque  $D^{k,k}$  debido a que sólo depende de sí mismo.
2. Fase 2: actualizar los bloques de la fila  $k$  de bloques ( $D^{k,*}$ ) debido a que cada uno de estos depende de sí mismo y de  $D^{k,k}$ .
3. Fase 3: actualizar los bloques de la columna  $k$  de bloques ( $D^{*,k}$ ) debido a que cada uno de estos depende de sí mismo y de  $D^{k,k}$ .
4. Fase 4: actualizar los bloques restantes  $D^{i,j}$  de la matriz porque cada uno depende de los bloques  $D^{i,k}$  y  $D^{k,j}$  de su fila y columna de bloques, respectivamente.

En la Fig. 2 se muestra gráficamente cada una de las fases de cómputo y las dependencias entre bloques. Los recuadros en verde son aquellos bloques



**Figura 2.** Fases del cómputo de FWB y sus dependencias

```

for  $k \leftarrow 0$  to  $N - 1$  do
  for  $i \leftarrow 0$  to  $N - 1$  do
    for  $j \leftarrow 0$  to  $N - 1$  do
      if  $D_{i,j} \geq D_{i,k} + D_{k,j}$  then
         $D_{i,j} \leftarrow D_{i,k} + D_{k,j}$ 
         $P_{i,j} \leftarrow k$ 
      end if
    end for
  end for
end for

```

**Figura 1.** Pseudocódigo del algoritmo de FW

que están siendo computados, los grises los que ya se procesaron, y los celestes representan los que faltan computar.

### 2.3. Trabajos Relacionados

La comparación de arquitecturas HPC en un tema ampliamente estudiado por su comunidad. En particular, existen diversos artículos que involucran estudios comparativos entre Xeon Phi KNL y NVIDIA Pascal en el ámbito del modelado económico a gran escala [14], del álgebra lineal [3], de la dinámica de fluidos computacional [11], o del aprendizaje automático profundo [5], entre otros. Sin embargo, hasta donde llega el conocimiento de los autores, este es el primero en considerar algoritmos de grafos, en particular el de FW para caminos mínimos entre todos los vértices.

## 3. Optimización de FW

### 3.1. Implementación en Xeon Phi KNL 7230

La implementación usada considera las siguientes optimizaciones:

- *Localidad de datos.* Al computar según FWB, no sólo es posible explotar la localidad de datos sino que al mismo tiempo se logra aumentar el paralelismo disponible en la aplicación.
- *Paralelismo a nivel de hilos.* A partir del uso de OpenMP se obtiene una versión multi-hilada. Tanto los bloques de la fase 2 como los de la fase 3 son repartidos entre los diferentes hilos mediante la directiva `for` con planificación `dynamic`. En el caso de la fase 1, al consistir de un único bloque, los hilos se reparten las iteraciones dentro del mismo.
- *Paralelismo a nivel de datos.* A través de la directiva OpenMP `simd` se vectorizan las operaciones del bucle más interno en el cómputo de cada bloque, lo que permite aprovechar las instrucciones AVX-512.
- *Loop unrolling.* Desenrollando en forma completa el bucle más interno y una única vez el bucle *i*.
- *Predicción de saltos.* Mediante la inclusión de la función integrada del compilador `__builtin_expect`, se lo ayuda a predecir mejor los saltos de la sentencia `if`.
- *MCDRAM.* Al ser una aplicación limitada por ancho de banda, se obtienen grandes beneficios por el uso de esta memoria especial. Las ejecuciones se realizan usando el comando `numactl`.

Resulta importante aclarar que se puede considerar a esta implementación como una versión optimizada de [12], ya que incluye además la paralelización intra-bloque de la fase 1 y la mejora en la predicción de saltos.

### 3.2. Implementación en NVIDIA Titan X

La implementación utilizada considera las optimizaciones conocidas para FWB al momento [7,8]. Entre ellas se pueden mencionar:

- *Concurrencia.* Se han desarrollado 3 kernels que son invocados una vez por cada ronda de cómputo de FWB. En la ronda  $k$ , el primer kernel computa el bloque  $D^{k,k}$  (fase 1) y se instancia con un único grid formado por un único bloque. A continuación, se invoca el segundo kernel que computa en forma combinada los bloques de las fases 2 y 3 ( $D^{k,*}$ ,  $D^{*,k}$ ), el cual se instancia con un único grid de  $2 \times (R - 1)$  bloques. Por último, se invoca al tercer kernel que computa los bloques restantes correspondientes a la fase 4 ( $D^{i,j}$ ). Este kernel se instancia con un único grid de  $(R - 1)^2$  bloques. En todos los casos, los bloques se conforman de  $TB \times TB$  hilos.
- *Explotación de jerarquía de memoria.* Para el cómputo de FWB, el uso de memoria compartida no sólo resulta conveniente sino también necesario, especialmente en las fases 1-3 ya que los hilos leen y escriben en los mismos bloques de la matriz debido a sus dependencias. En el caso de la fase 4, es posible aprovechar la memoria local para el bloque de escritura ( $D^{i,j}$ ), lo que mejora aun más los tiempos de acceso. Por último, los accesos a memoria principal fueron organizados de forma de que sean coalescentes.
- *Ocupación de recursos.* De forma de maximizar este aspecto, se varió el tamaño de los bloques de hilos usando  $TB = \{8, 16, 32\}$  para encontrar aquel que conduce al máximo número posible de *warps* activos.
- *Loop unrolling.* Desenrollando en forma completa el bucle que computa cada hilo.

## 4. Resultados Experimentales

### 4.1. Diseño experimental

Las pruebas se realizaron en dos plataformas diferentes <sup>6</sup>. Por un lado, un servidor Intel Xeon Phi KNL 7230 configurado en modo cluster *All-to-all* y de memoria *flat* (ICC v19.0.0.117). Por otro, un equipo Intel Core i7-7700 3.6 GHz y 16GB RAM, el cual integra una GPU NVIDIA Titan X (CUDA v9.0).

Para ambas plataformas se consideró la variación del tamaño de la matriz de distancias ( $N = \{4096, 8192, 16384, 32768, 65536\}$ ) y el tipo de dato (*float*, *double*). En el caso del KNL, se varió tanto la cantidad de hilos OpenMP  $T = \{64, 128, 192, 256\}$  como el  $TB = \{16, 32, 64, 128\}$ , encontrando como valores óptimos  $T_{float} = 128$ ,  $T_{double} = 64$  y  $TB = 64$ . En cuanto a la GPU, los mejores rendimientos se presentaron al usar  $TB_{float} = 32$  y  $TB_{double} = 16$ . Por último, cada prueba particular fue repetida 15 veces, calculando los promedios para reducir la variabilidad.

<sup>6</sup> Las características de cada una fueron descritas al final de la Sección 2.1

## 4.2. Resultados de Rendimiento y Eficiencia Energética

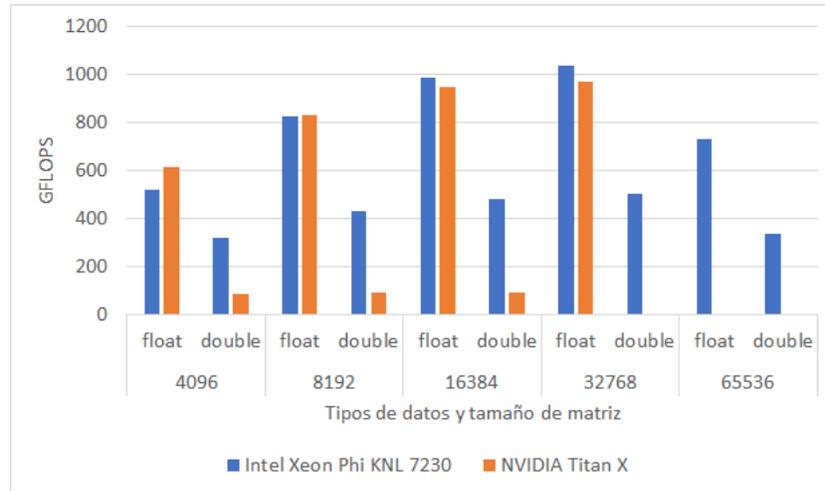
Para evaluar el rendimiento se emplea la métrica GFLOPS, utilizando la fórmula  $GFLOPS = \frac{2 \times N^3}{t \times 10^9}$ , donde  $N$  es el tamaño de la matriz de distancias,  $t$  es el tiempo de ejecución (en segundos) y el factor 2 representa la cantidad de operaciones en punto flotante requerida por cada iteración del bucle más interno.

En la Fig. 3 se presenta el rendimiento obtenido por cada implementación al variar la magnitud de la matriz y el tipo de dato usado. En SP (*float*), se puede observar que la GPU obtiene mejores rendimientos para los tamaños de matriz más chicos, siendo aproximadamente 19% superior cuando  $N = 4096$ . En estos casos, la GPU se adapta mejor en este caso que el KNL, el cual requiere cargas mayores de trabajo para alcanzar su máximo aprovechamiento. Este hecho se ve reflejado en el gráfico ya que a medida que se incrementa el tamaño de la matriz de distancias, es el KNL quien obtiene las mejores prestaciones alcanzando hasta un 7% adicional de diferencia. Resulta importante notar que con  $N = 65536$ , el KNL experimenta una pérdida de rendimiento de aproximadamente 30%, debiéndose al hecho de que con ese valor, se supera el tamaño disponible de la MCDRAM y se requiere del uso parcial de la DDR4, la cual posee un ancho de banda bastante menor. Aun así, resulta más conveniente que el caso de la GPU, quien no puede computar los casos con  $N > 32768$  debido a que se supera el espacio disponible en la memoria principal <sup>7</sup>.

En cuanto a DP (*double*), los resultados obtenidos son los esperables debido al débil soporte para esta clase de operaciones que tiene la Titan X. Mientras que la GPU obtiene un rendimiento casi constante de  $\sim 90$  GFLOPS, el KNL mejora su rendimiento a medida que  $N$  crece hasta superar el tamaño de la MCDRAM. En particular, el KNL obtiene cerca de la mitad de los FLOPS que en SP, pero que mejoran hasta  $4.3\times$  los de la Titan X. Por último, el problema del límite de memoria en la GPU, se ve agravado por el hecho de que el tipo *double* requiere más espacio (no fue posible computar los casos con  $N > 16384$ ).

Como se mencionó en la introducción, el rendimiento de la aplicación no es el único punto de interés a considerar, sino que también importa la eficiencia energética. La Tabla 2 presenta los cocientes de eficiencia energética teniendo en cuenta los picos de GFLOPS alcanzados y el TDP de las plataformas utilizadas. Se puede notar que el KNL resulta superior en ambos casos. En particular, KNL obtiene un cociente GFLOPS/Watt que resulta  $1.2\times$  mejor que la Titan X en SP. Sin embargo, este factor crece hasta  $5.5\times$  en DP, debido al débil soporte para estas operaciones de esta GPU. Se debe notar que en este análisis no se consideró el consumo del *host* por parte de la GPU, con lo cual las diferencias podrían ser aún mayores.

<sup>7</sup> Por supuesto que es posible desarrollar una implementación que procese la matriz de a partes y no tenga esta limitación de memoria. Sin embargo, la necesidad de hacer operaciones de E/S para cada ronda, perjudicarían el rendimiento significativamente.



**Figura 3.** Rendimiento obtenido al variar el tipo de dato y el tamaño de la matriz de distancias en Intel Xeon Phi KNL 7230 y NVIDIA Titan X

**Tabla 2.** Comparación de rendimiento y eficiencia energética entre plataformas

Plataforma	GFLOPS (pico)		TDP (Watt)	GFLOPS/Watt	
	SP	DP		SP	DP
<i>Xeon Phi KNL 7230</i>	1037	501	215	4,82	2,33
<i>NVIDIA Titan X</i>	972	90	250	3,89	0,36

### 4.3. Limitaciones

Resulta importante mencionar que este análisis se centra en dos modelos particulares de estas arquitecturas y que, como tales, algunos de los resultados encontrados podrían cambiar si se emplearan otros. Por ejemplo, la GPU NVIDIA GP100 (Pascal) posee un pico de rendimiento en SP levemente menor al de la Titan X (10.1 TFLOPS). Sin embargo, su soporte para DP es ampliamente superior, siendo de la mitad de SP (5 TFLOPS). En ese sentido, se considera que la comparación resulta igualmente valiosa para mostrar tendencias y características distintivas de cada arquitectura, aun cuando existan diferentes modelos de cada una que puedan presentar variaciones en sus especificaciones.

## 5. Conclusiones y Trabajo Futuro

Este trabajo se enfoca en la comparación de arquitecturas Intel Xeon Phi KNL y NVIDIA Pascal. Tomando como caso de estudio el algoritmo FW para cómputo de caminos mínimos en un grafo, se utilizaron implementaciones opti-

mizadas para comparar el rendimiento alcanzable en cada plataforma y así poder extraer algunas reglas generales. Entre ellas se pueden mencionar:

- A pesar de que en el análisis preliminar la Titan X se mostraba ampliamente superior al KNL, los rendimientos (SP) obtenidos para FW fueron comparables. Mientras que la GPU obtuvo mejores prestaciones para grafos pequeños, a medida que se incrementó el tamaño de la matriz de distancias, fue el KNL el que obtuvo un rendimiento superior. De este hecho se desprende la necesidad del KNL de grandes cargas de trabajo para poder aprovecharlo al máximo.
- En cuanto a la eficiencia energética, contrariamente a lo encontrado en el análisis preliminar, fue el KNL el que obtuvo el mejor ratio en SP. Este resultado contribuye al hecho de que el rendimiento sostenible y la eficiencia energética de los dispositivos varían de acuerdo a cada problema particular y a su correspondiente implementación en software.
- Más allá del aprendizaje específico de sus lenguajes, las GPUs pueden requerir esfuerzos de programación adicionales. Aunque experimentó una pérdida de rendimiento significativa cuando el tamaño de la matriz superó el de la MCDRAM, el KNL fue capaz de procesar todos los grafos. Por su parte, la Titan X no fue capaz de procesar aquellos que superaran el tamaño de su memoria RAM. Si bien es posible desarrollar una implementación que sí lo haga, también tendría una pérdida de rendimiento asociada y requeriría de un esfuerzo adicional de programación (ausente en el caso del KNL).

Entre los trabajos futuros, se pueden mencionar:

- Extender la implementación GPU para que soporte grafos más grandes que el tamaño de la memoria principal.
- Medir el consumo real en ambos dispositivos para hacer un estudio más preciso de la eficiencia energética, considerando que el primero puede diferir del estimado.
- Incluir otros modelos de las arquitecturas en cuestión (especialmente de las GPUs Pascal).

El desarrollo de estas actividades le otorgaría mayor robustez y representatividad al estudio realizado.

**Agradecimientos.** Los autores agradecen el soporte de la empresa NVIDIA por la donación de la GPU Titan X usada en esta investigación.

## Referencias

1. Codreanu, V., Rodríguez, J., Saastad, O.W.: Best Practice Guide - Knights Landing (2017), <https://bit.ly/2CEo1bR>
2. Deng, L., Bai, H., Zhao, D., Wang, F.: Kepler gpu vs. xeon phi: Performance case study with a high-order cfd application. In: 2015 IEEE International Conference on Computer and Communications (ICCC). pp. 87–94 (2015)

3. Deveci, M., Trott, C., Rajamanickam, S.: Multithreaded sparse matrix-matrix multiplication for many-core and gpu architectures. *Parallel Computing* **78**, 33 – 46 (2018). <https://doi.org/https://doi.org/10.1016/j.parco.2018.06.009>, <http://www.sciencedirect.com/science/article/pii/S0167819118301923>
4. Foley, D., Danskin, J.: Ultra-performance pascal gpu and nvl link interconnect. *IEEE Micro* **37**(2), 7–17 (2017)
5. Gawande, N.A., Daily, J.A., Siegel, C., Tallent, N.R., Vishnu, A.: Scaling deep learning workloads: Nvidia dgx-1/pascal and intel knights landing. *Future Generation Computer Systems* **108**, 1162 – 1172 (2020)
6. Giefers, H., Staar, P., Bekas, C., Hagleitner, C.: Analyzing the energy-efficiency of sparse matrix multiplication on heterogeneous systems: A comparative study of gpu, xeon phi and fpga. In: 2016 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). pp. 46–56 (2016)
7. Katz, G.J., Kider, Jr, J.T.: All-pairs shortest-paths for large graphs on the gpu. In: Proceedings of the 23rd ACM SIGGRAPH/EUROGRAPHICS Symposium on Graphics Hardware. pp. 47–55. GH '08, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2008)
8. Lund, B.D., Smith, J.W.: A multi-stage CUDA kernel for floyd-warshall. *CoRR abs/1001.4108* (2010), <http://arxiv.org/abs/1001.4108>
9. NVIDIA: NVIDIA Tesla P100., <https://bit.ly/20zrrk1>
10. Reinders, J., Jeffers, J., Sodani, A.: Intel Xeon Phi Processor High Performance Programming Knights Landing Edition. Morgan Kaufmann Publishers Inc., Boston, MA, USA (2016)
11. Robertsén, F., Mattila, K., Westerholm, J.: High-performance simd implementation of the lattice-boltzmann method on the xeon phi processor. *Concurrency and Computation: Practice and Experience* **31**(13) (7 2019). <https://doi.org/10.1002/cpe.5072>
12. Rucci, E., De Giusti, A., Naiouf, M.: Blocked all-pairs shortest paths algorithm on intel xeon phi knl processor: A case study. In: De Giusti, A.E. (ed.) *Computer Science – CACIC 2017*. pp. 47–57. Springer Int. Pub., Cham (2018)
13. Rucci, E., Garcia, C., Botella, G., De Giusti, A., Naiouf, M., Prieto-Matias, M.: Swifold: Smith-waterman implementation on fpga with opencl for long dna sequences. *BMC Systems Biology* **12**(5), 96 (Nov 2018). <https://doi.org/10.1186/s12918-018-0614-6>
14. Scheidegger, S., Mikushin, D., Kubler, F., Schenk, O.: Rethinking large-scale economic modeling for efficiency: Optimizations for gpu and xeon phi clusters. In: 2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS). pp. 610–619 (2018)
15. Venkataraman, G., Sahni, S., Mukhopadhyaya, S.: A Blocked All-Pairs Shortest-Paths Algorithm, pp. 419–432. Springer Berlin Heidelberg (2000). [https://doi.org/10.1007/3-540-44985-X\\_36](https://doi.org/10.1007/3-540-44985-X_36)
16. Véstias, M., Neto, H.: Trends of cpu, gpu and fpga for high-performance computing. In: 2014 24th International Conference on Field Programmable Logic and Applications (FPL). pp. 1–6 (Sept 2014). <https://doi.org/10.1109/FPL.2014.6927483>

# Workshop Tecnología Informática Aplicada en Educación

Coordinadores

Gladys Gorga (UNLP)  
Sonia Rueda (UNS)  
Alejandra Malberti (UNSJ)  
Claudia Russo (UNNOBA)

## **Sistema de Comunicación y Armado Multimedial de Material Pedagógico para Estudiantes Regulares y Diversos Funcionales en la Educación Superior.**

Vanina Cecilia Chiavetta<sup>1</sup>; Luis Mariano Mongelo<sup>1</sup>; Marcela Dávila<sup>1</sup>; María Laura Villarruel<sup>1</sup>; Gisela Diaz<sup>1</sup>; Néstor Ariel Pan<sup>1</sup>; Elías Biscaia<sup>1</sup>.

<sup>1</sup> Universidad Nacional de La Matanza, Departamento de Ciencias Económicas, Florencio Varela 1903,  
Buenos Aires, Argentina.

{Vanina.Chiavetta, Luis.Mongelo, Marcela.Davila, María.Villarruel, Gisela.Diaz, Nestor.Pan, Elías.Biscaia, [luis.mongelo@gmail.com](mailto:luis.mongelo@gmail.com)}

**Abstract.** La implementación de Tecnologías de la Información y la Comunicación en educación superior, exige, la puesta en marcha constante de nuevas técnicas de enseñanza, apoyadas con elementos audiovisuales, tales como videos, material interactivo y aplicaciones para celular. Por ello, es necesario contar con desarrolladores multimediales especializados, que interpreten las necesidades de contenidistas de las diferentes áreas educativas y las plasmen en los mismos, para que sirvan de apoyo a la tarea educativa. Articulando las áreas pedagógica, tecnológica y comunicativa, y a fin de brindar una educación de calidad, el propósito de este proyecto es diseñar y generar una herramienta de enlace entre los educadores y los desarrolladores multimediales, con un banco de contenidos digitales para seleccionar y utilizar, brindando especial atención a estudiantes con diversidad funcional motriz, auditiva y visual. Además, se busca ponerlo a disposición de los alumnos de nuestra universidad y de otras casas de altos estudios.

**Keywords.** Diversidad Funcional, Banco Multimedial, Software Educacional, Administración de Materiales Didácticos, Planificación de Contenidos.

### **1 Introducción**

El desarrollo tecnológico aporta grandes posibilidades para la mejora de la calidad de vida de las personas con diversidad funcional. Las nuevas tecnologías aplicadas al campo de la inclusión educativa de sujetos con diversidad física buscan facilitar la educación y comunicación y eliminar las barreras físicas y lógicas existentes en los ámbitos educativos. Su principal finalidad es la de potenciar el rendimiento y la autonomía de estos estudiantes en el proceso de enseñanza-aprendizaje.

Basados en varias investigaciones anteriores, y contando con un equipo de profesionales multidisciplinario, hemos acrecentado una cierta experticia en el tratamiento de materiales educativos para clasificarlos y almacenarlos, sirviendo como fuente de inspiración para generar una herramienta en la que el usuario contenidista o profesor, se relaciona con una interfase de selección y genera un informe paso a paso para hacérselo llegar al usuario creativo multimedial, que desarrollará el material didáctico adaptado para el uso con la población estudiantil.

## 2 Estado del arte

El estado actual del arte, en referencia a herramientas de software para la asistencia de problemáticas de los estudiantes diversos funcionales, no se encuentra demasiado desarrollado en nuestra región. La presente investigación tomará como punto de partida algunas investigaciones previas desarrolladas en nuestra casa de altos estudios, como así también trabajos realizados en otras universidades.

Desde lo exógeno, podemos nombrar el análisis de herramientas y dispositivos desarrollados en forma parcial para la materia Tecnología Educativa de la Universidad de Morón, por algunos de nuestros investigadores. Investigaciones y desarrollos llevados adelante por expertos en el área de informática educativa, como Robert Blackman, que sostiene: “Una necesidad creciente en el desarrollo de materiales educativos, para su soporte tanto en digital como físico; es el desarrollo de aplicaciones que nos permitan acopiar y catalogar materiales teóricos, prácticos y conceptuales para el desarrollo de contenidos multimedia aplicables al área de estudiantes diversos funcionales”... o desarrollos de clasificación y calificación de estos materiales en bancos de contenidos multimediales, como propone Mike Wall.

En lo que respecta a publicaciones, mencionamos el trabajo del Centro del Profesorado y de Recursos de la Universidad de Gijón; titulado *Guía multimedia de recursos educativos para alumnado con necesidades educativas especiales*. Esta obra es una guía de referencias a sitios y aplicaciones multimedia de toda España y otros países de habla hispana, que ofrece una descripción del objetivo y alcances del sitio, quiénes son sus desarrolladores, un mapa web y principales recursos que presenta la página para la atención de estudiantes que poseen algún tipo de diversidad funcional. Dicho trabajo, cuenta con links para descargar instaladores de programas de multimedia o recursos gráficos, de video o de sonido. El sistema de catálogos que posee para clasificar el material y el método utilizado en el reconocimiento de los recursos, resultó útil para proyectar el sistema de captura de requerimientos a utilizarse en nuestra aplicación educativa.

Con respecto a las interfaces de selecciones de materiales, en internet se dispone de bancos de imágenes y multimedia como PublicDomainArchives, DesignersPics o NewoldStock, dentro de los gratuitos o de uso público bajo licencias Creative Commons, o Image Bank o Pexels, dentro de los bancos de autor o pagos. Estos permiten descargar sus contenidos y usarlos como materia prima para nuestros materiales didácticos y sirven de inspiración en sus métodos de selección del material

y en la clasificación del mismo, incluyendo técnicas de ponderación y libre asociación para permitirnos encontrar el material apropiado.

En todos los casos, no hemos encontrado actualmente herramientas capaces de generar un paquete inteligente de los recursos, y una clasificación dinámica que permita generar una “hoja de ruta virtual” para describir el paso a paso de la construcción de recursos pedagógicos orientados a técnicas accesibles para diversos funcionales, como la lectura braille o la dinámica del audiolibro.

### **3 Desarrollo**

La génesis de nuestra aplicación comenzó con un relevamiento, evaluación y análisis de las diferentes propuestas tecnológicas en el área de bancos de contenidos digitales gratuitos o pagos en internet, orientados al contenidismo de materia educativa. A partir de ellas, se investigará cómo apropiarlas, clasificarlas, ponderarlas, calificarlas y eliminar las que no son aptas, para arribar a un pool o paquetes de contenidos, que califique para armar, cual un rompecabezas, materiales didácticos multimediales. El modelo de aplicación a generarse, deberá adaptarse a los valores de especificaciones y problemáticas propias de los estudiantes diversos funcionales, que también fueron investigadas en un estado del arte de la población educativa de nuestra casa de altos estudios y su región.

Seguidamente, se diseñó un software bajo la tecnología Java que articule la gestión de docentes y contenidistas en la generación de materiales didácticos multimediales, lo que permitirá nuevas experiencias de aprendizaje y la expansión de recursos.

Este software va a seleccionar en bancos de recursos multimediales, imágenes, sonidos, textos, videos, gráficos u otros, con el propósito de empaquetarlos y adjuntarlos a formularios de requerimientos que les permitan a desarrolladores multimediales, crear contenidos educativos para la utilización en el aprendizaje de estudiantes regulares y diversos funcionales.

Los sistemas educativos actuales acuden a los medios tecnológicos para mejorar el proceso enseñanza-aprendizaje. Uno de estos medios son los materiales multimedia. Determinados materiales tecnológicos afianzan cada vez más su presencia haciéndose, en muchos casos, imprescindibles.

Desde la consideración de los medios como materiales curriculares y didácticos, la cuestión clave estará en su utilización y su selección con la intención de aplicarlos convenientemente a las distintas situaciones educativas y, también, de aprovechar al máximo todas sus características técnicas y sus posibilidades didácticas.

Desde una perspectiva crítica, es preciso someter cualquier material didáctico o no, a un profundo análisis y criterios que orienten ese proceso de análisis y que den pautas para la adquisición, la selección y la aplicación de medios y materiales.

Los materiales didácticos multimedia son aquellos recursos electrónicos, que han sido diseñados con el objeto de facilitar los procesos de enseñanza-aprendizaje, o con una finalidad educativa. Según Gil y Vilches (2004), un material educativo multimedia está compuesto de los siguientes elementos:

-**Texto:** Sirve para presentar un tema, organizar ideas, apuntador de claves en la observación o un simple elemento para controlar el flujo de información. Si se diseña un material o proyecto que no utilice textos, su contenido puede resultar complejo, requiriendo imágenes y símbolos para guiar al usuario.

-**Sonido:** El sonido es quizás el elemento multimedia que más excita los sentidos, es el modo de hablar en cualquier lengua; bien escuchando música o con algún efecto especial.

- **Imagen:** Utilizar imágenes es muy importante pues permite al usuario orientarse visualmente, a la vez que puede transmitir ideas, conceptos, relaciones, etc. Este recurso motiva la atención, el descubrimiento y la comprensión, no solo es válida como un auxiliar de la palabra, sino que permite aclarar o reforzar lo que ésta manifiesta.

-**Animación y video:** La animación es uno de los recursos que puede dar mayor calidad a los productos multimedia y en términos generales, se puede decir que el video: Aumenta la sensación de realismo, mejora la autenticidad, credibilidad, y sintetiza los contenidos, con la finalidad de aprovechar las diferentes vías perceptivas.

En cuanto a sus características, los materiales educativos multimedia deben ser eficaces para facilitar el logro de los objetivos para los cuales fueron diseñados. A continuación, se ofrecen algunas de las características principales:

- **Flexibilidad:** posibilidad de tratamiento de la información desde diferentes puntos de vista.
- **Funcionalidad:** en cuanto a cuestiones como la adaptación al tipo de alumno y a la complejidad de los contenidos.
- **Multidimensional:** generando un ambiente mucho más atractivo que contribuye a la asimilación del conocimiento, así como favorece la creatividad y el desarrollo de la imaginación.
- **Dinamismo:** en cuanto que la información se puede cambiar de lugar, cambiar su presentación, estructurarla en diferentes niveles y, sobre todo, ofrecer diferentes caminos de interacción en función del interés del usuario.
- **Interactividad:** permite la posibilidad de dialogar con el programa, utilizando búsquedas, indagación, experimentación, ext.
- **Modulación de la información:** por cuanto se puede acceder a ella desde diferentes puntos del sistema.
- **Acceso multiusuario:** independientemente de factores espacio/temporales, pueden ser utilizados.

En cuanto a los materiales didácticos accesibles, Tim Berners-Lee, inventor del World Wide Web, sostiene: "... el poder de la web está en su universalidad. El acceso de todo el mundo con independencia de su discapacidad es un aspecto esencial". Esta idea se complementa con la consideración de que son accesibles los sitios que todos los usuarios podrán explorar de manera equivalente, cualquiera que sea su forma de acceso.

Es por esto que la Iniciativa de Accesibilidad a la Web (Web Accessibility Initiative) del Consorcio Mundial de la Web (World Wide Web Consortium) es aceptada en todos los ámbitos de la red para la elaboración de páginas de Internet accesibles para las

personas con discapacidad, accesibilidad universal y diseño para todos, denominado diseño universal para el diseño de productos, entornos, programas y servicios, que puedan utilizar todas las personas, sin necesidad de adaptación o diseño especializado. En ella se establecen principios generales para el diseño accesible, y se presentan catorce pautas asociadas a uno o más puntos de verificación, que describen cómo aplicarlas a las presentaciones de las páginas web.

Para la realización de la presente investigación se han tenido en cuenta entre otras, los criterios y principios de accesibilidad empleados en el diseño de documentos y herramientas accesibles a saber:

- a) Asegurar que el sitio Web o aplicación descargable de la institución cumpla con la normativa específica de accesibilidad.
- b) Certificar que cuando los estudiantes utilicen productos de apoyo el material en formato electrónico sea legible en el software de ampliación de pantalla, texto-a-voz, etc.
- c) Asegurar que los textos presentados en pantalla, gráficos, tablas, transparencias y presentaciones electrónicas estén disponibles en formato de audio.
- d) Garantizar que cuando en el material existen elementos de audio se toman las medidas necesarias para proporcionar demostraciones visuales alternativas y subtítulos para complementar a la interpretación en Lengua de Señas Argentinas (LSA).
- g) Usar argumentos cortos mejor que largos en textos largos.
- h) Utilizar tamaños diferentes de las fuentes (pero no fuentes distintas) para enfatizar los títulos (mejor que el subrayado).
- i) Escribir en un lenguaje claro y transparente, evitando frases largas y el uso excesivo de sub-oraciones.
- j) Proporcionar glosarios cuando sea necesario.
- k) Generar un buen contraste entre el fondo necesario y el texto (por ejemplo, texto blanco/amarillo sobre fondo azul).

Al tratarse de un proyecto de desarrollo tecnológico se reemplaza la hipótesis de trabajo por la propuesta de solución al problema de investigación mediante el diseño de un prototipo. Entonces, se propone el desarrollo de un sistema de comunicación multimedial que les permita a los docentes universitarios, administrar y organizar en forma intuitiva, un grupo de procesos necesarios para la selección, apropiación, diseño y puesta en uso de materiales físicos (impresos, en forma de apuntes teóricos o guías de trabajos prácticos) o multimediales (presentaciones multimedia como presentaciones digitales, videos, páginas web o guías interactivas) mediante un conjunto de formularios o guías de trabajos implementadas en forma automática por dicha herramienta.

## **4 Metodología**

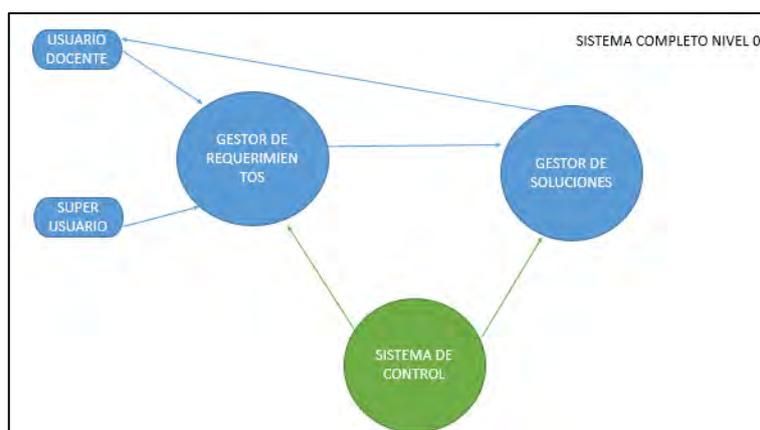
Para la investigación que se presenta se optó, en primera medida, por un método cuantitativo como es el análisis de documentos y la síntesis del análisis por lo que se aplicarán fichas de observación, destinadas a la recolección de información en torno al objeto del estudio. Asimismo, a través de un método analítico se abordarán los

materiales multimediales educativos concebidos aquí como documentos para determinar aquellos adecuados en el proceso de enseñanza – aprendizaje.

Los métodos propuestos implicaron las siguientes tareas:

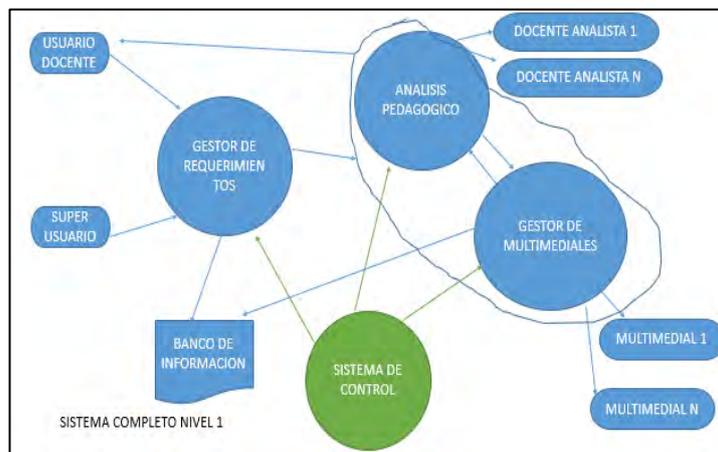
- Análisis de los nuevos adelantos en bancos de contenidos multimediales y herramientas de creación de contenidos para el soporte educativo.
- Comparación de metodologías de administración y comunicación de proyectos ya utilizadas a nivel internacional.
- Desarrollo de un grupo de herramientas de clasificación, calificación y administración de contenidos, para el armado de un paquete de recursos y un formulario explicativos automatizado, para la comunicación de objetivos a desarrolladores especializados.
- Realización de pruebas piloto con docentes y desarrolladores voluntarios de diferentes especialidades.

A partir de estos postulados, hemos generado un prototipo de herramienta de armado multimedial, denominado con la sigla HERA (Herramienta de Ejecución de Recursos Áulicos), basada en un diagrama básico de utilización, que divide a la misma en tres módulos principales. Su diagrama puede observarse en la siguiente figura.



**Fig. 1.** Diagrama del prototipo de la herramienta.

A partir de estos 3 subsistemas básicos, se procedió a subdividir el Gestor de Soluciones en dos componentes lógicos, el Sistema de Análisis Pedagógico y el Gestor de Diseñadores Multimediales; y a enumerar los componentes de cada módulo, arribando a un segundo nivel de diseño, que puede verse en la siguiente figura:



**Fig. 2,** Sistema Completo, Con Nivel 1 de Diseño.

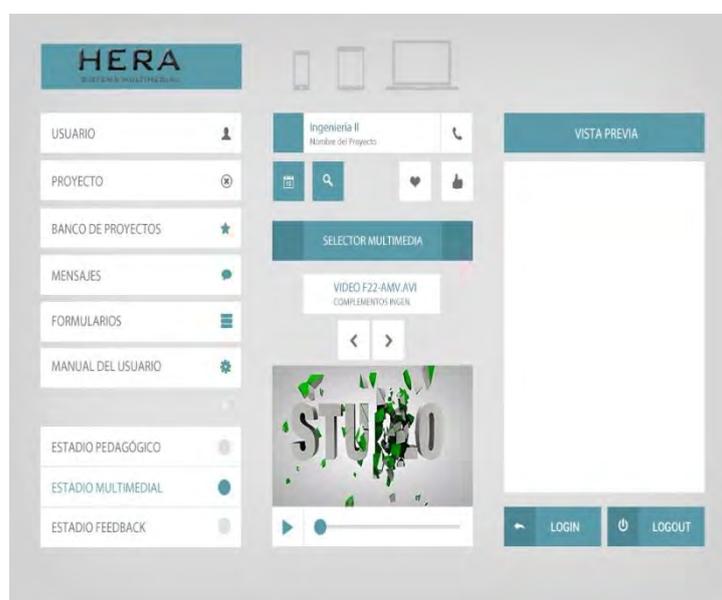
La siguiente enumeración permite describir las funciones de cada uno de estos componentes:

- **Gestor de Requerimientos:** Desde aquí se llevará a cabo la interacción con el usuario, donde se detallará lo que solicita, ingresará su material y se le informará el estado del/los proyectos. En este sistema se darán de alta los perfiles de usuarios y también se derivará al gestor de soluciones. Desde este sistema se emitirán estadísticas sobre incidencias de requerimientos. Asimismo, se indicará la necesidad o no de recursos a desarrollar para alumnos diversos funcionales.
- **Gestor de Soluciones:** retornará el material finalizado al Usuario Docente y esperará su feedback para dar por finalizado o realizar modificaciones. En este sistema, se asignará un docente analista por proyecto y los multimediales necesarios a tal fin. También se emitirán estadísticas de productividad y para ello interactúan dos subsistemas:
- **Análisis Pedagógico:** Recibido el requerimiento se realiza el matcheo del docente analista adecuado según tema y categoría y se asigna la tarea conjunta con el/los multimediales (contemplando analistas especializados en material pedagógico para diversos funcionales).
- **Gestor de Multimediales:** Recibido el requerimiento se realiza el matcheo del multimedial adecuado al proyecto (o varios si lo indica el requerimiento).

**Sistema de Control:** el sistema de control interactúa con todos los subsistemas. Su función principal es testear los tiempos de cada etapa y emitir las alertas pertinentes. También emitirá alertas sobre sistemas de Soluciones en caso de reingresos reiterados y objeciones. Desde este sistema se mostrarán estadísticas de las alertas. El Usuario del sistema es el docente que necesita el material pedagógico y el Banco de Recursos es un repositorio de materiales multimedia bajo licencias Creative Commons, con la que se provee de componentes audiovisuales, textos, gráficos y conceptuales para nutrir los proyectos a realizar.

## 5 Resultados

Lo anteriormente expuesto, se basa en el desarrollo de una aplicación ligera y adaptable, que puede distribuirse mediante Internet, DVD o CD, recibiendo el encargado del curso (Coordinador General, Organizador o Jefe de Cátedra) y procediendo a posteriori a cargar los datos que definirán su perfil de materia y el tipo de recurso a solicitar (presentaciones audiovisuales interactivas, apuntes teóricos, diagramas y herramientas conceptuales para el trabajo en aula, manuales o guías de aprendizaje paso a paso, etc.). Actualmente se encuentra en etapa de desarrollo el prototipo con un diseño de interfaz de acceso, similar al siguiente:



**Fig. 3.** Interfaz de acceso al sistema – Menú Principal.

En la figura se puede observar el acceso a los módulos, y el selector de material multimedia, que se retrotrae con pantallas deslizables que se superponen, obedeciendo a la barra de navegación lateral izquierda, mostrando entre otros contenidos, la vista previa del documento o las especificaciones técnicas de los componentes del documento a entregar.

También se puede visualizar rápidamente el estadio en que se encuentra el desarrollo del proyecto actual, según el módulo que está siendo procesado en cada momento. La aplicación se encuentra actualmente en proceso de desarrollo, para pasar del prototipo a su fase semi-operativa.

## Bibliografía

1. Blackman, R.: Nuevos Desarrollos para el Nuevo Mundo Digital. Ediciones Orbe. Ciudad de México (2009)
2. Deitel, P. y Deitel, H.: Como programar en Java – 9na Edición. Pearson-Addison Wesley. Naucalpan (2012)
3. DePirene, A.: Administración de la Educación Virtual. Publicaciones Planeta Inteligente. Ciudad de México (2008)
4. Eric Zabre, B. e Islas, P.: Evaluación de herramientas de hardware y software para el desarrollo de aplicaciones. Wiley-Interscience. Barcelona (2011)
5. Gil, D. y Vilches, A.: Una Alfabetización Científica para el Siglo XXI. Obstáculos y propuestas de actuación. Investigación en la escuela, 43 pp. 27-37 (2004)
6. Monjo Palau, T.: Diseño de Interfaces Multimedia. Edicions Universitat Oberta de Catalunya. Cataluña (2011)
7. Suárez Turbón, I. y Sueiras Rodríguez, E.: Guía multimedia de recursos educativos para alumnado con necesidades educativas especiales. Centro del Profesorado y de Recursos de Gijón. Principado de Asturias (2017)
8. Wald, M.: Creating Accesible Educational Multimedia. American Council for Education. Whashington (2016)

# Educación Virtual con Alumnos de Primer Año de Ingeniería en tiempos de aislamiento social obligatorio

Artemisa Trigueros<sup>1</sup>, Mabel Compagnoni<sup>1</sup>, Larisa Toro<sup>1</sup>,

<sup>1</sup> Universidad Nacional de La Matanza, Florencio Varela 1903,  
B1754 San Justo, Pcia. De Buenos Aires, Argentina  
(artemisa, mcompagnoni, lara, [@unlam.edu.ar](mailto:@unlam.edu.ar))

**Abstract.** La introducción de las TICs (Tecnologías de la Información y la Comunicación) con el objetivo de innovar el diseño curricular y las metodologías de enseñanza frente a las demandas de la sociedad del conocimiento, es un tema recurrente dentro del ámbito educativo, anterior a la situación de pandemia mundial. Hace tiempo que comenzó a considerarse la incorporación de las nuevas tecnologías en los modelos educativos como un componente clave. Sin embargo, las implementaciones de Blended Learning<sup>1</sup> en el dictado de materias de primer año de las carreras de Ingeniería en la Universidad Nacional de La Matanza (UNLaM), con anterioridad a esta investigación, ha sido la de reforzar o compensar las situaciones de enseñanza presencial; sólo ocasionalmente se plantearon actividades curriculares que permitieran crear un escenario cualitativamente distinto de enseñanza y aprendizaje con el uso de las TICs. Ante la necesidad de seguir impartiendo una educación de excelencia en el marco del aislamiento social, la UNLaM, implantó la enseñanza mediada por las TICs. Este artículo presenta información relativa a la experiencia del cursado de materias incorporando la modalidad virtual de enseñanza como consecuencia de la pandemia, durante el primer cuatrimestre del año 2020, mediante los resultados de una encuesta realizada a los alumnos de la cátedra Fundamentos de TICs correspondiente al primer año de las carreras de Ingeniería de UNLaM. Esta investigación específica la valoración que el alumnado asigna a las distintas herramientas pedagógicas provistas por las nuevas tecnologías, recoge sus opiniones respecto de las ventajas y desventajas de la educación virtual y analiza la importancia de la interacción sincrónica entre alumnos y docentes durante las clases. Se concluye que una propuesta pedagógica sustentada con un adecuado uso de las plataformas tecnológicas y sostenida por un acompañamiento sincrónico del docente, permite una educación de calidad en el entorno virtual de enseñanza.

**Keywords:** Educación, Enseñanza Virtual, Fundamentos de TICs, Ingeniería, Nuevas Tecnologías.

## 1 Introducción

El campo del diseño curricular muestra signos de avance con el paso del tiempo, y tiene además elementos que permiten caracterizar su producción, pero también enfrenta retos [1]. Se podría considerar como uno de ellos, la implementación de la educación mediada por las TICs, en alumnos de primer año de las cinco Ingenierías (Civil, Electrónica, Industrial, Informática y Mecánica) en la Universidad Nacional de La Matanza ante la situación de imposibilidad de la enseñanza presencial.

Como antecedente a la experiencia relatada por el presente artículo, se pueden mencionar las realizadas por las cátedras Análisis Matemático I [2], Álgebra I, Elementos de Programación e inclusive Fundamentos de TICs (Plan 1997) correspondientes al primer año de las carreras de Ingeniería de la Universidad Nacional de La Matanza, que incorporaron Blended Learning en el dictado de las respectivas materias.

<sup>1</sup>Blended Learning es el aprendizaje que combina el e-learning (encuentros asincrónicos) con encuentros presenciales (sincrónicos) tomando las ventajas de ambos tipos de aprendizajes.

Cabe destacar, que las posibilidades de las TICs en educación se sustentan en los siguientes aspectos: la creación de entornos de aprendizaje que posibilitan la ampliación de las capacidades cognitivas y la importancia de la interacción con los otros para la construcción conjunta de significados creando un contexto mental cada vez más compartido en situaciones de aprendizaje complejo [3].

Realizando un análisis desde una mirada socioconstructivista, se puede aseverar que la cognición ocurre dentro de un sistema de actividades y se distribuye entre los sujetos, las herramientas culturales que se utilizan (textos, programas computacionales, sitios web, etc.), las otras personas con quien se interactúa (profesor, otros alumnos, autores de textos, etc.) y las prácticas culturales en donde se participa mientras se realiza alguna actividad [4].

Teniendo en cuenta lo expresado con anterioridad y ante la nueva perspectiva de confeccionar un diseño tecno-pedagógico de enseñanza mediada, como es el caso del diseño de ambientes y situaciones de aprendizaje en las plataformas MIeL<sup>2</sup> (Campus Virtual de UNLaM) y Teams<sup>3</sup> brindadas por la Universidad Nacional de La Matanza, se consideraron la naturaleza y el tipo de actividades que los alumnos realizarían de forma no sólo individual sino colectiva; las características de las herramientas que se pusieron a disposición del alumnado; los usos previstos de dichas herramientas; y las reglas y prácticas educativo-culturales en las que se insertarían las actividades propuestas.

Desde el punto de vista de los recursos tecnológicos que se pudieron incorporar en el diseño de entornos virtuales de enseñanza y aprendizaje, se destacaron aquellos que brindaron la posibilidad de integrar varias prestaciones comunicativas y colaborativas que facilitaron la comunicación sincrónica y asincrónica entre compañeros y profesores; tal es el caso de los chats, foros, mensajería, videoconferencias, entre otros. No obstante, es importante señalar que por sí mismos, y como meros recursos tecnológicos, no resolvieron el problema de la construcción conjunta del conocimiento, dado que lo relevante resulta ser el tipo de prácticas educativas y comunicativas que se propiciaron entre los participantes, es decir, los usos que se hicieron de dichas herramientas. Así pues, uno de los primeros aspectos relevantes en el diseño de entornos digitales tiene que ver, justamente, con la selección de uso que se les dará a las plataformas tecnológicas que dan soporte a la educación en un entorno virtual [3].

Con el objetivo de favorecer el desarrollo de la modalidad virtual de enseñanza y que ella adscribiera las características de un verdadero modelo de innovación, se garantizó que el modelo pedagógico incorporado evidenciara el cumplimiento de las siguientes condiciones:

- Las actividades propuestas en la modalidad virtual debían tener un nivel de integración o de complementariedad con las actividades de tipo presencial para el buen desarrollo del diseño curricular.
- Los argumentos y criterios de las actividades llevadas a cabo en las plataformas virtuales, se sustentarían con contenidos y experiencias de enseñanza-aprendizaje (trabajo colaborativo, actividades de práctica, indagación en la red, discusión entre pares, retroalimentación mediante rúbricas, uso de listas de cotejo, etc.), que fuesen meritorias de ser puestas en práctica en la modalidad mediada.
- La propuesta pedagógica debía asegurar la promoción de una mayor cantidad y calidad de aprendizajes significativos sobre los contenidos curriculares, así como el desarrollo de habilidades cognitivas de alto nivel [5].

La cátedra, ante la continuidad del contexto de aislamiento social obligatorio y teniendo en cuenta las características vinculadas con la modalidad de cursado virtual adaptó su

<sup>2</sup> **MIeL**: Materias Interactivas en Línea

<sup>3</sup> **Microsoft Teams** es una plataforma unificada que permite la colaboración entre personas de un mismo equipo, compartiendo recursos y cuya función principal es la comunicación constante entre los miembros del equipo.

planificación, incorporando las herramientas pedagógicas necesarias para cumplir con los requerimientos de la currícula.

Surgió entonces la necesidad de analizar cómo había resultado esta nueva experiencia dando voz a los alumnos ingresantes, constituidos por aquellos jóvenes que cursan por primera vez una materia en la universidad y también a los alumnos recursantes, quienes se caracterizan por haber comenzado sus estudios universitarios con anterioridad al período analizado, a fin de incluir a todos los inscriptos en la materia. Para ello, se les solicitó su colaboración a través de la realización de una encuesta. En la misma, además de incluir preguntas relativas a aspectos particulares de los alumnos y de la materia, se incluyeron otras preguntas que guardaban correlato con los objetivos que guiaron esta investigación, ellos fueron: consultar al alumnado respecto a cómo valoraron a las distintas herramientas pedagógicas aportadas por las TICs, cuáles consideraban como ventajas y desventajas de la educación en la modalidad virtual y qué tipo de interacción con sus docentes preferían durante las clases.

La innovadora planificación de la propuesta pedagógica llevada a cabo por la cátedra Fundamentos de TICs durante el primer cuatrimestre del año 2020, permitió una trayectoria satisfactoria del proceso virtual de enseñanza, pero nos preguntamos si el aprendizaje hubiese sido significativo sin la interacción sincrónica de los alumnos con los docentes a través de las clases por videoconferencias.

## 2 Materiales y Métodos

Para llevar a cabo la investigación se implementó una encuesta no obligatoria y anónima, dirigida a la totalidad de los alumnos de los siete cursos de primer año de la carrera de Ingeniería de La Universidad de La Matanza en que se dictó la materia Fundamentos de TICs durante el primer cuatrimestre del año 2020.

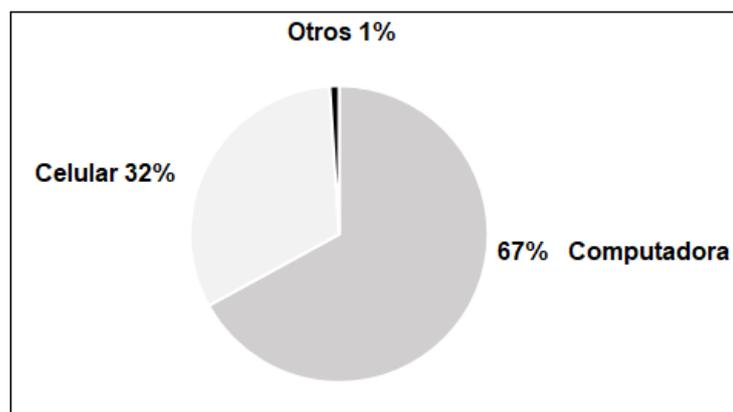
De los cursos consultados en las encuestas, tres de ellos estaban conformados por alumnos ingresantes y otros cuatro cursos más constituidos por alumnos que recursaban la materia. Siendo la cantidad de encuestas respondidas por un total de 241 alumnos (73% de los alumnos que completaron el cursado de la materia). Se realizó una valoración cuantitativa y cualitativa de las respuestas obtenidas como muestran los resultados.

## 3 Resultados

La encuesta contempló distintos aspectos del cursado de la materia.

Se comenzó formulando algunas preguntas con el propósito de contextualizar el ámbito virtual de estudio del alumno, considerando distintos aspectos que pudieron afectar el desempeño académico del educando.

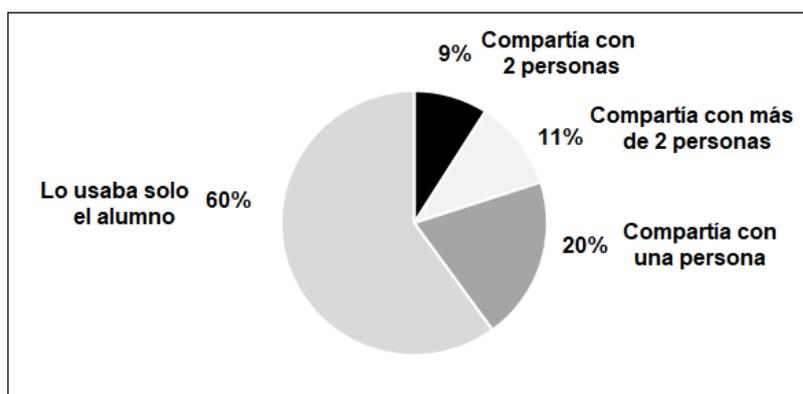
En la Fig. 1 se observan las respuestas de los alumnos al ser consultados acerca de qué dispositivo utilizaron para conectarse a sus clases virtuales.



**Fig. 1:** Dispositivo con que el alumno se conecta a sus clases virtuales.

Como puede observarse en la Fig. 1, la encuesta arrojó que un 67% de los alumnos se conectaron mediante una computadora, un 32% lo hicieron por celular y un 1% a través de una Tablet.

La Fig. 2 refleja las respuestas referidas a la cantidad de personas con quienes el alumno compartió el dispositivo con el cual se conectaba durante sus clases virtuales.

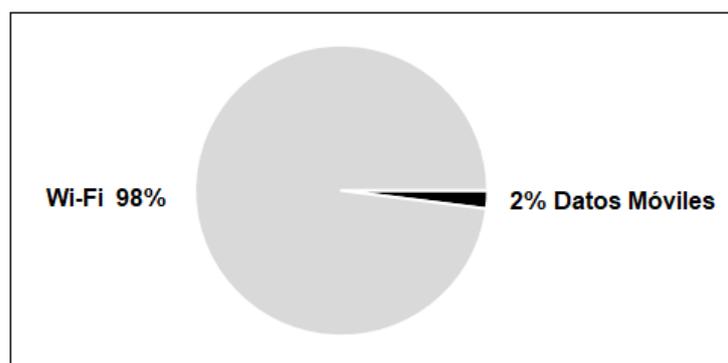


**Fig. 2:** Cantidad de personas con quienes el alumno compartía el dispositivo en su hogar

La Fig. 2, muestra que un 60% del alumnado contó con un dispositivo utilizado únicamente por él, un 20% de los consultados lo compartió con una persona más, en tanto que un 9% lo compartió con dos personas más y un 11% lo hizo con más de dos personas.

Estos porcentajes ponen de manifiesto que el 40% del alumnado, comparte su dispositivo, debiendo coordinar tiempos y actividades con otras personas.

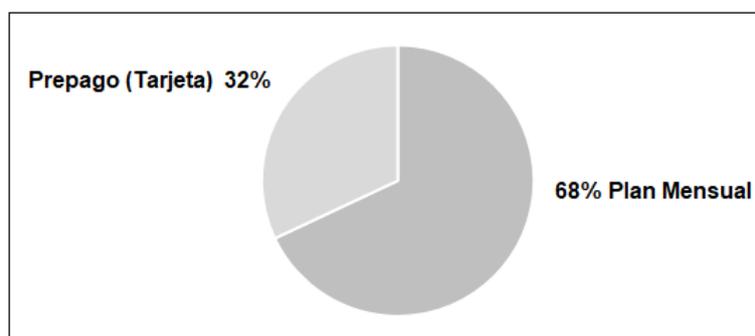
La Fig. 3 muestra el tipo de conexión a Internet que dispone el alumno.



**Fig. 3:** Tipo de conexión a Internet que dispone el alumno

Se puede observar en la Fig. 3, que el 98% del alumnado dispone de conexión Wi-Fi y sólo un 2% lo hace a través de datos móviles.

La Fig. 4 representa el tipo de servicio de celular que dispone el alumno: Prepago o abonando un Plan Mensual.

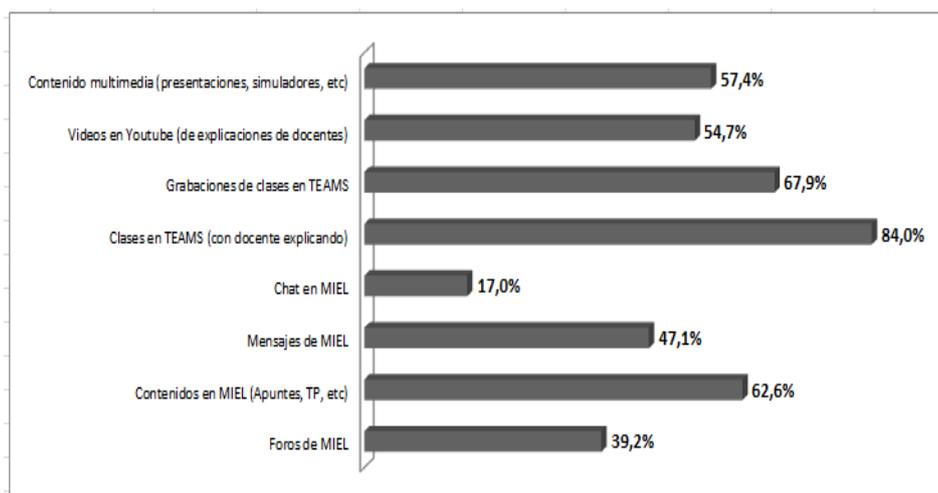


**Fig. 4:** Tipo de servicio de celular que dispone el alumno

El gráfico de la Fig. 4 muestra que el 68% del alumnado cuenta con un plan o abono mensual, mientras que el 32% opta por la modalidad de prepago.

A continuación, se solicitó al alumno que asignara un puntaje a cada una de las herramientas tecnológicas implementadas en cursado virtual en Fundamentos de TICs y en otras materias, concediendo el valor 5 (mayor puntaje) a la herramienta que les ha brindado más ayuda y el valor 1 (menor puntaje) a la que les ha brindado menos ayuda.

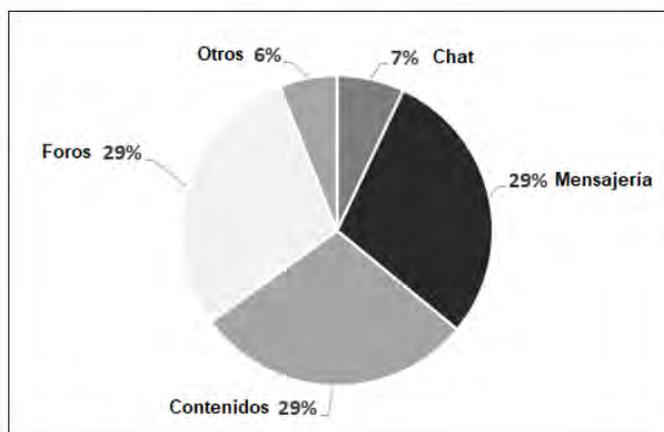
Los resultados obtenidos se pueden observar en la Fig. 5 donde se representa el puntaje asignado por los alumnos a las herramientas utilizadas en las clases virtuales. Cabe destacar que estos porcentajes, representan la cantidad de alumnos que le asignaron la mayor valoración a cada una de las herramientas.



**Fig. 5:** Puntaje asignado por los alumnos a las herramientas utilizadas en las clases virtuales. Se grafica el porcentaje obtenido por cada herramienta con la MAYOR puntuación.

En dicha figura Fig. 5 se observa que la herramienta que representó la mayor ayuda fueron las clases por videoconferencia en la plataforma Teams con el profesor explicando en vivo de manera sincrónica (84%), las grabaciones de las clases llevadas a cabo por videoconferencia (67,9%), los contenidos teóricos accedidos desde la plataforma MIEl (62,6%), los contenidos multimedia tales como presentaciones y simuladores (57,4%), videos en YouTube con explicaciones de los docentes de la cátedra (54,7%), mensajes en la plataforma MIEl (47,1%), foros en el entorno MIEl (39,2%) y Chats en la plataforma MIEl (17%).

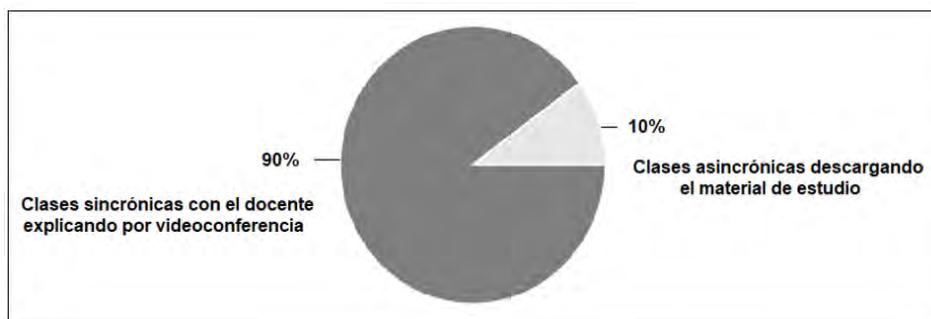
En la siguiente pregunta se consultó a los alumnos qué servicios utilizaron dentro de la plataforma MIEl, la Fig. 6 muestra los porcentajes de uso de cada uno de los servicios.



**Fig. 6:** Servicios de MIEl utilizados por el alumno.

Se puede observar en la Fig. 6 que los servicios más utilizados fueron foros (29%), mensajería (29%) y contenidos (29%), los tres servicios en la misma proporción. Mientras que los servicios menos usados fueron Chat (7%) y otros (6%).

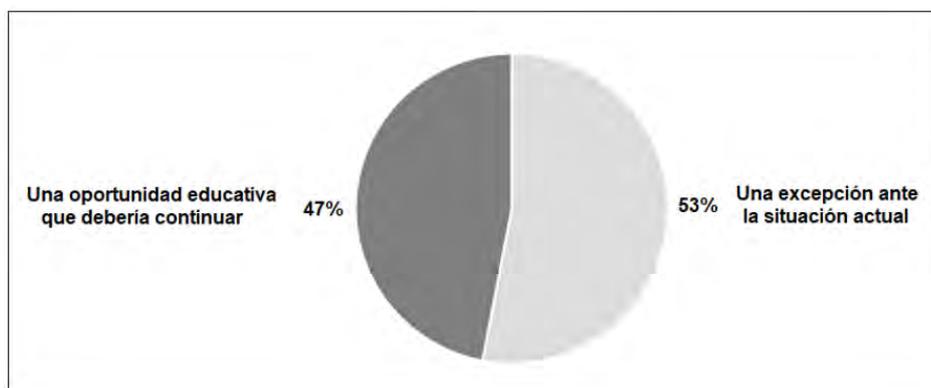
Luego se preguntó a los alumnos qué metodología de cursado de las clases virtuales preferían, pudiendo optar entre las clases on line con el Profesor (clases sincrónicas) explicando como si el docente estuviera en forma presencial o bajar el material (guías de estudio, videos, presentaciones, material multimedia, etc.) y estudiar en forma autónoma (clases asincrónicas), La Fig. 7 representa la preferencia de los alumnos por las clases sincrónicas con el docente explicando por videoconferencia.



**Fig. 7:** Preferencia del alumno respecto de la metodología de dictado de las clases virtuales.

Como puede observarse en la Fig. 7, el 90% prefirió las clases on line con el Profesor (clases sincrónicas) explicando como si el docente estuviera en forma presencial y sólo un 10% optó por bajar el material (guías de estudio, videos, presentaciones, material multimedia, etc.) y estudiar en forma autónoma (clases asincrónicas).

Posteriormente se consultó al alumnado acerca de cómo consideraban a la modalidad virtual de cursado, como se grafica en la Fig. 8.



**Fig. 8:** Consideración del alumno sobre la modalidad virtual.

La Fig. 8 muestra que el 53% de los alumnos la consideraban como una excepción ante la imposibilidad del cursado presencial y un 47% la considera una oportunidad educativa que debería continuar aun cuando finalice el aislamiento social obligatorio. El gráfico permite observar que ambas opciones recibieron valoraciones similares.

Por último se propuso a los alumnos brindar su opinión en relación a la modalidad virtual de cursado mediante cuatro preguntas de respuestas abiertas, que tenían como objetivo promover la reflexión desde diferentes perspectivas y favorecer la libre expresión de los educandos. Estas cuatro preguntas fueron:

**Pregunta 1:** Se solicitó a los alumnos que manifiesten tres aspectos positivos de la modalidad virtual en la materia Fundamentos de TICs. Se muestran a continuación los aspectos más citados en las respuestas:

- Destacan el compromiso y la buena predisposición de los docentes de la cátedra para explicar los temas y responder sus consultas.
- Manifiestan que las clases de los docentes dictadas en videoconferencias de Teams, los trabajos prácticos grupales y la claridad de los apuntes brindados, fueron de gran ayuda para la comprensión de los temas<sup>4</sup>.

<sup>4</sup> La totalidad de las respuestas obtenidas puede verlas en: <https://tinyurl.com/y8hxsxzp>

Pregunta 2: Se solicitó a los alumnos que manifiesten tres aspectos negativos de la modalidad virtual en la materia Fundamentos de TICs. Se muestran a continuación los aspectos más citados en las respuestas:

- La mayoría de las dificultades enunciadas por los alumnos se relacionan a problemas de conectividad.
- Se destaca que varios alumnos manifestaron no tener puntos negativos en relación a esta materia<sup>5</sup>.

Pregunta 3: Se solicitó a los alumnos que manifiesten las mayores dificultades de la modalidad virtual, en todas las materias cursadas este cuatrimestre. Se muestran a continuación los aspectos más citados en las respuestas:

- En los casos en los que no tuvieron clases en videoconferencias de Teams dictadas por los docentes, manifiestan la dificultad de comprensión de temas teóricos complicados, dificultad para resolver la parte práctica y no contar con un docente que responda sus dudas.
- Comentan también dificultades en las conexiones, así como los problemas de concentración por distracciones en el hogar y falta de organización de su tiempo.
- Varios alumnos citaron la necesidad de tener contacto personal, tanto con sus docentes como con sus compañeros<sup>6</sup>.

Pregunta 4: Se solicitó a los alumnos que manifiesten los principales puntos positivos de la modalidad virtual, en todas las materias cursadas este cuatrimestre. Se muestran a continuación los aspectos más citados en las respuestas:

- Comentan como beneficioso poder cursar desde sus hogares, y el aprovechamiento del tiempo que destinarían en viajar.
- Destacan como positiva la disponibilidad y facilidad de acceso a los materiales didácticos (teóricos, multimedia y grabación de las clases por Teams)<sup>7</sup>.

## 4 Discusión

Las propuestas de desarrollo del diseño curricular adoptadas por las cátedras universitarias, deberían partir del hecho de proponer el cursado de las materias como un “*entorno abierto de aprendizaje*” (el cual puede ser presencial, virtual o híbrido) y al mismo tiempo, como la configuración de una auténtica “*comunidad de aprendizaje*” [6].

Sin embargo el cursado de materias mediante la modalidad virtual no se había implementado en la totalidad de las cátedras de las carreras de primer año de Ingeniería con anterioridad a la imposibilidad del cursado presencial.

El análisis de la contextualización del ámbito virtual de estudio del alumnado brinda como información que los alumnos mayoritariamente contaron con recursos tecnológicos que permitieron el cursado virtual de sus clases, pero lamentablemente revelan que frecuentemente experimentaron dificultades de conectividad en las zonas de residencia, lo que refleja una deficiente red de comunicación para sostener una educación virtual ante la imposibilidad de la presencialidad.

<sup>5</sup> La totalidad de las respuestas obtenidas puede verlas en: <https://tinyurl.com/yaq2fk2l>

<sup>6</sup> La totalidad de las respuestas obtenidas puede verlas en: <https://tinyurl.com/y9bowvh4>

<sup>7</sup> La totalidad de las respuestas obtenidas puede verlas en: <https://tinyurl.com/y6v8dxnf>

En el primer año de las carreras de Ingeniería de la institución, se contaban con pocos antecedentes de educación mediada, si bien es deseable que en el nivel universitario se promueva el cursado de materias en un “*entorno abierto de aprendizaje*”, en otras palabras, aquellas propuestas pedagógicas en las que:

- a) Se plantean situaciones problemáticas complejas vinculadas a casos reales.
- b) Los alumnos tienen oportunidades para realizar proyectos y actividades constructivas de forma individual y colaborativa en los que analizan, interpretan e investigan.
- c) Se vinculan estrechamente las actividades constructivo-cognitivas para aprender y desarrollar destrezas complejas en contextos relativos al conocimiento aprendido;
- d) se utilizan una variedad de herramientas pedagógicas, recursos provistos por las nuevas tecnologías y recursos de información de diversas fuentes [7].

Entre los principales puntos positivos de la modalidad virtual se destaca que al poder cursar desde sus casas, los alumnos hacen un mejor aprovechamiento del tiempo, esto se debe a que el aprendizaje ubicuo<sup>8</sup> se relaciona también con una nueva forma de lectura, en la que se aprovechan los tiempos muertos de traslados o esperas, y se lee en espacios de tiempo denominados “*burbujas de ocio*” [8].

Otro aspecto señalado por el alumnado como positivo de la modalidad virtual enfatiza la disponibilidad y facilidad de acceso a los materiales didácticos (teóricos, multimedia y grabación de las clases por Teams). Esto se halla en concordancia con el aprendizaje ubicuo: “*la posibilidad de acceder a la información en cualquier lugar o cualquier momento, la interacción con pares y expertos eruditos genera oportunidades estructuradas de aprendizaje desde una variedad de fuentes*” [9].

La incorporación de nuevas herramientas tecnológicas puede aportar distintas modalidades y dinámicas para lograr nuevos conocimientos. Dado que “*...los medios son extensiones de las capacidades del hombre...*” y, como docentes, debemos promover acciones para extender esas capacidades [10]. Teniendo en cuenta que “*...a medida que se utiliza frecuentemente una tecnología se la naturaliza...*” [11].

El alumnado también resalta en sus respuestas abiertas, que los trabajos prácticos grupales y la claridad de los apuntes brindados, fueron de gran ayuda para la comprensión de los temas. Esto evidencia que la cátedra al elaborar la planificación pedagógica tuvo especial consideración que en la modalidad de enseñanza mediada, es de vital importancia la gestión (diseño, elaboración, uso y evaluación) de los materiales didácticos ya que “*...son uno de los ejes vertebradores de gran parte de las acciones de enseñanza desarrolladas en cualquiera de los niveles y modalidades de educación...*” [12].

En relación a la preferencia de los alumnos referida a la metodología de cursado de las clases virtuales mayoritariamente optaron por las clases on line con el profesor por videoconferencia (clases sincrónicas) explicando como si el docente estuviera en forma presencial, revelando la importancia de la interacción sincrónica entre docentes y alumnos durante las clases virtuales. Cabe destacar también, el consenso de los educandos al resaltar el compromiso y la buena predisposición de los docentes de la cátedra para explicar los temas y responder sus consultas. Demostrando que en este proceso de enseñanza-aprendizaje en un entorno virtual, el docente se desempeña como mediador y facilitador del trayecto formativo: “*... el docente deberá transformarse en un guía o facilitador de los aprendizajes de sus alumnos...*” [13].

<sup>8</sup> **Aprendizaje ubicuo** es el que se produce en cualquier lugar y momento.

## 5 Conclusión

El modelo pedagógico que se implementó en el dictado de la materia, consolidó que es posible generar una propuesta conjunta de experiencias sincrónicas llevadas a cabo a través de videoconferencias en la plataforma Teams y otras experiencias asincrónicas a través de la plataforma MIEl, en el sentido de que por su propia complementariedad se potencian mutuamente, creando así un entorno abierto con múltiples intercambios comunicativos entre profesores y alumnos que promovió una mayor profundización y apropiación del conocimiento. Se concluye que la modalidad virtual puede brindar una trayectoria exitosa en el cursado de las materias de primer año en las carreras de Ingeniería, produciendo una educación mediada de calidad.

## Bibliografía

1. Díaz Barriga, F. y Lugo, E.: Desarrollo del currículo. En: A. Díaz Barriga (Coord.). La investigación curricular en México. La década de los noventa, vol. 5, p.p. 63-123. México: Consejo Mexicano de Investigación Educativa (2003)
2. Scorzo R., Favieri A., Williner B.: Desarrollo de un espacio de enseñanza aprendizaje para realizar actividades con uso de software en una cátedra numerosa. Revista Iberoamericana en Educación y Educación en Tecnología (TE&ET), número 7, pág.77-83. En: <https://teyet-revista.info.unlp.edu.ar/TEyET/article/view/926> (2018)
3. Díaz Barriga F., Hernández G., Bustos Sánchez A., Romero V.: Experiencias educativas con recursos digitales: prácticas de uso y diseño tecnopedagógico, pp.121-140, Capítulo 5: Experiencias b-Learning enfocada a la construcción colaborativa de WebQuest con estudiantes de posgrado en Psicología y Pedagogía. Publicado por Universidad Nacional Autónoma de México (2011)
4. Daniels, H.: Vygotsky y la pedagogía. Barcelona. Paidós (2003)
5. Bustos, A.: Un modelo para Blended Learning aplicado a la formación en el trabajo. Compartimos practicas, ¿compartimos saberes? Revista Iberoamericana de Educación a Distancia, 1/2, 113-132  
En: [http://www.utpl.edu.ec/ried/index.php?option=com\\_content&task=view&id=370&Itemid=112](http://www.utpl.edu.ec/ried/index.php?option=com_content&task=view&id=370&Itemid=112) (2004)
6. Wenger, E.: Comunidades de práctica. Aprendizaje, significado e identidad. Barcelona: Paidós. (2001).
7. Hannafin, M., Land, S. y Oliver, K.: Entornos de aprendizaje abiertos: fundamentos, métodos y modelos. En Ch. Reigeluth. (Ed.). Diseño de la instrucción. Teorías y modelos Parte 1, p.p. 125-152. Madrid: Santillana (2000).
8. Igarza, R.: Burbujas de ocio. Nuevas formas de consumo cultural. Buenos Aires. Ediciones La Crujía (2009)
9. Burbules, N.: El aprendizaje ubicuo y el futuro de la enseñanza. Rencontres on Education. Vol. 13, pp. 3-14 (2012)
10. McLuhan, M.: Comprender los medios de comunicación. Las extensiones del ser humano. Barcelona. Ed. Paidós (1994)
11. Ong, Walter.: Oralidad y Escritura. Fondo de cultura económica. Capítulo 4 (1994)
12. Área Moreira, M.: Manual electrónico: Introducción a la Tecnología educativa, p24, Universidad de La Laguna (España) (2009)
13. Marín Díaz, V., & Romero López, M. A.: La formación docente universitaria a través de las TICS. Pixel. Bit. Revista de Medios y Educación. (2009)

## Formación de Competencias de la Industria 4.0 en Estudiantes de Primer Año de Ingeniería

Artemisa Trigueros<sup>1</sup>, Mabel Compagnoni<sup>1</sup>, Larisa Toro<sup>1</sup>, Sabrina Gómez<sup>1</sup>,

<sup>1</sup> Universidad Nacional de La Matanza, Florencio Varela 1903,  
B1754 San Justo, Pcia. de Buenos Aires, Argentina  
{artemisa, mcompagnoni, lara, sggomez, [@unlam.edu.ar](mailto:@unlam.edu.ar)

**Abstract.** La sociedad y la tecnología evolucionan vertiginosamente en estos tiempos. Los jóvenes actuales consideran a los teléfonos celulares como extensiones de su propio cuerpo y dominan su uso desde edades muy tempranas. Estamos transitando la Cuarta Revolución Industrial que requiere competencias y conocimientos encuadrados en la llamada Industria 4.0. Paralelamente las Ciencias de la Educación proponen adoptar metodologías que incluyan las herramientas tecnológicas propuestas por la Educación 4.0. Este trabajo muestra la incorporación de las tecnologías y los temas requeridos por la Industria 4.0, en una materia de Primer Año de Ingeniería de la Universidad Nacional de La Matanza (UNLaM), utilizando las metodologías propuestas por la Educación 4.0, para formar en los futuros ingenieros las competencias y conocimientos requeridos por la Industria 4.0. desde el primer año de sus carreras.

**Keywords:** Industria 4.0, Educación 4.0, Fundamentos de TICs, Competencias para Ingeniería.

### 1 Introducción

La materia Fundamentos de TICs (Tecnologías de la Información y la Comunicación), perteneciente al 1º Año de todas las Carreras de Ingeniería de UNLaM, constituye uno de los primeros contactos académicos formales, que tienen los alumnos con las TICs, cuya actividad profesional al graduarse, consistirá en desenvolverse en el marco de la “Cuarta Revolución Industrial” o Industria 4.0. Paralelamente, la Educación 4.0 es una respuesta pedagógica a los requerimientos de competencias específicas de la Industria 4.0, colocando al estudiante en el centro del ecosistema de la educación superior [1].

Los educandos deben ser protagonistas en sus procesos de aprendizaje, dado que toda persona aprende “haciendo”, por lo cual, se deben implementar herramientas pedagógicas que posibiliten potenciar y satisfacer los requerimientos que estamos viviendo en esta Cuarta Revolución Industrial promoviendo la innovación y las

tendencias y a partir de ahí preparar a los estudiantes para desenvolverse satisfactoriamente en este mundo globalmente conectado y tecnológico.

La Educación 4.0 propone el aprendizaje activo, la colaboración como base del proceso de aprendizaje, la interacción y retroalimentación entre docentes y alumnos, el aprendizaje competencial favoreciendo la apropiación de conocimientos para resolver problemas reales, utiliza el juego y los casos reales como motor de aprendizaje. Asimismo, la evaluación es un proceso continuo de retroalimentación que promueve el progreso del estudiante. En este entorno, las TICs están incluidas como herramientas educativas. Particularmente, en este trabajo, las TICs cumplen una doble función, por un lado, como herramientas de la Educación 4.0, por el otro son el objeto de estudio de la materia Fundamentos de TICs.

Estas tecnologías, incorporadas por la cátedra de Fundamentos de TICs en el desarrollo de las clases tienen como objetivo mejorar y actualizar la transmisión de contenidos, facilitando el aprendizaje, adquisición de nuevos conocimientos, pensamiento crítico, analítico, estrategias de aprendizaje, creatividad, originalidad e inteligencia emocional, propuestas por la Educación 4.0. La implementación de tecnologías como realidad aumentada, realidad virtual, diseño de modelos, problemas de inteligencia artificial, impresiones 3D, Big Data, resolución de problemas reales en forma grupal y colaborativa utilizando medios tecnológicos, etc. colocará al alumno frente a situaciones reales de trabajo, formando en él las competencias requeridas por la Industria 4.0.

## 2 Encuadre General

La asignatura Fundamentos de TIC's, pertenece al Ciclo General de Conocimientos Básicos (CGCB) de las cinco carreras de Ingeniería de la UNLaM: Ing. Civil, Electrónica, Industrial, Informática y Mecánica, formando parte del primer (o segundo) cuatrimestre de la currícula de Primer Año. La cátedra adscribe como principal objetivo promover en el alumno conocimientos informáticos y tecnológicos, introduciéndolo además, en temas que serán desarrollados en profundidad en años posteriores de la carrera, a fin de brindar un panorama general que será de gran utilidad en el perfil del ingeniero cuando egrese. Este objetivo apunta a generar actitudes que lleven al alumno a tomar contacto con la realidad de Mercado en un contexto tecnológico, iniciarlo en las nuevas tecnologías y sus características, lo que permitirá un adecuado desempeño en su actividad futura y estar en condiciones de interpretar la mayoría de los conceptos que normalmente se emplean en los ambientes dedicados a las nuevas tecnologías de la información y la comunicación [2]. Asimismo, el Libro Rojo del CONFEDII formula competencias de egreso genéricas y específicas comunes a todas las carreras de ingeniería y necesarias para asegurar el perfil de egreso, que son incorporadas en la currícula de la materia. [3].

### 3. Industria 4.0 y Educación 4.0

Industria 4.0: La llamada Industria 4.0, es el resultado de la llamada 4<sup>ta</sup>. Revolución Industrial, que tiene como objetivo, aumentar la productividad, donde las TICs cobran un rol preponderante, mediante la utilización de las redes de comunicación (Internet) que permitirán la comunicación entre hombres y máquinas en sistemas ciber-físicos, recolectando y analizando enormes cantidades de datos, que permitirán producir bienes tangibles e intangibles de mayor calidad a un costo más reducido.

La industria 4.0 se basa en las siguientes tecnologías: Big Data, Cyberseguridad, Fabricación por Adición, Fábricas digitalizadas que jerarquizan al factor humano, Internet de las Cosas, Realidad Aumentada, Robótica, Simulaciones, Software basado en la nube. “La generación y adopción de estas tecnologías aumenta la productividad y competitividad de toda la economía, con el consecuente impacto positivo sobre la creación de empleos” [4].

La flexibilidad tiende a ser la forma de actuar tanto de las personas como de empresas e industrias. Hay que adaptarse mucho más rápidamente de lo que se adaptaban las personas en otras épocas. Hoy quienes se incorporan a las empresas o a las fábricas, trabajan en ellas durante un período, aprenden lo que consideran adecuado y cambian de trabajo, requiriendo para eso mucha flexibilidad. Ese factor va a ser imperativo para los próximos tiempos [5]. “La capacidad de trabajar en equipo según los métodos de colaboración propios de entornos laborales menos jerárquicos y estructurados, más tecnológicos y dinámicos se vuelve fundamental. Lo que hay que tener claro, sin embargo, es que la cuarta revolución industrial va de personas, no de máquinas” [6].

Educación 4.0: El proceso de enseñanza - aprendizaje se ha transformado, evolucionado con respecto a improntas anteriores, centrando su atención en modelos dirigidos al aprendizaje [7]. Dentro de esta evolución las TICs, se convierten en herramientas poderosas para la comunicación directa o indirecta, entre docentes y alumnos, apoyada por medios auxiliares [8]. Dado que los estudiantes son “nativos digitales”, poseen “...una habilidad innata en el lenguaje y en el entorno digital. Las herramientas tecnológicas ocupan un lugar en sus vidas y dependen de ellas para todo tipo de cuestiones cotidianas como estudiar, relacionarse, mensajearse, comprar, informarse o divertirse” [9].

Las TICs sirven para motivar y estimular al estudiante a que sea protagonista del proceso de aprendizaje, interactuando con la realidad, analizando los resultados de esta interacción a fin de desarrollar habilidades de pensamiento crítico y creativo y de esta manera, facilitar la comprensión de lo que se ha aprendido de manera integral y dinámica [10]. Es necesaria la búsqueda e implementación de innovaciones educativas basadas en las tecnologías que consoliden nuevas actitudes y expectativas de los alumnos, así como sus nuevas competencias. “El docente tiene que explorar un nuevo desafío que plantea la educación actual. La baja dosis atencional de los estudiantes nos obliga a reelaborar nuestras clases repensando formatos que despierten la curiosidad y el interés del alumnado por los contenidos que les presentamos” [11].

#### 4. Estrategias Didácticas

La Educación 4.0 propone fomentar el pensamiento crítico y analítico, la creatividad, la originalidad y la inteligencia emocional y la metacognición como base del aprendizaje significativo. Las TICs están incluidas como herramientas pedagógicas y al mismo tiempo son el objeto de estudio de la materia Fundamentos de TICs, como por ejemplo: realidad aumentada, realidad virtual, diseño de modelos, problemas de inteligencia artificial, impresiones 3D, Big Data. Las estrategias propuestas son:

- La resolución de situaciones problemáticas reales en forma grupal por medios tecnológicos, simuladores, aplicaciones móviles, Campus Virtual, creación y publicación de videos presentando sus proyectos, etc. que coloca al alumno frente a situaciones reales de trabajo, formando en él las competencias requeridas por la Industria 4.0.
- La utilización de entornos de aprendizaje reales, estudio de casos, juegos, desafíos, para promover el aprendizaje de competencias por medio de resolución de problemas reales del ámbito ingenieril.
- La estrategia de aprendizaje se realizará en base a metodologías activas.
- La colaboración como base del proceso de aprendizaje afianzado la comunicación e interacción entre docentes y alumnos.
- El docente debe generar un contexto favorable al aprendizaje donde los estudiantes sean los protagonistas.
- La apropiación de conocimientos se logra como consecuencia del análisis y resolución de situaciones problemáticas teórico-prácticas, sumando la retroalimentación del docente que favorece la evolución del alumno.
- El espacio de reflexión siempre tiene un lugar fundamental.
- El trabajo en equipo para que los estudiantes resuelvan problemas y ejercicios.
- La inclusión de Rúbricas y/o Listas de Cotejo en cada Trabajo Práctico Obligatorio, explicitando al alumno y grupo de desarrollo, los contenidos y actitudes a desarrollar, los criterios de evaluación de cada contenido y los niveles de logro alcanzados.
- Charlas y demostraciones con los profesionales del polo Tecnológico en temas como Industria Aditiva (Impresión 3D), Realidad Aumentada y Realidad Virtual.

#### 5. Unidades

La materia Fundamentos de TICs cuenta con 6 (seis) Unidades Temáticas. Cada una de ellas comprende un tema específico de las TICs.

El desarrollo de contenidos comienza con una Introducción a la Teoría de Sistemas, continuando con Sistemas de Representación de la Información y Códigos para su transmisión segura, introduciendo al estudiante en el mundo digital y binario, Circuitos Lógicos, ya que por medio de éstos se diseñan los sistemas de procesamiento de datos y control de procesos, Hardware, integrando las unidades anteriores, Redes de Comunicación para interconectar sistemas y Software necesario para que dichos sistemas se desempeñen en forma automática.

## 6. Aplicaciones de Herramientas de la Educación 4.0 a las Unidades de Fundamentos de TICs para crear competencias requeridas por la Industria 4.0

Los temas curriculares de la materia Fundamentos de TICs, corresponden al Plan de Estudios 2009. En estos 11 años, la sociedad y la tecnología han realizado grandes cambios. Basta pensar en la difusión de los dispositivos móviles, que han cambiado el perfil del alumno, y de los docentes, al permitir tener en la palma de la mano, toda la información y los recursos comunicativos necesarios al instante.

Para lograr el objetivo de creación de competencias requeridas por la Industria 4.0, en los estudiantes de 1<sup>er</sup>. Año, la cátedra de Fundamentos de TICs, incorporó estudio de casos y ejemplos de la Industria 4.0 y metodologías activas de la Educación 4.0, que pertenecen a los grandes temas curriculares de la materia.

Estos cambios se basan fundamentalmente en la figura de Trabajos Prácticos Obligatorios (TPO), donde se plasman distintas tecnologías y metodologías y en charlas de transferencia profesional y exposiciones de Ingenieros y equipos de investigación pertenecientes al Polo Tecnológico de la UNLaM.

Los TPO, tienen la característica imprescindible de ser Trabajos Prácticos Grupales, ya que se persigue crear en los alumnos recién incorporados a la Universidad, las siguientes competencias sociales, políticas y actitudinales:

- Desempeñarse de manera efectiva en equipos de trabajo.
- Comunicarse con efectividad.
- Actuar con ética, responsabilidad y compromiso social, considerando el impacto económico, social y ambiental de su actividad en el contexto local y global.
- Aprender en forma continua y autónoma.
- Actuar con espíritu emprendedor.

Asimismo se busca comenzar a incorporar las siguientes competencias tecnológicas:

- Identificar, formular y resolver problemas de ingeniería.
- Concebir, diseñar y desarrollar proyectos de ingeniería.
- Gestionar, planificar, ejecutar y controlar proyectos de ingeniería.
- Utilizar de manera efectiva las técnicas y herramientas de aplicación en la ingeniería.
- Contribuir a la generación de desarrollos tecnológicos y/o innovaciones tecnológicas.
- Promover el aprendizaje activo a través del pensamiento estratégico aplicando herramientas tecnológicas.

A continuación, se describen, como ejemplo en Tabla 1, los TPOs pertenecientes a cada unidad temática. Es de destacar que dichos TPO, pueden incluir otros temas y/o casos de ejercitación y/o otras tecnologías a medida que se éstas van surgiendo o como inquietud de los alumnos.

Los TPO se realizan en clase y posteriormente se publican en Internet, para que todos los alumnos de la materia puedan acceder a todos los trabajos.

Tabla 1. TPO POR CADA UNIDAD

Unidad TPO	Situación Propuesta
Unidad 1 TPO 1	A. Desarrollo de un proyecto tangible simple que requiera aplicar las 5 etapas de desarrollo de cualquier sistema, de acuerdo a requerimientos de un supuesto cliente. B. Resolución de casos reales en los que el alumnado debe aplicar los temas vistos referidos a Sistemas de Información para identificar, formular y resolver problemas de ingeniería.
Unidad 2 TPO 2	A. Investigación sobre Realidad Aumentada y su utilización en la Industria 4.0. Confección de un texto expositivo con publicación en un foro grupal. B. Acorde a lo investigado en la Parte A, aplicación en un caso concreto, utilizando códigos BCD, códigos QR y realidad aumentada. (Aprendizaje Creativo basado en problemas concretos). Transferencia Profesional: Charla sobre Realidad Aumentada y Realidad Virtual en el Polo Tecnológico de la UNLaM
Unidad 3 TPO 3	A. Utilización de software Logisim para simular la creación de circuitos lógicos, requeridos por los docentes y para la simplificación de funciones algebraicas.
Unidad 4 TPO 4	A. Simulación del Procesamiento de instrucciones utilizando la aplicación por medio de SIMUPROC. B. Investigación sobre Periféricos. Armado de un Video colaborativo grupal y su publicación en una pizarra digital, con posterior retroalimentación por parte de sus pares. Transferencia Profesional: Charla sobre Impresoras 3D en el Polo Tecnológico.
Unidad 5 TPO 5	A. Análisis de un caso similar a los reales, que plantea la necesidad de armado de una Intranet y propone la búsqueda de la óptima solución que satisfaga el requerimiento. B. Confección de una Presentación dinámica aplicando herramientas tecnológicas de la Web 2.0 basada en la investigación grupal sobre aplicaciones de IoT (Internet de las Cosas) y Cyberseguridad. Posterior publicación en una pizarra digital y retroalimentación de sus pares.
Unidad 6 TPO 6	A. Investigación sobre casos de uso de Inteligencia Artificial, Robótica, Big data y Software Basado en la Nube por las empresas. Confección colaborativa de una infografía digital Posterior publicación en una pizarra digital y retroalimentación de sus pares.

Todo el material de cátedra, inclusive los TPO descritos en la Tabla 1, se encuentra publicado en MIE<sup>2</sup>[2]. Durante las clases, los alumnos utilizan activamente el teléfono celular para realizar ejercicios, plantear cuestionarios, mostrar procesos, etc. Las aplicaciones Kahoot<sup>3</sup> y PADLET<sup>4</sup> y Simuladores de distribución gratuita son utilizadas para practicar y compartir conocimientos.

<sup>2</sup> Material Interactivas en Línea

<sup>3</sup> Kahoot. Disponible en <https://kahoot.com/>

## 7. Anexo: ejemplos de trabajos prácticos obligatorios y testimonios de transferencia profesional

**Consigna del Proyecto (TPO 1. B):** Los alumnos se reunirán en grupos de 4 o 5 integrantes y dispondrán de 30 minutos para debatir en forma colaborativa acerca de las posibles soluciones para el estudio de caso planteado a través de las situaciones enunciadas a continuación, aplicando los conceptos analizados en la unidad 1 que consideren útiles para que un Ingeniero pueda cumplir en forma óptima su tarea. Luego cada grupo presentará por escrito lo producido. La Fig. 1 plantea la Situación 2, correspondiente a la siguiente consigna: Implementar algún mecanismo de desacoplamiento que solucione la situación actual de producción.

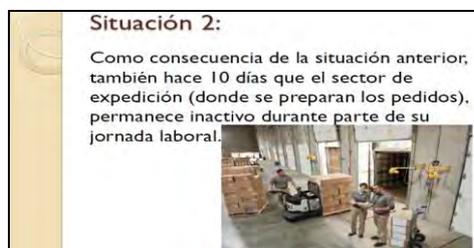


Fig. 1 TPO 1. B. Planteo de la Situación 2.

La Tabla 2, presenta las Rúbricas correspondientes a la Situación 2 del TPO 1.

**Tabla 2.** RÚBRICAS PARA TPO 1

Criterios	Niveles de Logro			
	1	2	3	4
Identificación e implementación de mecanismos de desacoplamiento.	Dificultad en la identificación de mecanismos de desacoplamiento.	Reconoce algunos mecanismos de desacoplamiento pero presenta dificultad en la implementación de los mismos.	Reconoce todos los mecanismos de desacoplamiento pero se implementan sólo algunos de ellos.	Reconoce todos los mecanismos de desacoplamiento y los implementa en forma satisfactoria.

Como puede observarse en la Tabla 2, se explicitan los Criterios y los requerimientos para alcanzar cada Nivel de Logro.

TPO 4. Parte B: Investigación sobre Periféricos. Armado de un Video colaborativo grupal y publicación del mismo en una pizarra digital, con posterior retroalimentación por parte de sus pares que se introducía como comentario del postcard incorporado en

<sup>4</sup> PADLET. Disponible en <https://es.padlet.com/auth/login>

la pizarra por cada grupo conteniendo el video colaborativo realizado. La Fig. 2 muestra los TPO 4 de un curso.



Fig. 2: TPO 4. Imagen de una de las pizarras realizadas por una comisión de la cátedra El TPO 4 incluye Transferencia Profesional sobre Impresoras 3D, (Fig.3).



Fig. 3: TPO 4. Charla sobre Impresoras 3D en el Polo Tecnológico y Brazo ortopédico. (Nov. 2019) Los TPO contaron con gran entusiasmo por parte del alumnado.

**8. Resultados**

2019 fue el primer año de implementación de la nueva metodología. Los resultados obtenidos fueron muy satisfactorios, ya que se logró un mayor porcentaje de aprobación y se disminuyó la cantidad de Ausentes, como muestra la Fig. 4.

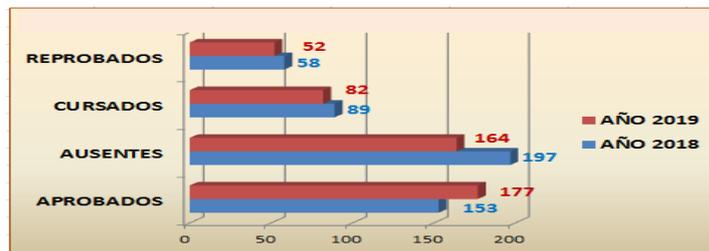


Fig. 4: Comparación de Porcentajes de 2º Cuatrimestre 2018 y 2019.

Como puede observarse en la Fig. 4, los porcentajes de la Cursada 2019, muestran que aumentó en un 6% el porcentaje de Aprobados, y disminuyó un 4% los Ausentes, con respecto al Año 2018, en el que no se realizaban los TPO.

Durante la cursada fue realizada una Encuesta de Cátedra, donde los estudiantes valoraron la nueva metodología en forma muy positiva. La Fig. 5 muestra los resultados de la encuesta por medio de un gráfico de sectores.

La pregunta realizada fue si “Los TPO mejoran la comprensión de los temas.”

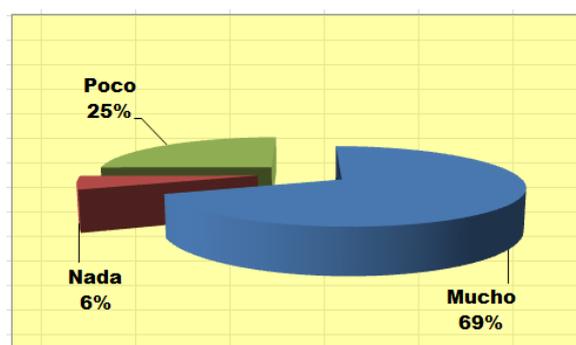


Fig. 5: Consulta a los alumnos sobre los TPO.

Como puede observarse en la Fig. 5 el 69% de los estudiantes respondió positivamente, y sólo un 6% consideró que no.

## 9. Conclusiones

La metodología de Educación 4.0 implementada, fue recibida y ejecutada en forma muy satisfactoria por los estudiantes y docentes, mejorando los resultados de años anteriores. El proyecto fue un éxito para la cátedra y los estudiantes.

A través del uso pedagógico de las TIC's, han surgido nuevos planteamientos pedagógicos que resaltan la prioridad del aprendizaje sobre la enseñanza; el papel clave del alumno y la consiguiente necesidad de acudir a metodologías en las que el estudiante universitario, tome las riendas de su aprendizaje con la dirección, la orientación y el apoyo de su profesor. Se privilegia el proceso formativo basado en una evaluación continua y en la continua propuesta de tareas [12]. El aprendizaje colaborativo, añade a los logros académicos repercusiones educativas como el compañerismo, la solidaridad y la capacitación para el trabajo en equipo.

Las herramientas tecnológicas de la Educación 4.0 tienden a favorecer un aprendizaje más complejo e integrador basado en las asociaciones, comparaciones y un pensamiento crítico y reflexivo promoviendo el desarrollo de la capacidad de análisis, lo que permite a los alumnos, conocer más profundamente y construir nuevos conocimientos a partir de otros previos y al docente adquirir información valiosa para reorientar el proceso de enseñanza y colaborar en el logro de aprendizajes.

## 10. Trabajo Futuro

La cátedra se propone seguir avanzando en la aplicación de nuevas herramientas y perfeccionando los casos planteados en los TPOs. Asimismo, se propone realizar nuevos cursos de capacitación docente sobre metodologías de Educación 4.0.

## Reconocimientos

La cátedra de Fundamentos de TICs, agradece al Ing. Daniel Cundari y equipo por las demostraciones sobre impresiones aditivas a los alumnos y al Ing. Mavrommatis por su capacitación en Técnicas Activas aplicadas al proceso de enseñanza y aprendizaje.

## Referencias

1. V. Moctezuma. La Universidad del Futuro y la Industria 4.0. <https://ilab.net/la-universidad-del-futuro-y-la-educacion-4-0>
2. MIEL. Materias Interactivas En Línea. UNLaM. <http://miel.unlam.edu.ar>  
Libro Rojo de CONFEDI.
3. [https://confedi.org.ar/download/documentos\\_confedi/LIBRO-ROJO-DE-CONFEDI-Estandares-de-Segunda-Generacion-para-Ingenieria-2018-VFPublicada.pdf](https://confedi.org.ar/download/documentos_confedi/LIBRO-ROJO-DE-CONFEDI-Estandares-de-Segunda-Generacion-para-Ingenieria-2018-VFPublicada.pdf)
4. G. Nieponice. <https://www.apertura.com/negocios/Que.es.la-Industria-4.0-y-cuanto-falta-para-que-llegue-a-la-Argentina-20180425-0007.html>
5. L. De Bernárdez. <https://www.infobae.com/def/development/2019/01/08/la-industria-4-0-una-revolucion-en-las-fabricas>
6. A. Raya. <https://www.lavanguardia.com/tecnologia/20190307/46893035712/industria-40-empleo-formacion-profesional-profesionales-mano-de-obra-espana.html>
7. Edel, R. El Concepto de Enseñanza-Aprendizaje. Revista Electrónica Red Científica. Ciencia, Tecnología y Pensamiento. 2004
8. Betancur, D., Moreno, J. & Ovalle, D. A. Modelo para la Recomendación y Recuperación de Objetos de Aprendizaje en Entornos Virtuales de Enseñanza/Aprendizaje. Revista Avances en Sistemas e Informática. 2009.
9. García, M. R. Interacción y Comunicación en Entornos Educativos: Reflexiones Teóricas, Conceptuales y Metodológicas. Revista de Asociación Nacional de Programas de Posgrado en Comunicación (E-Compós). Vol. 8. Disponible en <https://doi.org/10.30962/ec.143>. 2007.
10. Esteve, F.M. & Gisbert, M. (2011) El Nuevo Paradigma de Aprendizaje y las Nuevas Tecnologías. Revista de Docencia Universitaria. REDU. Vol.9 (3). Pag.s 55 – 73.
11. Lazzatti, Pablo. "Jugarle a la cabeza: una apuesta al aprendizaje lúdico y las neurociencias en la enseñanza del Derecho." Anales de la Facultad de Ciencias Jurídicas y Sociales (2017).
12. Pérez Justé, R. (Coord.) (2012). El portafolio. Aprendizaje, competencias evaluación. Universidad de Educación a Distancia (UNED), Madrid, España. Capítulo 4.

## Framework para configuración de batallas históricas: Generando a través del juego un espacio de aprendizaje

Christian Parkinson<sup>1</sup>, Roxana Martínez<sup>1</sup>, Pablo Vera<sup>1</sup>, Rocío Rodríguez<sup>1</sup>

Centro de Altos Estudios en Tecnología Informática (CAETI).  
Universidad Abierta Interamericana,  
Ciudad Autónoma de Buenos Aires, Argentina

{Christian.Parkinson, Roxana.Martinez, PabloMartin.Vera,  
RocioAndrea.Rodriguez}@uai.edu.ar

**Abstract.** Este trabajo se aboca al análisis de juegos serios en el ámbito académico y la implicancia de su contexto en lo que respecta el área de historia. Si bien la cantidad de videojuegos basados en acontecimientos históricos es amplia, se encuentran condicionadas al desarrollo de expansiones o actualizaciones para agregar nuevos escenarios, objetivos y recursos. En el presente trabajo se propone un Framework que permite crear nuevos escenarios de forma dinámica, dentro de un espacio centralizado de aprendizaje en el cual intervienen instituciones educativas con sus respectivos docentes.

**Palabras Claves:** Juegos Serios; Framework; Espacio de aprendizaje.

### 1. Introducción

Actualmente se están aplicando nuevas estrategias de enseñanza a través del uso de juego serios, involucrando a las universidades, colegios terciarios, secundarios y primarios. “Los juegos potencian el aprendizaje de conductas y actitudes necesarias para el eficiente desempeño de una actividad particular y también tienen la funcionalidad del ocio cultural. En la actual sociedad ese papel lo desempeñan los videojuegos, éstos proveen a los jugadores de habilidades y destrezas propias de la época y facilitan el aprendizaje de procesos complejos con eficacia” [1]. Por otra parte, “numerosos estudios profundizan en los objetivos de la retórica clásica: docere, delectare et movere (enseñar, deleitar y emocionar) o, lo que es lo mismo, indican que un entorno que da cabida a la diversión del individuo favorece el aprendizaje. Desde este convencimiento, y con el objetivo específico de transmitir conocimientos entre los jugadores, surgen los conocidos videojuegos serios” [2]. A continuación, trataremos el concepto de juegos serios y sus principales factores en el ambiente de la enseñanza.

### 1.1. Juegos Serios

Los Juegos Serios son aquellos cuyo principal objetivo no se centra en la diversión, sino en el aprendizaje o adquisición de un conocimiento o habilidad, hoy en día son utilizados para la formación de conocimientos dentro del ámbito militar, político, empresarial, salud y educación. Este “concepto de juegos serios busca potenciar el aprendizaje, la estimulación del pensamiento crítico, el entrenamiento, la alfabetización digital, cambios de actitud y generación de emociones, lo cual va más allá del componente lúdico propio de los juegos” [3]. Cabe destacar que también “se potencia el aprendizaje activo y se capacita en competencias complementarias como la toma de decisiones, el trabajo en equipo, habilidades sociales, liderazgo y colaboración” [4]. Dentro del ámbito de la educación los docentes saben que “la inasistencia a clase está asociada al fracaso académico y al abandono o prolongación de los estudios. Varios autores han demostrado la eficacia de los juegos serios en la docencia universitaria: aumento de la motivación, mayor concentración y consecución de los objetivos de aprendizaje marcados” [5], de modo que “Los videojuegos constituyen una excelente herramienta de multiestimulación cognitivo afectiva que acelera el aprendizaje, genera placer, y potencia las habilidades digitales, el pensamiento estratégico y la creatividad, dependiendo en mayor o menor medida del tipo o género de videojuego que más se juegue.” [6], por esta razón, a la hora de desarrollar videojuegos se puede pensar en temáticas que permitan implementarse como complemento a los materiales educativos vigentes, en marcando el desarrollo dentro de la categoría de juego serio.

Se puede afirmar que a la hora de desarrollar un videojuego (sea serio o no) se debe pensar en una idea, a que público va orientada, que dispositivos serán alcanzados (Computadora, teléfono celular, tablets, etc.) y qué tan robusto y sofisticado será el desarrollo, lo que lleva a la necesidad de seleccionar un motor que proporcione características y componentes para lograr el objetivo. González y del Carmen [7] determinan parte de los componentes a tener en cuenta a la hora de seleccionar un motor: 1.- *Renderización*, 2.- *Animación*, 3.- *Física*, 4.- *Inteligencia Artificial*. 5.- *Scripting*. Todas estas consideraciones a tener en cuenta no deben dejar que se pierda de vista que “los juegos proporcionan un ambiente motivador y envolvente donde los jugadores aprenden haciendo a través de sus propios errores, gracias a desafíos adecuados a su nivel de competencia y a una realimentación constante. Es un entorno controlado capaz de reforzar las capacidades de decisión, de trabajo en equipo, las habilidades sociales, de liderazgo y de colaboración” [8], por lo cual, la inclusión dentro de los ámbitos educativos resulta fundamental para ayudar a mejorar la calidad educativa, la integración de los docentes y alumnos a la tecnología, y establecer nuevas herramientas que proporcionen un enfoque diferente a la presentación de contenidos.

### 1.2. Elementos de los Juegos Serios

Los juegos serios tienen que contemplar determinados aspectos para mantener el interés del jugador, por ello, las mecánicas dentro de un juego son un factor fundamental dado que determina cual es el comportamiento que el jugador va a tener. Este comportamiento producirá cambios en uno o más elementos dentro del videojuego. Otro aspecto a tener cuenta son las dinámicas del juego, las mismas se encuentran entre

las reglas del juego y la experiencia del usuario. Por consiguiente, algunos componentes estructurales que deben tomarse en cuenta a la hora de desarrollar están previstos en el trabajo [9] donde los autores, proponen formalizar un modelo de diseño de juegos serios basados en los componentes detallados en la Tabla 1.

**Tabla 1** – Componentes propuestos por [9] para el diseño de juegos serios.

Componente	Descripción
Interacción	Surge de la mecánica y dinámica del juego, da lugar a todas las experiencias del jugador que surgirán como resultado de la retroalimentación inmediata ofrecida por el juego.
Desafío	El jugador se enfrentará a problemas para los que se buscará soluciones. El desafío es resolverlos, los criterios del final de juego, tanto parciales como generales, serán especificados en los resultados del aprendizaje
Reglas	Determinará el orden, los derechos y las responsabilidades de los jugadores, así como los objetivos a cumplir por cada jugador con el fin de alcanzar el reto que enfrenta.
Objetivos	Deben estar claramente definidos y conocidos por el jugador, serán explícitos en las competencias ejecutadas

## 2. Trabajos Relacionados

### 2.1. Juegos serios educativos

Proyecto SEGAN [10] es un consorcio abierto de intercambio de ideas y conocimientos, tal como lo afirman Cruz-Lara, Manjón y Carvalho [11], su “objetivo es estudiar, analizar y difundir información sobre los juegos serios”, dentro de su sitio web agrupan a desarrolladores de videojuegos serios, profesores e investigadores, junto con una serie de artículos orientados a la temática, proyectos, y una cantidad numerosa de juegos serios educativos enfocados en diferentes asignaturas para los grados escolares en Europa. Dentro de los tantos juegos serios que promocionan se puede destacar 0 A.D. [12], es un juego de estrategia que está orientado al aprendizaje de historia, específicamente abocado al estudio de las civilizaciones antiguas, pero con el valor agregado que es de código abierto, permitiendo que pueda agregarse contenidos, escenarios y otros recursos sobre la base del juego. Aplicando el mismo criterio de código abierto también promocionan otro juego de estrategia llamado FreeCiv [13], cuya trama de juego consiste en el aprendizaje de la evolución de las civilizaciones a partir de la prehistoria.

Otro trabajo para destacar es la digitalización del juego Trivial Pursuit propuesta por Vergara Rodríguez y Mezquita [14], bajo el nombre de Trivia Virtual, el juego “está planteado para ser utilizado en formato multijugador (tanto individual como por

grupos). Los usuarios moviendo a lo largo del tablero siguiendo la dinámica del juego original, intentando responder correctamente a las preguntas que le correspondan.” Si bien el diseño inicial fue pensado para preguntas sobre economía, existen variantes del mismo que se aplican a diferentes asignaturas.

La saga Europa Universalis [15] es una serie de juegos de estrategia ambientadas en los períodos de 1419 hasta 1821, donde el jugador no maneja un personaje, sino que, a partir de decisiones militares, políticas, y diplomáticas, pueda expandir y colonizar a diferentes ciudades y territorios no explorados, a través de una interfaz de usuario muy compleja con una gran cantidad de opciones y comandos. “Lo interesante de este videogame es su veracidad y profundidad histórica, ningún detalle está librado al azar, además de que su precisión sobre el pasado histórico es de gran utilidad para fines didácticos —por elección de los propios diseñadores de la plataforma, los denominados «sucesos históricos» aparecen automáticamente si las condiciones de ese pasado recreado en la jugada se asemejan a la situación histórica real, lo cual propondrá la toma de decisiones por parte del jugador” [16].

## 2.2. Juegos comerciales aplicados en la educación

La saga de juegos “Age of Empires” se basan en transitar diferentes etapas de las civilizaciones, donde el jugador administra, crea y gestiona su metrópolis, ejércitos y aldeanos, cumpliendo determinados objetivos puede ir avanzando de era, utilizando ejércitos para atacar a otros pueblos. “Un aspecto muy destacable del juego es que cuenta con editor de escenarios y campañas. Nosotros mismos podemos crear escenarios reales eligiendo la ubicación de los elementos físicos que aparecen en el mapa. Además, podemos crear campañas enlazando varios escenarios y eligiendo diferentes objetivos que reflejen un suceso histórico. Esto le da la posibilidad al docente de diseñar actividades creando escenarios o campañas donde intercalar conceptos históricos” [17]. Todas las versiones cuentan con diversas modalidades de juego, entre ellas, eliminar a todas las civilizaciones enemigas (tanto CPU como otros jugadores en modo multiplayer), también realizar campañas históricas con sus protagonistas, por ejemplo, Juana de Arco. Además, da la posibilidad de realizar partidas multijugador de hasta 8 jugadores en simultáneo.

La originalidad del juego produjo que se llevaran a cabo una serie de experimentos para evaluar el uso del juego en una serie de cursos, arribando a la siguiente conclusión “Los videojuegos producen una gran motivación en los alumnos, pero no siempre sus efectos son positivos para el aprendizaje. La sobreexcitación en el aula y el apasionamiento en la construcción de explicaciones se convierten en obstáculos para realizar las reflexiones oportunas que lleven a aprovechar didácticamente los elementos históricos que posee el videojuego. Así, alumnos fascinados por el videojuego no son capaces de superar un nivel narrativo en sus apasionadas explicaciones del mismo, lo que conduce, por ejemplo, a olvidar que se está en una clase de Historia, estudiando una época histórica concreta. En resumen, ese apasionamiento puede dificultar los oportunos procesos inductivos y deductivos que se buscan con la utilización del videojuego en el aula.” [18].

La saga queda excluida de la categoría de juegos serios porque el conocimiento de los acontecimientos históricos no es un impedimento para avanzar en el juego y sus batallas no representan sucesos reales. Por último, existen algunos MODS (extensión de software que modifica al videojuego) que agregan algunos ejércitos y características, entre ellos, un mod “desarrollado por un grupo de diseñadores brasileros, que, siendo jugadores, han incorporado una serie de opciones para hacer coincidir las posibilidades del juego con la Guerra del Paraguay. De esta manera, se incorporan las naciones de Argentina, Paraguay y Brasil, y el escenario de Asunción, para las modalidades de un jugador o de multijugador online.” [16]

La saga de juegos “Assassin’s Creed” de la empresa Ubisoft, centra su historia en la recreación de las memorias genéticas de un personaje, que lo sitúan geográficamente en una gran cantidad de ciudades y en diversos períodos, (Cruzadas, Renacimiento, Revolución Americana, etc.), junto con los personajes históricos, desde Ricardo “Corazón de León”, Leonardo da Vinci, Cristóbal Colón, entre otros. Cada una de las entregas poseen un desarrollo muy rico en detalles arquitectónicos de cada lugar representado, de la vestimenta, estandartes, banderas, políticas y cultos religiosos, logrando una excelente recreación. Su repercusión “como videojuego comercial que es, AC posee limitaciones evidentes como la oferta de una serie de conocimientos desestructurados, y en ocasiones, poco profundos, o la carencia de exactitud histórica plena.” [19], cabe destacar que las misiones y objetivos no tienen relación con los sucesos históricos, tampoco permite crear escenarios.

La saga CAESAR que recrea la expansión del Imperio Romano mediante la construcción de ciudades y administración de recursos, fue utilizado en un trabajo de investigación con alumnos de 2º año para el aprendizaje de Geografía, obteniendo los siguientes resultados: “

- Alta motivación e interés en torno a una época histórica y a unos conceptos geográficos relativos a la misma (Imperio Romano).
- Mejora en la comprensión de los conceptos geográficos e históricos.
- Mejor comprensión de la realidad circundante como consecuencia de la aplicación de los conocimientos geográficos e históricos a la actualidad.
- Mayor comprensión de las interrelaciones que se producen, o se han producido, en cualquier situación geográfica o histórica.” [20]

La saga de juegos RPG “Mount & Blade” recrean escenarios de las batallas, factores económicos, y evolución de una población histórica. Permite personalizar el personaje en su apariencia, vestimenta, etc., también posee con MODs instalables para expandir los ejércitos, por ejemplo, el ejército de los Andes en la batalla por la independencia.

Todos los juegos mencionados para expandir la cantidad de pantallas, ejércitos u objetivos requieren de la instalación de algún MOD, o actualización del programa. Este trabajo propone el desarrollo de un Framework basado en parámetros para expandir estos aspectos utilizando una tecnología cliente – servidor sin la necesidad de la instalación de mods, actualizaciones, ni compilación del programa.

### 3. Batallas Históricas

“Dentro del encuadre de la taxonomía de ‘aprendizaje’ en ‘Juegos Serios para la Educación’, ‘1810’ es una herramienta desarrollada para chicos de tercer año del colegio secundario, para mejorar la asimilación de conocimientos de la asignatura ‘Historia’, cuyos contenidos abordan el contexto, próceres, resultados y consecuencias de las batallas de la ‘Guerra de la Independencia’” [21], siendo dichas batallas comprendidas entre 1810 a 1824 el corazón del Juego “Proyecto 1810” [22]. En dicho contexto histórico se libraron aproximadamente 187 combates y batallas; Si se debiera diseñar una escena por cada contienda con sus objetivos, recursos, y matices, demandaría muchísimo tiempo de diseño, desarrollo, pruebas e implementación, por consiguiente, un alto costo de recursos, por lo cual, para suplir esta problemática se propone un Framework de desarrollo de batallas basado en parámetros, que posea un repositorio centralizado en la nube para almacenar las batallas y sus características, mientras que en el aplicativo cliente queden embebidos varios terrenos con elementos de decorado (árboles, ríos, mares, montañas, pasto, etc.), donde al ejecutarse verifique y descargue los recursos y las configuraciones en un archivo XML que se utilizará para generar las batallas dinámicamente, sin la necesidad de compilar un nuevo cliente.

#### 3.1. Parámetros de Configuración

Los parámetros de configuración son el ladrillo fundamental de la esta propuesta, a partir de ellos, el cliente los leerá y generará dinámicamente, ejércitos, escenarios, y batallas, en la figura 1 se puede apreciar de la forma en que se configura un nuevo ejército.

**Figura 1** - Gestión de ejércitos

Para la configuración de las batallas se consideraron los parámetros descriptos en la Tabla 2.

**Tabla 2 – Parámetros de configuración de batalla**

Parámetro	Descripción
Fecha	indica el momento exacto cuando comenzó la batalla.
Lugar	Se menciona el punto geográfico donde ocurrió.

Ejércitos	Ejército al que pertenece el jugador y el enemigo
Cantidad soldados ejército	Establece la cantidad de soldados al iniciar la partida, los soldados son generados dinámicamente.
Regeneración de soldados	Establece en caso de la muerte de un soldado (IA), si se crea automáticamente otro en algún punto específico de la pantalla, manteniendo la cantidad de soldados previamente establecida.
Resultado	Indica el resultado real de la batalla
Descripción	Establece un detalle explicativo del contexto histórico, político, geográfico y social que generó el enfrentamiento y sus consecuencias.
Objetivos	Establece los objetivos para poder a cumplir para ganar la pantalla: <ul style="list-style-type: none"> <li>• Temporales: Se establece un tiempo en el que el usuario debe sobrevivir o bien, para cumplir otros objetivos.</li> <li>• Supremacía: se establece el número de enemigos que el usuario debe vencer dentro de una partida.</li> <li>• Recupero: establece los elementos que el jugador debe recuperar para ganar la partida.</li> </ul>
Recursos y Accesorios	Establece todos los recursos (campamentos, carretas, caballos, cañones, etc.) que forman parte del escenario y sus ubicaciones.
Terreno	Se selecciona un escenario embebido en el ejecutable para recrear la batalla.

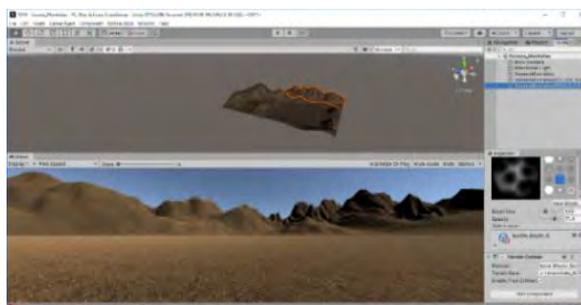
En la figura 2 se representa la pantalla de configuración de las batallas.

**Figura 2 - Gestión de batallas**

### 3.2. Diálogos y Escenarios

Dentro del Cliente ejecutable existen diversas topologías de terrenos (selva, montañas, ríos, etc.) que son utilizados como base para la generación nuevas batallas,

ver figura 3. Para reutilizar los terrenos se establecen puntos para la ubicación dinámica de los recursos (campamentos, cañones, casas, etc.).



**Figura 3** - Escenario de Montañas de Ejemplo

Los diálogos se pueden configurar en el repositorio, subiendo un archivo de sonido, seleccionando una imagen del personaje histórico al que se le adjudique la oratoria, acompañado del detalle escrito para ser posteriormente presentado en la escena. Al conectarse cada cliente, se realiza una consulta al servidor de forma automática y se descargan los recursos previamente configurados. Es importante tener en cuenta a la hora de configurar los diálogos, en qué momento debe ejecutarse, esto puede darse al inicio de una batalla, al subir de nivel el personaje, en algún segundo específico o aleatorio, al completar o presentar un nuevo un objetivo, como también al perder o ganar una partida.

#### 4. Espacio de Aprendizaje

El Framework posee cuatro tipos de usuarios con privilegios diferentes para llevar a cabo actividades de administración, gestión y participación.

**Usuario administrador:** posee todos los permisos, desde la creación de elementos del juego (batallas, ejércitos, objetivos, etc.) hasta gestionar los permisos especiales de otros usuarios, por ejemplo: asignar a un usuario los privilegios de institución educativa, para invitar a otros y así, proporcionar permisos de docentes, para su posterior seguimiento.

**Usuario docente:** puede crear cursos, e invitar a sus alumnos para participar del curso de comisión y adquirir desafíos propuestos, como así también, pueden crear sus propias batallas, fijar los objetivos y establecer premios.

**Proceso de Evaluaciones:** Dentro de cada temática el docente tiene la facultad de generar evaluaciones, que sirven como una herramienta extracurricular para realizar el seguimiento de una temática áulica. El proceso de evaluación consiste en la creación de temas, lo cual facilita el armado de preguntas que se almacenan a través del Framework y sus posibles respuestas. Cada pregunta tiene una puntuación establecida por el docente el que además establece el porcentaje de aprobación de la evaluación, configurando premios que recibirán los jugadores que consigan aprobar. Estos premios consisten en beneficios para mejorar la experiencia dentro juego, como ser, armas, ropas, municiones, puntos, distinciones, etc.

**Jugador:** Un jugador que pertenezca a un curso puede solicitar su baja en cualquier momento, realizar las evaluaciones desde la web como en el aplicativo, consultar el historial de evaluaciones previas y las vigentes. El resultado de los exámenes influye en el ranking del alumno, del curso y de la institución. Cada batalla explica de forma detallada oral y escrita todo su contexto, brindando a los jugadores un espacio de aprendizaje mientras juegan y se ordenan automáticamente en orden cronológico, y desbloqueándose las futuras batallas en la medida que se avanza, esto se muestra en la Figura 4, en la que se visualizan imágenes de las batallas del juego.

**Medición de Seguimiento/resultados:** Se podrá analizar mediciones sobre temas que transversalmente, hayan presentado mayor dificultad, y a su vez, evaluar el rendimiento individual de cada estudiante.

**Seguimiento de contenidos:** Básicamente, un objetivo primordial es evaluar el rendimiento de los cursos, ya que todos los contenidos creados por el docente, deben ser autorizados desde la plataforma por las autoridades del colegio, brindando la seguridad y la calidad requerida por la institución.



Figura 4 - Batallas de Ejemplo

## 5. Conclusiones

Ante el crecimiento exponencial de la industria de los videojuegos y su aplicación en el ambiente educativo, van apareciendo cada vez más disciplinas abordadas ante esta nueva tendencia, y en donde el aprendizaje en la temática de historia no queda exento. Cada vez son más los juegos que abordan estos conceptos, pero la necesidad de extender sus pantallas y recursos demanda la realización de la compilación de un nuevo ejecutable, o la instalación de un agregado que permita ampliar dicha funcionalidad.

En estos términos se llega a la conclusión de la importancia del desarrollo de un Framework basado en parámetros, el cual permita la generación dinámica de escenarios y objetivos, a partir de la descarga de un archivo con los parámetros, como ser para este trabajo de: batallas, ejércitos, objetivos y toda la distribución de recursos que se pueda generar automáticamente sin la necesidad de instalar un MOD o actualización, donde los docentes puedan generar y sus propios escenarios y así, gestionar sus instrumentos de aprendizaje y evaluación.

## Referencias

1. Lobo, J. C., & del Sur, M. (2014). Juegos serios: alternativa innovadora. *Conocimiento Libre y Educación (CLEDE)*, 2(2).
2. Calvo-Ferrer, J. R. (2018). Juegos, videojuegos y juegos serios: Análisis de los factores que favorecen la diversión del jugador. *Miguel Hernández Communication Journal*, (9), 191-226.
3. Delgado, J. C. S., & Sanz, C. V. (2020). Juegos serios para potenciar la adquisición de competencias digitales en la formación del profesorado/Serious Games to Enhance Digital Competencies Acquisition for Training Faculty. *Educación*, 44(1), NA-NA.
4. Petri, G., von Wangenheim, C. G., & Borgatto, A. F. (2017, May). Quality of games for teaching software engineering: an analysis of empirical evidences of digital and non-digital games. In 2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET) (pp. 150-159). IEEE.
5. García-Sánchez, F., García-Berna, J. A., Fernández-Alemán, J. L., & Toval, A. (2019). Aplicación de juegos serios en la enseñanza de Ingeniería de Requisitos y Gestión de Proyectos. *Actas de las Jornadas sobre Enseñanza Universitaria de la Informática*, 4, 7-14.
6. Lárez, B. E. M. (2006). Estimulación emocional de los videojuegos: efectos en el aprendizaje. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 7(2), 128-140.
7. González, T., & del Carmen, J. (2019). Análisis comparativo para la evaluación de motores de juegos en el desarrollo de juegos serios (Bachelor's thesis, Loja: Universidad Nacional de Loja).
8. Cruz-Lara, S., Manjón, B. F., & de Carvalho, C. V. (2013). Enfoques innovadores en juegos serios.
9. Bezanilla, M. J., Arranz, S., Rayón, A., Rubio, I., Menchaca, I., Guenaga, M., & Aguilar, E. (2014). Propuesta de evaluación de competencias genéricas mediante un juego serio. *New Approaches in Educational Research*, 3(1), 44-54.
10. Segán : Junio 2020 <http://seriousgamesnet.eu/>
11. Cruz-Lara, S., Manjón, B. F., & de Carvalho, C. V. (2013). Enfoques innovadores en juegos serios.
12. 0 A.D. Consultado: Julio 2020 [http:// play0ad.com/](http://play0ad.com/)
13. Free civ Consultado. Julio 2020 <http://www.freeciv.org/>
14. Rodríguez, D. V., & Mezquita, J. M. M. (2016). Diseño de juegos serios para reforzar conocimientos: una experiencia educativa en Secundaria. *Profesorado, Revista de Curriculum y Formación del Profesorado*, 20(2), 238-254.
15. Paradox Interactive (2020). Europa Universalis  
Sitio web: <https://www.paradoxplaza.com/europa-universalis-all/>
16. Irigaray, M. V., & del Rosario Luna, M. (2014). Cine y video en el aula: La enseñanza de la Historia a través de videojuegos de estrategia. *dos experiencias áulicas en la escuela secundaria. Clío & asociados*, (19), 411-437.
17. Ayén, F. (2010). Aprender Historia con el juego Age of Empires. *Clío: History and History Teaching*, 36.
18. Mugueta, I., Manzano, A., Alonso, P., & Labiano, L. (2015). Videojuegos para aprender Historia: una experiencia con Age of Empires. *Revista Didáctica, Innovación y Multimedia*, 32.
19. Alarcia, D. T., & Barco, D. I. (2014). Videojuegos y aprendizaje de la Historia: la saga Assassin's Creed. *Contextos educativos: Revista de educación*, (17), 145-155.
20. Muñoz, M. I. V., & Garrote, M. R. C. (2008). El videojuego como recurso didáctico en el aprendizaje de la geografía. Un estudio de caso. *Papeles de geografía*, (47-48), 249-261.
21. Parkinson, C. & Martínez, R. (2018). Propuesta para la participación e interacción con Juegos Serios: "Proyecto 1810". *CoNaIIISI 2018 – 6to Congreso Nacional de Ingeniería Informática – Sistemas de Información*. Universidad CAECE – Mar del Plata, Buenos Aires, Argentina.
22. Proyecto 1810 (2020). Sitio web: <https://1810.uai.edu.ar>

## Juego serio de realidad virtual para acercar figuras importantes de la historia de la Informática

Mariano Mazza<sup>1</sup>, Cecilia Sanz<sup>2,3</sup>, Verónica Artola<sup>2</sup>

<sup>1</sup> Facultad de Informática Universidad Nacional de La Plata,  
La Plata, Argentina

<sup>2</sup> Instituto de Investigación en Informática LIDI, CIC, Facultad de Informática, UNLP

<sup>3</sup> Investigador Asociado a la Comisión de Investigaciones Científicas de la Prov. De Buenos Aires  
{csanz, vartola}@lidi.info.unlp.edu.ar

**Abstract.** En este trabajo se presenta un juego serio de realidad virtual para dispositivos móviles, llamado Innovática, que tiene como objetivo educativo acercar algunas figuras centrales en la historia del desarrollo de la Informática. Innovática ha sido diseñado y desarrollado a partir de la consideración de un conjunto de heurísticas creadas para el desarrollo de juegos serios que sirvieron de guías desde el momento de definir el concepto de este juego. Se presentan aquí las consideraciones y guías de diseño abordadas en el proceso de diseño, así como la descripción de Innovática y su aplicación a docentes y estudiantes con los primeros resultados parciales obtenidos. Los participantes han mostrado gran interés en Innovática, destacando su preferencia por la historia de algunas figuras con relatos que desconocían. El juego cumplió su objetivo en este sentido, ya que logró acercar algunos relatos sobre los personajes y darlos a conocer.

**Keywords:** realidad virtual, juego serio, innovadores de la Informática.

### 1 Introducción

Los juegos serios son aquellos que tienen un objetivo caracterizante que va más allá del entretenimiento. En general, plantean un objetivo educativo o de entrenamiento [1], [2]. En la actualidad se encuentran numerosas investigaciones vinculadas al uso de este tipo de juegos en contextos educativos [3], [4]. En estos trabajos se analizan los efectos de los juegos serios sobre la motivación de los estudiantes, el aprendizaje, el rendimiento académico, y las actitudes, entre otros.

Al mismo tiempo, la realidad virtual constituye un paradigma de interacción que ofrece una inmersión del usuario en un entorno virtual. La inmersión puede ser parcial o completa, de acuerdo al tipo de realidad virtual que se esté implementado. Esto se vincula con los sentidos involucrados y estimulados durante el uso de sistema de realidad virtual, y los dispositivos físicos de entrada salida que se utilicen. “Dependiendo del grado de identificación e involucramiento del usuario en el sistema, puede pasarse de una interactividad o simulación débil –como, por ejemplo, la de la

imagen ficticia como construcción intencional de estructuras que tienen una experiencia real limitada, a la de los juegos, en la que el usuario mantiene la conciencia de su no-veracidad-, a una simulación fuerte o ficción, hecho que puede ser experimentado aisladamente o ser compartido por otros observadores que se encuentren en el mismo espacio ficcional” [5]. La investigación sobre las posibilidades de la realidad virtual para el ámbito educativo es un tema vigente [6], a lo que se suma su combinación con los juegos serios, dando lugar así a trabajos que ponen el foco en el diseño, desarrollo y evaluación de juegos serios educativos basados en realidad virtual [7], [8].

Este trabajo aborda esta temática en la que se estudian las posibilidades de los juegos serios basados en realidad virtual. Se explican los componentes considerados para el diseño del juego serio Innovática, los modelos de interacción seleccionados para la realidad virtual, y se analizan los resultados respecto de la usabilidad y la experiencia de un grupo de 14 usuarios (docentes y estudiantes) con este juego.

De aquí en más este trabajo se organiza de la siguiente manera: en la sección 2, se presentan algunas definiciones y componentes considerados para el diseño de juegos serios basados en realidad virtual, en la sección 3 se presentan algunos antecedentes de este tipo de juegos para contextos educativos, y en la sección 4 se describe el juego Innovática. Luego, la sección 5 presenta la evaluación llevada a cabo para estudiar la usabilidad y la experiencia de juego con Innovática, para finalmente presentar los resultados y las conclusiones del trabajo.

## 2 Realidad virtual y juegos serios

Las tecnologías de realidad virtual, están siendo utilizadas en escuelas y universidades alrededor del mundo [9]. Se insertan cada vez más en el sistema educativo debido a los beneficios que aportan, por ejemplo, permiten al estudiante no sólo conocer ciertas temáticas sino además vivenciarlas, a partir de experiencias que hasta podrían sólo existir en la imaginación del individuo. Es por esto, que se puede explotar este potencial y despertar interés en temas importantes, a partir de la riqueza de la experiencia virtual multisensorial, que además puede lograr que el aprendizaje resulte significativo. La realidad virtual es una tecnología especialmente adecuada para contextos educativos, debido a su facilidad para captar la atención de los estudiantes mediante su inmersión en mundos virtuales relacionados con las diferentes ramas del saber, lo que puede ayudar en el aprendizaje de los contenidos de cualquier materia. Al mismo tiempo, este tipo de herramientas puede emplearse también con el fin de promover el desarrollo de competencias con tecnologías digitales, lo que es requerido como una competencia necesaria en este siglo [10], [21]. Más allá de las posibilidades de la realidad virtual, cuando se diseña un sistema con este paradigma de interacción, requiere atender a ciertas variables importantes para alcanzar los objetivos esperados. Aspectos como: la inmersión, el sentido de presencia, la calidad de las imágenes de las escenas, la interactividad, los tipos de sentidos involucrados, y las formas de interacción planteadas influyen en la experiencia del usuario y en el logro de los objetivos que se proponen con el sistema [11], [12].

Por parte del diseño de juegos serios, se debe atender a diversos componentes también. Estos se ven limitados por reglas establecidas por la jugabilidad y ponen un límite a las acciones del jugador, que a su vez lo guían en una dirección, buscando llevarlo a cumplir un objetivo; en un contexto en el que el jugador mismo desee alcanzarlo [5]. Estos objetivos son llevados a cabo en un escenario dentro del cual se desarrollan las acciones. En estos escenarios el jugador puede encontrar distintos personajes que pueden cumplir la función de ser aliados y ayudarlos en su recorrido o cumplir el rol de enemigos y buscar así ser un desafío que el jugador debe superar. Una vez alcanzados los objetivos el jugador suele pasar a un siguiente nivel, normalmente de una mayor dificultad. Esto causa el progreso de una historia que guía la experiencia de juego, historia con la que las mecánicas se adaptan a un entorno de cierta verosimilitud [13]. Esta historia son los hechos que hilan y dan sentido a los eventos que se desarrollan, aquellos por los que pasa el jugador. A su vez la historia cuenta con una narrativa que determina cómo se relatan los sucesos descritos en la historia a lo largo del videojuego [13]. Así, la historia con su narrativa, y las mecánicas, son componentes que afectan a la jugabilidad.

Cuando se diseña un juego serio de realidad virtual, se deben definir inicialmente los objetivos del juego, que van más allá del entretenimiento. Tanto la historia como las mecánicas deben orientarse a alcanzar un equilibrio entre el entretenimiento y el logro educativo. Además, al tratarse de un juego con realidad virtual se deben considerar aspectos importantes a la hora del diseño, tales como la inmersión del estudiante, el sentido de presencia en el entorno simulado, la consideración de cuánto realismo se quiere lograr, qué emociones movilizar, qué interactividades serán parte de la experiencia, retroalimentaciones de interés para el aprendizaje, tiempos del juego, entre otros aspectos de interés. Es por esto que se está trabajando en el desarrollo de metodologías para este tipo de sistemas específicos que tienen la característica de ser juegos serios de realidad virtual [14]. En este contexto una de los conceptos clave en el diseño es la jugabilidad. Se trata de un término empleado en el diseño y análisis de juegos que describe la calidad del juego en términos de sus reglas de funcionamiento y de su diseño como juego. Se refiere a todas las experiencias de un jugador durante la interacción con sistemas de juegos. Puede ser definido también como aquello que hace que un juego resulte fácil y divertido de usar, poniendo énfasis en el estilo interactivo y en la calidad del *gameplay*<sup>1</sup> estando afectado éste por la usabilidad, la narrativa e historia, la intensidad interactiva, el grado de realismo, etc. [14], [15], [20]. Varios autores han presentado heurísticas relacionadas con el diseño de la jugabilidad y también para su evaluación. Algunos se han basado en el modelo de procesamiento descrito por [16], que establece una separación entre los mecanismos afectivos y cognitivos por medio de los cuales los seres humanos interactúan con sus medios, y que también es tomado como base para el diseño emocional. Este modelo considera tres niveles de procesamiento: visceral (bajo), conductual (intermedio) y reflexivo (alto). El visceral que forma parte de la naturaleza humana, del jugador en este caso, y se genera automáticamente a partir del estímulo de los sentidos; el conductual que se da como parte del comportamiento humano en

<sup>1</sup> Se toma el texto de [13] donde se lo define como la dinámica de las acciones que ocurren cuando un jugador juega al juego, y se manifiesta cuando el usuario interactúa con las mecánicas del videojuego.

general y que puede fomentarse a partir de las acciones y consecuencias de éstas durante el proceso de juego; mientras que el reflexivo constituye una capa superior, que influye en el nivel conductual a mediano y largo plazo. En el caso de los juegos, se puede trabajar a partir de los pensamientos y recuerdos que el jugador tiene, activa otras habilidades del pensamiento. De esta forma cuando se diseña el juego, de acuerdo a sus objetivos, se puede buscar estimular los tres niveles de procesamiento para alcanzarlos y equilibrar objetivos educativos y de entretenimiento. En este trabajo se han considerado estos aspectos y también las dimensiones y heurísticas presentadas en [13]. Las dimensiones son: jugabilidad intrínseca, mecánica, interactiva, artística, personal y social. En la sección 4 se explicarán la forma en que fueron utilizadas para el diseño de Innovática.

### **3 Antecedentes de juegos serios de realidad virtual**

En esta sección se presentan algunos antecedentes de juegos serios que se basan en realidad virtual.

Se estudió el caso de Labster que es un laboratorio de RV que busca incentivar a los estudiantes en el proceso educativo, mediante un acercamiento innovador a la enseñanza de las ciencias. Sus creadores se basaron en la idea de utilizar simuladores, y esto resultó en el desarrollo de laboratorios virtuales. En estos laboratorios, los estudiantes pueden realizar experimentos con resultados similares a lo que sucedería en un laboratorio real. En [17], se describe la utilización de Labster en dos casos de estudio en una universidad. Los resultados muestran que los estudiantes alcanzaron los objetivos educativos de esta práctica y mostraron alto nivel de satisfacción durante la experiencia. Es un caso de realidad virtual parcial ya que solo se estimulan vista y oído y se utiliza la pc para una navegación del ambiente simulado.

Nefertari es un juego desarrollado por la compañía Experius VR para la PC, pensado para ser utilizado con un casco de realidad virtual. El juego busca que todos los alumnos en un mismo salón de clases puedan visitar la tumba de la reina Nefertari, apreciar las ilustraciones, los jeroglíficos en las paredes e informarse y aprender respecto a este lugar y la mitología egipcia mediante narraciones, superficies interactivas y exploración. El jugador debe introducirse en la experiencia mediante el uso de cascos de realidad virtual de alta gama, tales como el HTC Vive, Valve Index o el Oculus [18]. Estos cascos hacen un seguimiento detallado de la posición del jugador, los movimientos de su cabeza, su rotación, e inclinación de su cuerpo. Se trata de un caso de realidad virtual completa en la que se puede interactuar con todo el cuerpo. Se experimenta una inmersión completa en la tumba de Nefertari. El lugar está recreado tal como puede verse en la realidad, incluso se mantienen las imperfecciones en la superficie causadas por el pasar del tiempo. El equipo explicó que toda la iluminación fue creada dentro del motor gráfico en tiempo real y fue optimizada para la realidad virtual. Fue utilizado en la Universidad de Melbourne como una experiencia compartida por múltiples usuarios simultáneos en el mismo entorno virtual, todos representados por sus avatares [17].

En [19] se presenta un juego serio de realidad virtual, en el que se cuenta una historia. El trabajo destaca la importancia de la estructura de la historia en el contexto

educativo, citando investigaciones que indican su impacto en la retención y recuperación de los contenidos. El estudio de los autores se enfoca en analizar cómo impacta la estructura de la historia y las interactividades diseñadas en el aprendizaje de los estudiantes. Se utiliza el juego “*The Chantry*”, disponible para la plataforma de VR de PlayStation. La aplicación cuenta la historia Dr. Edward Jenner y su invento de la vacuna contra el virus de la viruela (<https://jennermuseum.com/>). Trabajan con estudiantes de primaria, y se analizan dos modos de recorrido, uno guiado y estructurado y otro de libre exploración. En las conclusiones se discute que mientras en el entorno guiado los estudiantes demuestran mayor retención de los conocimientos trabajados a través de las preguntas que se le realizan, en el segundo modo, manifiestan más sentido de presencia, más involucramiento con la experiencia.

Estos antecedentes resultan de interés pues recuperan aspectos centrales del diseño de los juegos serios de RV y sus posibilidades educativas. Mientras que el primer caso presenta un diseño de juego de RV parcial basado en un entorno 3D en la web, posibilita el logro de prácticas riesgosas en laboratorios de ciencias. Además, cuenta con un alto nivel de interactividad a través de la experiencia, con feedback en las diferentes posibilidades de acción que tiene el estudiante. En el segundo ejemplo, se rescata la gran calidad del ambiente generado con la realidad virtual, lo que ha dado lugar a una inmersión casi como si se estuviera en el lugar físico de la tumba de Nefertari. En el tercer caso analizado, se atiende al componente de la historia, y su modo de interacción con ésta: libre o guiada, y cómo impactan en la actitud y el aprendizaje.

#### **4 Descripción de Innovática**

Innovática (Innovadores de la Informática) es un juego serio de realidad virtual basado en dispositivos móviles, creado en el marco de un trabajo de grado de la Facultad de Informática de la UNLP, y se enmarca en los temas de un proyecto de investigación e innovación del III LIDI, dentro de la misma facultad. Su objetivo educativo se centra en acercar a figuras importantes de la historia de la Informática a jóvenes. Específicamente se orienta a estudiantes de los últimos años de secundaria y primeros años de la universidad, de manera tal de vincularlos, más afectivamente, con estas figuras que han sido innovadores de la Informática. Al mismo tiempo, el objetivo del juego es reparar una anomalía de tiempo – espacio, que ha provocado afectar la historia de la Informática. El estudiante así se vuelve un protagonista de esta historia, y debe reparar hechos históricos de los personajes viajando a través de portales, y accediendo a los contextos específicos de estas figuras.

Como aspectos centrales del diseño de Innovática se consideró: la creación de historia, las mecánicas para alcanzar los objetivos del juego, el recorrido de una forma semiestructurada, con ciertos aspectos guiados y otros libre para lograr un mayor involucramiento y retención de información, la inclusión de feedback e información para el aprendizaje, y un premio al final de recorrido completo (como un aspecto de motivación para el cierre del juego). Los contextos de cada personaje se planificaron para que se acercaran a la situación real, aunque son simulados.

En cuanto a su implementación, fue desarrollada en Unity 2019.3.0f6. Se ha compilado para dispositivos Android y requiere contar con acelerómetro y giroscopio en el celular donde se ejecute, para que se detecten los movimientos de la cabeza del jugador. La aplicación estará disponible próximamente en Google Play.

A continuación se describen más detalladamente algunos aspectos contemplados para el diseño de este juego.

#### 4.1 Heurísticas y aspectos del diseño considerados

Para el diseño se utilizaron las heurísticas de jugabilidad presentadas en [13]. Se consideraron las dimensiones de jugabilidad: intrínseca, mecánica, interactiva, artística y personal aportadas por este autor. Para cada una se analizaron como atributos: la satisfacción, el aprendizaje, la efectividad, inmersión, motivación, emoción y el aspecto social o de vinculación con otros. Esto se llevó a cabo instrumentalizándolo con una tabla. La figura 1 muestra solo una de las dimensiones como ejemplo. Allí también se destacan las decisiones relacionadas con las emociones y estímulos relacionados a los tres niveles presentados por [16].

**Tabla 1.** Ejemplo de la tabla usada en este caso para la jugabilidad intrínseca y sus atributos.

Atributos	Aspectos de diseño según atributo
Satisfacción	Se involucra al estudiante como protagonista de la historia. Tiene un rol activo resolviendo la anomalía espacio tiempo y salvando la historia de la Informática. Se combinan hechos reales con la ficción.
Aprendizaje	La historia guía al estudiante para ir a salvar una parte de la historia de personajes destacados de la Informática, como Babbage, y Ada. Cuando viaja a través de portales experimenta anécdotas/situaciones que formaron parte de la vida de estas figuras, conoce sus contextos, lugares de trabajo o de sus casas. Al resolver la anomalía, regresa al punto actual y conoce información de algunas ideas innovadoras y aportes de estos personajes a la historia de la Informática.
Efectividad	El recorrido completo se cumple en un tiempo máximo de 15 minutos, para respetar aspectos de la atención del estudiante. La experiencia de aprendizaje y el entretenimiento se combinan a través de la resolución de desafíos simples, con relatos que acercan a los personajes de una manera más afectiva. La historia es parcialmente guiada, con posibilidad de elección de qué salas recorrer.
Inmersión	Se utilizan gafas de RV con dispositivos móviles que provocan que el estudiante esté inmerso visual y auditivamente en la historia. El recorrido se realiza a través de movimientos de la cabeza y no se requiere de otro dispositivo. La calidad del audio y de las imágenes aporta a la inmersión, así como también el involucramiento en primera persona del estudiante como protagonista de la historia.

Motivación	Para alcanzar la motivación del estudiante, se planificó un alto grado de interactividad, acciones como el cruce de los portales, los entornos de cada personaje, sus anécdotas, la exploración de los objetos de estos personajes, el premio obtenido al reparar algún invento del personaje en cuestión, y el cierre de la historia con un feedback afectivo. Además la propia historia y la atención a los 3 niveles de procesamiento del modelo de [13].
Emoción	Sorpresa al inicio de la historia cuando se presenta la anomalía. Alegría cuando se logra resolver cada desafío (nivel visceral). En la resolución de los desafíos, se espera a través de una exploración semiguída o semiestructurada, conducir al estudiante para conocer alguna vivencia o anécdota de uno de los personajes. Luego de reparar cada desafío se espera alcanzar un nivel más reflexivo, a partir de un relato específico sobre los aportes de ese personaje a la historia de la Informática.

## 4.2 La historia

La historia del juego Innovática inicia en un museo, en el que el estudiante participante se encuentra con un robot que, a través de sus movimientos y sus relatos en audio, lo orienta en el recorrido. El robot presenta inicialmente la posibilidad de ir a tres salas, cada una de las cuales introduce la historia de algunas figuras destacadas de la historia de la Informática. Se tiene libertad de elegir cualquier sala, aunque la numeración de la sala (1, 2 y 3), invita a hacer un recorrido cronológico, según el momento histórico de los personajes (ver Fig.1 izq.).

Al ingresar a una sala de pronto los cuadros se caen (ver Fig.1 derecha) y se busca sorprender al jugador frente a la situación. El robot anuncia que ha ocurrido una anomalía de espacio-tiempo, y que eso no debería estar pasando. Se invita al jugador a tomar un rol protagónico, ya que es éste quién deberá viajar a través de portales para reparar la historia de la Informática. Así al usar cada portal se accede a un escenario completamente diferente, que muestra el contexto de uno de los personajes (según la sala y el portal por el que se viaje). En la primera sala se presentan a: Charles Babbage y Ada Lovelace (ver Fig.2 izquierda y derecha, que muestran respectivamente, el lugar de trabajo de Babbage y una sala de la casa de Ada); en la segunda sala a: Alan Turing, John Von Neumann y a Grace Hopper; finalmente, en la tercera sala: se debe recuperar parte de la historia a través de figuras como Linus Tovalds, Vinton Cerf y Jack Kilby. Al finalizar el recorrido, si el estudiante logró reparar todas las historias de los personajes recibe un premio y una felicitación, apelando a un tipo de feedback afectivo.



**Fig. 1.** Sala central del juego (*izquierda*) – b. Momento en que se produce la anomalía en sala 1 (*derecha*)



**Fig. 2.** Lugar de Trabajo de Charles Babbage con objetos para explorar. Se ve el robot también que acompaña el recorrido (*izquierda*). Sala de la casa de Ada Lovelace (*derecha*)

### 4.3 Las mecánicas y su relación con los objetivos de aprendizaje

El juego propone la interacción con un robot guía, que se lo puede ver en la Fig.3 (*izquierda*). Los audios del robot orientan y presentan parte de la historia y los conocimientos que buscan ser acercados a los estudiantes. Los relatos en audio son cortos para alcanzar la efectividad y la atención. Al mismo tiempo, se ha elegido acercar a los personajes inicialmente de una manera más vivencial, con algún relato que vincule al estudiante con ese personaje de forma afectiva. Por ejemplo, en el caso de Grace Hopper se recupera el origen del término “bug” como referencia a un error en el sistema. El estudiante tiene que atrapar el insecto que da origen al desperfecto en la máquina que estaba utilizando Hopper y que dio lugar a decir es un “bug”. Así cuando se viaja por el portal, la experiencia es exploratoria y vivencial, con componentes afectivos para lograr la vinculación con el personaje y buscar que el estudiante retenga algún pasaje de su historia. La experiencia en cada viaje implica explorar objetos con la mira para resolver el desafío. Los objetos al ser enfocados con la mira tienen un feedback, un breve audio del robot que dicen qué es el objeto, cómo se vincula con el personaje en cuestión. Una vez resuelto el desafío, se regresa por el mismo portal a la sala del museo y allí aparece un invento/aporte del personaje a la historia de la Informática que funciona como premio por haber reparado la historia. Al interactuar con este objeto, se despliega un relato más formal sobre lo que ese innovador ha aportado a la Informática. Así tanto las mecánicas utilizadas para resolver los desafíos como su posterior reparación, constituyen oportunidades para el aprendizaje, a partir del uso de niveles diferentes de procesamiento.

## 5 Experiencia con Innovática para su evaluación y resultados

El juego Innovática estaba planificado para ser utilizado durante 2020 con estudiantes del primer año de las carreras en Informática de la UNLP. Debido a la situación de la pandemia se ha realizado una experiencia a distancia con algunos estudiantes y docentes de diferentes contextos, con el fin de evaluar su usabilidad y aspectos relacionados con la experiencia y el aprendizaje.

## 5.1 Metodología utilizada

Para llevar a cabo la experiencia se seleccionaron algunos estudiantes y docentes de las carreras de la Facultad de Informática, y del entorno cercano a los creadores del juego, y se los invitó a participar. Se buscó incluir docentes para tener su punto de vista respecto de la aplicación, ya que posteriormente pueden ser los motores para integrar a Innovática al contexto educativo. Para ello se envió un mail con la presentación del juego, en el que se indicó su objetivo, el contexto en el que se desarrolló, instrucciones para su utilización y se dio el enlace para descargar la .apk. Además, se les solicitó que en forma posterior a usar Innovática, se responda a una encuesta en Google Form, para ello se envió su enlace en el mail.

La encuesta utilizada se estructuró con 4 ítems para indagar datos del participante como edad, género, si es estudiante o docente, el contexto en el que estudia o trabaja; tres ítems para conocer la satisfacción con la experiencia; 2 ítems vinculados con conocer qué personaje le gustó más y el otro para recuperar los aprendizajes que quedaron de las historias; finalmente se integran los 10 ítems para la evaluación de la usabilidad del juego, tomados del instrumento SUS (*System Usability Scale*) [22]. El cuestionario SUS provee un método fiable y rápido para la medición de usabilidad. Consiste en un cuestionario de 10 puntos con escala de 5 posibilidades desde “totalmente en desacuerdo” a “totalmente de acuerdo” (siendo su valor para analizar los resultados de 1 a 5, respectivamente). El puntaje promedio del cuestionario es 68. Eso significa que un puntaje SUS superior a 68 está por encima del promedio y por debajo de 68 está por debajo del promedio [22].

Un total de 15 participantes respondieron a esta invitación, con 60% de género masculino y 40% femenino, con 13 participantes con edades comprendidas entre 21 y 35 años, y solo 2 participantes superaban la edad de 35 años. De estos 9 son estudiantes, y el resto docentes.

## 5.2 Resultados

En relación a la pregunta de qué fue lo que menos les gustó del juego orientada a analizar aspectos deficientes de la experiencia de usuario, algunos participantes hicieron referencia a la navegación cuando se trabajaba sin las gafas y se usaba solo el celular en las manos. Esto se debe a que como la experiencia se llevó a cabo a distancia, algunos no tenían gafas y usaron solo el celular.

Respecto de lo que más les gustó del juego, los participantes hicieron referencia a las anécdotas y contextos de los personajes con comentarios como: “*El hecho de conocer datos curiosos o de color acerca de los innovadores. No fue como leer un artículo de wikipedia sino que también aprendí datos curiosos y divertidos*”; “*Los ambientes están muy bien logrados, se ven llamativos e interesantes, se pueden notar las distintas personalidades de los personajes históricos y el tiempo en el que vivían, te invita a ver todo el lugar, el guía es de mucha ayuda y no desentona con la experiencia*”. También se hizo referencia a las mecánicas elegidas para entretener y aprender, como un aspecto positivo: “*Me gusta la idea en sí de distraerte con viajes en el tiempo y salvar al mundo para enseñarte de cualquier tema en particular*”.

Cuando se les consultó por el personaje de preferencia, las historias más elegidas fueron las de Ada Lovelace y Alan Turing. Siguiendo por Linus Torvalds y Grace

Hopper. En la pregunta sobre qué recuerdan de las historias de los personajes se manifestaron mayormente, cuestiones que les habían llamado la atención como la historia del término “*bug*” (4 personas refirieron a esto), las personalidades de Babbage y Von Neumann, la historia de la tormenta de Kilby, y la ambientación de la casa de Ada.

En la pregunta respecto a qué tan complejo pareció el juego, 13 de los 15 calificaron 1, totalmente en desacuerdo con que sea complejo. Se alcanzó este resultado al haberse apuntado a mecánicas de interacción y un planteo temático sencillos. La persona puede introducirse en el juego sin una gran curva de aprendizaje. Estas mecánicas de interacción sencillas prueban su efectividad, esto se observa a partir de los 9 usuarios que calificaron con 5 al preguntarles si las funciones del juego están bien integradas. Otros 5 usuarios calificaron con un 4 en esta categoría, una opinión igualmente muy favorable. Cuando se les preguntó si la mayoría de la gente aprendería a usar este juego en forma rápida, 5 usuarios indicaron estar totalmente de acuerdo, y otros 7 usuarios calificaron con un 4, es decir están de acuerdo. En cuanto a ser necesaria la ayuda de una persona con conocimientos técnicos para poder utilizar el juego, 7 se consideran absolutamente en desacuerdo y otros 5 calificaron con un 2, es decir en desacuerdo. El juego se distribuye en la forma de un APK para Android, es decir, un instalador descargable. Esta es una manera común y reconocida de instalar aplicaciones en la actualidad, y el usuario promedio no encuentra problemas mayores para hacerlo. Es por esto que al preguntarles si necesitaron aprender muchas cosas antes de ser capaz de usar este juego, 12 estuvieron absolutamente en desacuerdo.

Finalmente, cuando se procesaron estos ítems correspondientes al SUS, según la metodología descrita en la sección 5, se obtuvo un promedio de 83.975, el cual está muy por encima del promedio, y se considera un resultado muy positivo respecto de la usabilidad.

**Acknowledgments.** Se agradece al Proyecto 11/F023 del III LIDI y al RTI2018-096986-B-C31: "Pergamex: pervasive gaming experiences for all".

## 6 Conclusiones y trabajos futuros

Este trabajo realiza un aporte a la investigación sobre el diseño de juegos serios basados en realidad virtual y sus posibilidades educativas. Se presenta un recorrido por aspectos relacionados al diseño, donde se consideran una variedad de componentes vinculados a los juegos y a la realidad virtual. Los antecedentes han permitido obtener elementos claves que fueron considerados para el diseño y creación de Innovática. Este juego ha posibilitado acercar a docentes y estudiantes participantes de la experiencia con Innovática, a figuras destacadas de la historia de la Informática. Los aspectos de diseño tenidos en cuenta han sido percibidos positivamente por los participantes, con un muy buen valor arrojado por el SUS en cuanto a la usabilidad de la aplicación. Si bien se ha trabajado con una muestra reducida, se abre la puerta para profundizar este estudio con nuevos destinatarios. Al mismo tiempo, se indagarán en más profundidad cuestiones como las estrategias

utilizadas para el recorrido (se sigue el recorrido ordenado por sala 1, 2 y 3 o uno de preferencia), qué impacto tienen las historias en los estudiantes, entre otros.

## Referencias

1. Archuby F., Sanz C., Pesado P. Experience Analysis for the Use of Desafiante Serious Game for the Self-assessment of Students. *Comm. in Comp. and Inf. Science*, vol 1184. Springer, (2020)
2. Dörner, R., Göbel, S., Effelsberg, W., Wiemeyer, J. (eds.): *Serious Games*. Springer, Cham (2016). <https://doi.org/10.1007/978-3-319-40612-1>
3. Kiili, K., Ketamo, H: Evaluating cognitive and affective outcomes of a digital game-based math test. *IEEE Trans. Learn. Technol.* **11**(2), 255–263 (2018)
4. Chittaro, L., Buttussi, F.: Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE Trans. Visual Comput. Graphics* **21**(4), 529–538 (2015)
5. Giannetti, C. (2005). Estética de la simulación como endoestética. En: *Estética, ciencia y tecnología, creaciones electrónicas y numéricas*. I. Hernández García.
6. Kavanagh, S., Luxton-Reilly, A., Wuensche, B. A Systematic Review of Virtual Reality in Education, *Themes in Science & Technology Education*. 85-119. (2017).
7. Chirinos, Y., Sanz, C., Rucci, A., Comparato, G., Gonzalez, G. y Dapoto, S. HUVI: una aplicación de RVI para acercar el patrimonio argentino. *TEyET* (2020).
8. Schott, C., Marshall, S. Virtual reality and situated experiential education: A conceptualization and exploratory trial. *JCAL*. (2018)
9. Faaborg, A. Designing for virtual reality and the impact on education. (2015)
10. Merchant, Z., Goetz, E., Cifuentes, L., Keeney-Kennicutt, w., Davis, T. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis, *Computers & Education*, Volume 70, PP. 29-40, (2014)
11. Tsaramirsis, G., Buhari, S., AL-Shammari, K. O., Ghazi, S., Nazmudeen, M. S., & Tsaramirsis, K. Towards simulation of the classroom learning experience: Virtual Reality approach. In H. Mn (Ed.), *3rd Inter. Conference on Computing for Sustainable Global Development* (pp. 1343–1346). (2016).
12. Ritz, Leah T. Teaching with CAVE virtual reality systems: Instructional design strategies that promote adequate cognitive load for learners. (2015)
13. González Sánchez, J. y Gutierrez Vela, F. Jugabilidad: caracterización de la experiencia del jugador en videojuegos. Tesis doctoral de la Univ. Granada (2010).
14. Chirinos, Y., Sanz, C., y Almirón, M. Propuesta de Tesis Doctoral: Metodología para el diseño, desarrollo y evaluación de juegos serios con realidad virtual. Propuesta de tesis en desarrollo en la Facultad de Informática de la UNLP. (2020)
15. Usability First, Sitio web.
16. Normal, D. *El Diseno Emocional: Por Que Nos Gustan (O No) los Objetos Cotidianos*. Ed. Paidós. (2004)
17. Smith, C., Coleman, S. Using Labster to improve Bioscience student learning and engagement in practical classes, *Royal Society of Biology*. Univ. of Leicester. (2017)
18. Hilliker, H. A Guided Tour of Queen Nefertari's Tomb With a Docent in High Fidelity VR. Recuperado de [https://www.youtube.com/watch?v=yX\\_rDXOL2oA](https://www.youtube.com/watch?v=yX_rDXOL2oA)
19. Ferguson, C., van den Broek, E. y van Oostendorp, H. On the role of interaction mode and story structure in virtual reality serious games, *Computers & Education*, Volume 143, (2020), 103671, ISSN 0360-1315
20. Giannakos, M.. Enjoy and learn with educational games: Examining factors affecting learning performance (2013)
21. Kenwright, B.. *Virtual Reality: Ethical Challenges and Dangers* (2018)
22. Brooke, J. SUS: A Retrospective. *Journal of Usability Studies*. 8. 29-40. (2013).

## La Producción de Juegos Serios Móviles. Posibilidades y Desafíos para el Docente de Nivel Superior

Edith Lovos<sup>1</sup>, Iván Basciano<sup>1</sup>, Evangelina Gil<sup>1</sup>, y Cecilia Sanz<sup>2</sup>,

<sup>1</sup> Universidad Nacional de Río Negro, Sede Atlántica, CIEDIS, Viedma, Río Negro

<sup>2</sup> III-LIDI Facultad de Informática Universidad Nacional de La Plata

<sup>2</sup> Investigador Asociado de la Comisión de Investigaciones Científicas de la Pcia. de Bs. As.  
*elovos@unrn.edu.ar, vaanbas19@gmail.com, EvaEG19@gmail.com, csanz@lidi.unlp.edu.ar*

**Abstract.** Las posibilidades que los juegos digitales presentan en el marco de una propuesta didáctica, no resulta novedoso, sin embargo y en relación a las experiencias concretas de integración de los juegos digitales en los espacios de enseñanza y aprendizaje, se observa una limitación dada por las posibilidades de adaptación de los mismos a contextos específicos. Así, resulta importante avanzar en la búsqueda de herramientas preferentemente de acceso libre y gratuito que permitan a los docentes la creación de sus propios juegos o la edición de otros creados por terceros, a partir de perfiles con distinto grado de conocimientos sobre programación. En este artículo se presentan los resultados obtenidos a partir de la búsqueda, selección y análisis de herramientas entre ellos plataformas web, aplicativos de software, y/o frameworks, que posibilitan la creación de juegos serios en particular aquellos considerados móviles, y que incluyen interacciones usando realidad aumentada a través de códigos QR y/o geoposicionamiento, por parte de usuarios con perfiles técnicos de diferentes niveles.

**Keywords:** Juegos Serios, Herramientas de Autor, Frameworks, Realidad Aumentada, M-learning

### 1 Introducción

Michel & Chen [1] definen el término juego serio (JS) como una forma de combinar videojuegos y educación, donde el objetivo principal es la educación (en cualquiera de sus formas), y cuyas componentes principales son: objetivos, reglas, retos e interacción. Los JS habilitan otro mecanismo para llevar adelante la enseñanza y aprendizaje, a la vez que extiende los objetivos de entrenamiento y genera no solo condiciones para que el jugador (estudiante) aprenda sino que además pueda aplicar y demostrar lo aprendido [1]. Por otra parte, un juego serio móvil (JSM), es un JS que puede ser ejecutado en un dispositivo móvil como celulares o tablets o laptops,

dotando al jugador de la posibilidad de jugar en cualquier momento y lugar, es decir posibilitando la generación de entornos de aprendizajes menos rígidos, personalizados y ubicuos [2,3]. Sumado a ello, elementos propios de estos tipos de dispositivos como: cámaras, giroscopios, gps, etc., dan lugar al uso de tecnologías emergentes como es el caso de la realidad aumentada (RA), que permite enriquecer un contexto físico con información virtual. La RA se caracteriza por: (a) una combinación de objetos virtuales y reales en un escenario real, (b) usuarios interactuando en tiempo real y (c) una alineación entre los objetos reales y virtuales [4]. Existen diferentes niveles de RA, dependiendo del elemento que se use como disparador de ésta. Así el nivel 0 se usan códigos QR que solo posibilitan enlazar otros contenidos sin ningún tipo de interacción ni seguimiento, el nivel 1 usa marcadores (imágenes cuadrangulares en blanco y negro con dibujos esquemáticos) que permiten el reconocimiento de patrones en 2D y objetos en 3D, el nivel 2 incluye el reconocimiento de imágenes u objetos y/o de posicionamiento del usuario dando lugar a otro tipo de interacciones y a la superposición de elementos virtuales en el mundo físico captado, y en el último nivel se usan dispositivos especiales como lentes de RA u otros, que permite una integración más acabada entre el mundo físico y el virtual y así las experiencias que se pueden generar son más inmersivas y personalizadas.[5] En cuanto a su potencialidad en el plano educativo, la RA permite incorporar multimedia al proceso enseñanza aprendizaje, aumentando la riqueza del contexto físico y así, innovar en la práctica docente desde el diseño de actividades educativas que aporten a la comprensión de conceptos abstractos, de la espacialidad, a la motivación de los estudiantes y al aprendizaje basado en descubrimiento entre otros[3,5,6,7]. En este sentido, la combinación de RA y dispositivos móviles junto a las características propias de los juegos posibilitan un diseño instruccional que puede colaborar a alcanzar resultados de aprendizaje tanto afectivos como cognitivos [8,9].

### 1.1 Desarrollo de Juegos Serios

El desarrollo de un JS requiere combinar el diseño instruccional (DI) con el diseño del juego (características del juego, mecánicas, y jugabilidad), así avanzar en el desarrollo de un JS, exige contar con conocimientos sobre diseño de juegos, teorías de aprendizaje y dominio del contenido a abordar con el mismo. El DI permitirá definir los contenidos, las habilidades a desarrollar, las estrategias que se usarán para ofrecer los contenidos y los mecanismos de evaluación, todo ello en base a las necesidades de los estudiantes, sus características y el contexto de aplicación[10]. Sin embargo, aunque los docentes tienen experticia en el DI, en muchos casos no cuentan con los conocimientos técnicos, en particular de programación que les permitan avanzar en el desarrollo de un JS a la medida del contexto de su práctica docente. Sobre este punto, en la literatura [11,12,13,14,15] sobre el tema aparecen algunas posibilidades a través de las denominadas herramientas de autor (AT) y también frameworks que pueden asistir en el proceso de diseño y desarrollo de JS con poco o casi sin conocimientos de programación.

El concepto de herramienta de autor (AT) está asociado a la idea de software que posibilita la creación y edición de contenido de aprendizaje en formato multimedia, de manera que pueda ser utilizado en propuestas de enseñanza y aprendizaje en

diferentes soportes, dando lugar a procesos de e-learning y m-learning. En este sentido, una AT es un software que facilita la producción aplicaciones, que pueden ser utilizado por perfiles que no necesariamente van a contar con conocimientos técnicos. Para ello, el software estará dotado de un conjunto de funcionalidades y características (interfaces amigables, plantillas o modelos, sistemas de tutorías, interoperabilidad con otros recursos, entre otras) que permitan avanzar en esa línea [16,17].

Desde el punto de vista de la ingeniería del software [18] un framework representa una abstracción esquelética de soluciones a una serie de problemas que son similares. En un framework, se describen los pasos o etapas a seguir para implementar la solución sin entrar en detalles sobre las actividades que se realizarán en cada uno de los pasos. En el caso específico del desarrollo de videojuegos, los frameworks permiten con mínimos o en algunos casos sin conocimientos de programación, acelerar el proceso de desarrollo ofreciendo soluciones predefinidas con posibilidades de extensión y/o personalización. Existen frameworks como el caso de Construct2, que utilizan el sistema drag-and-drop (arrastrar y soltar) permitiendo crear la lógica del juego a través de scripts pre-armados que se asocian a los elementos del juego que se va construyendo de manera visual. Más allá de la facilidad de uso, una característica deseable de un framework para el desarrollo de JS es como señala Paiva[20], la posibilidad de cambiar el contenido (educativo) re-utilizando la estructura del juego, sin necesidad de comenzar desde cero.

En este trabajo, se presenta un listado de AT y frameworks recolectados a través de un proceso de revisión bibliográfica, así como también los criterios establecidos para la evaluación de los mismos. Luego se describen los resultados alcanzados con la evaluación de aquellos seleccionados y, por último, se presenta una discusión y las conclusiones alcanzadas junto con los trabajos a futuro. Es importante resaltar que este trabajo se llevó adelante en el marco de un proyecto de investigación, acreditado y financiado por la Universidad Nacional de Río Negro (PI-UNRN-40C-750), en el contexto de la unidad ejecutora CIEDIS. A través del mismo, se busca generar conocimiento sobre el diseño, desarrollo y aplicación de juegos educativos móviles con realidad aumentada, en espacios de enseñanza y aprendizaje de nivel medio y superior en el contexto de la UNRN.

## 2 Selección de Herramientas de Autor y Frameworks

En el proyecto de investigación mencionado en el apartado anterior, se llevó adelante una revisión de bibliografía con la intención de recuperar herramientas de autor y frameworks que puedan utilizarse para el desarrollo de JSM en particular aquellos que permitan incluir interacciones usando RA (dónde los elementos disparadores sean códigos QR y/o la ubicación geográfica del jugador). La revisión se limitó al periodo 2015-2019 y buscó responder a las siguientes preguntas:

- Qué Infraestructura tecnológica es necesaria para su uso?
- Requiere conocimientos de programación? De qué nivel?
- Qué tipos de juegos se pueden construir con estas herramientas?
- Qué nivel de RA se puede implementar?

A partir de este proceso de revisión se recolectaron las herramientas que se presentan en la tabla 1, y que han sido categorizadas de acuerdo a su tipo (AT o Framework), forma de acceso, plataforma a la que puede exportarse el juego desarrollado, tipo de elemento que se utiliza para activar la RA, tipos de actividades que pueden incluirse en el juego usando RA (TARA) e idioma.

**Tabla 1.** Herramientas de Autor y Frameworks recolectados

	Acceso	Plataforma	Activador de RA	TARA	Idioma
ARIS [18]	Libre(open source)	Móvil (iOS)	QR GPS	Recorridos interactivos Puzzles	Inglés
ARLEARN [19]	Libre(open source)	Móvil (Android)	GPS	Recorridos interactivo, Puzzles	Inglés
I_learnTest[20]	Libre con autenticación	web	No disponible	Quizzes Asociación	Inglés
FJSU[21]	Libre	Móvil (Android, iOS)	No disponible en la versión actual	Juegos adaptativos Disparos RPG Puzzles	Portugués
MAGIS[22]	Libre	Móvil (Android, iOS)	Marcadores Fiduciales, GPS	Aventura Recorridos interactivos Puzzles	Inglés
MOLE[11]	Prototipo	Móvil (Android, iOS)	QR	Recorridos Puzzles	Español
U-Adventure[23]	Libre	Móvil (Android, iOS), Consola, Principales SO	GPS QR	Aventura Recorridos interactivos Puzzles	Español
VEDILS[13]	Libre	Móvil (Android)	Imágenes Marcadores Fiduciales Texto	Recorridos Puzzles	Español

Se realizó un filtrado de la selección de tabla 1, teniendo en cuenta los siguientes criterios: acceso libre y gratuito, posibilidad de exportar el juego a dispositivos móviles con sistema operativo Android, posibilidad de incluir actividades con realidad aumentada usando códigos QR y/o geoposicionamiento. Y por último, y atendiendo al contexto de aplicación (UNRN), la posibilidad de usar la herramienta en idioma español y/o en forma visual. Esto redujo las posibilidades a las AT: MOLE [11], U-Adventure [22] y VEDILS [13] sin embargo la primera fue descartada porque al momento de realizar la evaluación, la herramienta aún estaba en versión de prototipo. En el caso de los frameworks, iLearnTest[20] y FSJU[21] se han descartado por no presentar soporte para el desarrollo de juegos que incluyan interacciones con RA.

## 2.1 Criterios de Evaluación

Para la evaluación de las AT, se han seguido los criterios establecidos en otras investigaciones [15,24,25] así como también se han incorporado otros establecidos

por los autores de este trabajo. Los aportes de [15], permiten evaluar las características de las herramientas de autor para el desarrollo de materiales educativos que incluyan realidad aumentada. En el caso [24], se propone llevar adelante la evaluación de herramientas de autor, desde la perspectiva denominada Factor Crítico de Éxito (CSF; por sus siglas en inglés), analizando factores como: infraestructura (requerimientos tecnológicos), usuario (niveles, habilidades y conocimiento) y comunidad (de práctica), y otros aspectos como estrategias y estilos de aprendizajes, métodos de enseñanza, presencia social (interacción, engagement y realismo), el grado de compromiso que genera el producto desarrollado y la diversión proporcionada tanto desde el punto de vista de quien está llevando adelante el diseño como del usuario final, en este contexto docentes y estudiantes.

A partir de esto, se presentan en la tabla 2 los criterios de evaluación con los que se analizarán las herramientas de autor y frameworks seleccionadas para este estudio.

**Tabla 2.** Criterios de evaluación

Herramienta de Edición	Nombre	Detalle
Criterios de Análisis (CAX)		
CAX0	Licencia	Costo de acceso y uso de la herramienta de autor
CAX1	Requerimientos técnicos	Requerimientos técnicos (hardware, software, conectividad) que requiere la herramienta
CAX2	Edición compartida	Es posible realizar la edición del juego en forma compartida/colaborativa
CAX3	Perfil de Usuario	Competencias digitales necesarias para utilizar la herramienta
CAX4	Plantillas/Moldes	Plantillas o modelos provistos por la herramienta para la generación de juegos a partir de los mismos.
CAX5	Tipos de Juegos	Tipo de juegos soportadas
CAX6	Nivel de RA soportada	QR, marcadores fiduciales, geoposicionamiento, imágenes,
CAX7	Información aumentada soportada	Tipos de información aumentada soportada (imágenes, videos, audio, objetos 3D, etc.)
CAX8	Accesibilidad	Posibilita el desarrollo de juegos que puedan ser considerados accesibles
CAX9	Distribución del contenido	Formas provistas por la herramienta para la distribución del contenido desarrollado
CAX10	Analíticas	Permite recolectar y registrar información sobre las interacciones de los jugadores

## 2.2 Análisis y discusión de las AT y Frameworks seleccionados

A continuación se describen y presentan los resultados de las herramientas seleccionadas para el análisis.

**U-Adventure.** Es un editor de JS que se distribuye y funciona como una extensión del motor de juegos Unity, y según sus creadores, el Grupo de investigación e-UCM de la Universidad Complutense de Madrid, permite la producción de juegos a perfiles de usuarios que no necesariamente dispongan de conocimientos de programación. La herramienta se constituye sobre la base de una versión previa llamada e-Adventure[25]. Busca solucionar problemas de obsolescencia de ésta, así como también abrir la posibilidad a la generación de juegos multiplataformas[14]. Entre las potencialidades de U-Adventure, destacan la potencialidad de incluir en el juego

interacciones usando RA tanto outdoor, a través de geoposicionamiento como indoor usando como disparadores códigos QR (esta última opción formaba parte de la versión e-adventure). Asimismo, es posible definir sobre qué eventos del juego recolecta información durante su ejecución, para luego realizar un análisis de ésta como parte del proceso de evaluación de los aprendizajes. En relación a la generación de juegos para dispositivos móviles, aunque no se requiere conocimientos de programación, el proceso de exportación de los proyectos al sistema operativo Android no es transparente para el usuario, ya que requiere de la instalación y configuración de Android Studio. En este sentido, se observa que aunque la herramienta presenta las características deseables para una AT, es necesario contar con distintos perfiles de usuario durante el proceso de producción del JSM. En la tabla 3 se presentan los resultados del análisis de la herramienta en base a los criterios establecidos.

**Tabla 3.** U-Adventure según criterios de evaluación definidos en 2.1

Herramienta de Edición	U-Adventure
Criterios de Análisis (CAX)	Detalle
CAX0	Descarga instalación y uso gratuito desde <a href="https://github.com/e-ucm/uAdventure">https://github.com/e-ucm/uAdventure</a> . No requiere registro de usuario
CAX1	Requiere tener instalado el motor de desarrollo de juegos Unity versión 2017.3 para la versión 1.0
CAX2	No soporta la edición compartida
CAX3	Aunque no requiere conocimientos de programación para usar la herramienta, es necesario tener conocimientos técnicos sobre cómo vincular la U-Adventure con el motor Unity y también es necesario saber instalar y configurar Android Studio para realizar la exportación de los proyectos en formato Android.
CAX4	Dispone a través del sitio <a href="https://github.com/e-ucm/uAdventure">https://github.com/e-ucm/uAdventure</a> de un manual de usuario y de una guía de instalación, que actualmente se encuentra en idioma inglés. No cuenta con tutoriales
CAX5	No cuenta con plantillas
CAX6	Aventuras clásicas
CAX7	QR, Geoposicionamiento
CAX8	Imágenes (jpg y png), videos( mp4), audio (mp3), enlaces a páginas web
CAX9	Es posible incluir texto y audio para una mayor accesibilidad en el desarrollo de los proyectos.
CAX10	Integrada a la herramienta a través de la Experience API (xAPI), posibilitando su uso tanto en línea como fuera de ella [27]. Cuenta con un editor específico para seleccionar los eventos a registrar.

**VEDILS.** Fue desarrollada en la Universidad de Cádiz (España), y sus autores [13] la definen como un entorno visual que permite diseñar escenarios de aprendizaje interactivos, en los cuales es posible incluir tecnologías como RA entre otras. VEDILS toma como base el entorno de desarrollo App2Inventor del MIT y posibilita la producción de contenidos aumentados para dispositivos móviles (celulares y tablets), a través de una plataforma en la nube, previo registro y autorización de uso. La herramienta, que actualmente va por la versión 1.6, se compone de dos partes: una donde se trabaja sobre el diseño de la aplicación a desarrollar, y otra denominada blocks donde es posible definir la lógica de la aplicación usando un lenguaje visual. En este sentido, el usuario de la herramienta, debe contar como se indica en otras investigaciones [11] con conocimientos de programación usando bloques. VEDILS

cuenta con una componente llamada ActivityTracker que posibilita recolectar información de las interacciones que se dan como parte de la ejecución del juego, aunque hay que tener en cuenta que las mismas se registran y procesan usando el servicio Google Fusion Tables y MongoDB. Aunque su uso podría no resultar sencillo para un usuario novel, VEDILS ofrece vídeos y tutoriales que pueden acompañar el proceso de configuración de los componentes, así como también cómo llevar adelante el proceso de análisis de la información. En relación a la exportación del proyecto, la misma es transparente para el usuario y puede elegir entre exportar el proyecto como APK (descargarlo en la computadora) y luego distribuirlo de la forma que resulte conveniente, o generar un código QR para acceder al mismo, en este caso el código tendrá una duración de 2 horas. Por otra parte, es recomendable descargar los proyectos generados ya que estos pueden ser removidos de la plataforma por razones de mantenimiento.

**Tabla 4.** VEDILS según criterios de evaluación definidos en 2.1

Herramienta de Edición	VEDILS
Criterios de Análisis (CAX)	Detalle
CAX0	Requiere de registro de usuario y de autorización para el uso de la plataforma
CAX1	Es necesario contar con conectividad a internet y un usuario registrado y autorizado
CAX2	No disponible
CAX3	Aunque no requiere conocimientos de programación para el diseño de la interface, si es necesario tener conocimientos de programación con bloques para poder explicitar la lógica de las interacciones con los elementos del proyecto[11].
CAX4	Cuenta con varios tutoriales y explicaciones paso a paso de aplicaciones desarrolladas con la herramienta
CAX5	Es posible importar otros desarrollos (.aia) para su uso en un nuevo proyecto
CAX6	Rompecabezas
CAX7	QR, Geoposicionamiento, reconocimiento de texto e imágenes
CAX8	Imágenes (png y jpg), videos( mp4), audio (mp3), modelos 3D ( obj, 3ds, mds)
CAX9	Es posible incluir texto y audio para una mayor accesibilidad en el desarrollo de los proyectos.
CAX10	Permite recolectar y registrar información de las interacciones de los jugadores usando un servicio no propio (Google Fusion Tables). Su uso podría no resultar sencillo para un usuario novel

**MAGIS.** Es un framework diseñado y creado por un grupo de investigadores de la Universidad Ateneo de Manila en Filipinas, como extensión del motor Unity. El mismo busca simplificar el diseño de juegos basados en narrativa que usan la ubicación del jugador, como por ejemplo recorridos históricos o de museos.[22] El desarrollo de un juego usando MAGIS implica la creación de escenas dónde para cada uno de los objetos que se incluyen en la misma, se genera un guion en forma de script en base a comandos propios del framework (y no en el lenguaje C#) y con el formato TVS(*Tab Separated Values*) y la extensión .txt. Sobre este punto, como señalan los propios autores crear en forma manual este tipo de archivos, puede resultar complejo y propenso a errores, y para evitar esto han desarrollado una AT de acceso libre, que intentan automatizar la tarea, sin embargo presenta una interfaz que no es visualmente amigable, lo cual puede resultar desalentador. Asimismo, la generación de los marcadores fiduciales que se pueden utilizar para activar la RA deben ser generados con otros aplicativos como por ejemplo Vuforia. En cuanto a la posibilidad de

registrar y procesar la información que se produce como parte de las interacciones del jugador, sus autores indican que cuenta con un subsistema de análisis que puede activarse de dos formas: a través de los eventos que tienen lugar durante el juego (ejemplo el escaneo de un marcador) o a través de los scripts que definen la lógica del juego usando los comandos específicos @analytics. El juego almacena los datos recolectados en cache para luego enviarlos al servidor de análisis, dónde se registran para su posterior visualización y análisis. Sobre este punto la guía de usuario de MAGIS no provee información específica.

**Tabla 5.** MAGIS según criterios de evaluación definidos en 2.1

Herramienta de Edición	MAGIS
Criterios de Análisis (CAX)	Detalle
CAX0	Libre bajo licencia GNU
CAX1	Unity 5 (2016)
CAX2	No disponible
CAX3	Requiere contar con conocimientos de programación
CAX4	Cuenta con una guía de usuario del framework y un manual para generación de scripts usando la AT de scripting
CAX5	Aventuras basadas en posicionamiento
CAX6	Marcadores, Geoposicionamiento
CAX7	Imágenes (png), audio (ogg), modelos 3D (FBX)
CAX8	En algunos casos se podría utilizar el audio como un recurso para brindar accesibilidad
CAX9	Usando las funcionalidades provistas por Unity versión personal se puede distribuir como APK
CAX10	Cuenta con un subsistema de análisis, pero no se especifica su uso en el manual de usuario

### 3 Conclusiones y trabajos futuros

En el marco de una propuesta didáctica las posibilidades que ofrecen los JSM son variadas y permiten atender diferentes aspectos: cognitivos, emocionales y motivacionales entre otros, y en el caso particular de aquellos juegos que incluyen interacciones usando RA, estas posibilidades permiten aprovechar el contexto físico en el que se llevan adelante, dando lugar a un aprendizaje contextualizado y por descubrimiento. Sin embargo, no siempre los juegos existentes se adaptan a las necesidades del contexto de aplicación, y así surge la necesidad de contar con herramientas que posibiliten a los docentes con diferentes niveles de conocimientos técnicos avanzar en la producción y posterior re-uso de JSM. En el caso de las herramientas que se analizaron en el apartado anterior, VEDILS se presenta como una alternativa cuando se desea realizar la edición desde la nube y a partir de otros juegos, y dónde la distribución del juego producido se lleve adelante de forma transparente para el usuario, sin embargo será necesario contar con conocimientos de programación usando bloques. En el caso de U-Adventure y MAGIS se presentan como alternativas para perfiles de usuarios con conocimientos técnicos (de configuración en el caso del primero y de programación en el caso del segundo) que

les permitan avanzar con el uso de un motor de generación de juegos como es el caso de Unity. En el caso de MAGIS, puede generar confusión para un usuario novel en Unity el lenguaje usado para armar los script de las escenas del juego. En relación a la posibilidad de contar con funcionalidades específicas para llevar adelante analíticas de aprendizaje sobre las interacciones que se dan durante el juego, cualquiera de los casos analizados exige un esfuerzo extra por parte del usuario, aunque en el caso de U-Adventure contar con el editor específico integrado a la herramienta, puede resultar de soporte para el proceso. En conclusión para que los docentes puedan avanzar en la producción de JSM que incluyan interacciones usando RA, será necesario no solo contar con conocimientos sobre diseño instruccional sino también con conocimientos técnicos que nos permitan maximizar el uso de las herramientas (sean AT o frameworks) y ponerlas al servicio de la propuesta pedagógica. Si bien las herramientas analizadas, mejoran las posibilidades de intervención en la adaptación y de desarrollo de JSM por parte de usuarios no técnicos, se vuelve necesario trabajar en equipo. Se propone como trabajo futuro, abordar el diseño y la producción de un JSM con una de estas herramientas, analizando los roles e intervenciones de un equipo en el que los docentes tengan una participación activa en este proceso.

## Referencias

1. Michael, D. R., & Chen, S. L. (2005). *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade.
2. Demir, K., & Akpinar, E. (2018). The Effect of Mobile Learning Applications on Students' Academic Achievement and Attitudes toward Mobile Learning. *Malaysian Online Journal of Educational Technology*, 6(2), 48-59.
3. De la Torre Cantero, J., Martín-Dorta, N., Pérez, J. L. S., Carrera, C. C., & González, M. C. (2015). Entorno de aprendizaje ubicuo con realidad aumentada y tabletas para estimular la comprensión del espacio tridimensional. *Revista de Educación a Distancia*, (37).
4. Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). *Recent advances in augmented reality*. Naval Research Lab Washington Dc.
5. Melo, I. M. (2018). Realidad aumentada y aplicaciones. *TIA*, 6(1), pp. 28-35.
6. Fonseca Escudero, D., Redondo Domínguez, E., & Valls, F. (2016). Motivación y mejora académica utilizando realidad aumentada para el estudio de modelos tridimensionales arquitectónicos. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 17(1).
7. Salazar Mesía, N., Gorga, G., Sanz, C. V. (2015). EPRA: Herramienta para la enseñanza de conceptos básicos de programación utilizando realidad aumentada. In *X Congreso sobre Tecnología en Educación & Educación en Tecnología (TE & ET)(Corrientes, 2015)*.
8. Schmitz, B., Klemke, R., & Specht, M. (2013). Effects of mobile gaming patterns on learning outcomes: a literature review.
9. Chang, C. Y., & Hwang, G. J. (2019). Trends in digital game-based learning in the mobile era: a systematic review of journal publications from 2007 to 2016. *International Journal of Mobile Learning and Organisation*, 13(1), 68-90.
10. Fernández-Robles, J. L., & Hernández-Gallardo, S. C. Diseño instruccional de un juego serio que facilite a niños de tercer grado de primaria el ejercicio de operaciones matemáticas básicas

11. Dal Bianco, P. A., Mozzon Corporaal, F., Llitas, 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)
12. Firsova, Olga; Vogel, Cathrin; Brouns, Francis; Diegel, Noëlle; Forsman, Panu; Stracke, Christian M. (2019). Designing for Virtual Mobility : Potentials and Caveats. In Fessl, Angela; Zdolšek Draksler, Tanja (Eds.) EC-TEL Practitioner Proceedings 2019 : 14th European Conference on Technology Enhanced Learning, Delft, Netherlands, September 16-19, 2019., CEUR Workshop Proceedings, 2437. Springer International Publishing. <http://ceur-ws.org/Vol-2437/paper4.pdf>
13. Mota, J.M., and Ruiz-Rube, I. (2017). VEDILS: a toolkit for developing Android mobile apps supporting mobile analytics. Seventh European Business Intelligence & Big Data Summer School (eBISS 2017)
14. Perez-Colado, I. J., Perez-Colado, V. M., Martínez-Ortiz, I., Freire-Moran, M., & Fernández-Manjón, B. (2017, April). UAdventure: The eAdventure reboot: Combining the experience of commercial gaming tools and tailored educational tools. In 2017 IEEE Global Engineering Education Conference (EDUCON) (pp. 1755-1762). IEEE.
15. Moralejo, L. (2014). Análisis comparativo de herramientas de autor para la creación de actividades de realidad aumentada (Doctoral dissertation, Facultad de Informática).
16. Cubillo, J. (2014). ARLE: una herramienta de autor para entornos de aprendizaje de realidad aumentada (Doctoral dissertation, UNED. Universidad Nacional de Educación a Distancia (España))
17. Mnkandla, E. (2009). About software engineering frameworks and methodologies. In AFRICON 2009 (pp. 1-5). IEEE.
18. <https://arisgames.org/editor/>
19. <http://portal.ou.nl/web/topic-mobile-learning/home/-/wiki/Main/ARLearn>
20. Paiva, A. C., Flores, N. H., Barbosa, A. G., & Ribeiro, T. P. (2016). iLearnTest–framework for educational games. *Procedia-Social and Behavioral Sciences*, 228, 443-448.
21. Silveira Júnior, G. D. (2019). FJSU: Um Framework para o desenvolvimento de jogos sérios ubíquos.
22. Vidal Jr, E. C. E., Ty, J. F., Caluya, N. R., & Rodrigo, M. M. T. (2019). MAGIS: mobile augmented-reality games for instructional support. *Interactive Learning Environments*, 27(7), 895-907.
23. Pérez-Colado, V. M., Pérez-Colado, I. J., Freire-Morán, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). uAdventure: Simplifying Narrative Serious Games Development. In 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT) (Vol. 2161, pp. 119-123). IEEE.
24. Teoh, K.K. & Sliuzas, R. (2016). Multimedia development of online literacy objects: An evaluation of authoring tools for novice authors. In H. Hamerton and C. Fraser (Eds.), *Te tipuranga – Growing capability: Proceedings of the 2015 National Tertiary Learning and Teaching Conference* (pp. 70-76). Tauranga, New Zealand: Bay of Plenty Polytechnic.
25. Christopoulou, E., & Xinogalos, S. (2017). Overview and comparative analysis of game engines for desktop and mobile devices.
26. <http://e-adventure.e-ucm.es/>
27. Ivan Perez-Colado, Víctor Manuel Pérez-Colado, Manuel Freire, Iván Martínez-Ortiz, Baltasar Fernández-Manjón (2017): Integrating learning analytics into a game authoring tool. ICWL 2017: 16th International Conference on Web-based Learning. Cape Town, South Africa, 20 September -22 September 2017

## Prácticas educativas STEAM para fomentar el desarrollo de competencias científico-tecnológicas

Silvina Manganeli<sup>1,2</sup>

<sup>1</sup> ITU, Instituto Tecnológico universitario, Universidad Nacional de Cuyo

<sup>2</sup> FCEN, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo

[smanganeli@fcen.uncu.edu.ar](mailto:smanganeli@fcen.uncu.edu.ar)

**Resumen.** Las profesiones, negocios y oportunidades laborales más valiosas en las próximas décadas están relacionadas con ciencia, tecnología, ingeniería y matemática (STEAM), por lo cual gran parte de los estudiantes se desempeñarán en profesiones y labores STEAM que hoy en día no existen. ¿Cómo se puede educar a los jóvenes en temas que hoy se desconocen y prepararlos para tomar puestos y carreras que aún no se han inventado? Esta es la razón por la cual debemos formarlos en las competencias necesarias para afrontar los desafíos del siglo XXI. El presente trabajo plantea la necesidad de ampliar las prácticas educativas, las cuales deberían incluir el desarrollo de proyectos tecnológicos interdisciplinarios basados en tecnologías emergentes, que fomenten, el desarrollo de competencias científico-tecnológicas en los alumnos pertenecientes a carreras del área de las ciencias exactas e ingeniería, con el objeto de renovarlas y optimizarlas, acorde a los desafíos del siglo XXI.

**Palabras Claves:** stem, steam, maker, didáctica, competencias, tecnologías emergentes, competencias del siglo XXI, pensamiento computacional

### 1 Introducción

Las competencias requeridas por los ciudadanos para asumir un papel activo en la sociedad actual, son consideradas como competencias para el siglo XXI [1]. Con el inicio de este siglo, los modelos educativos y los marcos de competencia se han ajustado y reformado para ayudar a insertar a los estudiantes en el ámbito profesional y apoyarlos durante toda la vida, de acuerdo con las demandas contemporáneas. Desde principios del presente siglo, diversos estudios están reflejando un descenso notable en la proporción de estudiantes que optan por las disciplinas científico-tecnológicas. En contraste con ello, se observa que está aumentando la demanda de profesionales cualificados en estos campos. Si ambas tendencias se mantienen, en un futuro cercano, el sistema educativo no va a formar todas las personas que se precisan para cubrir estas necesidades del mercado [2].

Por estos motivos, el desarrollo de programas o iniciativas que fomenten y desarrollen las vocaciones científico-tecnológicas y, en general, las habilidades y competencias relacionadas con la innovación, se han convertido en uno de los objetivos fundamentales de la planificación educativa. En este contexto es donde se ubica la educación STEAM [3].

La educación STEAM se podría enmarcar dentro del aprendizaje basado en problemas o proyectos (ABP), con la particularidad de que la solución al problema

planteado suele ser un objeto tecnológico (un dispositivo, un programa, etc.) [3]. Entre las metodologías activas, las más adecuadas para el desarrollo de las competencias STEAM son: el trabajo por proyecto, el movimiento maker y el pensamiento computacional como estrategias de enseñanza aprendizaje.

En esta comunicación se proponen y describen, dos practicas educativas que integran TICs emergentes (plataforma de hardware de código abierto Arduino y el pensamiento computacional). Se trata de dos proyectos tecnológicos que involucran metodologías activas como el Aprendizaje Basado en Problemas, el movimiento maker y el pensamiento computacional. En cada práctica participan dos espacios curriculares. Para llevar a cabo el desarrollo de cada actividad, se revisaron cuidadosamente los descriptores, programas y planificación educativa de los espacios involucrados, con el objeto de seleccionar una temática en común, que permitiera el desarrollo de una práctica que aborde los contenidos curriculares, objetivos de aprendizaje y competencias deseadas por ambas cátedras.

### 1.1 Objetivos

Esta propuesta de actividades tiene como objetivo, desarrollar en los estudiantes habilidades transversales como: la investigación, el pensamiento crítico, la resolución de problemas, la creatividad, la comunicación y la colaboración. Como un subgrupo de habilidades se encuentra el Pensamiento Computacional, una práctica intelectual muy importante en el desarrollo de las competencias, habilidad necesaria para una sociedad cada vez más tecnológica.

De forma adjunta, las actividades propuestas, tienden también a contribuir al aumento de la motivación de los estudiantes, el deseo de aprender y a otorgar un mayor dinamismo a la planificación áulica. Lo ideal en este sentido sería llegar a considerar la tecnología como el idioma global del futuro, ya que la digitalización está y estará muy presente en el día a día de todos, y tendrá un fuerte impacto tanto en el desarrollo personal como profesional de los estudiantes.

## 2 Propuesta de actividades STEAM: Oportunidades de la didáctica de las ciencias para el despliegue STEAM.

Los objetivos STEM requieren el desarrollo de la competencia científica y la competencia tecnológica y es, desde la didáctica de las ciencias donde se pueden hacer aportaciones. La emergencia de espacios y proyectos STEM genera la necesidad de evaluar qué aportaciones a las distintas dimensiones de la competencia científica y tecnológica realizan esas actividades y de qué modo pueden convertirse en vías para su despliegue [4].

La enseñanza de las ciencias basada en la indagación (ECBI) propone que el alumnado aprenda ciencias emulando el proceso investigativo que sigue la ciencia para crear conocimiento: formular una pregunta investigable, diseñar experimentos y/o recoger y analizar datos, sacar conclusiones de datos, formular una explicación [5].

A continuación se describe una actividad STEAM basada en la indagación y la experimentación científica que incorpora el uso de una TIC muy emblemática dentro del mundo maker, la plataforma de hardware de código abierto, Arduino [6].

#### 1. Actividad 1 *Experimentación Científica: Testeo de Humedad*

Esta primera actividad integra a dos espacios curriculares que forman parte del ciclo básico de carreras del área de las ciencias exactas. Estos son:

##### *Ciencias de la Tierra*

- Unidad 2: Climatologías, geomorfología y pedología
- Subunidad: Pedología, la formación y estructura del suelo.
- Unidad 4: Protección ambiental y riesgos naturales.
- Subunidad: Riesgo Hídrico

##### *Informática*

- Unidad 6: Procesador de textos científicos: Lyx
- Unidad 7: Nuevas herramientas tecnológicas aplicadas a la ciencia
- Subunidad: Herramientas aplicadas a la investigación y/o simulación

El tema de la clase común a las 4 unidades de los espacios curriculares mencionados, será *Interpretación del contenido de humedad del suelo para evitar riego excesivo utilizando sensores de humedad*.

La misma consiste en llevar a cabo una experiencia científica, basada en el testeo de humedad. El objetivo principal es hacer Ciencia, utilizando placas Arduino y un sensor de humedad, luego se realizarán mediciones para llevar a cabo un experimento científico. Se desea probar que un cobertor de plástico, mejorará la retención de humedad en una maceta.

##### *Objetivos de la actividad:*

- Afianzar los conceptos y contenidos de ambos espacios curriculares para llevar a cabo la experiencia científica propuesta.
- Fortalecer la apropiación de los contenidos necesarios para lograr los objetivos de aprendizaje planteados en la unidad temática a la cual pertenece la experimentación científica propuesta.
- Estimular el desarrollo de las competencias científicas necesarias para resolver situaciones problemáticas similares
- Estimular vocaciones STEAM en los alumnos
- Estimular el desarrollo de competencias necesarias para abordar los desafíos del siglo XXI

##### *Estrategia de indagación*

- Tema (contenido a desarrollar en clase): Interpretación del contenido de humedad del suelo para evitar riego excesivo utilizando sensores de humedad.
- Curso (donde se imparte la clase): 1<sup>er</sup> Año, ciclo básico de carreras del área de ciencias exactas
- Espacios curriculares participantes: Ciencias de la Tierra, Informática
- Año (de ejecución de la clase): 2021
- Nombre del observador: Silvina Manganeli

### 1.1. Formulación de la hipótesis

La primera tarea de esta actividad consiste en formular la hipótesis: si se realiza el experimento en una maceta con cobertor plástico y en otra sin cobertor plástico, se tendrá una mayor retención de humedad en la primera maceta, en la maceta que cubrimos con el cobertor plástico.

### 1.2. Experimentación

La segunda tarea será diseñar el experimento, para ello se utilizarán 3 envases con tierra que simularán las macetas, luego se los impregnará a dos de ellos con agua (100 ml). De esta forma se realizarán dos tratamientos y un control: un envase con tierra y cubierta de plástico, otro envase con tierra y sin cubierta, y como control, un envase con tierra (sin agregar agua).

La variable de control será la humedad del suelo.

### 1.3. Mediciones

La tercera tarea será realizar mediciones, es aquí donde se introducirá la placa Arduino y un sensor de humedad, para determinar cuanta humedad existe en la tierra.

### 1.4. Creando código con Arduino

La cuarta tarea será desarrollar el código o programa en Arduino correspondiente para el testeo de humedad

A continuación el código con CIRCUITPYTHON<sup>1</sup>, se utilizará como editor de código a MU<sup>2</sup>.

```
# Se lee desde entrada analógica, el nivel de humedad, desde un sensor
conectado a GND, 3.3V, y al pin A6
# Imprime el resultado en la consola serial
import time
import board
import analogio
analogin = analogio.AnalogIn(board.A6)
def getVoltage(pin):
    return (pin.value * 3.3) / 65536
while True:
    print("Analog Voltage: %f" % getVoltage(analogin))
    time.sleep(0.5)
```

Ahora veamos el código con Arduino IDE en Arduino UNO

```
/*De 0 a 300 superficie seca
* 300 a 700 superficie húmeda
* 700 a 950 superficie mojada
*/
void setup (){
  Serial.begin (9600);
}
void loop() {
```

<sup>1</sup> CircuitPython es un lenguaje de programación diseñado para simplificar la experimentación y aprendizaje de programar en microcontroladoras de bajo costo. Está basado en Python

<sup>2</sup> Mu es un simple editor de código que funciona con las tarjetas CircuitPython de Adafruit. Está escrito en Python y trabaja en Windows, MacOS, Linux y en Raspberry Pi.

```

Serial.print ("Valores detectados: ");
Serial.println (analogRead(0));
delay(100);
}

```

#### 1.5. Registro y evaluación de mediciones

La quinta tarea será registrar las lecturas y luego de 24h medir nuevamente la humedad con el sensor, también comparar las mediciones efectuadas de cada día, durante al menos una semana.

#### 1.6. Emisión de juicios y conclusiones, redacción de un informe en un procesador científico como Lyx

Los alumnos reflexionarán sobre las mediciones efectuadas, y sobre sus hallazgos, durante toda la semana. Realizarán observaciones, comparaciones, analogías, clasificaciones y finalmente emitirán juicios y conclusiones, las cuales volcarán a un informe desarrollado en Lyx

#### 1.7. Retos científicos y tecnológicos

Finalmente se proponen varios retos o desafíos a los alumnos, tanto científicos como tecnológicos.

##### *Retos científicos*

- a. Hacer el experimento con más réplicas
- b. Aplicar estadística para determinar si las diferencias no se deben al azar
- c. Comparar el suelo en el jardín y el suelo en una maceta ¿hay diferencias y que explican esas diferencias?
  - a. ¿Se esperarían obtener resultados diferentes si los envases de tierra tuvieran una planta?
  - b. Comparar días lluviosos y no lluviosos (simplemente colocar el sensor en el jardín)
  - c. Medir la humedad del suelo y comparar con los cambios de temperatura en el ambiente ¿Qué sucede? (días de viento, de mucho calor, de mucho frío, días nublados, días de mucho sol, días nublados)
  - d. Ahora varíe con diferentes métodos para poder mantener o no la humedad (sol-sombra, diferentes tipos de tierra, etc.)

##### *Retos tecnológicos*

- a. Cambie el programa, aplique una alarma que avise cuando a la planta le falta agua y cuando esté demasiado húmeda para evitar el riego (usar el buzzer)
- b. Investigar otros fenómenos, utilizando otros sensores (luz, distancia, color, gases, presión, humedad, temperatura, etc.)
- c. El sensor sirve también para detectar la cantidad de agua en un tanque. ¿Qué proyecto podrían construir?

## 2. *Actividad 2 Resolución de Triángulos Rectángulos con Python*

Esta segunda actividad está enfocada en la resolución de problemas, específicamente el alumno deberá resolver dos problemas: el problema matemático, la resolución de una figura formada por triángulos rectángulos y el problema computacional, el desarrollo de un algoritmo que represente una de las posibles soluciones del problema planteado y permita el desarrollo de un producto o programa codificado en un lenguaje de programación, como puede ser Python.

Esta segunda actividad integra a dos espacios curriculares que forman parte del ciclo básico de carreras del área de las ciencias exactas. Estos son:

*Introducción a las matemáticas*

- Unidad 5: Trigonometría
- Subunidad: Trigonometría en triángulos rectángulos

*Informática*

- Unidad 7: Nuevas herramientas tecnológicas aplicadas a la ciencia
- Subunidad: Herramientas aplicadas a la investigación y/o simulación

Objetivos de la actividad:

- Afianzar los conceptos matemáticos necesarios para resolver la situación problemática planteada
- Fortalecer la apropiación de los contenidos matemáticos necesarios para lograr los objetivos de aprendizaje planteados en la unidad temática a la cual pertenece el problema
- Estimular el desarrollo de las competencias matemáticas necesarias para resolver situaciones problemáticas similares
- Estimular vocaciones STEAM en los alumnos
- Estimular el desarrollo de competencias necesarias para abordar los desafíos del siglo XXI
- Concebir el pensamiento computacional como una herramienta en Matemática que puede propiciar el despliegue de formas lógicas de pensamiento, necesarias para la construcción de distintos algoritmos que podrían usarse para resolver los problemas.
- Promover el uso del pensamiento computacional no sólo como un objeto de enseñanza en sí mismo, sino como una herramienta que permita mejores aprendizajes matemáticos, favoreciendo también la motivación de los estudiantes

2.1. Resolución del problema matemático

La primera tarea de esta actividad consiste en llevar a cabo la resolución del problema matemático (Fig.1). Se trata de una figura formada por 4 triángulos rectángulos, para poder resolverla, el alumno hará uso de sus conocimientos de trigonometría, teorema de Pitágoras y razones trigonométricas de triángulos rectángulos.

2.2. Resolución del problema computacional

2.2.1. Diseño del algoritmo

La segunda tarea será diseñar el algoritmo correspondiente que represente una de las soluciones posibles al problema.

2.2.2. Desarrollo del pseudocódigo

La tercera tarea será desarrollar el pseudocódigo correspondiente que facilitará el armado y codificación del programa y dará la libertad de elegir el lenguaje de programación deseado no sólo Python

2.2.3. Codificación del producto o programa

Finalmente se llevará a cabo la codificación del producto terminado, programado en Python, aunque puede ser codificado en cualquier lenguaje de programación de alto nivel.

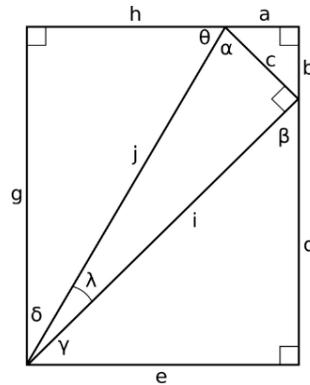


Fig.1 Problema Trigonométrico con Triángulos Rectángulos

```

Algoritmo ProblemaTrigonométrico (pseudocódigo)
Escribir 'RESOLUCIÓN DE TRIANGULOS RECTANGULOS'
Escribir 'Ingrese el lado a de la figura (no nulo, ni
negativo) '
Leer a
Escribir 'Ingrese el mismo valor para el lado b de la
figura'
Leer b
Escribir 'Ingrese ángulo lambda'
Leer lambda
Si a>0 Y b>0 Y lambda>0 Y a=b Entonces
Escribir 'RESOLVIENDO LA FIGURA'
// Aplicando Pitágoras
c <- rc(a^2+b^2)
Escribir 'El lado c mide: ',c
// Aplicando razones trigonométricas de triángulos
rectangulos
lambdarad <- lambda*PI/180
j <- c/sen(lambdarad)
Escribir 'El lado j mide: ',j
// Aplicando teorema de angulos complementarios
alfa <- 90-lambda
Escribir 'El ángulo alfa mide: ',alfa
// Aplicando teorema de la suma de angulos de un
triángulo
// Aplicando teorema del triángulo rectángulo
isosceles
epsilon <- 45
tita <- 180-alfa-epsilon
Escribir 'El ángulo tita mide: ',tita
// Aplicando razones trigonométricas de triángulos
rectangulos
titarad <- tita*PI/180
g <- sen(titarad)*j
Escribir 'El lado g mide: ',g
h <- cos(titarad)*j
Escribir 'El lado h mide: ',h

```

```

// Para obtener e, usamos la definición de un
rectángulo, sus pares de lados opuestos son iguales
// Contamos con h y a, por lo cual podemos obtener
e
e <- h+a
Escribir 'El lado e mide: ',e
// Para obtener d, también usamos la definición
de un rectángulo, sus pares de lados opuestos son iguales
d <- g-b
Escribir 'El lado d mide: ',d
// Para obtener delta, aplicamos teorema de
ángulos complementarios
delta <- 90-tita
// Para obtener gamma, aplicamos teorema de
ángulos complementarios
gamma <- 90-delta-lambda
// Para obtener Beta, aplicamos teorema de ángulos
complementarios

beta <- 90-gamma
Escribir 'El ángulo delta mide: ',delta
Escribir 'El ángulo gamm mide: ',gamma
Escribir 'El ángulo Beta mide: ',beta

SiNo
Escribir 'Ingrese un número válido, no nulo, no
negativo y los lados a y b deben ser iguales'
FinSi
FinAlgoritmo

```

## Consideraciones Finales

En la actividad 1, los alumnos podrán apreciar cómo funciona la ciencia. La ciencia comprende la capacidad de experimentar, la capacidad de razonar, realizar argumentaciones lógicas, racionales. La ciencia comienza siempre con una idea, esa idea genera una proposición concreta, una propuesta de la cual se obtendrán consecuencias, estas consecuencias se transformarán en hipótesis, que se podrán probar.

Si esa idea o teoría, plasmada en una proposición concreta, es Verdadera, entonces existe una serie de consecuencias, o sucesos que van a ocurrir, a partir de ahí surgirán muchas hipótesis que se desearán demostrar.

En la actividad 2, los alumnos podrán apreciar que el pensamiento computacional no es sinónimo de programación. El pensamiento computacional es un proceso de formulación y resolución de problemas que utiliza conceptos de la informática. La programación es una de las herramientas para poner en práctica el pensamiento computacional. Ambos conceptos comparten procesos cognitivos parecidos, son un medio para descomponer y resolver problemas, ponen en práctica la construcción de distintos algoritmos que podrían usarse para la resolución de problemas. Sin embargo la programación, limita sus procesos de resolución de problemas al ámbito de la

informática (para ello, hace uso de diferentes lenguajes de código, de variables o instrucciones de control como bucles y condicionales). Mientras que el pensamiento computacional es universal, pretende resolver problemas del mundo que nos rodea aplicando conceptos informáticos [7].

Por otro lado se eligió a Python como lenguaje de programación, por diversos motivos:

- Lenguaje de alto nivel, presenta gran calidad en su sintaxis
- Puede ser utilizado en múltiples plataformas
- Es libre, gratuito y nos ofrece código abierto
- Python es utilizado en una gran variedad de ámbitos académicos, sobre todo en machine learning y data science. Su ventaja frente a Matlab (otra herramienta utilizada para machine learning) es que esta última no es gratuita.

Pero sobre todo porque fue un pedido común de los alumnos, en las encuestas realizadas a los mismos, tanto el primer día de clase (rapport para evaluar el conocimiento e inquietudes tecnológicas y de aprendizaje por parte de los alumnos) o al finalizar el cursado (comentarios en la evaluación de desempeño docente) Se tuvieron en cuenta estos pedidos, con el fin de incorporar futuras mejoras para la cátedra.

Finalmente cabe destacar que estas actividades forman parte de una prueba piloto, correspondiente a un proyecto de investigación y a un proyecto de posgrado, la cual se realizará con alumnos pertenecientes al ciclo básico de carreras del área de ciencias exactas e ingeniería.

#### *Instrumento de Evaluación*

A continuación se detalla los instrumentos de evaluación que se utilizarán para evaluar el nivel o grado de desarrollo de las competencias científicas presente en los alumnos del ciclo básico de carreras del área de las ciencias exactas e ingeniería. Estos instrumentos se aplicarán tanto en su modalidad de pre-test como pos-test. Se han desarrollado dos instrumentos de evaluación, uno para medir el desarrollo de las competencias científico-tecnológico (Fig.2) y otro para medir las competencias necesarias para afrontar los desafíos del siglo XXI (Fig.3).

#### *Instrumento para evaluar competencias científicas*

##### Competencias a evaluar

- Explorar hechos y fenómenos
- Analizar problemas
- Formular Hipótesis
- Observar, recoger y organizar información
- Utilizar diferentes métodos de análisis
- Evaluar los métodos
- Compartir los resultados

#### *Instrumento para evaluar competencias del siglo XXI*

##### Competencias a evaluar

- Pensamiento Crítico
- Colaboración
- Resolución de problemas
- Creatividad
- Pensamiento computacional

Clasificaremos el grado de desarrollo de estas competencias presentes en los alumnos en 5 niveles:

1. Inferior
2. Medio Bajo
3. Medio
4. Media Alto
5. Superior

**Instrumento para evaluar competencias científicas**

		1. Inferior	2. Medio Bajo	3. Medio	4. Media Alto	5. Superior					
COMPETENCIA CIENTIFICA	INDICADOR	1	2	3	4	5	DESCRIPCIÓN DE LA EVIDENCIA				
Explorar hechos y fenómenos	1. Lee y o escucha para explorar el fenómeno										
	2. Utiliza diferentes fuentes para explorar fenómenos										
	3. Hace inferencias para establecer el fenómeno y la situación problema										
Analizar problemas	1. Utiliza diferentes fuentes para analizar un problema										
	2. Propone y construye en equipo soluciones a los problemas planteados										
	3. Recoge información significativa para resolver el problema										
Formular Hipótesis	1. Elabora conjeturas preliminares										
	2. Resume los elementos que someterá a estudio										
	3. Explica la forma de abordar la relación entre los elementos contenidos en la hipótesis										

**Fig.2.** Extracto del instrumento de evaluación propuesto para evaluar competencias científicas

**Instrumento para evaluar competencias del siglo XXI**

		Inferior	2. Medio Bajo	3. Medio	4. Media Alto	5. Superior					
COMPETENCIA	INDICADOR	1	2	3	4	5	DESCRIPCIÓN DE LA EVIDENCIA				
Pensamiento Crítico	El pensamiento crítico es la capacidad de desarrollar la reflexión crítica independiente. El pensamiento crítico permite el análisis de las ideas, los conocimientos y los procesos con un sistema de valores y juicios propios. Se trata de un pensamiento responsable, basado en criterios y sensible al contexto y a los demás.										
	1. Componente 1 (C1): Identifica los componentes de una idea o trabajo. 2. Componente 2 (C2): Explora diferentes perspectivas y posiciones relacionadas con una idea o trabajo. 3. Componente 3 (C3): Toma posición en comparación con una idea o trabajo.										
Resolución de problemas	La resolución de problemas es la capacidad de identificar una situación problemática, para la cual, el proceso y la solución no se conocen de antemano, y que implica determinar una solución, desarrollarla y ponerla en práctica de manera efectiva.										
	1. Componente 1 (C1): Analiza los elementos de la situación. 2. Componente 2 (C2): Explora una variedad de soluciones y desarrolla una solución eficaz teniendo en cuenta el contexto de la situación problemática inicial. 3. Componente 3 (C3): Pone a prueba las posibles soluciones; las evalúa y adopta un funcionamiento flexible										

**Fig.3.** Extracto del instrumento de evaluación propuesto para evaluar competencias del siglo XXI

## Agradecimientos

La autora agradece la ayuda financiera recibida de la Secretaría de Investigación, Internacionales y Posgrado de la UNCUYO, para llevar a cabo el Proyecto Bienal SIIP Tipo 1 2019-2021, T003 “STEAM un nuevo enfoque didáctico para la formación científica de alumnos pertenecientes a carreras del área de las ciencias exactas, necesarias para afrontar los desafíos del siglo XXI” (aprobado por resolución N° 3922/2019-R). También a las unidades académicas ITU y FCEN Uncuyo, involucradas en el desarrollo del proyecto, por sus prestaciones y excelente predisposición para hacer posible la ejecución de las actividades propuestas.

## Referencias

1. El desafío de las vocaciones STEM, <https://www.digitales.es/wp-content/uploads/2019/09/informe-el-desafio-de-las-vocaciones-stem-digital-af.pdf>
2. Competencias del siglo XXI en proyectos co-tecnocreativos [https://www.researchgate.net/publication/323549661\\_Competencias\\_del\\_siglo\\_XI\\_en\\_proyectos\\_co-tecnocreativos](https://www.researchgate.net/publication/323549661_Competencias_del_siglo_XI_en_proyectos_co-tecnocreativos)
3. Sánchez Ludeña, E. La educación STEAM y la cultura «maker». Padres Y Maestros / Journal of Parents and Teachers, (379), 45-51(2019).
4. Domènech-Casal, J. Aprendizaje Basado en Proyectos en el marco STEM. Componentes didácticos para la Competencia Científica. *Ápice. Revista de Educación Científica*, 2(2), 29-42 (2018).
5. Domènech Casal, J. STEM: Oportunidades y retos desde la Enseñanza de las Ciencias, *Universitat Tarraconensis. Revista de Ciencias de la Educación*, 155-168 (2019)
6. Experimentación Científica (Circuit Playground) <https://www.instructables.com/id/Experimentaci%C3%B3n-Cient%C3%ADfica-Circuit-Playground/>
7. Zapata-Ros, M. *Revista De Educación a Distancia (RED)*, (46). Recuperado a partir de <https://www.um.es/ead/red/46/zapata.pdf> (2015).

# Herramientas de model finding para asistir en la Construcción de Especificaciones Formales

Sonia Permigiani<sup>1</sup>, María Marta Novaira<sup>1</sup>, Gastón Scilingo<sup>1</sup>, and Marcelo Arroyo<sup>1</sup>

Dpto. de Computación, FCEFQyN, Universidad Nac. de Río Cuarto, Argentina  
{spermigiani,mnovaira,gscilingo,marcelo.arroyo}@dc.exa.unrc.edu.ar

**Resumen** Este trabajo plantea el desafío técnico de asistir en la construcción de especificaciones en el contexto de la enseñanza de la programación, en materias introductorias, a través de la provisión automática de información de análisis, basada en constraint solving relacional. Se presentan a través de ejemplos concretos, por un lado, la traducción de especificaciones que involucran expresiones cuantificadas al lenguaje Alloy, y por otro la definición de consultas de análisis que, mediante las herramientas Alloy y DynAlloy, brindan *feedback* automático a los estudiantes. Se discuten además algunas dificultades técnicas que se presentan al construir modelos de especificaciones con dominios numéricos, para su utilización en herramientas de model finding.

## 1. Introducción

Los sustanciales avances en técnicas de análisis automático de software han dado lugar a poderosas herramientas que comienzan a ser usadas en contextos industriales reales, explotando mecanismos como la generación automática de tests, el constraint solving y la ejecución simbólica. Estas herramientas permiten encontrar defectos sutiles en software, y pueden brindar *feedback* de gran utilidad a los desarrolladores. No es sorprendente entonces que también hayan encontrado aplicaciones en contextos educativos, donde pueden emplearse para brindar *feedback* automático a estudiantes, en sus procesos de construcción de programas. Herramientas como KeY (<https://www.key-project.org/>), Pex [16], y herramientas como Infer en su versión accesible desde CodeBoard (<http://codeboard.io>), son ejemplos de herramientas con sofisticados análisis, orientadas a contextos de enseñanza de la programación. Asimismo, otras técnicas menos elaboradas, como la provisión *manual* de tests, suelen usarse para contrastar soluciones de estudiantes contra las expectativas de los docentes (véase su uso, por ejemplo, en plataformas como HackerRank).

En todos los ejemplos antes mencionados el énfasis está en la construcción de *programas*, y su respectivo análisis. No son pocos, sin embargo, los cursos introductorios a la programación que ponen especial énfasis en las *especificaciones*. Enfoques introductorios a la programación como los presentados en [2,3] son sólo algunos ejemplos de estrategias de enseñanza que se centran en especificaciones.

En éstas y otras metodologías relacionadas, el proceso de *construcción* de una especificación es muy difícil, y los estudiantes suelen contar con prácticamente ningún tipo de asistencia automática, a través de herramientas.

Nuestro objetivo es justamente cubrir esta falencia. En [18] comenzamos a delinear una metodología basada en el uso de *model finders*, como los subyacentes a lenguajes del estilo de Alloy y DynAlloy, para asistir automáticamente a estudiantes en el proceso de construcción de especificaciones, por ejemplo, proveyendo escenarios que sus especificaciones candidatas no capturan, o capturan erróneamente. En este trabajo, mostramos pasos concretos en esta dirección. Tomando como punto de partida el lenguaje de expresiones cuantificadas introducido en [3], mostramos una caracterización indirecta del mismo en el lenguaje formal Alloy, con el objetivo de aprovechar el model finder Alloy Analyzer, asociado a este último. La traducción es indirecta porque, como mostramos en el artículo, requiere pasar por el lenguaje DynAlloy, una extensión dinámica de Alloy, con el objetivo de poder capturar fielmente algunas expresiones del lenguaje origen. Mostramos las dificultades que se dan en el análisis del lenguaje origen a través de model finders, en especial debido al fuerte énfasis del mismo en el uso de aritmética. Finalmente, mostramos algunos ejemplos de la herramienta resultante, con casos específicos de problemas de especificación, y un análisis del feedback provisto por nuestra herramienta.

## 2. Alloy y DynAlloy

*Alloy* [8] es un lenguaje de especificación de primer orden basado en el uso de relaciones. Está orientado a la especificación de propiedades estructurales de sistemas e intenta resolver la complejidad en la expresión de algunas propiedades comunes en lógica de primer orden utilizando operadores relacionales, definidos en la lógica relacional, el formalismo subyacente a Alloy. Las *signaturas* permiten describir dominios y sus estructuras, siguiendo el estilo de Z [12]. De manera similar, se pueden especificar operaciones mediante expresiones que relacionan el estado previo a la aplicación de una operación y el estado posterior correspondiente. Para ilustrar la definición de operaciones en Alloy, considere, por ejemplo, una operación que especifica la escritura de un valor en una dirección de memoria:

```
pred write(m, m': Memory, d: Data, a: Addr) {
    m'.map = m.map ++ (a -> d)
}
```

La operación *write*, descrita a través de un predicado, transforma una memoria, sobre-escribiendo el valor asociado a una dirección *a* con un dato *d*. Este predicado, como los esquemas para operaciones en Z, no está directamente asociado a ninguna signatura. Las variables primadas son usadas para denotar estados posteriores a la ejecución de la operación, aunque esto es simplemente una convención, ya que no se refleja en la semántica.

*DynAlloy* [20] es una extensión del lenguaje Alloy, que agrega nuevas características que permiten expresar mejor los cambios de estado. Una de las deficiencias más importantes de Alloy es que carece de una forma sencilla para expresar propiedades dinámicas de sistemas. *DynAlloy*, inspirado en la lógica dinámica, consigue especificar propiedades de trazas de manera más apropiada. También se caracteriza por la sencillez y rapidez de poder traducir modelos *DynAlloy* a Alloy sin requerir intervención por parte del usuario, para luego realizar validaciones utilizando el Alloy Analyzer. La sintaxis de *DynAlloy* extiende la de Alloy agregando una cláusula para definir programas:

```
formula ::= . . . {formula} program {formula}
```

Con la incorporación de programas, los predicados que representaban operaciones sobre el modelo ahora se pueden definir como acciones dentro de un programa:

```
{ true }
write(m : Memory, d : Data, a : Addr)
{ m'.map = m.map ++ (a -> d) }
```

En la especificación anterior, la variable  $m'$  representa el valor de la variable  $s$  después de la ejecución de la acción `write`, y se asume que aquellas variables que no ocurren en la post-condición de forma primada, retienen su valor inicial.

### 3. El Lenguaje de Expresiones Cuantificadas

La comprensión precisa de un problema a resolver es un primer paso, ampliamente entendido como ineludible en la Ingeniería de Software moderna y las ciencias de la computación, en el camino para el desarrollo de software de manera efectiva. Es por esto que es importante que existan técnicas automáticas de software que den soporte al tipo de problemas que se dan al intentar capturar la especificación de un problema. Esta importancia se observa, por ejemplo, en el lugar que ocupan las técnicas basadas en lenguajes formales para especificación de programas en varias currículas de carreras de computación. En esta sección presentamos un lenguaje para especificaciones formales basado en una versión ecuacional de lógica de predicados con cuantificadores [2], utilizado en el marco de cursos de programación que introducen el concepto de especificación.

Las *expresiones cuantificadas* resultan útiles para expresar la reiteración de aplicaciones de un cierto operador sobre una expresión. Esta notación resulta muy práctica para especificar los problemas usuales que se plantean en los cursos introductorios a la programación. Consideremos un ejemplo sencillo: la suma de los  $n$  primeros números impares. Utilizando una notación matemática estándar, podemos describir el problema con la siguiente expresión:

$$\sum_{i=0}^{n-1} 2 \times i + 1 \quad (1)$$

En esta expresión distinguimos varias componentes: el operador usado (la sumatoria) se corresponde a la operación binaria suma,  $i$  es una variable con un rango de variación asociado ( $0$  a  $n - 1$ ), y por último la expresión  $(2 \times i + 1)$ , que denota cuáles van a ser los términos de la sumatoria.

En el lenguaje aquí utilizado, este mecanismo de definición de expresiones es generalizado a operadores binarios asociativos y conmutativos. La noción de cuantificación aparece dada por variables formales definidas con un alcance delimitado específicamente por los paréntesis, y éstas se podrán usar para construir expresiones dentro del alcance especificado. La forma general de las especificaciones será:

$$(\oplus i : R : T) \quad (2)$$

Ésta consta de tres partes, delimitadas por “:”. En la primera, el símbolo  $\oplus$  representa el operador (e.g.,  $\vee$ ,  $\wedge$ ,  $+$ ,  $Max$ ,  $Min$ , etc), donde  $i$  es la variable cuantificada (no hace falta definir el tipo de la misma). La segunda parte es el rango de la expresión cuantificada, denotado con el predicado  $R$ . Y finalmente  $T$  es una función que denota el término de la cuantificación. Utilizando esta notación, podemos, por ejemplo, escribir la expresión (1) de la siguiente manera:

$$\left(\sum i : 0 \leq i < n : 2 * i + 1\right) \quad (3)$$

El resultado de esta sumatoria depende del valor de  $n$ , una variable fuera del alcance de la expresión. Cabe destacar que una expresión cuantificada puede involucrar más de una variable ligada. Por ejemplo, si queremos expresar la propiedad que indica que una secuencia es capicúa, podríamos hacerlo a través de la siguiente especificación:

$$(\forall p, q : 0 \leq p \wedge 0 \leq q \wedge p + q = N - 1 \wedge N \leq \#xs : xs.p = xs.q) \quad (4)$$

donde  $p$  y  $q$  son dos variables cuantificadas y  $xs$  es una secuencia (el punto es el operador de indexación para secuencias, i.e.,  $xs.p$  será el elemento de  $xs$  en la posición  $p$ ).

#### 4. De Expresiones Cuantificadas a (Dyn)Alloy

En [18] se plantea la utilización de herramientas como Alloy y DynAlloy para la asistencia automática en la construcción de especificaciones, y se introducen ejemplos en los que se muestra cómo detectar errores comunes, tomados de exámenes y prácticas. El proceso de detección requiere de una “versión dorada”, una especificación correcta (provista por el docente) del problema en cuestión, contra la cual contrastar la especificación candidata del estudiante, y por supuesto una *traducción* a Alloy, para poder contrastar las especificaciones automáticamente, usando Alloy Analyzer. En este trabajo estudiamos justamente esta traducción. El objetivo es poder brindar una traducción automática de expresiones cuantificadas a Alloy, preservando la semántica de las fórmulas originales. Como veremos, esta tarea implica una serie de desafíos técnicos, que intentamos sortear en el artículo.

#### 4.1. Cuantificadores Booleanos

El lenguaje Alloy provee soporte directo para las cuantificaciones universal ( $\forall$ ) y existencial ( $\exists$ ), de la lógica de primer orden, a través de sus operadores `all` y `some`. Sin embargo, la traducción de las expresiones cuantificadas booleanas de un lenguaje a otro no es directa, ya que requiere el uso de propiedades lógicas que permitan pasar el predicado del *rango* al *término*. Mostramos la traducción de estas expresiones, las más simples del lenguaje, a través de ejemplos ilustrativos.

Para el cuantificador universal, el rango se reduce al *antecedente* de una implicación, que tiene al *término* como consecuente. Por ejemplo, la expresión cuantificada siguiente:

$$(\forall i : 0 \leq i < \#xs - 1 : xs.i = xs.(i + 1))$$

que indica que todos los elementos de una secuencia de enteros son iguales, se captura en Alloy de la siguiente manera<sup>1</sup>:

```
all i: Int | 0 <= i < #xs-1 implies xs[i]= xs[i+1]
```

Para la traducción de expresiones cuantificadas existencialmente, la traducción convierte al rango en una conjunción que acompaña al término. Por ejemplo, la expresión cuantificada siguiente:

$$(\exists i : 0 \leq i < \#xs : xs.i = 0)$$

que captura el hecho de que la secuencia `xs` contiene al menos un cero, se traduce en la siguiente expresión Alloy:

```
some i: Int | 0 <= i < #xs and xs[i]= 0
```

Como vemos, para el caso de las expresiones con dominio en los booleanos, la traducción a Alloy es bastante directa.

#### 4.2. Cuantificadores Numéricos

La utilización de expresiones con sumatorias y productorias son naturales para los estudiantes, dado que las trabajan de manera intensiva en materias de matemática discreta y cálculo. El lenguaje de expresiones cuantificadas las introduce como una cuantificación más, incorporando una forma natural de capturar programas en esta notación. Esto genera sin embargo un problema al intentar utilizar Alloy como herramienta de análisis, ya que el lenguaje no provee soporte para sumatorias y productorias.

Para expresar sumatorias y productorias en Alloy, recurriremos a construir expresiones equivalentes que *computen* las sumas y productos correspondientes. Para este *cómputo*, utilizaremos el lenguaje DynAlloy, que extiende Alloy con la posibilidad de escribir *programas*, expresiones pseudo-declarativas que manejan

<sup>1</sup> En los ejemplos con los que se describe la traducción de las diferentes expresiones, se hace abuso de notación en la sintaxis Alloy a los fines de legibilidad.

estado. Esencialmente, dada una sumatoria o productoria, generaremos automáticamente un programa DynAlloy que “calcule” la expresión respectiva; esto estará oculto, por supuesto, al estudiante, dado que es parte de la traducción a Alloy (las sumatorias y productorias, capturadas a través de expresiones cuantificadas, seguirán siendo el “front-end” para el estudiante).

Para ilustrar la traducción, consideremos el siguiente programa DynAlloy, que calcula la sumatoria de los elementos de una secuencia:

```

program sumatoria[s:seq Int, suma:Int] var [indexSeq: Int] {
  indexSeq := 0;
  suma := 0;
  while (#s.elems > 0 and indexSeq <= lastIdx[s]) do {
    suma := s[indexSeq] + suma;
    indexSeq := indexSeq + 1
  }
}

```

Como vemos, la representación de una sumatoria en este caso está expresada con un programa, que recorre la secuencia y va acumulando en una variable la suma de los elementos en cada índice de la misma. La traducción sigue esta idea. Para capturar una expresión cuantificada aritmética arbitraria, nuestra traducción construye los elementos de la secuencia que sumará el programa DynAlloy, considerando el rango de las variables involucradas, las restricciones sobre la expresión y además el término que representa una expresión aritmética.

A continuación, a través de un ejemplo, describimos los detalles de la traducción de una especificación cuantificada con dominio numérico a DynAlloy. Consideremos la siguiente especificación, que captura la suma de los elementos pares de una secuencia:

$$\left(\sum i : 0 \leq i < \#xs \wedge \text{par}(xs.i) : xs.i\right)$$

Para traducir esta especificación introducimos en DynAlloy un predicado para representar el rango:

```

pred rango[xs: seq Int, i: Int] {
  0 <= i and i < #xs and par[xs.i]
}

```

De manera similar, capturamos el término de la expresión cuantificada como una función:

```

fun termino[xs: seq Int, i: Int]: one Int {
  xs[i]
}

```

Finalmente, mediante un predicado, se construye la secuencia con los valores tomados de la función término, aplicada a las variables que cumplen con el rango:

```

pred buildSeq[xs: seq Int, res: seq Int] {
  #xs = 0 => #res = 0 else
  ( #xs <= 1 and #res <= #xs and
    (some k: inds[xs] | rango[xs,k]) implies #res > 0) and
  ( all i: inds[xs] |
    ( rango[xs,i] implies
      some j: Int | res[j]=termino[xs,i] and j <= i
      and all e: res.elems | some k: inds[xs] | termino[xs,k] = e
      and # { z:inds[xs] | rango[xs,z] } = #inds[res]
    ) and
    !rango[xs,i] implies # {z:inds[xs] | rango[xs,z] } = #inds[res]
  )
)
}

```

El predicado que construye la secuencia resultante considera primero el caso de una secuencia vacía, para el cual no construye ningún elemento. En caso que existan elementos para analizar, aparecen una serie de conjunciones de predicados que restringen los modelos con aserciones tales como que si algún índice cumple el rango, el resultado tiene elementos, para cada elemento del rango, el término correspondiente debe estar en el resultado, etc.

Continuando con el ejemplo, escribimos una aserción para comprobar que al aplicar el programa sumatoria a la secuencia resultante se obtiene lo esperado. Para esto definimos la siguiente aserción:

```

assert sumatoria[s, res: seq Int, suma: Int] {
  pre { buildSeq[s, res] }
  program { call sumSeq[res,suma] }
  post { par[suma'] }
}

```

Si la construcción de la secuencia es correcta, el resultado de la sumatoria de elementos pares de una secuencia debería dar como resultado un número par. Para comprobar si la aserción produce errores usamos el siguiente comando:

```
check sumatoria for 4 int, 2 seq lurs 4
```

Al ejecutarlo, obtenemos el siguiente resultado:

```

Executing "Check sumatoria for 4 int, 2 seq lurs 4 default"
Solver=sat4j Bitwidth=4 MaxSeq=2 SkolemDepth=4 Symmetry=20
4890 vars. 243 primary vars. 14379 clauses. 63ms.
No counterexample found. Assertion may be valid. 94ms.

```

Como es de esperar, la aserción para una especificación correcta no muestra contraejemplos. Como parte de la evaluación de nuestra caracterización, sin embargo, observamos problemas que tienen que ver con la naturaleza no tipada de Alloy. Uno de estos surgió al analizar diferentes tipos de errores típicos de

estudiantes, en la construcción de especificaciones, como por ejemplo la definición de rangos inválidos. Por ejemplo, al indexar una secuencia, si un valor admisible por el rango para el índice no pertenece a la secuencia, el predicado que construye la secuencia con los elementos simplemente ignora los casos (las expresiones indefinidas se anulan en Alloy). Esto da como resultado que algunos errores no sean detectados. Para resolver este problema, incorporamos un programa DynAlloy adicional, que se encarga de recorrer los índices del rango y verificar que todos dan lugar a valores válidos del término:

```

program checkTermino[xs: seq Int, flag: Bool ]
  var [index,cantElems,i : Int, r: set Int] {
    flag := False;
    r := {z: Int | rango[xs,z]};
    index := min[r];
    i:=0;
    cantElems := #r;
    while ( lt[i,cantElems] and (False = flag) ) do {
      if ( error [xs,index] ) {
        flag := True
      } else {
        skip
      };
      r := r - {min[r]};
      i := add[i,1];
      if (! no r ) {
        index := min[r]
      }
    }
  }
}

```

Utilizando este programa y el chequeo de la siguiente aserción:

```

assert terminoValido[xs: seq Int, flag: Bool] {
  pre { flag = False and (#s > 0) }
  program { call checkTermino[s,flag] }
  post { flag' = False }
}

```

podemos distinguir los casos en los que la especificación permite indexar una secuencia en un índice inválido. Para ver cómo funciona el análisis, supongamos que la especificación tuviera el siguiente rango incorrecto:

```

pred rango[xs: seq Int, i: Int] {
  0 <= i and i <= #xs and par[xs.i]
}

```

Como vemos, el índice de la variable cuantificada abarca  $\#xs$  mientras que los índices válidos de la secuencia no lo incluyen. Cuando chequeamos los términos válidos utilizando este nuevo rango, de la siguiente manera:

```
check terminoValido for 4 int, 3 seq
```

obtenemos lo siguiente:

```
Executing "Check terminoValido for 4 int, 3 seq"
  Solver=sat4j Bitwidth=4 MaxSeq=3 SkolemDepth=4 Symmetry=20
  3791 vars. 269 primary vars. 10600 clauses. 124ms.
  Counterexample found. Execution. Assertion is invalid. 78ms.
```

Como vemos, la herramienta detecta un contraejemplo con instancias para las cuales, luego de ejecutar el programa que recorre los términos, la “bandera” se activa, indicando que existe un valor del rango para el cual el término está indefinido (y retornando un modelo que lo refleja). Esto permite detectar problemas de tipo rango inválido, divisiones por cero, etc., obteniendo situaciones concretas que lo ejemplifican, para que el estudiante las pueda analizar.

## 5. Conclusiones y Trabajo Futuro

En contextos educativos se hace cada vez más necesario contar con herramientas de análisis automático para asistir a los estudiantes en las diferentes etapas del desarrollo de software. En este trabajo tomamos un lenguaje formal de especificaciones, basado en la construcción de expresiones cuantificadas, y presentamos una traducción automática al lenguaje relacional Alloy, con la finalidad de utilizar el analizador automático Alloy Analyzer, un model finder, para analizar las especificaciones originales. El trabajo se basa en las ideas introducidas en [18], y muestra algunos desafíos técnicos en la construcción de la traducción, que demandó la utilización de un lenguaje operacional, DynAlloy, para capturar ciertas expresiones y propiedades de las mismas. Mostramos ejemplos en los cuales las herramientas de análisis proveen feedback sobre especificaciones, al estilo de rangos inválidos y problemas similares, equivalentes en cierta medida al tipo de feedback que se recibe al analizar programas. Cabe resaltar que las herramientas de análisis automático tienen limitaciones que están atadas a la expresividad de los lenguajes analizados. En nuestro caso, al trabajar sobre lenguajes de primer orden, no decidibles, es evidente que no se puede garantizar la equivalencia de dos especificaciones, ni garantizar la validez de ciertas fórmulas. El análisis aquí utilizado no es exhaustivo, sino acotado, pero aún así es sumamente útil, ya que en la práctica permite encontrar violaciones a propiedades esperadas dentro de las cotas de análisis (véase la *Hipótesis de la Cota Pequeña* [9] como argumento sobre este punto).

En cuanto a trabajo futuro, y de acuerdo a la experiencia obtenida en la representación de las cuantificaciones numéricas y los problemas que surgen durante la traducción a DynAlloy, consideramos importante analizar algunas herramientas tradicionales para el análisis de aritmética, como SMT (satisfiability modulo theories) solvers, que posiblemente tengan mejor soporte para el tipo de análisis que necesitamos.

## Referencias

1. J. Abrial, *The B-Book: Assigning Programs to Meanings*, Cambridge University Press, 2005.
2. R. Backhouse, *Program construction: calculating implementations from specifications*, Wiley, 2003.
3. D. Barsotti, J. O. Blanco, S. Smith, *Cálculo de Programas*, Universidad Nacional de Córdoba, 2008.
4. A. Biere, M. Heule, H. van Maaren y T. Walsh, *Handbook of Satisfiability: Volumen 185*, Frontiers in Artificial Intelligence and Applications. IOS Press, Amsterdam, Holanda. 2009.
5. D. Chaudhari y O. Damani, *Automated Theorem Prover Assisted Program Calculations*, en Proc. de IFM 2014, LNCS, Springer, 2014.
6. E. Clarke, O. Grumberg y D. Peled, *Model Checking*, MIT Press, 2000.
7. N. Eén, N. Sörensson, *An Extensible SAT-solver*, en Proc. de SAT 2003, LNCS 2919, Springer, 2004.
8. D. Jackson, *Software Abstractions: Logic, Language, and Analysis*, The MIT Press, 2006.
9. D. Jackson, *Alloy: a lightweight object modelling notation*, in ACM Transaction Software Engineering and Methodology (ACM TOSEM), vol. 11, Nro. 2, 2002.
10. M. Frias, J.P. Galeotti, C. López Pombo y N. Aguirre, *DynAlloy: upgrading alloy with actions*, en Proc. de International Conference on Software Engineering ICSE 2005, St. Louis, USA. ACM Press, 2005.
11. M. Frias, C. López Pombo y M. Moscato, *Alloy Analyzer+PVS in the Analysis and Verification of Alloy Specifications*, en Proc. de TACAS 2007, LNCS 4424, Springer, 2007.
12. J. Jacky, *The Way of Z: Practical Programming with Formal Methods*, Cambridge University Press, 1997.
13. J. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, 1992.
14. A. Levitin, M. Papalaskari, *Using puzzles in teaching algorithms*, en Proc. de SIGCSE 2002, USA, 2002
15. M. W. Moskewicz, C. F. Madigan, Y. Zhao, L. Zhang, S. Malik, *Chaff: Engineering an Efficient SAT Solver*, en Proc. de DAC 2001, ACM, 2001.
16. N. Tillmann y J. de Halleux, *PEX: White Box Test Generation for .NET*, en Proc. de TAP 2008, LNCS 4966, Springer, 2008.
17. T. Xie, N. Tillmann y J. de Halleux, *Educational software engineering: where software engineering, education, and gaming meet*, en Proc. de 3rd International Workshop on Games and Software Engineering: Engineering Computer Games to Enable Positive, Progressive Change, IEEE Press, 2013.
18. C. Cornejo, M. Politano, F. Raverta, S. Permigiani, P. Ponzio, G. Regis, N. Aguirre, *Analizando el uso de (Dyn)Alloy como herramienta educativa*, In Proc. de CACIC 2015.
19. R. Singh, S. Gulwani and A. Solar-Lezama, *Automated Feedback Generation for Introductory Programming Assignments*, in Proceedings of the 34th ACM SIGPLAN Conference on Programming Language Design and Implementation, 2013.
20. G. Regis,; C. Cornejo, S.G. Brida, M. Politano, F. Raverta, P. Ponzio, N. Aguirre, J.P. Galeotti, and M. Frias, *DynAlloy analyzer: a tool for the specification and analysis of alloy models with dynamic behaviour*, in Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering, ESEC/FSE 2017, Paderborn, Germany, September 4-8, 2017, pages 969–973, 2017. ACM.

# Experiencia Educativa con Tecnologías Emergentes

Mariela Zúñiga<sup>1</sup>, Claudia Liendo<sup>1</sup>, María Rosas<sup>1</sup>, Graciela Rodríguez<sup>1</sup>,  
Nicolás Jofré<sup>1</sup>, Yoselie Alvarado<sup>1</sup>, Myriam Villegas<sup>2</sup>, Jacqueline Fernandez<sup>1</sup>,  
and  
Roberto Guerrero<sup>1</sup>

[1] Laboratorio de Computación Gráfica (LCG)

[2] Departamento de Física

FCFMyN, Universidad Nacional de San Luis,

Ejército de los Andes 950

Tel: 02664 420823, San Luis, Argentina

{mezuniga,caliendo,mvrosas,gbrodriguez,npasinetti,ymalvarado,mvilleg,  
jmfer,rag}@unsl.edu.ar

**Resumen** Las Nuevas Tecnologías de la Información y las Comunicaciones (NTIC) se han convertido en poderosos motores de crecimiento en los distintos órdenes sociales, transformando la estructura del mundo en que vivimos. En particular, las tecnologías emergentes inmersivas permiten generar experiencias inmersivas que interpelan las estrategias de enseñanza y aprendizaje y revolucionan al sistema educativo. En este trabajo se presenta una experiencia de introducción de tecnologías emergentes en las aulas como generadoras de experiencias educativas inmersivas. Las actividades se focalizaron en la ejecución de prácticas por parte de los estudiantes con las tecnologías de Realidad Virtual, Realidad Aumentada y Holografía. El éxito de los resultados obtenidos incentiva a continuar trabajando, extendiendo las actividades a docentes y directivos.

**Palabras claves:** Tecnologías Emergentes, Pedagogías Emergentes, Realidad Virtual, Realidad Aumentada. Holografía.

## 1. Introducción

La actual sociedad de la información y del conocimiento ha provocado la necesidad de cambios profundos en la educación que favorezcan el desarrollo de nuevas habilidades y competencias requeridas para asegurar a las generaciones actuales el acceso a nuevas oportunidades. En Latinoamérica y, en particular en Argentina, existe una urgente necesidad de dichos cambios que adapten la formación del capital humano competitivo en sintonía con la Cuarta Revolución Industrial y la Agenda 2030 de la ONU [1].

La existencia de computadoras y su conexión a Internet en el aula ó en la escuela, provocan una redefinición del aula como espacio pedagógico. En con-

secuencia, es difícil sostener la enseñanza tradicional frontal, simultánea y homogénea, dando espacio a otro tipo de organización. Mas aún, las nuevas pedagogías exigen una didáctica centrada en el estudiante, en la diversidad y sensible a las diferencias culturales, demandando al docente el ejercicio de una práctica educativa innovadora, que incluya estrategias de aprendizaje considerando las NTIC como parte integral del contexto, con base en las pedagogías constructivistas y multiculturales como punto de partida para el desarrollo de entornos altamente interactivos y participativos, donde el estudiante es capaz de modificar, construir, probar ideas e involucrarse activamente en la resolución de un problema [29].

En [15] Burbules establece que la escuela es el centro de aprendizaje más importante para los estudiantes, pero no el único. Debido a la existencia de los dispositivos móviles cada vez es más difícil marcar el límite entre lo escolar y lo no escolar. La mayoría de los estudiantes dispone de estos aparatos ubicuos que permiten “estar y no estar” en varios lugares a la vez, favoreciendo la desaparición de los límites de los espacios físicos y de sus reglas de uso. Esto no debe entenderse como una “invasión” del espacio escolar, sino como una ampliación de su radio de acción. Como consecuencia, el espacio del aula se está redefiniendo tanto en su estructura material como en sus formas de interacción.

Otro aspecto a considerar son los espacios de afinidad que surgen a partir de las redes sociales y las plataformas de juego. Estos permiten explorar temas de interés y adquirir competencias y desempeños avanzados, reconocen y valoran el conocimiento tácito que se construye entre todos durante el desarrollo de la tarea, a diferencia con la interacción en el aula tradicional donde el conocimiento tácito es poco valorado y reconocido.

Los responsables de la educación deben tomar en cuenta este escenario y actuar en consecuencia a través de una visión estratégica y proactiva que parta de la base de prever este escenario educativo, antes que seguir una estrategia adaptativa que reaccione a los nuevos requerimientos a medida que van apareciendo. En este sentido, el reconocer a las pedagogías emergentes como un conjunto de enfoques e ideas pedagógicas todavía no bien sistematizadas, generan en los docentes una cierta resistencia a la nueva cultura del aprendizaje [2]. Es por ello que las tecnologías emergentes tienen un rol particular en la concreción de estos cambios. Es imposible imaginar un sistema de enseñanza y aprendizaje sin una utilización competente e importante de las NTIC.

Por consiguiente, el gran desafío es formular experiencias educativas que permitan integrar la tecnología, para transformar la educación y proveer de algunos recursos que favorecen su implementación. La Realidad Virtual (RV) y la Realidad Aumentada (RA) son dos de las tecnologías que están revolucionando la educación al brindar la oportunidad de poder “vivir en primera persona” diferentes experiencias y llevar a cabo un aprendizaje mucho más atractivo e interactivo [4]. No obstante, a pesar de la gran cantidad de investigaciones realizadas durante las dos últimas décadas, existe la percepción generalizada de que las NTIC no han impactado tan favorablemente como deberían. Algunos autores

consideran que es posible que se esté produciendo un cambio de paradigma en la teoría de la educación y que una nueva epistemología esté surgiendo.

El Laboratorio de Investigación y Desarrollo en Computación Gráfica (LCG) tiene como uno de sus fines esenciales promover la transferencia tecnológica y la asistencia pedagógica y técnica al medio. El LCG realiza desarrollos en diversas áreas, que incluyen la Realidad Virtual, Realidad Aumentada, Interacción Natural Humano-Computadora, entre otras [17][18][19][5][28][6][27]. Por lo tanto, estudiar el impacto que las mismas provocan y acercar dicho conocimiento y tecnologías a las escuelas del medio posibilitaría llevar a cabo experiencias educativas innovadoras. En este contexto se definieron dos líneas de acción: Línea Docentes y Línea Estudiantes. El trabajo aquí presentado se centra en la línea de estudiantes para explorar algunas características presentes en el ecosistema educativo local e intentar responder a dos preguntas fundamentales, como punto de partida de investigación:

- ¿Cuánto conocen los alumnos de estas tecnologías?.
- ¿Cuánto es posible mostrar e introducir en el aula en una jornada escolar?.

Las siguientes secciones desarrollan la experiencia en escuelas locales a la ciudad de San Luis. La sección 2 introduce tres tecnologías emergentes inmersivas de actualidad. La sección 3 desarrolla la experiencia y los resultados obtenidos. La sección 4 brinda algunos conclusiones y lineamientos futuros.

## 2. Tecnologías Emergentes Inmersivas

Confucio dijo: “*Veo y olvido. Escucho y recuerdo. Hago y entiendo*”. En [12] Dale representó en un cono de experiencia, el aprendizaje logrado a través del uso de diferentes medios de transmitir la información. El cono describe desde su vértice la progresión de la experiencia de aprendizaje, donde las *experiencias directas con propósito* se encuentran en la base, proporcionando el mejor soporte para la comprensión. Más aún, según la teoría de la *cognición corporizada* [7], cuantos más sentidos estén involucrados con aquello que se desea aprender, más profundo, integrado, completo y significativo será el aprendizaje.

Actualmente, las nuevas tecnologías son cada vez más accesibles a las nuevas generaciones. Según estudios, las tecnologías emergentes se centran fundamentalmente en tres megatendencias: la *Inteligencia Artificial*, las *Experiencias Inmersivas* y las *Plataformas Digitales* [16]. En particular, las experiencias inmersivas permiten no sólo mejorar el aprendizaje sino también favorecer nuevas metodologías de evaluación de dicho aprendizaje. Del conjunto de tecnologías emergentes que permiten experiencias inmersivas la Realidad Virtual, la Realidad Aumentada y la Holografía se destacan por su popularidad y potencial. Ellas nutren de herramientas didácticas que facilitan los procesos de enseñanza y de aprendizaje a través de experiencias inmersivas.

### 2.1. La Realidad Virtual en Educación

La Realidad Virtual es la tecnología de inmersión por excelencia. Las aplicaciones de RV sumergen al usuario en un ambiente generado por computadora que simula

la realidad mediante el uso de dispositivos interactivos que envían y reciben información (gafas, cascos, guantes, trajes). Los sistemas de RV pueden variar enormemente en cuanto a los niveles de inmersión y la experiencia del usuario que ofrecen, en función de las características del sistema y el contexto de los entornos virtuales. La adición de información abstracta indirecta como texto, símbolos y multimedia directamente en la escena y en los objetos virtuales puede conducir a una comprensión más eficiente que la que se puede lograr en el mundo real.

Las experiencias de este tipo suelen caracterizarse por la ausencia de reflexión deliberada, ya que la acción surge directamente a partir de la percepción del mundo que tiene cada uno. Además, a menudo este aprendizaje se realiza de forma implícita, ya que no se es consciente del aprendizaje que se está produciendo. El gran aporte de la RV a la educación consiste en la utilización de elementos digitales que facilitan la interacción con situaciones que se caracterizan por ser de difícil acceso en contextos reales, generando así una experiencia memorable que impacte directamente en el aprendizaje. Su uso en el campo de la educación es muy notable especialmente en: Ciencias de la Computación [33], Ciencias Sociales [20], Ingeniería [35], Medicina [32], Matemáticas [23], y Psicología [25].

## 2.2. La Realidad Aumentada en Educación

La Realidad Aumentada supone la introducción de elementos virtuales en el mundo real. Se trata de generar objetos, seres, contextos, imágenes y textos virtuales, entre otros, por medio del ordenador, que puedan superponerse o “incluirse” en el mundo real. Es decir, el usuario visualiza a través de una o más cámaras el contexto real en el que se encuentra combinado con los elementos virtuales. El objetivo central de las aplicaciones de RA es que los elementos virtuales que se superponen en el mundo real proporcionen información adicional y relevante a la imagen final que está visualizando el usuario del sistema con el objetivo de aumentar su cognición.

La RA también se ha convertido en un importante foco de investigación en los últimos años [3]. Numerosas investigaciones sugieren que la RA refuerza el aprendizaje e incrementa la motivación del estudiante por aprender, al mismo tiempo que adquieren autonomía sobre el aprendizaje. Más aún, además de posibilitar contenidos didácticos que son inviables de otro modo, ayuda a que exista una continuidad en el hogar y ofrece interactividad, juego, experimentación, colaboración, entre otras. Asimismo se aprovechan las capacidades de los estudiantes en el manejo de la tecnología, convirtiéndolos en más pro-activos, participativos, creativos y autónomos.

Una de las razones más importantes por las que la tecnología de RA se utiliza tan ampliamente es que ya no requiere un hardware caro ni equipos sofisticados. Actualmente, su uso es muy simple y puede ser utilizada con computadoras o dispositivos móviles. Hoy en día se aplica en todos los niveles de la enseñanza, desde nivel inicial [10] hasta el nivel universitario [14], mostrando su mayor efectividad en experimentos de laboratorio [11], matemáticas [31] y geometría [21], geografía y ecología [9], aprendizaje de idiomas [22], el arte visual y la apreciación de la pintura [8], entre otras áreas.

### 2.3. La Holografía en Educación

La holografía es una técnica de captura de la realidad que al proyectarse lumínicamente sobre un objeto transparente permite generar un efecto tridimensional. La holografía trata de ser una fotografía que aporta una información más completa, al otorgar datos tridimensionales sobre el objeto retratado. A medida que el espectador se mueve alrededor de la figura holográfica es capaz de percibir sus diferentes perspectivas, así como su profundidad, aunque se debe tener en cuenta que la visualización de un holograma de forma adecuada siempre se da desde un punto determinado.

La evolución de las computadoras ha permitido la democratización de tecnologías como estas en cualquier ámbito y últimamente en el aula. El uso de hologramas ha resultado ser un agente motivante que actúa como factor extrínseco que ayuda a potenciar la motivación intrínseca de los alumnos y, por lo tanto, a que estén más predispuestos hacia el aprendizaje, debido a que el alumno tiene la sensación de que el elemento que está estudiando se encuentra realmente presente, que no se trata de una representación. Es así como esta tecnología ha sido utilizada para la exposición de contenido relacionado a las Ciencias Naturales [26,13], enseñanza de la Geometría [24] y Física[34] tanto en nivel secundario como universitario, representación del patrimonio histórico[30], entre otros.

## 3. Experiencia con Hologramas

### 3.1. Confección de la experiencia

#### Contexto

En este estudio exploratorio se trabajó con estudiantes de la Escuela N° 5 Bartolomé Mitre y el Centro Educativo N° 5 Senador Alfredo Bertín. La Escuela N° 5 Bartolomé Mitre se encuentra ubicada en la ciudad capital de San Luis, es un establecimiento con alta demanda por su ubicación céntrica estratégica. El Centro Educativo N° 5 Senador Alfredo Bertín se encuentra ubicado la localidad de El Trapiche, a 45 km de la ciudad de San Luis. Asisten alumnos de distintas localidades y parajes cercanos. Es la única institución con nivel secundario de la región. Posee una matrícula significativamente menor a la Escuela Bartolomé Mitre.

Ambas son escuelas públicas y mixtas y la población estudiantil tiene características similares, como así también han sido beneficiadas con equipamientos en el marco de diversos programas nacionales y provinciales, el cual han mantenido en forma sostenida lo que garantiza la disponibilidad de computadoras en la escuela.

#### Población

En la Escuela N° 5 Bartolomé Mitre se trabajó con 5° año B (turno tarde) de 28 alumnos (con edades entre 16 y 18 años) y en el Centro Educativo N° 5 Senador Alfredo Bertín se trabajó con 4° año A de 12 alumnos (con edades entre 15 y 18 años).

### Procedimiento

En cada escuela se destinó una jornada escolar completa para desarrollar la actividad planificada.

### Metodología

Se optó por organizar la actividad en modalidad taller por ser un ámbito de aprendizaje que favorece la integración de la vivencia con la reflexión y la conceptualización como síntesis del pensar, del sentir y el hacer, y construye el conocimiento a través de un intercambio social muy rico y del procesamiento individual. A modo innovativo para los estudiantes, del conjunto de tecnologías mencionadas en la sección 2, en esta oportunidad se seleccionó a la Holografía, como una estrategia para despertar la motivación de los estudiantes ante algo por ellos desconocido.

### Estructura del taller

El taller se conformó de los siguientes momentos:

1. *Presentación*  
Se realizó la presentación de los integrantes del equipo de trabajo y del espacio del “Laboratorio de Investigación y Desarrollo en Computación Gráfica (LCG)”.
2. *Indagación previa*  
Se solicitó a los estudiantes que respondieran una breve encuesta para indagar sobre los conceptos previos que poseen de determinadas tecnologías emergentes.
3. *Fundamentación del holograma*  
Se comenzó con la proyección de parte de un recital de un cantante haciendo dúo con el holograma de un músico ya fallecido. La mayoría de los alumnos reconocían a ambos músicos y sabían de los hologramas pero desconocían cómo se genera y de lo que se necesita para lograrlo (pirámide + dispositivo de visualización + imagen).
4. *Construcción de la pirámide*  
Con materiales de fácil acceso y de bajo costo, cada uno de los estudiantes construyó una pirámide en escala. Se les permitió trabajar en grupos de dos integrantes pero algunos optaron por hacerlo en forma individual. Se trabajó con un texto instructivo y en las directivas se incluyó conceptos de matemática.
5. *Construcción de la imagen*  
Con pocas directivas se les solicitó que construyeran la imagen que luego sería proyectada con ayuda de la pirámide. En esta instancia la didáctica propuesta se basó en promover el aprendizaje por indagación. Los alumnos disponían de un catálogo de diferentes personajes populares y debían seleccionar las diferentes vistas que conformarían al objeto 3D a ser proyectado. Utilizaron las netbooks que poseen las escuelas. El tamaño de las mismas les facilitaba el recostarlas y utilizar la pantalla como dispositivo de visualización en posición horizontal.

#### 6. *Conceptualización*

Una vez que los estudiantes comprobaron que la imagen se proyectaba en la pirámide, se construyeron los conceptos de reflexión y refracción de la luz en forma colectiva.

#### 7. *Encuesta final*

Como cierre de la jornada los alumnos respondieron una encuesta anónima donde evaluaron la experiencia y otros aspectos de interés.

### 3.2. Resultados de la Experiencia

En el proceso de búsqueda y obtención de datos descriptivos se realizaron dos encuestas para registrar las características de los grupos, la disponibilidad de tecnología y determinar el impacto que pudo haber provocado la actividad realizada durante la jornada.

- En ambas escuelas el taller tuvo una duración de 2 módulos de 80 minutos cada uno con un recreo.
- En el Centro Educativo N° 5 “Senador Alfredo Bertín” participó la totalidad los estudiantes de la población. Las actividades se desarrollaron en la Sala de Computación de la escuela.
- En la Escuela N° 5 “Bartolomé Mitre” participó la totalidad de los estudiantes de la población. Las actividades se desarrollaron en el aula. Las computadoras fueron llevadas al aula.
- La cantidad de alumnos del Centro Educativo N° 5 se refiere aproximadamente a la mitad de alumnos de la Escuela Mitre.
- Ambos cursos contestaron de manera similar sobre los conocimientos previos. Es decir, en relación a la cantidad de estudiantes en cada curso, aproximadamente la misma proporción declaró no tener conocimiento previos en RV. La misma relación se presentó a la hora de responder si tenían conocimientos previos en RA. Llamativamente si bien todos reconocían al holograma, desconocían el concepto y la técnica subyacente que la genera.
- En referencia a si disponen de acceso a Internet en sus hogares, solamente un 3 % responde no poseer conexión, un 5 % no responde a la pregunta, por lo tanto el 92 % de los estudiantes de ambos cursos dispondrían de acceso a Internet.
- Sobre si disponían de celular, el 77 % responde poseer uno con tecnología smartphone, el 15 % responde tener uno sin esa tecnología y el 8 % no responde a la pregunta.
- De la población total de estudiantes, el 30 % son mujeres (Ver Fig. 1).
- En el Centro Educativo N° 5 el 100 % de la población se sintió motivada por la experiencia, mientras que la motivación en la Escuela Mitre fue del 95 % (Ver Fig. 1).
- Ningún alumno de la Escuela Mitre encontró dificultades en la realización del taller, mientras que el 10 % de los estudiantes del Centro Educativo expresaron haber tenido dificultades (Ver Fig. 1).

- El 80 % de los alumnos de la Escuela Mitre y el 60 % del Centro Educativo repetirían la experiencia sin modificación lo cual expresa el atractivo de la experiencia (Ver Fig. 1).
- Aproximadamente el 90 % de la población de cada escuela respondió que le gustaría que se repitiera la experiencia (Ver Fig. 1).
- La totalidad de la población de cada escuela considera que el utilizar la computadora en la escuela no les complica el aprendizaje (Ver Fig. 2).

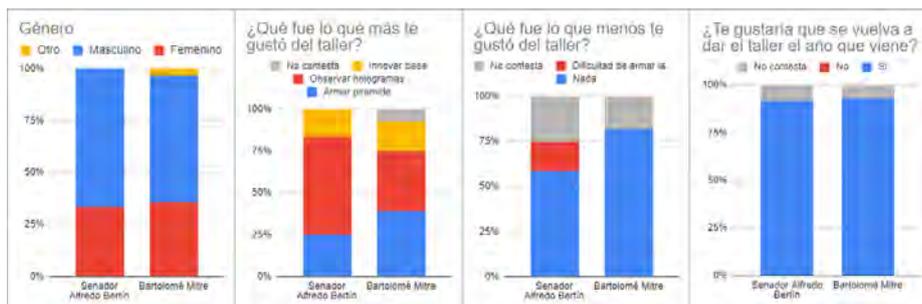


Figura 1: Características de la población y valoración de la experiencia.

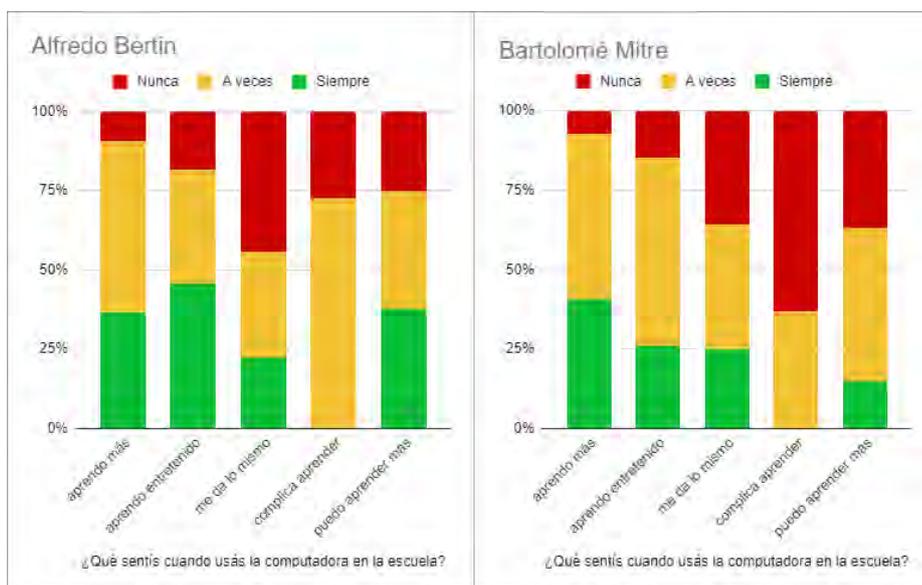


Figura 2: Valoración de la incorporación de la Tecnología.

## 4. Conclusiones

Las políticas educativas que se vienen llevando adelante se han centrado en proveer acceso a las NTIC en las escuelas, pero no han considerado los cambios en las formas, aproximaciones y prácticas de los docentes como así tampoco la calidad de uso de la tecnología que hacen los estudiantes ni las habilidades que ponen en juego.

En el presente trabajo se ha realizado un análisis de la experiencia llevada a cabo en dos escuelas de la provincia de San Luis (Argentina) donde se han presentado conceptos teóricos fundamentales para la aplicación e implementación en la práctica docente diaria de la RV, RA y Holografía. En particular la experiencia se ha focalizado en la realización de un experimento de holografía por ser el más desconocido por los alumnos y que no requería mayor disponibilidad de equipamiento especializado y por su facilidad de implementación. El objetivo de la investigación era poder determinar los conocimientos previos de los estudiantes en relación con las tecnologías mencionadas y poder valorar el esfuerzo requerido para la realización de una experiencia de este tipo en una jornada. Como consecuencia de las actividades se podría valorar el grado de motivación que provoca en los alumnos la metodología de la actividad y la tecnología a utilizar en la misma.

En este estudio, el trabajo planificado en sus orígenes, tanto con los docentes como con los estudiantes, consistía de dos etapas: la primera etapa se desarrollaría en cada una de las escuelas y la segunda se desarrollaría en el Laboratorio LCG. Como consecuencia del aislamiento social, preventivo y obligatorio por COVID-19 las actividades planificadas en el laboratorio debieron ser postpuestas.

En la fundamentación de la metodología propuesta para desarrollar este tipo de actividad, se planteó que, por su concepción, una de las principales características de esta actividad es el empleo de métodos participativos y de elaboración conjunta que garanticen una participación activa de los estudiantes y la creación de un ambiente de trabajo colectivo. En consecuencia, fue interesante trabajar con una escuela de la ciudad y con otra de zona desfavorable y, así, estudiar el impacto que provocaba en cada grupo.

De los resultados de las encuestas se puede concluir que las experiencias resultaron altamente motivadoras logrando un trabajo en colaborativo y en equipo captando la atención de aquellos alumnos que en general no demuestran interés ni participación en actividades cotidianas.

Como estímulo a futuras acciones queda por resolver las actividades que permitan medir el impacto de la inclusión de las tecnologías emergentes inmersivas en la práctica docente.

## Referencias

1. Transformación estructural, cuarta revolución industrial y desigualdad: desafíos para las políticas de ciencia, tecnología e innovación, 2019.
2. J. Adell and L. Castañeda. *Tecnologías emergentes, pedagogías emergentes?*, pages 13–32. 01 2012.

3. M. Akcayr and G. Akcayr. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20:1 – 11, 2017.
4. A. P. Altomari. *Realidad virtual y realidad aumentada en la educación, una instantánea nacional e internacional*. ECONOMA CREATIVA, 2017.
5. Y. Alvarado, N. Jofré, R. Guerrero, and G. Rodríguez. *Interacción Gestual para Ambientes de Realidad Virtual mediante un Guante de Datos*. XXV Jornadas de Jóvenes Investigadores, Universidad Nacional de Itapúa, Encarnación, Paraguay, 2017.
6. Y. Alvarado, N. Jofré, M. Rosas, and R. Guerrero. *Aplicaciones de Realidad Virtual y Realidad Aumentada como soporte a la enseñanza del Dibujo Técnico*. Libro de actas TE&ET - 2019, San Luis, Argentina, 2019.
7. J. Canal, A. Salazar, M. Garavito, and E. Quesada. *Cognición corporizada y embodiment*. Cognición, moral y desarrollo psicológico. Corporacion Universitaria Minuto de Dios, 2011.
8. K. Chang, C.-T. Chang, H.-T. Hou, Y.-T. Sung, H.-L. Chao, and C.-M. Lee. Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. *Computers & Education*, 71:185–197, 02 2014.
9. C.-p. Chen and C.-H. Wang. Employing augmented-reality-embedded instruction to disperse the imparities of individual differences in earth science learning. *Journal of Science Education and Technology*, 24, 05 2015.
10. T. H. Chiang, S. J. Yang, and G.-J. Hwang. Students' online interactive patterns in augmented reality-based inquiry activities. *Computers & Education*, 78:97–108, 2014.
11. J. Chiu, C. DeJaegher, and J. Chao. The effects of augmented virtual science laboratories on middle school students' understanding of gas properties. *Computers & Education*, 85, 07 2015.
12. E. Dale. *Audio-visual Methods in Teaching*. Dryden Press, 1969.
13. R. V. F. Elizabeth. Diseño de un modelo holográfico en el ciclo del agua para atraer la atención y mejorar el aprendizaje de los alumnos de tercer año de educación general básica en la unidad educativa pichincha en el año lectivo 2017-2018, 2019.
14. J. Ferrer-Torregrosa, J. Torralba, M. Jimenez, S. García, and J. Barcia. Arbook: Development and assessment of a tool based on augmented reality for anatomy. *Journal of Science Education and Technology*, 24(1):119–124, 2015.
15. S. Gvirtz, C. Necuzzi, V. Tatti, and A. M. de Educación. *Educación y tecnologías* :. Anses,, Buenos Aires :, 2011.
16. S. Hansson. *The Ethics of Technology: Methods and Approaches*. Philosophy, Technology and Society. Rowman & Littlefield International, 2017.
17. N. Jofre, G. Rodríguez, Y. Alvarado, J. Fernández, and R. Guerrero. *Natural Interface for Recycling Activity*. 6th International Symposium on Innovation and Technology (ISIT2015), Mar del Plata, Argentina, 2015.
18. N. Jofre, G. Rodríguez, Y. Alvarado, J. Fernández, and R. Guerrero. *La comunicación no verbal interfaz usando un guante de datos*. IEEE Congreso Argentino de Ciencias de la Informática y Desarrollos de investigación (Cacidi), Ciudad Autónoma de Buenos Aires, Argentina, 2016.
19. N. Jofré, G. Rodríguez, Y. Alvarado, J. Fernández, and R. Guerrero. *El uso de la Realidad Virtual Inmersiva en terapias motrices*. XIX Workshop de Investigadores en Ciencias de la Computación, Instituto Tecnológico, Buenos Aires (ITBA), Argentina, 2017.

20. J. Legault, J. Zhao, Y.-A. Chi, W. Chen, A. Klippel, and P. Li. Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages*, 4(1):13, 2019.
21. H.-C. Lin, M.-C. Chen, and C.-K. Chang. Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. *Interactive Learning Environments*, 23:1–12, 12 2013.
22. P. Liu and M. Tsai. Using augmented-reality-based mobile learning material in english composition: An exploratory case study. *British Journal of Educational Technology*, 44, 01 2013.
23. K. L. P. Nathal, M. E. P. Nathal, H. B. del Toro, M. A. G. Solano, and J. C. M. Sandoval. The immersive virtual reality: A study in three-dimensional euclidean space. *American Journal of Educational Research*, 6(3):170–174, 2018.
24. L. O. Palma, C. Jordan-Lluch, and Á. A. Magreñán. Uso del holograma como herramienta para trabajar contenidos de geometría en educación secundaria. *Pensamiento Matemático*, 8:91–100, 2018.
25. T. D. Parsons, G. Riva, S. Parsons, F. Mantovani, N. Newbutt, L. Lin, E. Venturini, and T. Hall. Virtual reality in pediatric psychology. *Pediatrics*, 140(Supplement 2):S86–S91, 2017.
26. A. M. C. Regalado. Impacto del uso de medios virtuales sobre el aprendizaje de la anatomía cardíaca en estudiantes de tercer semestre del programa de medicina en la universidad de ciencias aplicadas y ambientales, 2016.
27. G. Rodríguez, N. Jofré, Y. Alvarado, M. Zúñiga, M. Rosas, J. Fernández, and R. Guerrero. *Sistemas motivadores del Sujeto basados en Realidad Virtual y Realidad Aumentada*. XXI Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional de San Juan, San Juan, Argentina, 2019.
28. G. Rodríguez, Y. Jofré, N. Alvarado, and R. Fernández, J. and Guerrero. *Gestural Interaction for Virtual Reality Environments through Data Gloves*. Advances in Science, Technology and Engineering Systems Journal, ASTES, 2017.
29. M. Roussou. Learning by doing and learning through play: An exploration of interactivity in virtual environments for children. *Computers in Entertainment*, 2:10, 01 2004.
30. R. Serra Toledo and A. Moreno Yeras. La holografíA en la divulgaciÓy conservaciÓdel patrimonio cultural cubano. *Arquitectura y Urbanismo*, 34:86 – 89, 12 2013.
31. P. Sommerauer and O. Müller. Augmented reality in informal learning environments: A field experiment in a mathematics exhibition. *Computers & Education*, 79:59 – 68, 2014.
32. K. Stepan, J. Zeiger, S. Hanchuk, A. Del Signore, R. Shrivastava, S. Govindaraj, and A. Iloreta. Immersive virtual reality as a teaching tool for neuroanatomy. In *International forum of allergy & rhinology*, volume 7, pages 1006–1013. Wiley Online Library, 2017.
33. J. Stigall and S. Sharma. Virtual reality instructional modules for introductory programming courses. In *2017 IEEE Integrated STEM Education Conference (ISEC)*, pages 34–42. IEEE, 2017.
34. R. Toledo, G. Cruz, A. Zaldo, J. Lunazzi, and D. Magalhães. Fundamentación del holograma como un medio de enseñanza de la física. *Latin-American Journal of Physics Education*, 01 2008.
35. D. Vergara, M. Rubio, and M. Lorenzo. On the design of virtual reality learning environments in engineering. *Multimodal Technologies and Interaction*, 1(2):11, Jun 2017.

## **MEHI – Material Educativo Hipermedial Interactivo para un curso de Química General. Resultados de una experiencia áulica**

García Ricardo<sup>1</sup>, Gorga Gladys<sup>2</sup>, Bertone Rodolfo<sup>2</sup>

<sup>1</sup> Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

<sup>2</sup> Instituto de Investigación en Informática LIDI. Facultad de Informática – Universidad Nacional de La Plata

garciaricardo2010@yahoo.com.ar

{ggorga, pbertone}@lidi.info.unlp.edu.ar

### **Abstract.**

La incorporación e integración de material multimedia en escenarios educativos, abre la posibilidad de cambio y renovación en el funcionamiento del aula, en los procesos didácticos involucrados, en la infraestructura necesaria, en la actividad del docente y el rol del alumno. El uso de estos recursos incide sobre los procesos cognitivos, incrementando la capacidad para codificar, almacenar y procesar la información. En este marco, se muestra relevante el diseño y desarrollo de materiales educativos digitales para integrarlos a diferentes contextos y paradigmas. En el trabajo se revisan aspectos vinculados con el diseño e implementación de un material educativo hipermedial interactivo (MEHI) orientado al aprendizaje de contenidos de Química General. Además, se pone énfasis en la evaluación de resultados de una experiencia de aula desde el punto de vista de los docentes y alumnos participantes. Se inicia con la presentación de la metodología de trabajo llevada adelante y se describen las características del material educativo. Luego, se describen los instrumentos de recolección de datos utilizados para su evaluación y finalmente se presenta un análisis de los resultados obtenidos a partir de la experiencia realizada.

**Keywords:** material educativo hipermedial, Química General, instrumentos de recolección de datos, encuestas docentes, encuestas alumnos

## **1 Introducción**

La incorporación e integración de material multimedia en escenarios educativos, abre la posibilidad de cambio y renovación en el funcionamiento del aula, en los procesos didácticos, en la infraestructura necesaria, en la actividad del docente y el rol del alumno [1].

La adecuada complementación entre los conocimientos tecnológicos, disciplinares y didáctico-pedagógicos de los diseños, las tendencias y experiencias que permiten validar su efectividad, son consideraciones a tener en cuenta a la hora del diseño de esos materiales multimediales.

Es manifiesta la tendencia creciente al uso de materiales educativos digitales orientados a la enseñanza y el aprendizaje de las Ciencias en general y de la Química en particular, con el propósito de asistir al docente en el tratamiento de determinados contenidos [2].

Las investigaciones en el campo de la didáctica de las ciencias, muestran la utilidad de estos recursos para hacer frente a las dificultades de aprendizaje de contenidos abstractos difíciles de abordar o para el desarrollo de laboratorios virtuales y remotos como escenarios de sumo interés en el campo educativo. Se logra así superar las limitaciones físicas y de equipamiento, permitiendo mejorar las formas de razonamiento espontáneo, que actúan como obstáculos epistemológico y metodológico del pensamiento. De esta forma, se evita la fijación funcional, que lleva al aprendizaje memorístico que no favorece el análisis, la reflexión y el pensamiento creativo de una situación, hecho o problema y la reducción funcional que no permite correlacionar distintas variables o causas de un fenómeno [3] [4].

La aplicación en el ámbito educativo de recursos multimedia proporciona una gran riqueza de datos. Esto incluye no sólo texto, sino también otros medios como imágenes estáticas (fotografías, gráficos o ilustraciones, imágenes en movimiento (videos o animaciones) y audio (música o sonidos), dotando de mayor flexibilidad a la expresión de los contenidos desarrollados en forma multimedial [5]. Por otra parte, el hipertexto aporta una estructura que permite que los datos puedan presentarse y explorarse siguiendo distintas secuencias. De modo que la integración de hipertexto y multimedia, permite la accesibilidad a los contenidos multimediales, según las necesidades o intereses del usuario. El uso de un medio multidimensional más natural y amigable, no limitado a presentaciones lineales, permite al alumno elegir los hiperenlaces que desee en cada momento. Esta libertad de elección estimula la curiosidad, la posibilidad de administrar su propio proceso de aprendizaje, convirtiéndose en una herramienta que facilita el aprendizaje autónomo [6] [7].

En este contexto, se llevó a cabo un trabajo realizado en dos etapas. La primera etapa, contempla una revisión bibliográfica de antecedentes y experiencias de uso de materiales hipermediales para la enseñanza y aprendizaje de la Química y de las herramientas de autor para el diseño de materiales digitales. La segunda etapa se enfocó en el diseño e implementación de un material educativo digital para el aprendizaje de contenidos de Química General.

Para la producción del material se tuvieron en cuenta los lineamientos de la Teoría Cognitiva del Aprendizaje Multimedia (TCAM), la Teoría Situativa y de la Teoría de la Carga Cognitiva (TCC) que conforman el marco conceptual para comprender y mejorar el aprendizaje apoyado en la tecnología informática y se consideraron los Principios para el diseño de materiales digitales propuestos por como justificación para el diseño de materiales digitales de enseñanza [8] [9].

Para la evaluación de la experiencia, se realizó un estudio de caso en el contexto de una investigación descriptiva transversal, con una metodología cuali-cuantitativa que involucra la implementación de diversos instrumentos de recolección de datos como encuestas a estudiantes, encuestas docentes, observación participante y análisis de producciones escritas de los alumnos en las evaluaciones parciales, entre otros.

El presente artículo se organiza en 6 secciones. En la sección 2 se describe la metodología de trabajo desarrollada y en la sección 3 se detallan las principales características del material educativo y se muestran a modo de ejemplo algunas de las

actividades interactivas incluidas en el material. En la sección 4 se describen los instrumentos utilizados para la recolección de datos y luego en la sección 5 se presentan los resultados obtenidos de la experiencia realizada en el aula. Finalmente, en la sección 6 se presentan las conclusiones del trabajo y se describen las líneas de trabajo futuro.

## 2 Metodología de trabajo

Para la producción de un material educativo hipermedial, se planteó la siguiente secuencia metodológica:

**1-** Revisión bibliográfica de propuestas realizadas por otros autores para el desarrollo de estas competencias en diversas experiencias y diferentes contextos de uso. Esto permite aprovechar las fortalezas encontradas, corregir sus debilidades e incorporar elementos que resulten relevantes y que no hayan sido considerados en las propuestas revisadas.

**2-** Evaluación de los modelos del **diseño instruccional** más útiles para una planificación sistemática que incluye la valoración de necesidades, desarrollo, evaluación, implementación y mantenimiento de materiales para el aprendizaje de contenidos. Se utiliza el modelo **de ADDIE** designado por el acrónimo de los términos Análisis, Diseño, Desarrollo, Implementación y Evaluación que constituyen las 5 fases del modelo propuesto y puede aplicarse al diseño de cursos o materiales didácticos [10].

**3-** Evaluación de las Herramientas de Autor de uso libre disponibles en la WEB (Ardora, Constructor, Cuadernia, EdiLim, Exelerning, JClic, Malted y Educaplay) que permiten crear materiales didácticos atractivos, como actividades interactivas, con incorporación de videos, audios, fotografías, animaciones, etc. sin ser un especialista informático. Se evaluaron por la Calidad, Eficacia, Eficiencia, Pertinencia e Impacto de los materiales producidos a partir de ellas. Se optó por Educaplay, que es un proyecto para la creación de actividades interactivas, que permite elaborar materiales de forma online y permanecen en la plataforma para ser utilizados y compartidos. Las actividades generadas con Educaplay se pueden exportar como paquetes SCORM para integrarlas en plataformas o aulas virtuales y permite realizar un seguimiento de las calificaciones de las actividades desarrolladas.

**4-** Confeción del guión multimedia. Se estructuran y ajustan las actividades de aprendizaje y de evaluación.

**5-** Producción del material hipermedial interactivo, del sistema de gestión de contenidos y planificación de las evaluaciones.

En el marco de este trabajo se presenta una descripción del material hipermedial y los instrumentos de recolección de datos para la evaluación de la experiencia áulica.

## 3 Descripción del material hipermedial

El material educativo hipermedial interactivo (MEHI) está estructurado en 4 capítulos y 6 secciones en las cuales se desarrollan un conjunto de estrategias para

ayudar al alumno en la comprensión de los temas relacionados con los temas de cada capítulo. En Fig1 se presenta la portada del material, donde puede verse en el recuadro de la izquierda arriba, los cuatro capítulos que componen el material y en la parte superior, con diseño en pestañas, las seis secciones en las que se desarrollan contenidos de cada capítulos.



Fig1. Portada del MEHI

Las secciones en las que se desarrolla cada uno de los capítulos son las siguientes: 1. Antes de empezar, 2. Contenidos, 3. Ejercicios, 4. Autoevaluación, 5. Para enviar al tutor y 6. Para saber más.

En la sección **Antes de empezar** se presentan los contenidos a desarrollar, en **Contenidos** y **Para saber más** se plantean con distintos grados de profundidad los principales conceptos teóricos que trata el capítulo. En las secciones **Ejercicios** y **Autoevaluación** se presentan actividades para que el alumno aplique los conceptos presentados. El material presenta 46 actividades interactivas: 11 actividades de completar texto, 6 de video quiz, 5 de relacionar columnas, 10 de mapa interactivo, 7 test y 6 presentaciones. En la sección **Para enviar al tutor** se proponen que integran los temas incluidos en el capítulo. Estas evaluaciones una vez resueltas deben ser enviadas al tutor.

El desarrollo completo del material educativo MEHI se encuentra disponible en [11]

A modo de ejemplo, en la Fig 2 se muestra uno de los materiales interactivos. En este caso, una actividad de completar destinada a fijar contenidos principales. El alumno debe elegir de una lista, las palabras faltantes de un texto convenientemente seleccionado. En estas actividades, el sistema hace una devolución, indicando errores y respuestas correctas. En la opción **Comprobar**, se muestran los errores, se pueden hacer cambios para corregir los errores y se puede asignar puntaje.

En la Fig 3 se muestra otra de las actividad interactivas denominada **Videos quiz**. Se trata de una plantilla que permite transformar un video, de elemento expositivo unidireccional a material interactivo. A partir de uno o más videos o parte de ellos se genera un nuevo video, en el que se pueden intercalar preguntas durante su desarrollo, que permiten mantener o focalizar la atención en un determinado aspecto, o verificar si el alumno tomó en consideración puntos importantes de la exposición. Las preguntas que se introducen en el video pueden tener una o varias opciones correctas. También se pueden incluir imágenes y feedback.

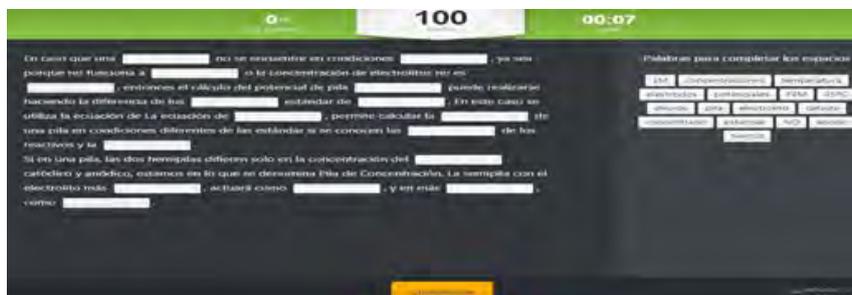


Fig 2. Ejercicio de completar texto



Fig3. Videoquiz utilizado en la sección Ejercicios del tema Aplicaciones tecnológicas

La plantilla de creación de mapas interactivos resulta muy útil en esta propuesta porque permite definir sobre una imagen (fotografía, mapa o esquema) una serie de puntos que deben luego ser identificados con su nombre. Se ha utilizado este recurso para fijar contenidos en el tema Pilas y Celdas Electrolíticas. La actividad consiste en asignar los nombres de los componentes de una pila sobre un esquema. Al finalizar se muestran aciertos y errores y asigna puntaje (Fig 4).



Fig4. Mapa interactivo de la sección Autoevaluación del tema Esquema y Componentes de una pila.

En relación al sistema de gestión de contenidos, es un gestor de contenidos estático, HTML, Javascript y JSON; carente de base de datos y lenguaje de programación compilado, no necesita procesamiento especial en servidores. Puede ser alojado en cualquier hosting gratuito, recurso compartido, CDN (content delivery network), o incluso en cualquier drive virtual como Google Drive o Dropbox. Es

interpretado por cualquier navegador web, sin necesidad de servidores especiales y costosos. Permite la creación, edición y gestión de contenido digital multimedia a partir de un administrador web. Dispone de una plantilla con un formato predefinido que permite una carga simple del material por parte de un usuario con conocimientos básicos de Informática. El acceso al gestor se realiza generalmente a través del navegador web.

## 4 Evaluación

El diseño metodológico que se elaboró para la evaluación del MEHI propone realizar un estudio de caso en el contexto de una investigación descriptiva transversal. Consiste en la realización de una experiencia áulica, en la cual los alumnos de una comisión de la asignatura Química General e Inorgánica de UNNOBA, utilizan el material hipermedial MEHI en el contexto de una modalidad Flipped Classroom o de aula invertida. En este caso, los alumnos reciben el material hipermedial previo a la clase, lo utilizan para la comprensión de los contenidos y luego en clase, realizan actividades en las cuales se involucran procesos cognitivos de mayor complejidad. En las experiencias de Flipped Classroom, es posible consolidar durante la clase presencial, con la asistencia del docente, el aprendizaje previo realizado por el alumno [12].

Para realizar la experiencia se seleccionó la comisión 2 de la asignatura Química General e Inorgánica (2do cuatrimestre 2019) que cursan alumnos de las carreras de Licenciatura en Genética y las Ingenierías Industrial, Mecánica y Alimentos. La comisión elegida, está integrada por 30 alumnos y por otro lado, se incluyen en el estudio otras dos comisiones, que se utilizarán como grupo control integradas por igual cantidad de alumnos.

En general, la actitud de los alumnos ante la propuesta de uso del MEHI (previo a la clase) despertó cierta curiosidad frente a la innovación y se mostraron proactivos e interesados en participar de la propuesta. Manifestaron una adhesión mayoritaria al uso del material digital. Ninguno de los participantes manifestó inconvenientes en el acceso al material digital previo a la clase. Algunos alumnos consultaron si la herramienta digital reemplazaría a la bibliografía propuesta por la cátedra y si esta actividad se evaluaría en parciales o finales. Estas consultas se reflejaron en las encuestas.

Se diseñaron los instrumentos para la recolección de datos de la evaluación del MEHI:

**A) Evaluación de las fases de producción:** Se elaboró un cuestionario para la evaluación de cada fase de producción del material, para ser respondido por los integrantes del equipo de producción.

**B) Encuesta a docentes:** Se decidió solicitar a docentes expertos en los contenidos curriculares abordados, su opinión sobre los aspectos pedagógicos, funcionales, estéticos, técnicos del MEHI.

**C) Encuesta a alumnos:** Se relevó la opinión de los alumnos luego de finalizada la experiencia áulica, sobre aspectos funcionales, estéticos y de usabilidad del MEHI, a través de una encuesta elaborada para tal fin.

**D) Evaluaciones con devolución al tutor:** Los alumnos realizaron las evaluaciones previstas en el MEHI y enviaron la evaluación con devolución al tutor a una dirección de correo indicada en el material.

**E) Examen parcial:** Se analizaron las producciones de los alumnos en el ejercicio correspondiente a la temática abordada por el MEHI en el segundo examen parcial de la asignatura Química General e Inorgánica. Se realizó un análisis comparativo de los resultados de los parciales de los alumnos de los grupos involucrados en la experiencia.

**F) Observación participante:** Se realizó un registro de los hechos significativos sucedidos durante la experiencia áulica, teniendo en cuenta las preguntas o intervenciones que realizan los alumnos, o las situaciones dentro de los grupos cuando se disponen a discutir y resolver alguna de las actividades propuestas.

## 5 Resultados

### 5.1 Evaluación de las fases de producción

Se analizó la actividad realizada en cada una de las fases de producción del material, basada en el modelo de ADDIE, dando respuesta al cuestionario de evaluación de resultados planteados previamente para cada fase. Se concluye que los datos recogidos para realizar la propuesta fueron relevantes, se dispusieron de los recursos necesarios, que coincide con el proyecto de la institución donde se va aplicar, que el análisis de la población a la que va destinada fue adecuado, que los contenidos abordados son precisos actuales y contextualizados, las estrategias de aprendizaje implementadas coinciden con lo que se había propuesto, que el material es amigable y motivador, los instrumentos de evaluación son adecuados y las devoluciones y el trata tratamiento del error adecuado.

### 5.2- Encuesta a docentes

La encuesta, fue remitida a siete docentes que habían aceptado su participación en la evaluación del MEHI. La encuesta utilizada tiene preguntas abiertas y cerradas, como puede verse en el esquema (Fig 5).



Fig5. Esquema de la encuesta a docentes

En las preguntas cerradas se utilizó un cuestionario tipo Lickert con las opciones: Muy bueno/Bueno/Regular/Malo. Para realizar una interpretación cuantitativa de los

resultados de las preguntas cerradas, se asignó puntajes a las respuestas a saber: Muy bueno 4/Bueno 3/Regular 2/Malo 1.

En el bloque de Aspectos Técnicos y Estéticos (11 preguntas), a partir del puntaje asignado a cada opción de respuesta posible y teniendo en cuenta que se recibieron 7 encuestas, el puntaje máximo posible para este bloque es de 308 puntos. Se obtuvieron 283 puntos que corresponde al 92% de los puntos posibles ( $283/308 = 92\%$ ).

En el bloque de Valoración de Aspectos didácticos, la encuesta a docentes incluye dos partes, la primera (12 preguntas) orientada a que el evaluador detecte habilidades cognitivas puestas en juego en la gestión del material, para lo cual a manera de encuesta cerrada, se listan una serie de habilidades que debe marcar si la considera presente. La segunda parte (13 preguntas) está destinada a evaluar los recursos didácticos utilizados en el MEHI, son preguntas con cuatro opciones de respuesta: Muy bueno/Bueno/Regular/Malo. Para tabular esta parte de la encuesta se asignó puntaje a las respuestas con el mismo criterio utilizado en el bloque anterior (Muy bueno 4/Bueno 3/Regular 2/Malo 1). Teniendo en cuenta que fueron siete encuestas el máximo puntaje a obtener de la sumatoria de todas ellas es 364. Se obtuvieron 286 puntos que corresponden al 79% ( $286/364 = 79\%$ )

En relación a las preguntas abiertas, en todos los casos los encuestados responden que fue una experiencia positiva con distintas justificaciones, concluyen que el material facilita el aprendizaje, que favorece tanto alumnos con mayor actitud como a los que tienen dificultades, afirman que lo utilizarían y lo extenderían a otras unidades temáticas

### 5.3 Encuesta a alumnos

La encuesta fue contestada por 27 alumnos del total de los 30 seleccionados para la experiencia. La encuesta tiene preguntas abiertas y cerradas (Fig. 6).



Fig 6. Estructura de la encuesta a alumnos

En lo que respecta a las preguntas cerradas, en el bloque de Aspectos Técnicos y Estéticos se incluyeron 6 preguntas que teniendo en cuenta las 27 encuestas y el puntaje asignado a cada respuesta, su puntaje máximo posible de obtener es de 648 puntos. Se obtuvieron 540 puntos que corresponde al 88% de los puntos posibles ( $570/648 = 88\%$ ). En el bloque Aspectos didácticos se incluyeron 13 preguntas. Tomando el mismo criterio que en el bloque anterior para cuantificar las respuestas, se encuentra que el máximo de puntos posible de obtener es de 1.404 puntos. Se obtuvieron 1.094, que representa el 78 % ( $1.093/1.404=78\%$ ). Si bien el porcentaje es

menor que en el bloque anterior (88%), también debe considerarse que el número de preguntas realizadas es superior (6 en lugar de 13).

En este caso las preguntas abiertas fueron referidas a inconvenientes durante el uso del software y aspectos a destacar. Los alumnos coinciden que el material presentado fue facilitador y motivador del aprendizaje y también orientador en el estudio de los temas involucrados.

#### **5.4 Evaluaciones con devolución al tutor**

El MEHI consta de cuatro capítulos, para cada capítulo se diseñó una prueba de evaluación con devolución al tutor que incluye los temas tratados en el capítulo. El porcentaje de respuestas correctas a los cuatro cuestionarios osciló entre 65% y 95% y se observa que estos porcentajes guardan relación con la complejidad de los contenidos evaluados. El número de evaluaciones remitida por los alumnos para cada capítulo, varió entre 20 y 27 evaluaciones.

#### **5.5 Examen parcial**

Se evaluó la producción de los 27 alumnos de la comisión 2, que participaron de la experiencia. Como grupo control se eligieron al azar 27 parciales de alumnos de las comisiones 1 y 3, de la misma cohorte, que rindieron la misma evaluación. Se analizaron las producciones de los alumnos de ambos grupos en relación a dos ejercicios del examen y se compararon los resultados obtenidos (Fig 7y Fig 8).

La suma de puntos expresado como porcentaje respecto del puntaje máximo posible para el ejercicio A, reflejó que la comisión 2 obtuvo 41,1% mientras que la comisión 1, el 15,5% y la comisión 3 el 15,9%.

En tanto que, la suma de puntos expresado como porcentaje respecto del puntaje máximo posible para el ejercicio B, reflejó que la comisión 2 obtuvo 30% mientras que la comisión 1, el 21,1 % y la comisión 3 el 16,3%.

En ambas producciones se observan resultados favorables para el grupo que realizó la experiencia con el material propuesto.

#### **5.6 -Observación participante**

Otra fuente de recolección de datos, fue la observación participante. El objetivo de esta estrategia fue registrar los eventos producidos durante la experiencia áulica, para evaluar las actitudes de los alumnos frente a la nueva propuesta didáctica en relación a la clase tradicional. La mayoría de los alumnos se integraron a la propuesta. Algunos se mantenían menos activos, conducta que también ocurre en la clase tradicional, posiblemente por propia determinación o por menos actividad o comprensión del material digital.

## 6 Conclusiones y trabajos futuros

Se han presentado las principales decisiones de diseño de un material educativo hipermedial, los instrumentos de recolección de datos para su evaluación por parte de docentes y alumnos y los resultados obtenidos de su aplicación en el aula.

Se propuso realizar una experiencia en una modalidad de aula invertida y se articularon técnicas cualitativas y cuantitativas que permitieron dar cuenta de los alcances y limitaciones de la estrategia didáctica implementada y se utilizó grupos control para comparar los resultados.

Los datos obtenidos resultaron convergentes, y permitieron una mejor comprensión de lo sucedido en el aula. Esta experiencia puede resultar valiosa a la hora de pensar acerca de las distintas posibilidades para modificar la práctica cotidiana del aula, con la incorporación de materiales hipermediales.

Analizando los datos obtenidos tanto los estudiantes como los docentes, califican como positiva a la propuesta didáctica de uso del MEHI per se y en contexto real, considerándola una herramienta motivadora para la enseñanza y facilitadora de los aprendizajes.

Se considera de interés la profundización y actualización continuas del marco teórico de la enseñanza y aprendizaje de la Química con la incorporación de materiales digitales.

Asimismo se considera de interés explorar la inclusión de actividades experimentales a través de laboratorios virtuales de Química y simuladores virtuales educativos que aportan interesantes ventajas, desde lo motivacional, permitiendo realizar mayor número de prácticas ante las limitantes de tiempo, lugar físico, insumos, docentes, etc., la eliminación del riesgo de accidentes, la inclusión de prácticas experimentales en cursos de la modalidad a distancia, la disminución de los costos, y la gestión del aprendizaje en propuestas asincrónicas o en contextos colaborativos.

Al mismo tiempo, y para poder sacar un máximo provecho de estos materiales, es necesario continuar trabajando sobre la estrategia de su inserción en el aula, extendiéndolo a otros temas y unidades.

Por otra parte, es posible proponer una línea de investigación que aborde cuestiones vinculadas a la optimización de los diseños de cursos de características similares que requieran una mínima conectividad y puedan ser utilizados en dispositivos móviles.

## Referencias

1. Salinas, J. (2004). Innovación docente y uso de las TIC en la enseñanza universitaria. Revista de Universidad y Sociedad del Conocimiento (RUSC). UOC. Vol.1,nº 1.
2. Dell'Arciprete, R. (2013). Propuesta de Especialización en TIAE: Relevamiento de Materiales educativos digitales para la enseñanza y el aprendizaje de la Química. Facultad de Informática. UNLP.
3. Furió, C., & Calatayud, M.L. (1996). Difficulties with the geometry and polarity of molecules: beyond misconceptions. *Journal of Chemical Education*, 73, 37-41.

4. Furió, C., & Furió, C. (2000). Dificultades conceptuales y epistemológicas en el aprendizaje de los procesos químicos. *Educación Química* 11(3), 300-308
5. Jiménez, G. & Lliujós, A. (2006). Cooperación en entornos telemáticos y la enseñanza de la química. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 3(1); 115-133
6. Deslie, N., & Schwartzs, M. D. (1989). Collaborative Writing with Hypertext. *IEEE Transactions on Professional Communication*, 32 (3), pp. 183-188.
7. Whittington, C. D. (1996). MOLE: Computer-Supported Collaborative Learning. *Computers & Education*, 26(1/3), pp. 153-161.
8. Mayer, R. E. (2014). Introduction to multimedia learning. R. E. Mayer (Ed.) *The Cambridge handbook of multimedia learning* (2nd ed., pp. 1-25). Cambridge, U.K.; New York: Cambridge University Press.
9. Artino, A. R. (2008). Cognitive Load Theory and the Role of Learner Experience: An Abbreviated Re-view for Educational Practitioners. *Association for the Advancement of Computing In Education Journal. AACE Journal*, 16 (4), 425-439.
10. Moreno, F. & Santiago, R. (2003). Formación online. Guía para profesores universitarios. Universidad de la Rioja (España).
11. <https://ricardogarciaquimica.github.io/electroquimica/>
12. Bergmann, J. & Sams, A. (2012). *Flip your Classroom: Reach Every Student in Every Class Every day*. International Society for Technology in Education. SBN-10: 1564843157; ISBN-13: 978-1564843159.

# Workshop Computación Gráfica, Imágenes y visualización

Coordinadores

Silvia Castro (UNS)  
Roberto Guerrero (UNSL)  
Oscar Bría (INVAP)

# Simulation of High-Visual Quality Scenes in Low-Cost Virtual Reality

Matías N. Selzer, M. Luján Ganuza, Dana K. Urribarri,  
Martín L. Larrea, and Silvia M. Castro

VyGLab Research Laboratory, UNS-CICPBA, Bahía Blanca, 8000, Argentina.  
Institute of Computer Sc. and Eng. (CONICET-UNS), Bahía Blanca, Argentina.  
{matias.selzer,mlg,dku,mll,smc}@cs.uns.edu.ar  
<http://vyglab.cs.uns.edu.ar>

**Abstract.** With the increasing popularity of virtual reality, many video games and virtual experiences with high-visual quality have been developed recently. Virtual reality with a high-quality representation of scenes is still an experience linked to high-cost devices. There are currently low-cost virtual reality solutions by using mobile devices, but in those cases, the visual quality of the presented virtual environments must be simplified for running on mobile devices with limited hardware characteristics. In this work, we present a novel Image-Based Rendering technique for low-cost virtual reality. We have conducted a performance evaluation of three mobile devices with different hardware characteristics. Results show that our technique represents high-visual quality virtual environments with considerably better performance compared to traditional rendering solutions.

**Keywords:** Virtual Reality, Low-Cost VR, Navigation, Image-Based Rendering

## 1 Introduction

Virtual Reality (VR) is a technology already established in many different markets and platforms around the globe. One of the main factors in the rise of VR was the possibility of experiencing VR with low-cost mobile devices. This low-cost VR, when properly implemented, can not only enable people to experience this technology by using their mobile phones but also brings a new world of possibilities to those users who cannot afford a high-end VR system.

The games and experiences designed for this type of VR are usually graphically very simple, and rendering complex 3D scenes in these devices is a challenging problem. This is partly due to mobile device limitations in hardware and graphical processing (compared to modern gaming computers).

In this paper, we introduce a novel technique to display high-visual quality 3D scenes in low-cost VR devices. More specifically, the work presented here yields the following benefits and contributions:

- Our technique simulates a 3D environment by using a spherical panorama 3D texture matrix and dynamic image warping.
- We designed a novel image warping technique that enables a smooth transition between image points.
- Our technique was compared to traditional rendering of the same scene, obtaining better performance for multiple hardware configurations.
- Our technique allows a fully 3D VR navigation.

The rest of the paper is organized as follows. Section 2 gives an overview of the related work. Section 3 presents the details of the proposed technique and the proposed algorithm description. The experimental setup is described in Section 4. The experiment results are shown in Section 5 and discussed in Section 6. Finally, we draw our conclusions and point out the future work in Section 7.

## 2 Related Work

In recent years there has been tremendous growth in the number and variety of GPU-intensive mobile applications, enabling users to interact and navigate in high-visual quality virtual environments. To get a suitable system for mobile phones, many approaches vary both in the degree of realism of the virtual environment and the navigation technique.

The most popular render techniques for this purpose are those belonging to traditional geometry rendering and those oriented to image-based rendering (IBR). In the first case, the most viable alternatives to having high-visual quality virtual worlds are those that integrate progressive meshes, caching and restricted objects resolution techniques, thus allowing to overcome, to different degrees, the limitations of mobile devices in regards to processing, memory and power consumption [1–5]. In those cases, rendering high-visual quality 3D scenes on limited mobile devices obtained a very low performance.

Currently, Image-Based Rendering (IBR) is emerging as the most promising approach since the rendering process requires less computational resources than traditional geometry rendering. These techniques emerged in the late 1990s to overcome the severe limitations that exist for the representation of photo-realistic 3D scenes in real-time [6–8].

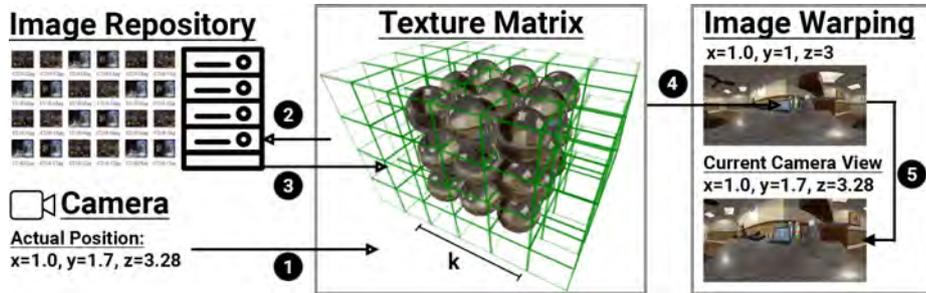
While more recent methods, running on PC, can synthesize new views based on captured panoramas from both video and images, the new generated views do not allow free navigation. Furthermore, they are based on the capture of a few panoramas and the synthesis of the intermediates is done by computationally intensive warping and interpolation algorithms [9, 10]. More recent approaches based on image interpolation for free navigation were proposed in the literature [11–15] but they only run on PC.

To represent high-visual quality 3D scenes with mobile devices, IBR is more suitable because it uses images as input and the rendering cost does not depend on the scene complexity, but on the final image resolution. Different approaches have been proposed that provide a variety of image types and qualities [6,

16]. More recently approaches, based on high-quality synthetic scenes generated in the server, have been presented in [17,18]. However, in all these cases, a significant latency is observed.

### 3 Our Proposal

In this article, we present a novel Image-Based Rendering technique based on spherical panoramic images to simulate high-visual quality 3D environments for low-cost VR devices. An overview of the technique's main components and their interaction is depicted in Figure 1. This process is continuously repeating as the user is moving, creating the sensation of a smooth transition through the virtual environment. All these components are detailed next.



**Fig. 1.** Technique overview. (1) The camera sends its current position to the Texture Matrix component. (2) A set of images is requested to the Image Repository. (3) The Image Repository sends the requested images. (4) The nearest image that corresponds to the camera current position is sent to the Image Warping component. (5) The Image Warping component calculates and generates the image that corresponds to the camera exact current position.

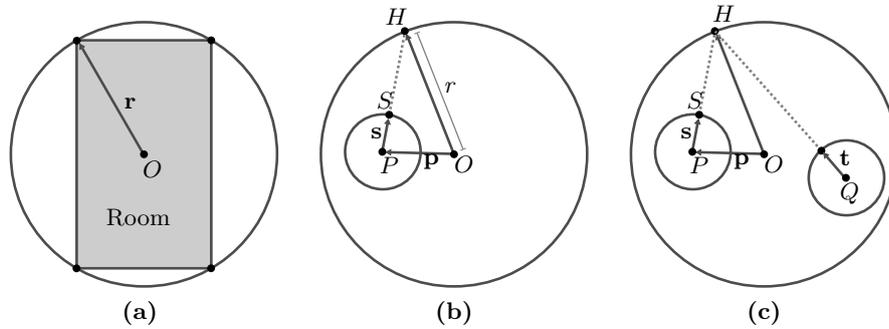
#### 3.1 Requirements

Our technique requires a set of  $360^\circ$  panoramic images captured from the original 3D environment to be simulated. The environment has to be sampled by capturing  $360^\circ$  panoramic images, each one from a different and specific position. If the environment has a dimension of  $dim_x$  in the x-axis,  $dim_y$  in the y-axis, and  $dim_z$  in the z-axis, and all images are separated by a distance  $\Delta$ , then all the images can be represented as an abstract 3D matrix with dimension  $N \times M \times R$  as shown in Figure 3 (right), where  $N = dim_x/\Delta$ ,  $M = dim_y/\Delta$ , and  $R = dim_z/\Delta$ . These  $360^\circ$  panoramic images are then stored in the Image Repository, which provides the images to the Texture Matrix component of our technique.

### 3.2 Local Image Management - Texture Matrix

The Texture Matrix component is in charge of loading, storing, and managing the  $360^\circ$  panoramic images recovered from the Image Repository. The number of cells in the Texture Matrix depends on the size of the original 3D environment and the distance  $\Delta$  between samples. The Texture Matrix dimension is  $N \times M \times R$ , where  $N$ ,  $M$ , and  $R$  are specified in Section 3.1. Therefore, this matrix is eventually capable of storing every sample of the scene.

In our approach, we provide efficient management of images. At any time, the Texture Matrix maintains a small number of images loaded in memory that correspond to the samples captured in a delimited neighborhood regarding the current position of the camera. As the camera moves, the Texture Matrix is updated. Finally, the matrix is regularly cleaned by removing those images captured in positions located at a certain distance of  $k$  from the current position of the camera.



**Fig. 2.** Cross-section of a room centered at  $O$  and two unitary spheres centered at  $P$  and  $Q$ . (a) Cross-section of the sphere with center  $O$  and radius  $r$  containing a room. (b) Point  $H$  over the room corresponding to the projection  $S$ . The unitary sphere is centered at  $P$ , where the equirectangular panoramic image  $\mathcal{I}_P$  was taken. (c) Projection of point  $H$  over the unitary sphere centered at  $Q$ , where no image was taken.

### 3.3 Image Warping

In order to avoid the user noticing jumps between the samples, we implemented a function to smoothen that transition. Let us assume that two contiguous spherical panoramic images  $\mathcal{I}_A$  and  $\mathcal{I}_B$  were taken at points  $A$  and  $B$ , then if the user stands at one of those points the system shows the corresponding image. However, while the user is moving from point  $A$  to  $B$  the system should show an intermediate image that smoothly switches from  $\mathcal{I}_A$  to  $\mathcal{I}_B$ .

We consider that the spherical panoramic images correspond to an equirectangular projection of a circular room that fully contains the original one (see Figure 2a). Let  $O$  be the center of the circle of radius  $r$  that contains

the room,  $Q$  be the point where the user stands and  $P$  be the point where the nearest panoramic image  $\mathcal{I}_P$  was taken. The strategy is to warp  $\mathcal{I}_P$  to match the panoramic image that should have been taken from the point of view of  $Q$ . If  $(u, v) \in [0, 1]^2$  is a texture coordinate over the panoramic image  $\mathcal{I}_P$ , we invert the projection to obtain the point  $S$  over a unitary sphere centered at  $P$  that projects onto  $(u, v)$  (see Figure 2b). Since the panoramic image is the result of a equirectangular projection, the vector  $\mathbf{s}$  between  $P$  and  $S$  is:  $\mathbf{s} = (x, y, z) = (\sin(\phi) \cos(\theta), \sin(\theta) \sin(\phi), \cos(\phi))$ , where  $\theta = 2u\pi$  and  $\phi = v\pi$ . Now, let  $H$  be the point over the room of radius  $r$  that projects on  $S$ . Then,  $H = k\mathbf{s} + \mathbf{p}$  where  $\mathbf{p} = Q - O$ ,  $|k\mathbf{s} + \mathbf{p}| = r$  and  $k = \frac{-(\mathbf{p} \cdot \mathbf{s}) + \sqrt{(\mathbf{p} \cdot \mathbf{s})^2 - |\mathbf{s}|^2(|\mathbf{p}|^2 + r^2)}}{|\mathbf{s}|^2}$ .

To warp the image to match the point of view of  $Q$ , we find the projection  $\mathbf{t}$  of  $H$  over the unitary sphere centered at  $Q$  (see Figure 2c),  $\mathbf{t} = \frac{H-Q}{|H-Q|}$ . Finally, the 2D projection  $(u, v)$  of  $\mathbf{t} = (x, y, z)$  over the equirectangular texture is  $(u, v) = (\frac{\theta}{2\pi}, \frac{\phi}{\pi})$  where  $\theta = \text{atan2}(y, x)$  and  $\phi = \text{atan2}(\sqrt{x^2 + y^2}, z)$ .

## 4 Experimental Setup

To study the benefits of our technique compared to a traditional 3D rendering approach, a performance experiment was conducted. Since performance depends on the used hardware, we designed an experiment to test our technique with three different hardware configurations. Here, the same scenario is presented in two different treatment conditions: a traditional rendering approach and by using our technique. This section details the experiment designed to evaluate our technique.

### 4.1 Virtual Environment

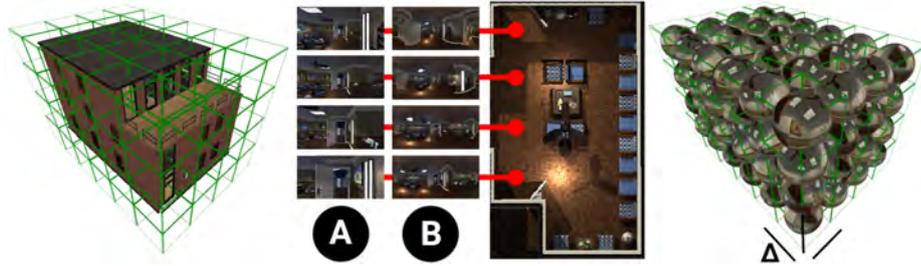
We used ‘‘Doctor’s Office’’ scenario[19] as the 3D virtual scenario for the test. This scenario is designed to run in Unity3d[20] in which we developed our technique. It consists of a representation of a doctor’s office with 2.1 million triangles, 2.3 million vertices, and it is rendered with high-level graphics techniques.

### 4.2 Images Dataset

The image dataset was created by using Unity360ScreenshotCapture[21]. This tool generates a 360° panoramic image based on the position of the virtual camera. The whole dataset is created automatically by iterating along the desired volume. Figure 3 shows an example of the sampling process.

### 4.3 Hardware Configurations

The proposed technique is specifically designed for low-cost VR, i.e., by using mobile phones. Since we want to show that our technique works for both low-end and high-end mobile phones, the current study evaluates and compares



**Fig. 3.** Image capture example. For each specified camera position, the application captures a 360° panoramic screenshot. (Left) The whole virtual environment is sampled. (Center) Column A shows the camera current view. Column B shows the 360° picture captured at that position. (Right) All the environment is sampled and an image corresponding to each cell of an abstract 3D matrix is generated.

the performance of our technique in the three different hardware configurations presented in Table 1.

In the rest of the paper, we use the letter H (high-end) to refer to the Le Eco Max 2 mobile phone, the letter M (mid-end) to refer to the Motorola Moto G mobile phone, and the letter L (low-end) to refer to the Google Nexus 7 tablet.

**Table 1.** Hardware specification of the three mobile devices used in the experiment.

Name	Le Eco Max 2	Motorola Moto G 2	Asus Google Nexus 7
OS	Android 6.0.1 / API-23	Android 6.0 / API-23	Android 6.0.1 / API.23
Screen	2560x1440@59Hz	1280x720@60Hz	1280x800@60Hz
Graphics API	OpenGL ES 3.1	OpenGL ES 3.0	OpenGL ES 2.0
GPU	Adreno (TM) 530	Adreno (TM) 305	NVIDIA Tegra 3
VRAM	2048 MB	256 MB	256 MB
Max texture size	16384px	4096px	2048px
Shader level	50	35	30
CPU	Quad-core 2.1 GHz Snapdragon-820	Quad-core 1.2 GHz Cortex-A7	Quad-core 1.2 GHz Cortex-A9
RAM	5774 MB	890 MB	971 MB

#### 4.4 Variables

In this study, we consider one independent variable, which is the technique used to render the virtual environment. This variable takes two values and the two treatment conditions of this study are explained as follows:

- Treatment Condition 1 (TC1) - Traditional Rendering: The virtual scenario is exported and rendered with a traditional graphic configuration.
- Treatment Condition 2 (TC2) - Our Technique: The application only contains the required elements for our technique.

The system response time is the time between the user's actions and the perceived response [22]. Several studies suggest that this latency degrades the sensation of presence [23,24], and it is suggested that the latency should not be greater than 100 ms [25]. However, more recent studies suggest that latency should be under 20 ms [26]. The frames-per-second (FPS) are a measure of how many frames the hardware can render in one second. Since the FPS represent a good measure of performance, this measure has been widely used to compare the performance between two or more different graphical techniques. For this reason, this study uses FPS as a measure of performance in both treatment conditions.

#### 4.5 Procedure

Even though both treatment conditions allow the user to freely navigate the virtual environment, to compare the performance between both treatment conditions, a specific path through the virtual environment was defined. This path has a total distance of 30 meters and it is shown in Figure 4.

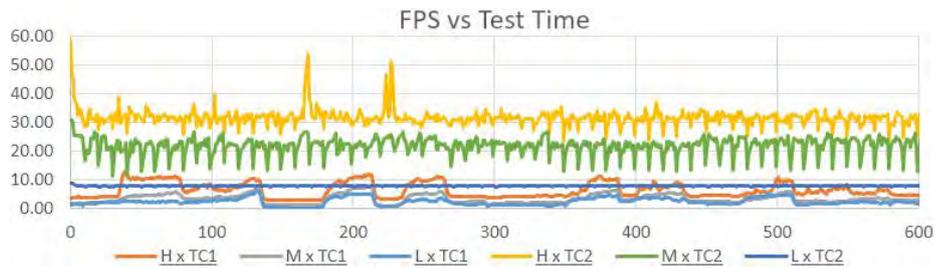


**Fig. 4.** The path followed by the virtual camera in the experiment. The camera travels from point A to point B in 300 seconds.

The duration time for the camera to get to point B is 5 minutes (or 300 seconds). At every 0.5 seconds, the current FPS value is recorded. Thus, for each test run, a total of 600 samples are generated. This test was performed by the three different hardware configurations in both treatment conditions. The results are presented in the next section.

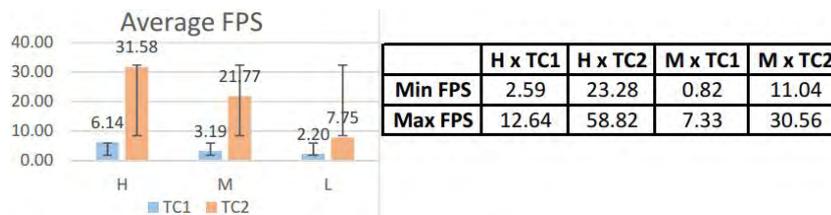
## 5 Results

The results of the performance experiment among the treatment conditions are depicted in Figure 5. Figure 6 presents the average, lower, and higher FPS scores obtained for both treatment conditions for the three hardware configurations. Based on these values, the H configuration obtained a performance improvement of 514%; the M configuration a performance improvement of 682%; and the L configuration a performance improvement of 352%.



**Fig. 5.** Performance Results Overview. The plot’s horizontal axis depicts the measured FPS during the 600 time samples of the total duration of the trial (300 seconds, one sample every 0.5 seconds) for each one of the hardware configurations. The histogram shows the average FPS of each trial.

To test the limits of the image warping effect beyond the 4 cm distance between samples, some informal user tests were also performed. Results showed that using a distance between samples up to 40 cm still proves an acceptable user experience. Higher values produced higher deformations in the images that induced cybersickness symptoms in the user. The implications of these results are discussed in the next section.



**Fig. 6.** (Left) Average FPS scores. (Right) Lower and Higher FPS scores comparison grouped by hardware configuration.

## 6 Discussion

The results obtained in the performance test are remarkable. For each one of the hardware configurations, there is a clear difference in FPS measured during the test. Each one of the hardware configurations obtained a performance improvement of at least 352% with an average FPS increment of 516%. These results suggest that our technique outperforms the traditional rendering technique regardless of the used mobile phone. This performance improvement was expected since a traditional rendering approach invests a lot of resources on rendering vertices, polygons, meshes, materials, calculating dynamic illumination

and shadows, among other things. In our technique, on the contrary, all such things are implicitly included in the images used to simulate the 3D environment.

Figure 6 shows the higher and lower FPS values obtained of each treatment condition and for each hardware configuration. For each one of the used mobile phones, the higher FPS value obtained while using traditional rendering was still lower than the lower value obtained by using our technique. If for example, we consider the middle hardware configuration used in this study, the average FPS value measured during the traditional rendering test was 3.19, which is equivalent to 1 image every 310 ms approximately. That is something unfeasible for VR in which the images presented to the user have to be displayed as fast as possible. On the contrary, by using our technique, the average measured FPS value was 21.77, which is equivalent to 1 image every 46 ms approximately. These results are more suitable for VR and satisfy the suggested values.

## 7 Conclusions and Future Work

This study presents a new Image-Based Rendering technique specially designed to enable the representation of high-quality virtual environments in low-cost VR devices. This work evaluated the performance of our technique compared to a traditional rendering approach. For this reason, image size optimizations were not investigated. Future work should, therefore, evaluate image optimizations for improving the performance.

One of the main benefits of our technique is that it does not depend on the visual quality or complexity of the virtual environment. Furthermore, this technique will also work with real-world pictures. If we are able to take 360° panoramic spherical pictures of the real world separated by a defined distance, this technique will allow virtual navigation through a total realistic environment. Future work should investigate the use of this kind of image dataset.

## References

1. Klaus Engel, Ove Sommer, and Thomas Ertl. A framework for interactive hardware accelerated remote 3d-visualization. In *Data visualization 2000*, pages 167–177. Springer, 2000.
2. JianPing Ma, Qiang Chen, Bo Chen, and Hong Wang. Mobile 3d graphics compression for progressive transmission over wireless network. In *11th IEEE Int. Conf. on CAD and Computer Graphics*, pages 357–362. IEEE, 2009.
3. Martin Isenburg and Peter Lindstrom. Streaming meshes. In *VIS 05. IEEE Visualization, 2005.*, pages 231–238. IEEE, 2005.
4. Zun Shen, Juan Liu, Yuhui Zheng, and Lu Cao. A low-cost mobile vr walkthrough system for displaying multimedia works based on unity3d. In *14th Int. Conf. on Computer Science & Education (ICCSE)*, pages 415–419. IEEE, 2019.
5. Ghada Fathy, Hanan Hassan, Walaa Sheta, and Reem Bahgat. Efficient framework for mobile walkthrough application. *Pervasive Mob. Comput.*, 18:40–54, 2015.
6. Shenchang Eric Chen. Quicktime vr: An image-based approach to virtual environment navigation. In *Proc. of the 22nd annual Conf. on Comp. graphics and interactive techniques*, pages 29–38, 1995.

7. Yuval Noimark and Daniel Cohen-Or. Streaming scenes to mpeg-4 video-enabled devices. *IEEE Computer Graphics and Applications*, 23(1):58–64, 2003.
8. Jimmy Chim, Rynson WH Lau, Hong Va Leong, and Antonio Si. Cyberwalk: a web-based distributed virtual walkthrough environment. *IEEE T. Multimedia*, 5(4):503–515, 2003.
9. Adarsh Kowdle, Sudipta N Sinha, and Richard Szeliski. Multiple view object cosegmentation using appearance and stereo cues. In *European Conf. on Computer Vision*, pages 789–803. Springer, 2012.
10. Gaurav Chaurasia, Sylvain Duchene, Olga Sorkine-Hornung, and George Drettakis. Depth synthesis and local warps for plausible image-based navigation. *ACM Transactions on Graphics (TOG)*, 32(3):1–12, 2013.
11. Christian Lipski, Christian Linz, Kai Berger, Anita Sellent, and Marcus Magnor. Virtual video camera: Image-based viewpoint navigation through space and time. In *Computer Graphics Forum*, volume 29, pages 2555–2568. Wiley Library, 2010.
12. Timo Stich, Christian Linz, Christian Wallraven, Douglas Cunningham, and Marcus Magnor. Perception-motivated interpolation of image sequences. *ACM Transactions on Applied Perception (TAP)*, 8(2):1–25, 2011.
13. Yanxiang Zhang and Ziqiang Zhu. Walk-able and stereo virtual tour based on spherical panorama matrix. In *Int. Conf. on Augmented Reality, Virtual Reality and Computer Graphics*, pages 50–58. Springer, 2017.
14. Feng Dai, Chen Zhu, Yike Ma, Juan Cao, Qiang Zhao, and Yongdong Zhang. Freely explore the scene with 360° field of view. In *IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)*, pages 888–889. IEEE, 2019.
15. Chao Liu, Susumu Shibusawa, and Tatsuhiko Yonekura. A walkthrough system with improved map projection panoramas from omni directional images. In *7th Int. Conf. on Ubiquitous Intelligence & Computing and 7th Int. Conf. on Autonomic & Trusted Computing*, pages 45–51. IEEE, 2010.
16. Yu Lei, Zhongding Jiang, Deren Chen, and Hujun Bao. Image-based walkthrough over internet on mobile devices. In *Int. Conf. on Grid and Cooperative Computing*, pages 728–735. Springer, 2004.
17. Juergen Doellner, Benjamin Hagedorn, and Jan Klimke. Server-based rendering of large 3d scenes for mobile devices using g-buffer cube maps. In *Proc. of the 17th Int. Conf. on 3D Web Technology*, pages 97–100, 2012.
18. Bernhard Reinert, Johannes Kopf, Tobias Ritschel, Eduardo Cuervo, David Chu, and Hans-Peter Seidel. Proxy-guided image-based rendering for mobile devices. In *Computer Graphics Forum*, volume 35, pages 353–362. Wiley Library, 2016.
19. Brick Project Studio. Doctorsoffice, 2020. <http://sojaexiles.com/>, online July 2020.
20. Unity. Unity3d, 2020. <https://unity.com>, online July 2020.
21. yasirkula. Unity3d, 2020. <https://github.com/yasirkula/>, online July 2020.
22. Jonathan Freeman, Jane Lessiter, and Wijnand IJsselsteijn. An introduction to presence: A sense of being there in a mediated environment. *The Psychologist*, 14:190–194, 2001.
23. Stephen R Ellis, Nancy S Dorigi, Brian M Menges, Bernard D Adelstein, and Richard H Jacoby. In search of equivalence classes in subjective scales of reality. *Advances in human factors/ergonomics*, pages 873–876, 1997.
24. Richard M Held and Nathaniel I Durlach. Telepresence. *Presence: Teleoperators & Virtual Environments*, 1(1):109–112, 1992.
25. Nathaniel I Durlach. Virtual reality-scientific and technological challenges. *National Academy Press*, pages 213–220, 1995.
26. Kjetil Raaen and Ivar Kjellmo. Measuring latency in virtual reality systems. In *Int. Conf. on Entertainment Computing*, pages 457–462. Springer, 2015.

# NATUS. A Physically-Based Rendering Engine with Real-Time Global Illumination

Ignacio del Barrio<sup>1</sup>, María Luján Ganuza<sup>1,2</sup>, and Silvia M. Castro<sup>1,2</sup>

<sup>1</sup> VyGLab Research Laboratory, DCIC, Universidad Nacional del Sur,  
[ignacioa.del.barrio@gmail.com](mailto:ignacioa.del.barrio@gmail.com), [{mlg, smc}@cs.uns.edu.ar](mailto:{mlg, smc}@cs.uns.edu.ar)  
<http://vyglab.cs.uns.edu.ar>

<sup>2</sup> Inst. for Computer Science and Engineering, ICIC (CONICET-UNS),  
San Andrés 800, 8000 Bahía Blanca, Argentina

**Abstract.** Real-time rendering applications are difficult software engineering development projects, due to the complexity of the implemented algorithms required to achieve interactive frame rates. To provide a faster development process, the rendering engines implement many of these algorithms to support developers. In this article, we present a real-time physically-based rendering engine supporting global illumination. It supports diffuse indirect illumination, glossy reflections, refractions (transparency), soft shadows, and light emitted from emissive surfaces at interactive frame rates.

**Keywords:** physically-based rendering, real-time rendering, rendering engine, global illumination, directx, cone tracing, voxels

## 1 Introduction

Real-time rendering is the process of synthesizing three-dimensional data as images on a computer at interactive frame rates [1]. For high-quality virtual rendered scenes, Global Illumination (GI) techniques are required to simulate the light exchanges of indirect illumination in those scenes. While these techniques allow a lot of realism to be added, calculating that GI, in real-time, is a difficult challenge.

Traditionally, indirect illumination has been too costly to compute under real-time constraints. Like most real-time rendering applications, access to the graphic accelerator cards (GPUs) is required. Rendering engines provide an abstraction layer on top of the graphic APIs and, in most cases, extra features that speed up the development of graphic applications. These characteristics should facilitate the engine to synthesize high-quality images with GI. By adhering to Physically-Based Rendering (PBR) models, a more accurate representation of how light interacts with surfaces could be provided.

In this paper, we present NATUS, an advanced rendering engine to assist developers in the design of realistic graphic applications. This is based on a PBR model, that allows users to easily configure the material appearance of the objects in their scenes, being assured that these will behave correctly under different lighting conditions. To achieve a high level of realism, the engine also provides a GI model. Its implementation allows dynamic scene objects and light sources, that would affect the final result of the indirect illumination.

Its architecture makes it easily extensible. New scenes can be created and the rendering pipeline modified. NATUS Engine is implemented in C++ and executable demos are available at <http://natus.io/project/natus-engine>.

## 2 Related Work

In the last couple of years several hybrid approaches have been developed, many of which are executed on the GPU, and have been implemented in commercial graphic engines like Unreal, CryEngine and Unity3D. Unreal Engine 4 [2] is the current iteration of one of the first major game engines to have come out to the public [3]. It uses DirectX, OpenGL, as well as WebGL. Unreal Engine 4 supports Screen Space Global Illumination (SSGI), a feature that aims to create natural-looking lighting by adding dynamic indirect lighting to objects within the screen view. CryEngine is another industry-level game engine [4]. It is known to produce state-of-the-art graphics and performance and supports Voxel-Based Global Illumination and soft shadows. Unity 3D [5] uses also a middleware for Realtime GI. At the moment, this solution is deprecated and Unity 3D is developing a new solution for real-time global illumination [6].

Most of the mentioned Engines require desktop GPUs with large amounts of memory and are suited for high-end desktop systems. But despite these limitations, it's thanks to the use of simplified reconstructions of the 3D scene coupled with clever algorithms and optimizations, that simulation of global illumination is practicable in the real-time graphic applications of today.

## 3 NATUS Rendering Engine Architecture

Like most engines, our system is built in layers [7]. Its different functionalities are carefully partitioned by topic and level of complexity. The upper layers of the system depend on the functionalities defined on the lower layers to create a more sophisticated and complex set of features. An overview of the architecture is shown in Fig. 1, and the details are described below.

### Middleware

The middleware is the lowest layer of the system architecture and provides the interface to the GPU via the DirectX 12 API. It contains all the third party libraries and APIs that make up the foundation of a windows-based rendering engine. Two of the most important programming interfaces are the DirectX 12 API and the Win32 API [8]. In addition, the application supports the libraries Tiny OBJ [9], to load 3D Meshes from files, and Dear ImGUI [10], for the construction of the application UI.

### Low-Level Renderer

The low-level renderer manages the drawing of the primitives by communicating with the graphics API. At this level, the design is focused on rendering a collection of *render items* as quickly as possible, without much regard to visibility issues. According to Gregory [7], this renderer should be completely agnostic as to the type of spatial subdivision or scene graphic used. This allows for the design of a system that is specifically suited to the needs of different types of applications. This module is integrated by six components.

The *Graphics Device Interface* main task is to initialize DirectX (configure the rendering pipeline, prepare the render buffers, etc.) for our engine framework and to manage the communication with the GPU.

The main components of a scene, *Cameras*, *Lights*, 3D objects and *Materials*, are defined in this engine layer. It supports two standard types of cameras available, a *Free Camera*, and a *Spherical Camera*, three different types of light sources, that is, point, spot and directional lights, and *Meshes* to represent

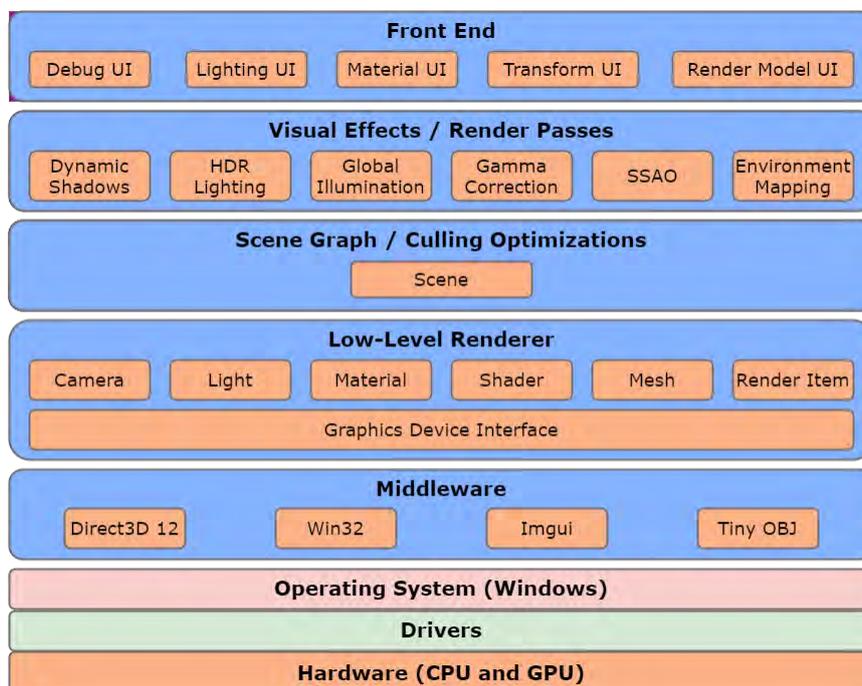


Fig. 1: The layered architecture of the rendering engine

the 3D objects. The *Materials* are lightweight components that describe how a surface should react under different light conditions, by including references to the texture resources it uses, color tints, and other physical properties. A detailed description of the material model is given in section 4. Materials only hold the attributes that describe the behavior given their interaction with light, but the actual computation of this behavior is done on the shaders.

The rendering engine supports four types of *Shaders*, that is, vertex, geometry, pixel, and compute. There is a predefined collection of shaders for different purposes, including texture filtering and clearing, geometry voxelization, voxels rendering, sky rendering, and global illumination. Finally, the *Render Item* is a lightweight component that keep rendering information from each scene object, such as a reference to the mesh and the material objects, the transformation matrix, and so on, in order to submit a draw call to the graphics API.

### Scene Graph

This layer manages which contents will be submitted to the low-level renderer based on some sort of visibility determination. The Scene class holds the scene information, that is, the main camera, collections of 3D Meshes, Render Items, textures and materials, and Light Sources. When the Rendering Engine starts up, it initializes the current scene instance by configuring the camera location followed by loading of textures and 3D Meshes, construction of materials, assembling of render items, and finalizing with lights configuration.

### Visual Effects

This layer is where all of the advanced rendering effects that our engine is capable of simulating are implemented, based on the functionality provided by the low-level renderer. These effects can be enabled or disabled on the custom scenes defined by the programmer. The most relevant effects, *Ambient Occlu-*

sion (AO) and *Global Illumination (GI)*, are described in more detail in section 5. The engine also implements *Dynamic Shadows* to calculate shadows from directional light sources with the traditional shadow map scheme proposed by Williams[11]. In addition to these, it provides *Gamma Correction*, performed at the end of all lighting calculations to ensure that the values in the final image are properly gamma-corrected, and *Environment Mapping*, to simulate reflections and refractions coming from the distant background. NATUS also implements *HDR Lighting*, in order to have a wider range of color values to calculate the final illumination. Once finished, it transforms the HDR values back to the LDR using the Hable [12] algorithm.

### Front End

This is the uppermost layer defined in the engine architecture, that is, the 2D user interface. To develop this interface we use a third party library for C++, Dear ImGui. It is fast, portable, renderer agnostic, and has no external dependencies, particularly suited for integration in game engines and real-time 3D applications. The UI includes the *lighting panel*, the *material panel* and *transform panel*.

The *Lighting Panel* panel is the main control point for the engine lighting features and is divided into two sections, one for the GI settings and another for configuring the light sources in the scene. The *Lighting Panel* is shown in Fig. 2. The GI section has settings for the scene voxelization, different effects

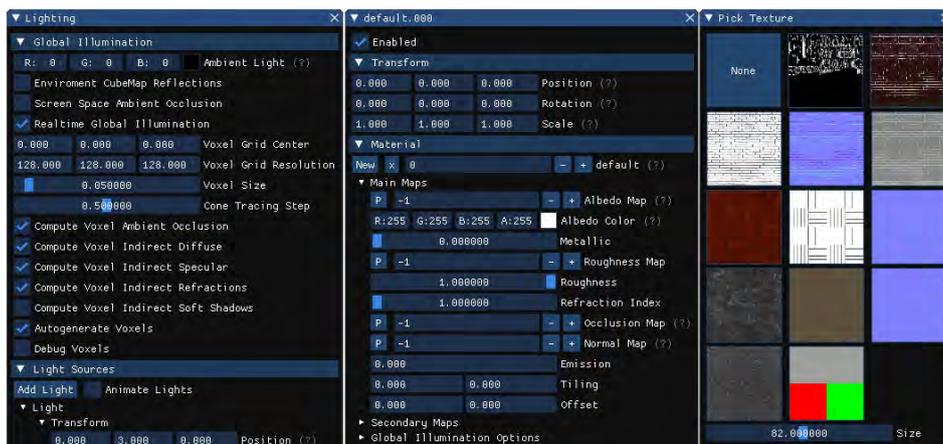


Fig. 2: The lighting panel (left). Material and transform panels (center and right).

computed via voxel cone tracing, environment reflections, and ambient light. The light sources section contains a list of light sources in the scene, whose properties can be modified via this panel. Lights can be added or removed from the scene through this panel. The *Material Panel* allows the user to control the appearance of objects in the scene, by manipulating the material properties of the selected object. This panel and its texture picking window are shown in Fig. 2. The *Transform Panel* controls the position, orientation and scale of objects, and lights in the scene relative to the world. It is shown in Fig. 2.

## 4 Physically-Based Materials

Our rendering engine materials follow a physically-based model implemented by our *standard shader*, which incorporates the advanced lighting algorithms and cal-

culations to simulate realistic surface lighting interaction. Physics-based rendering (PBR) is a methodology with no defined standard. The two most common workflows are metal/roughness and specular/glossiness. One of the first defined workflows, the metallic workflow, has been explored by Disney [13] and Adobe [14] and has been evolving, now emerging as a standard. We adopted this workflow, also used for real-time rendering by Epic Games in the Unreal Engine 4 [15]. This workflow is defined by a set of physical attributes, which are fed to the standard shader as textures or values. These attributes are diffuse albedo, metallic, and roughness:

- **Albedo.** The albedo or base color map is an RGB texture map that can contain the diffuse reflected color for dielectrics and the reflectance values for metals [16, 14]. It should not contain any lighting information. The alpha value of the albedo component color controls the transparency level for the material.
- **Metallic.** This map allows to define which areas of a material denote raw metal. As a grayscale map, it describes which areas in the base color should be interpreted as the reflected color (dielectric) and which ones as metal reflectance values, representing 1.0 (white) a raw metal.
- **Roughness.** This map describes the surface irregularities that cause light scattering. While rougher surfaces will have larger and dimmer-looking reflections, smoother surfaces will have concentrated specular reflections. On this map, white (1.0) represents a rough surface and black (0.0) a smooth one. The roughness map plays an integral part in the Cook-Torrance illumination model [1].

Other maps are usually attached. NATUS also supports *ambient occlusion (AO) map*, *Normal map* and *Emission* property. The *AO map* takes into account how much of the ambient environment lighting is accessible to a surface point. It only affects the diffuse contribution and should not occlude the specular contribution. In our implementation, the indirect diffuse lighting provided by the GI is multiplied by the AO. The *Normal map* is integrated into the PBR Material in order to simulate surface details. The *Emission* property controls the intensity of light emitted from the surface; when a material has a non-zero emission value it appears to be self-illuminated. During the real-time GI process, the emission value is used for the indirect diffuse lighting calculations to affect the illumination of nearby objects.

## 5 Voxel-Based Global Illumination

The voxel cone tracing is a GI technique introduced by Crassin et al. [17]. It calculates an approximation of the indirect lighting, generated by one or two bounces of light, for fully dynamic scenes in interactive applications, allowing both diffuse and glossy reflections with very realistic visual results. Instead of working on the actual geometry, a scene voxel representation is created and stored in the GPU. This representation can be created once for static geometry and per frame for dynamic objects. This algorithm is the central point of the GI Pipeline whose implementation is presented in Fig. 3.

**Shadow Mapping** The first step in the GI pipeline is to generate the corresponding shadow maps, that are going to be used later in the light voxelization step. We use *percentage closer filtering* [18], an improvement of the classic shadow mapping technique that allows smoothing the shadow edges. The map resolution is a trade off between speed and quality, and we use a screen resolution of 2048x2048 to produce an image without visual artifacts.

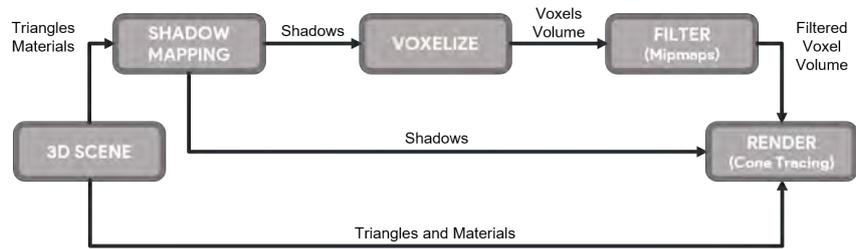


Fig. 3: Overview of the global illumination (GI) pipeline

**Voxelization** During the *voxelization step*, we use geometry shaders to output a voxelized representation of the lighting in the original scene geometry. We use the GPU hardware rasterizer to voxelize the scene geometry into a regular grid of voxels. Based on the Crassin et al. [17] approach, the voxel grid is stored in a 3D Texture. The scene illumination data is then saved into the grid voxels to effectively store the direct lighting from the scene. Each 3D texture voxel is represented by a single RGBA16 value; the alpha channel is used for opacity.

**Filtering** In order to perform voxel cone tracing of the 3D texture for full GI, it is necessary to filter it to generate a *mipmap chain*. The *filtering pass*, use the filtering GPU capabilities to create that mipmap. The 3D mipmap level generation is an extension of the 2D implementation by Nils Daumann [19].

**Voxel Cone Tracing** In this step, paths are traced through the filtered voxel structure in order to gather and approximate the indirect lighting for a point. A cone ray samples the volume on each step. The sampled volume is traced from the apex of a cone, located in a point  $p_0$  on the object surface, and along its axis oriented in the desired direction ( $p_d$ ) with an aperture angle  $\theta$ . The sampling region is increased on every step, based on the diameter of the cone. This expansion of the volume being sampled is approximated by taking samples from different mipmap levels generated during the filtering stage; the traced distance  $t$ , allows deciding the mipmap level to sample [17]. The whole process of tracing cones is done entirely on the pixel shader. Opacity  $\alpha_i$  and color  $c_i$  of each sample  $i$  from the voxel structure are integrated front-to-back along the cone to approximate the incoming indirect light at point  $p_0$  and the corresponding occlusion value  $\alpha$ .

**Render** The *illumination* at a point  $p$  combines the direct illumination, calculated using the physically-based Cook-Torrance BRDF, and the indirect diffuse illumination, reflections, and refractions from voxel cone tracing.

The *indirect illumination* at a surface point  $p$  is described by the hemisphere integral of the rendering general equation. To compute this integral efficiently,

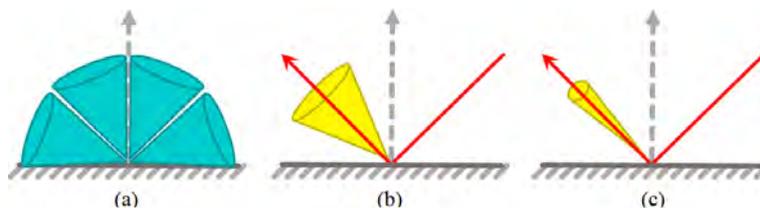


Fig. 4: Cone distributions approximate different phenomena. a) Multiple cones approximate indirect diffuse light. b) Wide cone approximates rough specular reflection. c) Narrow cone approximates fine specular reflections.

Crassin et al. [17] observed that the hemisphere can be partitioned into a sum of integrals, each of them gathering the indirect light from the scene for a cone. We use nine cones oriented over the hemisphere in order to gather all the indirect light from the scene for the diffuse reflection, as shown in Fig. 4. For the specular reflection and refraction we use a single cone in the reflected, or the refracted, direction, and its aperture is specified based on the surface material roughness. Specular cone apertures for different material roughness are shown in Fig. 4.

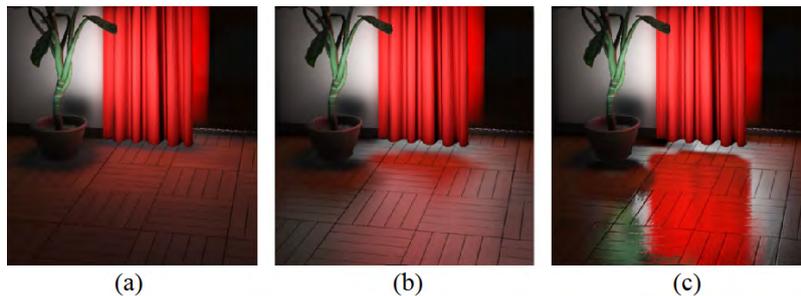


Fig. 5: Different material roughness on the same surface. (a) Completely rough surface with diffuse reflection only. (b) Slightly rough surface with diffuse and specular reflections. (c) Very smooth surface with high specular reflection.

The visual appearance of the cone tracing configurations are shown in Fig. 5. An advantage of voxel cone tracing is that the same cones distribution can be used to approximate the indirect diffuse reflection (combining color values), and the ambient occlusion (combining occlusion values). The accumulated occlusion value  $\alpha$  returned from cone tracing represent the final occlusion value. Soft shadows can be casted by tracing a cone from each surface point  $p$  towards the light sources.

## 6 Evaluation

The application was tested on a mid-range desktop PC. We used a NVIDIA GeForce GTX 960 (2GB) GPU, an Intel Core i5 8400 CPU and 16GB of RAM.

The **default scene** used to evaluate our rendering engine is based on the classic Cornell Box scene. We placed a point light slightly above the center, illuminating the entire box. The objects placed inside have different geometrical complexity and their appearances were configured to showcase different material properties (Table 1).

Object	Diffuse RGBA format	Metallic (0.0-1.0)	Roughness (0.0-1.0)	IOR $\geq 0.0$	Emission $\geq 0.0$
Dragon	[ 1.0, 1.0, 1.0, 1.0 ]	1.0	0.1	1.0	0.0
Buddha	[ 0.0, 0.0, 0.0, 0.0 ]	0.0	0.1	3.0	0.0
Suzanne	[ 1.0, 1.0, 1.0, 1.0 ]	0.0	0.1	1.0	0.0
Teapot	[ 0.0, 0.5, 0.0, 1.0 ]	0.0	1.0	1.0	0.0

Table 1: Material settings for each object in the test scene. IOR: Index Of Refraction.

The default scene was tested using a  $128^3$  voxel grid, and a  $1024 \times 768$  screen resolution. We tested the following lighting settings, with whom the scenes in Fig. 6 were rendered:

- **Test 1.** Only direct illumination (No GI).
- **Test 2.** GI, including reflections and refractions.
- **Test 3.** GI and soft shadows.
- **Test 4.** GI, soft shadows and ambient occlusion.
- **Test 5.** GI, with only emissive wall.

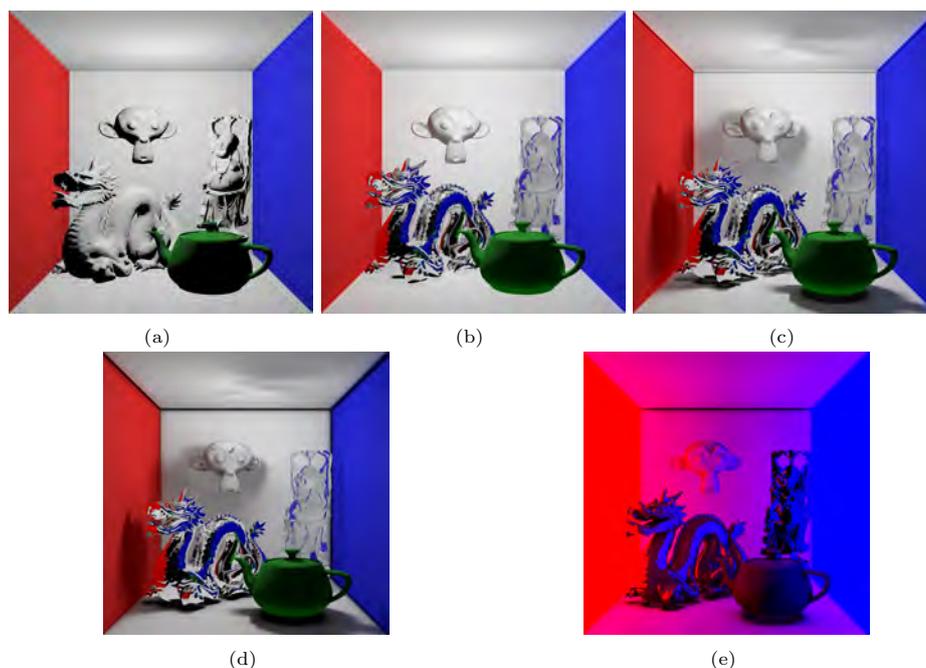


Fig. 6: Cornell Box scene with (a) only basic direct lighting, (b) direct and indirect lighting, (c) full GI, including soft shadows, (d) full GI and ambient occlusion, and (e) emissive wall materials, and no other light sources.

## 6.1 Results and discussion

We ran the tests described in section 6. All of them performed very well, with frame rates over 30 fps even on full HD (1920x1080) resolutions and all GI effects enabled. These results differ significantly from those obtained with Blender's Cycles renderer. Although the visual results of path tracing are slightly superior in Blender, the performance penalty compared to our engine is very significant. It takes about 20 ms to our application to generate one frame with full GI. Under the same hardware, Cycles needs 18 minutes to generate it. This is around 67,500 times higher. Fig. 7 shows the comparison between both rendering engines at 1024x768 resolution. AO is not enabled on our application to more closely match the reference.

Table 2 shows the average time to render one frame for different voxel grid resolutions and fixed 1024x768 screen resolution. The rows marked with an asterisk are the averages when voxelization is not run on each frame. This Table shows the performance implications of running GI in our rendering engine. It can be seen that the automatic voxelization of the scene on each frame has an impact on the performance, specially visible on Test 1 where no GI is being calculated. Automatic voxelization could be disabled for scenes with no dynamic lighting or



Fig. 7: (a) Voxel Cone Tracing in NATUS (20ms). (b) Path Tracing in Blender (18 min)

Voxel Resolution	Test 1	Test 2	Test 3	Test 4
256	6.62	24.05	29.02	29.01
256(*)	1.72	19.36	24.26	24.26
128	4.09	18.23	22.36	22.34
128(*)	1.79	16.23	20.36	20.36
64	4.92	15.41	18.45	18.44
64(*)	1.77	12.45	15.53	15.52

Table 2: Average time (ms) per frame for each test scene under different voxel grid resolutions. Rows with asterisk indicate that voxelization does not run on each frame.

objects. That allows to gain a small performance boost. Anyway, we get real-time performance (over than 33fps) in all test cases. We observe a strong performance decrease when GI is activated on Test 2, which is completely reasonable. Adding traced soft shadows in Test 3 also increases the render times by an average of 3 ms, but there is no penalty for enabling traced AO. As it is obtained using the same cones traced for indirect diffuse lighting, both effects can be calculated at the same time.

Voxel Resolution	Shadow Map	Voxelize	Filter	Cone Trace
256	0.31	3.01	1.72	28.96
128	0.30	1.96	0.23	22.83
64	0.30	3.14	0.06	16.22

Table 3: Average time (ms) per step for test scene 4 under several voxel grid resolutions

Table 3 shows the average time it takes to perform each step of the full GI algorithm, with indirect diffuse lighting, reflections, refractions and soft shadows. It can be seen that the most expensive step in the GI algorithm is cone tracing the voxel structure. It is interesting to see that voxelization of a  $64^3$  voxel grid takes more time than that of a  $128^3$  voxel grid. Generating the mipmap chain on the filtering step is quite fast, and the shadow mapping step does not really depends on voxel resolution which is why it doesn't really change.

## 7 Conclusions and Future Work

In this article, we present a real-time rendering engine capable of synthesizing high quality images with full Global Illumination. It was designed following a layered

architecture, where each layer has a specific responsibility within the application which also allows it to be easily extended.

Main features of our engine are the introduction of PBR materials coupled with a simulation of full GI. Our PBR material model is based on the metallic workflow which is a *de-facto* standard on most of the real-time rendering applications used nowadays, allowing for the authoring of materials in third party applications like Substance Designer [20].

The voxel cone tracing allows a wide range of GI effects such as indirect diffuse reflections, ambient occlusion, glossy reflections, transparency and soft shadows to be achieved in real-time with high quality results. While this method yields very good results for dynamic scenes, there are still a few issues that need to be addressed. The major downside of this technique is that it only performs well under small scenes. A similar technique, used for shadow mapping, known as Cascaded Voxelization [21] could be used to extend this solution for bigger scenes.

## References

1. T. Akenine-Möller, E. Haines, and N. Hoffman. *Real-time rendering*. CRC Press, 2019.
2. Epic Games. Unreal engine. [www.unrealengine.com/](http://www.unrealengine.com/), 1998. Last acc. July 2020.
3. A. Andrade. Game engines: a survey. *EAI Endorsed Trans. on Serious Games*, 2(6), 11 2015.
4. Crytek. Cry engine, 2002. <http://cryengine.com/>, online July 2020.
5. Unity Technologies. Unity engine, 2005. <https://unity.com>, online July 2020.
6. Unity3D.com. Realtime global illumination using enlighten, 2002. <https://docs.unity3d.com/Manual/realtime-gi-using-enlighten.html>, Acc Jul 2020.
7. J. Gregory. *Game engine architecture*. CRC Press, 2018.
8. Windows Developer Center. Windows api. <https://docs.microsoft.com/en-us/windows/win32/apiindex/windows-api-list>, 2020. Last acc. June 2020.
9. S. Fujita. Tinyobj:tiny but powerful single file wavefront obj loader. <https://github.com/syoyo/tinyobjloader>, 2020. Last acc. December 2019.
10. O. Cornout. Dear imgui, 2019. [github.com/ocornut/imgui](https://github.com/ocornut/imgui), online June 2020.
11. Lance Williams. Casting curved shadows on curved surfaces. In *Proce. of the 5th annual Conf. on Computer Graphics and Int. Techniques*, pages 270–274, 1978.
12. J. Hable. Uncharted 2: Hdr lighting. In *Game Developers Conf.*, page 56, 2010.
13. B. Burley. Physically-based shading at disney, part of practical physically based shading in film and game production. In *SIGGRAPH Courses*, 2012.
14. W. McDermott. *The PBR Guide*. Allegorithmic, 2018.
15. B. Karis. Real shading in unreal engine 4. *Proc. Physically Based Shading Theory Practice*, 4:3, 2013.
16. S. Lagarde. Feeding a physically-based shading model. <https://seblagarde.wordpress.com/2011/08/17/feeding-a-physical-based-lighting-mode/>, 2011. Last acc. June 2020.
17. C. Crassin, F. Neyret, M. Sainz, S. Green, and E. Eisemann. Interactive indirect illumination using voxel cone tracing. In *Computer Graphics Forum*, volume 30, pages 1921–1930. Wiley Online Library, 2011.
18. Randima Fernando et al. *GPU gems: programming techniques, tips, and tricks for real-time graphics*, volume 590. Addison-Wesley Reading, 2004.
19. N. Daumann. D3d12 texture mipmap generation. <https://sli-dev.com/d3d12-texture-mipmap-generation/>, 2017. Last acc. July 2020.
20. J. Jantunen. Creating procedural textures for games: with substance designer. 2017.
21. J. McLaren. Cascaded voxel cone tracing in the tomorrow children. In *Computer Entertainment Developers Conference*, 2014.

## Towards defining a CAVE like system performance evaluation

Nicolás Jofré, Graciela Rodríguez, Yoselie Alvarado,  
Jacqueline Fernandez, and Roberto Guerrero

Laboratorio de Computación Gráfica (LCG)  
Universidad Nacional de San Luis,  
Ejército de los Andes 950  
Tel: 02664 420823, San Luis, Argentina  
{npasinetti, gbrodriguez, ymalvarado, jmfer, rag}@unsl.edu.ar

**Abstract.** One of the main goals of Virtual Reality is to provide immersive environments that take participants away from the real life into a virtual one, this is how *Cave Automated Virtual Environment (CAVE)* came about many years ago. Nowadays there are many of this kind of room-sized systems providing a superior Virtual Reality experience and are used for research into a wide range of disciplines including archaeology, architecture, art, biology, engineering, geometry, geology, medicine and healthcare, meteorology and physics. Nevertheless, for a good Virtual Reality user experience, it is necessary to have a processing system optimized for visual computing (based on CAVE-related features, Interaction, Application, etc.). In this work we propose an evaluation methodology for our Cave-like multi-VRmedia System. The proposal is based on three generic criteria: *Performance*, *Usability* and *Relevance*. The strategy tries to prove how assertive a system is when it comes to solving a problem.

**Key words:** Virtual Reality, CAVE system, Computer Graphics, System Analysis, CAVE like Evaluation, Multidimensional Evaluation

### 1 Introduction

*Virtual Reality* (VR) started about 60 years ago in a form that would be recognized as today such as head-mounted display (HMD), head tracking, computer graphics generated images, among others, although the hardware was completely different. Over the last decade a huge amount of research has nevertheless been carried out across a vast range of applications – from medicine to business, from psychotherapy to industry, from sports to travel.

On the other hand, scientists, engineers, and people working in industry carried on with their research and applications using and exploring different forms of VR, and in turn developing VR systems or creating laboratories. In relation to VR systems, there are two main groups – head-mounted devices and CAVE-type installations, an advanced VR solutions based on specially configured set of projection screens.

Particularly, the *Cave Automatic Virtual Environment* is a room-sized, immersive virtual reality system developed in the early 1990s, which continues to provide a superior VR experience to those offered by other VR systems, such as HMDs. In spite of its sizable volume and cost, the popularity of the CAVE system as a research tool for science has not declined over time. The CAVE was envisioned from the outset as a device to enable distance collaboration among viewers immersed in shared computer-generated scenes – a kind of 3D telephone booth, a technique called tele-immersion [1].

There are several types of CAVE installations in the world, each of them differ in dimensions or number of screens on which images are projected, supporting architectural technology and technical installations. All of them, are improvements to the classic CAVE, mainly as a result of commercialization.

Software for CAVEs, or more generally, for large display systems, has also evolved hand-in-hand with developments in hardware. A common feature of most CAVE-type systems is that a highly immersive experience is afforded to users by surrounding them with large screens on which stereoscopic images are projected. All participants wear active stereo glasses to see and interact with complex 3D objects. One participant wears a six degree-of-freedom location and orientation sensor called a tracker so that when user moves within the CAVE, the perspective of the images is automatically adjusted to a viewer's eyes by the tracking system. The user in a CAVE room interacts with VR objects through a portable controller, which is sometimes called a *wand*.

In particular, the *Laboratorio de Computación Gráfica* (LCG) [2] has so far focused on the development of immersive environments that support multimodal and especially gestural interaction. The result of our efforts has been *IVI CAVE* (Immersive Virtual Innovation CAVE), a low-cost and powerful engine to manage real-time interaction within an immersive environment [3].

*IVI CAVE* has been used to give solutions to real problems of the society like: traffic accidents [4], car driving [5], civic rules teaching [6], stroke rehabilitation [7], multiple sclerosis rehabilitation [8], among others. Now, the aim is to validate the immersive environment achieved through years of hard work. The analysis would allow to determine if the solutions obtained to address real problems are improved by using the platform. Additionally, determining if our environment follows conventional cave standards and identifying any improvements is desired.

Our research questions are related to the added value of any system in the context of a CAVE-type system:

- To which extent the system brings a better sense of experience in virtual immersion?
- What is the precision of a CAVE like system versus others immersive system?

The following sections attempt to show some aspects to take into account when defining a systematic method of analysis of a CAVE like system. Section 2 names the existing basic criteria for any generic system. Section 3 attempts to define a test set for any CAVE like system according to the basic criteria of section 2. Section 4 provides a small discussion and future guidelines.

## 2 Typical Systems' Evaluation Criteria

Visual computing includes the traditional types of graphics applications plus many new applications. For a good virtual reality user experience, it is necessary to have a processing system optimized for visual computing. When talking about a CAVE system, components, machines and systems must be analyzed and optimized in order to give a better user experience, while at the same time depicted scenarios are simulated and processed at a very short time, and interactions are resolved quickly and easily. In brief, a CAVE like system features can be summarized as follows:

- A CAVE is used to visualize 3D data, transporting viewers directly into the system or machine environment.
- Implementing a system based on CAVE technology requires in-depth knowledge in many different disciplines, such as sensing and tracking technologies, stereoscopic displays, multimodal interaction and processing, computer graphics and geometric modeling, dynamics and physical simulation, performance tuning, etc.
- Input/output, screening and full recognition of the viewer's being and actions (speech, non-verbal utterances, and gestures) must be provided in real time with no great user encumbrances (special glasses, headphones, head tracking, among others).

All these features must be linked together giving a supporting structure to any application system suited for use in CAVEs with near-zero latency and no noticeable artifacts.

The construction of a CAVE like system is an ambitious, complex, and essentially interdisciplinary process. This is inevitable given the depth, sophistication, and many modalities of the products we seek to create. More over, when trying to analyze a CAVE like system, we dealt with different scientific difficulties:

- *CAVE-related questions*, like intrusiveness, visuo-haptic synchronization, and space perception. CAVEs cause specific issues that HMDs do not. Objects cannot be visually intrusive, physical and virtual parts are mixed within the environment (especially the body of the subject), space perception can be biased, anthropological parameters have an influence, etc. All of them can be synthesised in the pair hardware and software relationship.
- *Interaction system related questions*, specific to the modality (kinesthesia, cutaneous, sensory substitution, etc.) or any device implemented.
- *Application system related questions*, specific to the addressed problem.

There are different strategies in the literature to evaluate how assertive a system is when it comes to solving a problem. Generally speaking, systems are usually evaluated in accordance with 3 criteria.

- *Performance*. It should be sufficient to allow effective interactions and feelings of confidence.
- *Usability*. Industrial usage requires the system in study can be measured by taking into account the context of use of the system in order to be accessible to a large spectrum of users, and compatible with a large spectrum of use-cases.

- *Relevance*. Each system modality can respond to additional use-cases which must fit the context.

The following sections develop these concepts in detail for a CAVE like system.

## 2.1 System Performance

When designing a VR system, the ultimate hardware design goals are: *scalability, usability, potential to hold several users and/or be network connected, extended service intervals and easy access for maintenance, power-efficient, low cost*, among others.

On the other hand, desirable features include [9][10]:

- *High resolution*, so virtual images are seen as sharply and in as much detail as in reality.
- *High brightness and contrast*, so colors are vibrant and not washed out or dim.
- *3D realism*, production of computer graphics and the display of captured imagery in a way equivalent to or exceeding human visual acuity, in 3D with the correct viewer-centric perspective rendering for every viewer.
- *Whole interaction*, Extended input and full recognition of the viewer's or viewers' being and actions, including speech, non-verbal utterances/noise making, and gestures.

In addition, and if it is useful to the task, the following features are desired:

- *Audio* (sonification) at or exceeding human aural acuity, fully surround, listener centered and focused.
- *Touch* (tactile) input from the user and touch output from the VR system, allowing haptic input and feedback, for all users [9].
- *Olfactory* (smell) output delivered to each user, and input recognition as well.
- *Taste* output and input recognition.

Both design goals and desirable features involve hardware and software aspects. As currently there are a wide assortment of hardware and software configurations assembling a CAVE like system, a comprehensive analysis of their components must be done.

## 2.2 System Usability

This section address methods for evaluating it as artefacts according to their usability, in an attempt to analyse the quality of the system.

Usability is a quality attribute that assesses how easy user interfaces are to use. The word “usability” also refers to methods for improving ease-of-use during the design process. Usability is an empirical concept, which means that it can be measured and evaluated. In fact, usability is a quality attribute whose formal definition is the result of listing the different components or variables

through which it can be measured. Usability is defined by 5 quality components: *Learnability, Efficiency, Memorability, Errors, Satisfaction* [11].

One of the best ways to evaluate the usability of a product or application is to test it with real users, a method known as **user testing**. By observing how users are confronted with interactive tasks, we can objectively quantify design usability by counting the number of errors they make (*effectiveness*) or by measuring the time it takes them to complete them (*efficiency*). In addition, by asking users once they have completed their tasks, we can measure subjective or perceived usability, that is, how users rate the design or what their level of satisfaction is.

Considering generic usability evaluations, there is a wide variety of evaluation methods that are classified into four main methods [12].

- Usability Inspection Methods: used by experienced practitioners. While these methods do not involve users directly, they can provide some useful insights. However, the goal is to use them to supplement, not replace, direct user involvement in testing designs and systems; e.g. *Pluralistic Walkthrough, Heuristic Evaluation, Cognitive Walkthrough, Heuristic Walkthrough, Metaphors of Human Thinking (MOT), Persona Based Inspection*.
- Usability Testing with users: *Usability Testing, Benchmark Testing, Competitive Usability Testing, Summative Usability Testing, Remote Evaluation, Think Aloud Testing, Wizard of Oz*.
- Evaluate Usage of an Existing System: *Critical Incident Technique (CIT), User Edit, Web Analytics*.
- Questionnaire and Survey Methods: *Rating Scales, Satisfaction Questionnaire, System Usability Scale (SUS)*.

In particular, VR is thus often referred to as immersion technology. VR systems may vary greatly in levels of immersion and user experience they offer based on the system characteristics and context of virtual environments. One of the major evaluation problems is the lack of a suitable criteria for this type of systems. Immersion is a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences. An environment that produces a greater sense of immersion will produce higher levels of presence. Presence, is the subjective experience of being in one place or environment, even when one is physically situated in another. So *immersion* and *presence* are significant features to be defined as metrics:

- Immersion: measures the degree of visual simulation that a VR interface provides for the viewer - the degree of suspension of disbelief. Factors that affect immersion include *isolation from the physical environment, perception of self-inclusion in the virtual environment, natural modes of interaction and control, and perception of self-movement*. Immersion also refers to a level of sensory fidelity which depends on measurable system attributes such as *Field Of View (FOV); display size; stereoscopy; display resolution; head-tracking; or input devices*, among others [13, 14]. According to the literature

- some of the techniques to measure the degree of immersion are: *Immersive Tendencies Questionnaire* (ITQ) [15] and *Biosensors* [16].
- Presence: in [15] the authors state that presence is a normal awareness phenomenon that requires directed attention and is based in the interaction between sensory stimulation, environmental factors that encourage involvement and enable immersion, and internal tendencies to become involved. There are multiple mechanisms that attempt to address the perception of presence: *Presence Questionnaire* (PQ) [15, 17, 18], *Slater-Usoh-Steed questionnaire* (SUS) [19, 20], *Igroup Presence Questionnaire* (IPQ) [21–23], *ITC-Sense of Presence Inventory* [24].

### 2.3 System Relevance

In [25] Saracevic identified five types of relevance, namely: (1) topical or cognitive relevance; (2) algorithmic relevance; (3) pertinence or intellectual relevance; (4) situational relevance, and (5) motivational or affective relevance. The concept of relevance has long been studied in diverse fields including philosophy, communication, logic, psychology, artificial intelligence, natural language processing, documentation, information science, and information retrieval.

In particular, information science emerged as the third subject, along with logic and philosophy, to deal with relevance - an elusive, human notion. The concern with relevance, as a key notion in information science, is traced to the problems of scientific communication (type 3), due to relevance is considered as a measure of the effectiveness of a contact between a source and a destination in a communication process. Specifically, in information retrieval or data quality, relevance is the ability of an information retrieval system to retrieve material that satisfies the needs of the user. The information retrieval community has emphasized the use of test collections and benchmark tasks to measure topical relevance [26–29].

In contrast to this focus solely on topical relevance, the information science community has emphasized user studies that consider user relevance (type 1 and 5). These studies often focus on aspects of human-computer interaction [30–32]. According to the literature, all existing metrics in the area, so far, have been defined for information sciences and mostly related to information retrieval. This means that the relevance of CAVE systems is a possible gap in existing scientific knowledge.

## 3 Defining a basic CAVE like system evaluation

In section 2 a group of criteria (*performance*, *usability* and *relevance*) and the corresponding evaluations methods for any kind of system had been mentioned. Nevertheless, a more specific evaluation approach discerning the most relevant aspects of a CAVE like system is desirable. In sections below we will try to define a set of methods that can be used to assist with the analysis process of a CAVE like system following the referred criteria.

### 3.1 Performance

In order to evaluate the performance of the CAVE configuration, and for simplicity, components will be evaluated separately as hardware components and software components.

In the core of a any VR system events are processed and the corresponding images generated, transmitted, and depicted on screens; this is called a *graphic pipeline*. The greatest opportunity for any visual computing system is to “accelerate the graphics pipeline”. Among all the hardware components, graphics card is the most important, processing each image and deciding where it is going to be depicted (which screen will show them). However, there is no point in generating high quality images if the used display device does not truly depicts the generated graphic information to users. Moreover, allowing the user to affect or interact with 3d scenes will require space perception and visuo-haptic synchronization. In brief, mostly improvement in features named in subsection 2.1 can be achieved focusing on quality of three specific hardware component:

- *Graphic Cards*: Nowadays, graphic cards are manufactured to work specifically in applications that require processing scenes with millions of polygons with good refresh rates (real time). In addition, they have the power to work with multiple outputs, mostly, with three or more displays. These features enable different architectures of CAVE systems. It makes possible to build systems where the work of generating and distributing each image is done by a single graphic card or even combining several graphic cards, increasing the number of available outputs [33].
- *Display Devices*: Immersion and space perception require that images depicted by a CAVE system show the high quality information generated by graphics cards. Since a CAVE user is so close to screens, the perceived realism is obtained with high image resolutions (2k o more). Nowadays, high-resolution screen walls could be delivered by monitors or multimedia projectors. They differ in brightness, contrast, and covered space. While High Dynamic Range (HDR) monitors emit a high range of colors, they need a complex structure to get the required space. On the other hand, multimedia projectors enable to easily paint complete walls but colors are quickly washed out and dimmed.
- *Interaction*: Ideally, immersion and presence will be more likely obtained depending on how a CAVE system could recognize the user through different communications channels, that is, the use of verbal and non-verbal recognition technology, such as motion sensors, tracking, GPS, pointers, sound/voice recognition, among others.

From a software perspective, graphics libraries have evolved to maximize the power usage of both CPUs and GPUs. Some libraries have gained popularity like *DirectX* but continue to be OS-dependent. Other libraries like *Vulkan* (OpenGL) have little commercial impact but are multi-platform. The optimal coding under these libraries maximizes the use of graphics cards, allowing a reduction in work times and providing real time. The system performance will be affected by the

selected graphics library and its compatibility with hardware components. Finally, the whole CAVE like system performance will be the resulting of sum of each component performance.

### 3.2 Usability

Considering that usability is a quality attribute defined by five components (*Learnability, Efficiency, Memorability, Errors and Satisfaction*), and bearing in mind that a CAVE like system must prioritize *psychological/cognitive intrusiveness, visuo-haptic synchronization, space perception and functionality*, we are interested in evaluating our system from a *user interface point of view* (infrastructure use) and from a *design point of view* (infrastructure design). For both point of view evaluation, the method of *user testing* will be applied. Taken into account the average sample size used for questionnaires in the literature, a group of 30 participants will be design. To ensure heterogeneity in the samples, participants should be selected ensuring both age and gender assortment.

From the user interface point of view, the *perceived* or *subjective usability* of the user (evaluation of the level of satisfaction) will be computed by performing small specific tasks that will test the total architecture of the CAVE. Given the features of our system, the best adapted techniques are: *Critical Incident Technique (CIT), Satisfaction Questionnaire and Think Aloud Testing*.

From the design point of view, the efficiency/effectiveness of the design will be computed. Given the features of our system, the best adapted techniques are: *Rating Scales and System Usability Scale(SUS)*.

### 3.3 Relevance

Relevance has been most fundamentally studied in epistemology (theory of knowledge). Different theories of knowledge have different implications on what is considered relevant and these fundamental views have implications on all other fields as well. This subsection looks into the relevance of a CAVE system in the context of the emerging opportunities brought by the 4th industrial revolution and trying to fill a gap in the study of relevance in CAVE systems.

In a CAVE system the measurement of relevance should enable to determine what is the contribution of the system to the personal experience of the individual (user). The idea is to engage users in a context and not just display them something, for that matter, the most important features ensuring that users are involved in a cave systems are *immersion* and *presence*.

In [26, 34] authors said that relevance is naturally cognitive (type 1 of subsection 2.3) and the greater the cognitive effects, the greater the relevance, and the smaller the processing efforts to derive these effects, the greater the relevance (*Ease-of-use* and *Learnability*). As a consequence, we propose to evaluate the cognitive relevance of a CAVE system taking into account *immersion, presence, easy of use* and *easy to learn* as the minimum metrics that should be taken. Most of these metrics have been considered in subsection 2.2 (usability criteria), therefore the tests already defined could be considered to evaluate each metric.

In particular, for *immersion experience* and *presence*, the *Immersive Tendencies Questionnaire* (ITQ) and *Presence Questionnaire* (PQ) questionnaires will be used, due to they were defined jointly and are suitable for joint evaluation, allowing that the efforts required for both the respondents and the measurement collection to be reduced.

*Ease of use* and *learning* are standard metrics of usability, therefore they are covered by proposed tests in the subsection 3.2. Also, the basis established for user testing participants selection are copied from this subsection.

## 4 Discussion and Future works

Today, providing real-time visual interaction with a unified graphics (computed objects via graphics, images and video) and computing architecture that serves as both a programmable graphics processing and a scalable parallel computing platform combining hardware and software to form heterogeneous systems, is the proposal of any CAVE like system. Nevertheless, even when this is achieved that not means the obtained system is really assertive at the moment of solving a problem. Over time, it has become necessary to generate a methodology to evaluate a complete cave system.

In this paper we have presented a evaluation methodology for CAVEs systems. The criteria analysis and its evaluations produced an evaluation methodology that will be used to analyze, and if necessary modify and improve our CAVE like system. In summary, our research has produced evaluations on three levels:

- *Performance*: Graphic Cards, Display Devices, Interaction, and Graphic Library.
- *Usability*: Critical Incident Technique (CIT), Satisfaction Questionnaire and Think Aloud Testing, and Rating Scales and System Usability Scale (SUS).
- *Relevance*: Immersive Tendencies Questionnaire (ITQ) and Presence Questionnaire (PQ).

The research framework described in this document has been done, so far, only to analyze the characteristics and criteria that enable the evaluation of CAVE like systems. Based on this, our system will be evaluated in a future work. We hope that the methodology outlined in this paper will provide a starting point for techniques that allow immersive VR system developers create immersive environments that are usable, useful, and engaging. Finally, we believe that this type of analysis will allow us to outline new lines of work that will answer some unanswered questions, such as:

- Which is the best conceptual model for a CAVE like system?.
- Which is the best prototype technique?.
- How understand the underlying influence of each device.
- Which are good practices to minimize the lag between input and output devices used in the CAVE?.
- How to minimize the use of I/O devices that constraint the participant?.
- Can we use prop devices to convey real devices?.

## References

1. Robert V Kenyon. The cave (tm) automatic virtual environment: characteristics and applications. 1995.
2. Laboratorio de Computación Gráfica. <http://www.lcg.unsl.edu.ar/>.
3. Yoselie Alvarado, M Moyano, David Quiroga, Jacqueline Fernández, and Roberto A Guerrero. A virtual reality computing platform for real time 3d visualization. In *XVIII Congreso Argentino de Ciencias de la Computación*, 2012.
4. Guerrero R. et al. Uso de la realidad virtual para la prevención de accidentes de tránsito, 2013.
5. Guerrero R. et al. Conduce, 2016.
6. Yoselie Alvarado, Roberto Guerrero, and Francisco Serón. Be civic: An immersive serious game. *EAI Endorsed Transactions on Serious Games*, 4(15), 2018.
7. Nicolás Jofré, Graciela Rodríguez, Yoselie Alvarado, Jacqueline Fernández, and Roberto Guerrero. Natural user interfaces: A physical activity trainer. In *Argentine Congress of Computer Science*, pages 122–131. Springer, 2017.
8. Guerrero R. et al. Esclerogym, 2017.
9. Antoine Lassagne, Andras Kemeny, Javier Posselt, and Frederic Merienne. Performance evaluation of passive haptic feedback for tactile hmi design in caves. *IEEE Transactions on Haptics*, 11(1):119–127, 2017.
10. Thomas A DeFanti, Daniel Acevedo, Richard A Ainsworth, Maxine D Brown, Steven Cutchin, Gregory Dawe, Kai-Uwe Doerr, Andrew Johnson, Chris Knox, Robert Kooima, et al. The future of the cave. *Central European Journal of Engineering*, 1(1):16–37, 2011.
11. Norman Nielsen. Usability 101: Introduction to usability. [urlhttps://www.nngroup.com/articles/usability-101-introduction-to-usability/](https://www.nngroup.com/articles/usability-101-introduction-to-usability/), 2012.
12. User Experience Professionals’ Association. Usability body of knowledge. [urlwww.usabilitybok.org](http://www.usabilitybok.org), 2012.
13. Mel Slater and Sylvia Wilbur. A framework for immersive virtual environments five: Speculations on the role of presence in virtual environments. *Presence: Teleoper. Virtual Environ.*, 6(6):603–616, December 1997.
14. D. A. Bowman and R. P. McMahan. Virtual reality: How much immersion is enough? *Computer*, 40(7):36–43, 2007.
15. Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoper. Virtual Environ.*, 7(3):225–240, June 1998.
16. Bryn Farnsworth. Measuring the power of virtual reality immersion [a case study], 2019.
17. Rod McCall, Shaleph O’Neil, and Fiona Carroll. Measuring presence in virtual environments. In *CHI ’04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’04, page 783–784, New York, NY, USA, 2004. Association for Computing Machinery.
18. Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI ’19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.
19. Mel Slater and Anthony Steed. A virtual presence counter. *Presence: Teleoper. Virtual Environ.*, 9(5):413–434, October 2000.

20. Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. Using presence questionnaires in reality. *Presence: Teleoper. Virtual Environ.*, 9(5):497–503, October 2000.
21. Holger Regenbrecht and Thomas Schubert. Real and illusory interactions enhance presence in virtual environments. *Presence: Teleoper. Virtual Environ.*, 11(4):425–434, August 2002.
22. Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoper. Virtual Environ.*, 10(3):266–281, June 2001.
23. Thomas W. Schubert. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie*, 15(2):69–71, 2003.
24. Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence: Teleoper. Virtual Environ.*, 10(3):282–297, June 2001.
25. Tefko Saracevic. Relevance reconsidered. In *Information science: Integration in perspectives. Proceedings of the Second Conference on Conceptions of Library and Information Science, Copenhagen (Denmark)*, 1996.
26. Saman Koswatte, Kevin McDougall, and Xiaoye Liu. Relevance assessment of crowdsourced data (csd) using semantics and geographic information retrieval (gir) techniques. *ISPRS Int. J. Geo-Information*, 7:256, 2018.
27. Hansi Senaratne, Amin Mobasheri, Ahmed Loai Ali, Cristina Capineri, and Mordechai Haklay. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31:139 – 167, 2017.
28. J. ODonovan, B. Kang, G. Meyer, T. Höllerer, and S. Adalii. Credibility in context: An analysis of feature distributions in twitter. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, pages 293–301, 2012.
29. Saman Koswatte, Kevin McDougall, and Xiaoye Liu. Vgi and crowdsourced data credibility analysis using spam email detection techniques. *International Journal of Digital Earth*, 11:520 – 532, 2018.
30. Yunjie Xu and Hainan Yin. Novelty and topicality in interactive information retrieval. *Journal of the American Society for Information Science and Technology*, 59, 01 2008.
31. Roman Y. Shtykh and Qun Jin. A human-centric integrated approach to web information search and sharing. *Human-centric Computing and Information Sciences*, 1(1):1–37, December 2011.
32. Yunjie Calvin Xu and Zhiwei Chen. Relevance judgment: What do information users consider beyond topicality? *J. Assoc. Inf. Sci. Technol.*, 57:961–973, 2006.
33. Marián Hudák, Štefan Korečko, and Branislav Sobota. On architecture and performance of lirkis cave system. In *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 000295–000300. IEEE, 2017.
34. Howard D. White. Relevance theory and citations. *Journal of Pragmatics*, 43(14):3345 – 3361, 2011.

# ARPOT: an Augmented Reality Platform for Outdoor Teaching

Matías N. Selzer<sup>1,2</sup>, Dana K. Urribarri<sup>1,2</sup>, Nicolás F. Gazcón<sup>1,2</sup>, Martín L. Larrea<sup>1,2</sup>, Claudia Sereno<sup>3</sup>, Antonela Volonté<sup>3,4</sup>, Aldana Mastrandrea<sup>3</sup>, and Silvia M. Castro<sup>1,2</sup>

<sup>1</sup>VyGLab Research Laboratory, UNS-CICPBA, Bahía Blanca, 8000, Argentina.

<sup>2</sup>Institute of Computer Sc. and Eng. (CONICET-UNS), Bahía Blanca, Argentina.

<sup>3</sup>Departamento de Geografía y Turismo, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina.

<sup>4</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Ciudad Autónoma de Buenos Aires, Argentina.

{matias.selzer,mlg,dku,ml1}@cs.uns.edu.ar  
{csereno,antonela.volonte,aldana.mastrandrea}@uns.edu.ar  
smc@cs.uns.edu.ar

**Abstract.** In the last years, there has been interest in introducing Augmented Reality (AR) into education. However, users who are not technical specialists require tools to facilitate the creation and customisation of AR experiences. Since educational tours and excursions connect students with the concepts taught in classroom, and AR increases the students' interest and motivation, we developed ARPOT, an AR platform for outdoor teaching that integrates tools for the teacher to easily create AR content and tools for students to view and interact with such AR content without internet connection. ARPOT was tested by 240 students of different educational levels at the wetlands of Villa del Mar, Argentina, and 12 new students at the campus of our university. In both case studies, the students reported a more motivating and engaging experience, and the educators reported that ARPOT was a powerful tool for content creation.

**Keywords:** Augmented Reality; Education; Mobile Learning; Outdoors; Location-based Mobile Augmented Reality

## 1 Introduction

Augmented Reality (AR) is a technology that merges the real world with virtual objects and, thanks to the advances in technology, it is becoming increasingly popular in many disciplines. AR is particularly relevant for education because it is directly related to Constructivist concepts and situated learning [1, 2], and it provides visual elements and interaction that enhance the learning process.

AR has vast implications and numerous benefits for the augmentation of teaching and learning environments. According to [3], AR has the potential to:



Fig. 1: ARPOT in use. The icons are always fixed in the corresponding place, even when the user walks and rotates. A small label also shows the distance between the user and the point-of-interest. (Left) Student looking for a particular point-of-interest. (Right) Device view.

engage, stimulate, and motivate students to explore class materials from different angles; help teach subjects where students could not feasibly gain real-world first-hand experience (e.g. Geography); enhance collaboration between students and teachers as well as among students; foster student creativity and imagination; help students take control of their learning at their own pace and on their own path, and create an authentic learning environment suitable to various learning styles [4, 5].

Most schools around the world take their students on educational trips or excursions to educative or cultural heritage places. Generally, the information about those places is provided in very limited pamphlets or by hiring a special person to guide the students, which cannot always be afforded by the school. Considering the current technology and the fact that nowadays most students own a mobile phone, a free and accessible way to create and provide more information is needed.

To the best of our knowledge, only a few researchers have addressed this need considering the currently available technology. Most of the existing AR frameworks and applications for mobile phones include proprietary and commercial engines [6]. However, none of them provides insights into their functionality or a way to customise the application to a specific purpose. In addition, only some of them allow the creation of new content. Furthermore, the available systems require an active internet connection during the tour, restricting eligible places to deploy such systems.

The inclusion of technology in the classroom has never been more feasible. In this context, we present ARPOT, an open-source, free, easy to use, and internet-free location-based AR platform for educative purposes. By using ARPOT's web application, the teachers are able to easily create data about a particular educative place and, with ARPOT's mobile application, students can use their own smartphone to visualise the AR content created by their teachers during the tour (see figure 1).

Results from the conducted study confirm that ARPOT is a very useful tool for teachers to create AR content about any educative outdoor place.

Furthermore, as [7] suggests, the use of innovative technologies improve the students' learning skills by providing a more motivating and engaging experience.

The rest of the paper is organised as follows: in the next section, the related work about AR platforms and applications for outdoor teaching is presented. The following section introduces the platform, the actors involved, and the activities they can perform. It also includes the description of the platform's main components: the web and the mobile applications. The section "Case studies" presents two different case studies designed to test the system; in the first one, the students use ARPOT's first prototype and visit the wetlands of Villa del Mar, Argentina; in the second one, the students test an improved version of ARPOT in the campus of our university. Finally, the discussion and conclusions for this work and an outline of the future work are presented.

## 2 Related Work

The research community is constantly working on the inclusion of AR in education [8–11], not only into the classroom but also into field trips. AR provides an opportunity for both learners and teachers to engage with the educative process in a very innovative way. According to teachers' opinion, learning improves when students are interested, creative and inspired [7]. Furthermore, pedagogy is paying much attention to learner's creativity and productivity.

Mobile AR has been studied for more than a decade. Henrysson et al. [12] and Pence et al. [13] outline a broad vision and discuss various possible applications. Wagner et al. [14] consider the challenges, strategies, and limitations to overcome while building mobile AR systems.

In this context, the literature presents some AR mobile applications for teaching outdoors, but only for very specific and particular contents. Chiang et al. [15] presents an augmented reality-based mobile learning system for conducting inquiry-based learning activities with students between 9 and 10 years old. These activities were developed following a five-step design methodology (ask, investigate, create, share, and reflect) and the experiment was conducted using a natural science unit on aquatic animal and plants. The proposed system relies on the internet for most of the steps. A more recent platform for creating AR content is Metaverse <sup>1</sup>. This platform allows users to create and share AR content and, based on the information presented on their web page, it is being used in many educational institutions. However, Metaverse does require an active internet connection. For this reason, it is not suitable for outdoor teaching in places with no wifi or mobile internet connection. Chou et al. [16] conducts an experiment on their university campus to evaluate the exploration of the campus via an AR mobile application based on wifi or 3G connection. They report a significant gain in motivation, attention or satisfaction by the students. In general, researches regarding the usage of AR in education confirm the benefits of this technology to the motivation and engagement of students during learning activities [17, 18, 9].

<sup>1</sup> <https://gometa.io/>

### 3 ARPOT Workflow and Components

ARPOT is a platform to provide augmented information about different real world surrounding Points of Interest (POIs) in real time. There are different actors involved in the usage of the platform and each one is able to perform different activities. Moreover, the platform consists of components designed specifically for each actor and its activities (A video showing ARPOT's functionality can be found at <sup>2</sup>).

#### 3.1 System Actors and Activities

Two actors are involved in the use of ARPOT: the teacher who creates the tour and the student who, along with the teacher, experiences the tour. Within the use of our system, there are three stages whose responsibilities go from teacher to student. For each actor and stage, there are activities that must be executed. Figure 2 presents the workflow of ARPOT and each stage is detailed next.

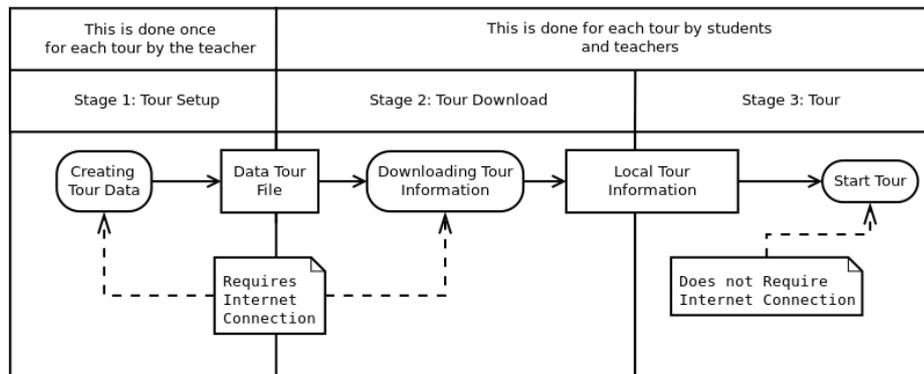


Fig. 2: The ARPOT Platform workflow is represented in this diagram. We identify three stages and two actors. Only the first two stages require an internet connection.

**Stage 1: Tour Setup** The tour setup focuses on the creation of a file with data for a particular tour. It is meant to be used by the teacher who wants to assemble a tour in a particular educative place. However, this is not restricted only to teachers; for example, the guides of an educative place may want to create a data file to assist them with more information about a particular tour. In this stage, the teacher must use the web application to create the data file that should be used later by the students. The web application was successfully tested in major web browsers, namely Chrome, Microsoft Edge, Microsoft IE, and Mozilla Firefox. The web application is currently available at <sup>3</sup>.

<sup>2</sup> <https://vimeo.com/290800417>

<sup>3</sup> <http://vyglab.cs.uns.edu.ar/arpot/index.html>

**Stage 2: Tour Download** Once the students have the data file downloaded on their smartphones, they can download all the tour information by using the mobile application. The mobile application automatically detects all the data files with ARPOT extension on the phone and asks the student to select the data to be downloaded.

One of the main objectives of ARPOT is that no internet connection is required to use the mobile application in the field. For this reason, the mobile application downloads all the information once and stores that information in the phone's local memory. This stage also requires internet connection, but not a web browser. The mobile application is responsible for downloading the required information. Once this stage is completed, the students will be able to access all that information in a transparent way without requiring an internet connection.

**Stage 3: Tour** With the data already downloaded and stored in the phone's local memory, the students are now ready to travel to the selected educative place and use the mobile application. The mobile application is able to display POIs from the students' surrounding on the smartphone's screen. Specifically, the POIs are drawn based on the phone's orientation and position.

The development of a mobile AR application involves some challenges. In order to augment the image captured by the phone's camera with virtual information about the surrounding POIs, location-based calculations regarding the GPS location must be performed. Furthermore, an efficient and reliable technique for calculating the distance between any two positions is required. The distance and orientation between the student's current position and all the POIs' position should be calculated in real-time. On the other hand, the application must deal with many phone's sensors in order to determine the orientation and position of the phone. Finally, the angle of view of the phone's camera must be calculated in order to display the virtual objects on the respective position of the camera view on the screen.

The mobile application was developed by using Unity3d<sup>4</sup>, a very popular and free game engine. Unity3d allows the developer to export applications to many different operating systems in a simple manner, including Android. In addition, Unity3d contains integrated tools to access most of the phone's sensors, which avoid the use of other external libraries.

By using this application, the students are able to view all the data generated by the web application, consisting of several POIs and their respective information. The students are not only able to see where these POIs are in terms of location and distance, but they are also able to see all the information associated to each particular POI (text, images, videos, etc.).

## 4 Case Studies

In order to test the usability and the performance of ARPOT, two case studies were carried out. First, teachers and researchers from the Department of

<sup>4</sup> <https://unity3d.com>

Geography and Tourism created a dataset of POIs about the Wetland of Villa del Mar and 240 students of different educational levels (160 students aged between 14 and 15 years –93 males and 67 females– and 80 students aged between 9 and 11 years –49 males and 31 females–) took the tour and used our mobile application. This location and these students were selected in the context of a pre-planned trip to Villa del Mar for students from the nearby schools, organised by the researchers of the Department of Geography and Tourism. Then, based on this first experience, ARPOT was improved and a second case study was planned, where 12 new students used our platform on the campus of our university. Both case studies are detailed next.

#### **4.1 Case Study 1: Wetland of Villa del Mar**

The coastline of Villa del Mar, Buenos Aires, Argentina, presents a diversity of landscapes resulting from the relationship between the elements of nature and the constructions that society performs. In this area, there are environments valued by their ecological functions and unique landscape; such is the case of the coastal wetlands of Villa del Mar. Therefore, acquiring knowledge and valuation of these local environments through formal and non-formal educational proposals encourages learning and awareness of their conservation. Every year, many students from nearby schools visit this place to learn more about its ecosystem.

In order to test the platform, teachers and researchers from the Department of Geography and Tourism used the web application to create a dataset with several points of interest about the wetland of Villa del Mar. After the dataset creation, the teacher downloaded the data file and sent it to the students via email. Every student then downloaded the data file and opened it with the mobile application. The mobile application automatically detects the data file and start downloading the corresponding content. This download process is performed just once, so no internet connection is required in the future for the same dataset. Therefore, the students performed this procedure at home with a regular internet connection. Finally, all the students arrived at the tour with the mobile application installed on their phones and with the tour's data already downloaded.

Villa del Mar was visited in October 2017 and each group was able to use the application to observe animal and plant species (see figure 3). The activity included four observation sites: Rocas Beach, the Municipal Beach, FRAAM (Foundation for the Reception and Assistance of Marine Animals) and an interpretive trail related to the wetland. After the visit, the students returned to the classroom to work on the gathered information, emphasising those aspects or problems of interest.

#### **4.2 Case Study 2: University Campus**

Every year, more than 4000 new students arrive on the campus of our university. The campus is growing each year with the addition of new departments, faculties, and classrooms; however, signs and maps are widely dispersed and very



Fig. 3: Using the mobile application. (Left) Student close to the location of the POI that represents “Balneario Municipal Villa del Mar”. (Right) Information presented after tapping on the “Gaviota Cangrejera” POI icon.

confusing. Those new students usually get confused with colours representing the buildings, the campus division into sectors, and some incorrectly marked roads.

Generally, the new students get lost, arriving late for classes and losing precious time and energy. For this reason, this case study presents the use of ARPOT for the creation of a dataset containing all the relevant information for students, especially new ones. This information contains POIs with the location of the classrooms, the departments, the buffet, the ATM, the bus stop, among others. In addition, every POI also contains pictures to help students to identify the place.

A group of 12 new university students between 17 and 18 years old was recruited to perform a tour around specific locations of the campus. Half of them used ARPOT mobile application, and the other half only used the available physical signs around the campus.

The students started at the campus' entrance and they were asked to go to a number of places in a specific order. First, they were asked to go to a particular classroom, then to the buffet, followed by another particular classroom, followed by the Department of Computer Science and Engineering, and finally to the bus stop. Figure 1 shows a student looking for the buffet POI. The student can tap on any POI icon in order to get more information about it.

## 5 Discussion

Regarding Case Study 1, back in the classroom, the students were informally interviewed about the experience and about what they have learned. Many reported that the experience using ARPOT was more engaging and that they were very motivated to explore and learn about the place, the flora, and the fauna. They also reported that they could remember the information about most POIs because they were encouraged to physically reach them. Since the POIs'

physical location represents a memory cue related to spatial cognition learning, walking through the POI helped the students to remember more information about such POIs. Even though in traditional trips the POIs are also associated with physical locations, the students are not encouraged to find and reach them, thus reducing their motivation.

Some POIs for plants or animals were intentionally placed where the actual plant or animal usually is. This was really surprising for many students because they could see a connection between the mobile application and the real world. For example, when the students approached the location where the mobile application indicated that there were crabs, they found actual crabs in the real physical location (see figure 4).



Fig. 4: POI indicating an area where actual crabs are expected to be found.

Regarding Case Study 2, the students who did not use the ARPOT reported that it was extremely hard to find the specified places. The existing signs around the campus are not very clear and the students needed to ask for directions or started searching the Internet for information. Four students mentioned that it would have been impossible for them to find some of the places if they had not asked for directions. The other students agreed with this.

The students who used the ARPOT mobile application reported that they could find each location very easily and that the general experience was motivating and fun. They mentioned that the inclusion of the buildings' facade pictures was really helpful. When they were walking from one location to another, some of the students were learning more about the surrounding places by looking at all the available information in the application.

## 6 Conclusions

ARPOT is an AR platform for outdoor teaching which provides a powerful tool for educators by helping them to create motivating and engaging experiences for their students. Considerable progress has been made regarding the inclusion of

AR technology in everyday education but many AR frameworks require a user with technical skills [6].

We created a platform so that people without a technical background can create educative AR tours. In addition, in contrast to most location-based AR systems for outdoors, ARPOT does not require an internet connection while the students use the mobile application in the tour. This is a novel and necessary feature because, for example in Argentina, most educative outdoor places are far away from internet coverage zones.

The present work has shown that ARPOT allows students to be protagonists of the teaching and learning process, interact with the elements that surround them, and visualise their closest environment from different perspectives. They become active participants by obtaining information that deepens on the knowledge of the complex social reality, thus contributing to the learning-by-discovery process. This set of tools presents an innovative potential to carry out educational practices that consider the exploration, the exchange of experiences and the active participation of students for the joint construction of knowledge. In addition, it is important to mention that thanks to the guidance and advice of the teachers, the students recognised the elements of the physical and natural environment, as well as the most significant problems of the addressed sector. This contributes to the understanding of complex geographic processes.

## References

1. Jean Lave and Etienne Wenger. *Situated Learning. Legitimate Peripheral Participation*. Cambridge University Press, 1991.
2. Dee Vyas. Increasing student engagement using augmented reality. *The Journal of Educational Innovation, Partnership and Change*, 1(2), 2015.
3. Steve Chi-Yin Yuen, Gallayanee Yaoyuneyong, and Erik Johnson. Augmented reality: An overview and five directions for ar in education. *Journal of Educational Technology Development and Exchange (JETDE)*, 4(1):119–140, 2011.
4. Neven A.M. El Sayed, Hala H. Zayed, and Mohamed I. Sharawy. Arsc: Augmented reality student card. *Computers & Education*, 56(4):1045 – 1061, 2011.
5. Lucinda Kerawalla, Rosemary Luckin, Simon Seljeflot, and Adrian Woolard. “making it real”: exploring the potential of augmented reality for teaching primary school science. *Virtual Reality*, 10(3):163–174, Dec 2006.
6. Teemu H. Laine. Mobile educational augmented reality games: A systematic literature review and two case studies. *Computers*, 7(1), 2018.
7. John Sener. In search of student-generated content in online education. *e-mentor*, 4(21):90–94, 2007. URL <http://www.e-mentor.edu.pl/artykul/index/numer/21/id/467>.
8. Kangdon Lee. Augmented reality in education and training. *TechTrends*, 56(2):13–21, March 2012.
9. Iulian Radu. Augmented reality in education: a meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, 18(6):1533–1543, Aug 2014.
10. Fatih Saltan and Ömer Arslan. The use of augmented reality in formal education: A scoping review. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(2):503–520, 2017.

11. Hsin-Kai Wu, Silvia Wen-Yu Lee, Hsin-Yi Chang, and Jyh-Chong Liang. Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, 62:41 – 49, 2013.
12. Anders Henrysson and Mark Ollila. Umar: Ubiquitous mobile augmented reality. In *Proceedings of the 3rd International Conference on Mobile and Ubiquitous Multimedia*, MUM '04, pages 41–45, New York, NY, USA, 2004. ACM.
13. Harry E. Pence. Smartphones, smart objects, and augmented reality. *The Reference Librarian*, 52(1-2):136–145, 2010.
14. D. Wagner and D. Schmalstieg. Making augmented reality practical on mobile phones, part 1. *IEEE Computer Graphics and Applications*, 29:12–15, 05 2009.
15. Tosti H. C. Chiang, Stephen J. H. Yang, and Gwo-Jen Hwang. An augmented reality-based mobile learning system to improve students' learning achievements and motivations in natural science inquiry activities. *Journal of Educational Technology & Society*, 17(4):352–365, 2014. URL <http://www.jstor.org/stable/jeductechsoci.17.4.352>.
16. Te-Lien Chou and Lih-Juan Chanlin. Location-based learning through augmented reality. *Journal of Educational Computing Research*, 51(3):355–368, 2014.
17. M. Bower, C. Howe, N. McCredie, A. Robinson, and D. Grover. Augmented reality in education – cases, places, and potentials. In *2013 IEEE 63rd Annual Conference International Council for Education Media (ICEM)*, pages 1–11, Oct 2013.
18. Amy M. Kamarainen, Shari Metcalf, Tina Grotzer, Allison Browne, Diana Mazzuca, M. Shane Tutwiler, and Chris Dede. Ecomobile: Integrating augmented reality and probeware with environmental education field trips. *Computers & Education*, 68:545 – 556, 2013.

# Workshop Bases de Datos y Minería de Datos

Coordinadores

Rodolfo Bertone (UNLP)  
Hugo Alfonso (UNLPam)  
Nora Reyes (UNSL)

# Seguimiento de rodeos en establecimientos agropecuarios

Pablo Rafael Palmero, Edilma Olinda Gagliardi, Maria Gisela Dorzán

Universidad Nacional de San Luis  
Ejército de Los Andes 950, San Luis, Argentina  
{prpalmero, oli, mgdorzan}@unsl.edu.ar

**Resumen** El presente artículo muestra una plataforma para el seguimiento espacio temporal de rodeos y control de las posiciones espaciales de elementos que integran un establecimiento agropecuario. Mediante el uso de las Tecnologías de la Información y la Comunicación y de los modelos de Bases de Datos Avanzados, la plataforma mantiene características orientadas a la recolección de datos de diferentes fuentes, almacenamiento compartido, integración progresiva de diversas funcionalidades, explotación y visualización de la información, y accesibilidad mediante tecnología móvil y de escritorio.

**Keywords:** Modelos Avanzados de Bases de Datos, Bases de Datos Espaciales, Bases de Datos Espacio Temporales, Geometría Computacional, TIC.

## 1. Introducción

Como un dominio de aplicación de las Tecnologías de la Información y la Comunicación (TIC), investigaciones recientes denotan que en la actividad agropecuaria promueven nuevas experiencias, transformando las prácticas de producción, los modos de comunicación, la gestión y organización en establecimientos agropecuarios.

Con el objetivo general de promover un ámbito de investigación, intercambio y desarrollo de TIC, que aporte a la cooperación científico tecnológico y a las prácticas sociales, productivas y comerciales de los actores de la producción agropecuaria, se articula y fomenta un proyecto denominado *Campo Conectado* [1,2]. El mismo se promovió entre la Estación Experimental Agropecuaria San Luis del Instituto Nacional de Tecnología Agropecuaria (INTA) y la Universidad Nacional de San Luis.

En este desarrollo se trabaja multidisciplinariamente entre el Proyecto de Investigación *Tecnologías Avanzadas de Bases de Datos - Línea Bases de datos Espaciales y Espacio Temporales* de la UNSL, el Proyecto de Investigación *Diseño jerárquico de arquitectura de red de sensores inalámbricos para aplicaciones en agricultura y monitoreo ambiental* de la UNSL, personal especialista del INTA y productores asesores de la zona.

Desde el ámbito de Base de Datos, se trabaja en el diseño y desarrollo de una plataforma soporte de diversos eventos y sistemas de información orientada al seguimiento espacio temporal de individuos y rodeos y gestión espacial de los elementos que integran un establecimiento agropecuario [10]. Las características principales son su disponibilidad en la web con accesibilidad mediante tecnología móvil y de escritorio, la posibilidad de recolección de datos de diferentes fuentes, su almacenamiento y compartición, la integración progresiva de diversas funcionalidades, la explotación y la visualización de información, capacidades de generación de reportes, estadísticas, mensajerías, entre otros.

Por tanto, en este artículo presentamos el desarrollo de una herramienta de apoyo a la gestión en establecimientos agropecuarios mediante el uso de las TIC, modelos de Base de Datos Avanzados y recursos de la Geometría Computacional.

Esta presentación está organizada de la siguiente manera. En la Sección 2 se presenta la Arquitectura de la plataforma junto con su descripción correspondiente. Luego, en la Sección 3, se muestra el marco teórico y de aplicación de las Bases de Datos Avanzadas (Espaciales y Espacio Temporales) y las herramientas de Geometría Computacional (Cierre Convexo, Diagrama de Voronoi, Triangulaciones, entre otras) que definen funcionalidades en la plataforma. En la Sección 4, se presenta la aplicación (móvil y de escritorio) y funcionalidades de la misma. Finalmente, se brindan las conclusiones y visión de futuro.

## 2. Arquitectura de la plataforma

La arquitectura de la plataforma presentada a continuación representa el diseño del nivel superior de la estructura de un sistema. El mismo está pensado como un desarrollo escalar e integral respecto de las funcionalidades y administración de bases de datos heterogéneas.

Si sólo se considerara una arquitectura de tres capas, se tendrían las capas de Presentación, de Lógica de Negocio y de Datos [7]. Este modelo tiene características útiles a los efectos del diseño y desarrollo de la plataforma propuesta para la gestión agropecuaria. Pero, en el contexto del Proyecto Campo Conectado, la captura de los datos tiene un rol protagónico, de diversas complejidades, que amerita distintos protocolos y, en consecuencia, conforman una cuarta capa denominada Captura.

Por tanto, esta plataforma conlleva a una arquitectura de múltiples capas. A continuación se describen cada una de ellas.

- *Presentación.* Consiste de una interfaz de usuario con diferentes características: (i) gráficas, a los efectos de visualizar el mapa del campo, con las correspondientes componentes de tratamiento y las consultas sobre las Bases de Datos Avanzadas; (ii) estadísticas, a los efectos de resumir en gráficos estadísticos la información relevante, de apoyo a la gestión agropecuaria; y (iii) de mensajería, a los efectos de enviar avisos, alarmas o mensajes de interés al usuario, que contribuyan a la toma de decisiones.
- *Lógica de Negocio.* Considerando las diferentes funcionalidades, está prevista la explotación de la información en pos de dar respuestas a los requerimientos

del usuario. En una personalización, con datos reales obtenidos del accionar diario de un establecimiento agropecuario, está prevista la aplicación de minería de datos sobre tipos tradicionales y complejos.

- *Datos.* Se administran bases de datos de diversa naturaleza y tecnología, con acceso y conectividad entre los diferentes repositorios. De esta forma se integran y se da respuesta a variados requerimientos y descubrimiento de información. Los modelos de bases de datos avanzados (bases de datos espaciales, temporales, espacio temporales, multimedia, entre otras) tienen un rol protagónico. Los objetos de estudio requieren almacenar datos actuales e históricos acerca de su variabilidad en el tiempo y en su forma.
- *Captura.* Se establecen las comunicaciones y sincronizaciones con los diferentes dispositivos electrónicos de recolección de datos, los cuales deben estar instalados en el establecimiento agropecuario. Consiste sustancialmente de un canal de comunicación, que recibe datos en diferentes formatos con frecuencias determinadas, generados y adquiridos a través de diversas fuentes de emisión y recepción. En este sentido, se requiere de una provisión de infraestructura en el establecimiento agropecuario, que permita mantener un canal de comunicación para recibir y transmitir los más variados tipos y cantidades de datos.

### 3. Bases de Datos Avanzadas y Geometría Computacional

Los requerimientos de aplicaciones actuales de Bases de Datos en el contexto del Agro imponen la necesidad de representar y operar con tipos de datos tradicionales y complejos.

En particular, para los objetos de tratamiento considerados en este dominio de aplicación, se requiere almacenar y consultar información histórica y actual, relacionada a la posición espacial y/o a los cambios de forma o posición que tuvieron los objetos de estudio en diferentes escenarios a lo largo del tiempo.

Respecto de la forma de un objeto complejo se puede modelar y representar como un objeto geométrico: punto, poligonal o polígono; y respecto de su posicionamiento espacial se requiere de un sistema de coordenadas como marco referencial [7,12,13]. Para el almacenamiento y tratamiento de objetos complejos se utilizaron Bases de Datos Espaciales y Espacio Temporales. Para los datos de tipo tradicional, se utilizan Bases de Datos Relacionales.

#### 3.1. Bases de Datos Espaciales

El objeto espacial tiene asociado una forma geométrica y una posición espacial, invariantes en el tiempo. Ejemplos de objetos espaciales en este dominio de aplicación pueden ser bebederos, parcelas de tierra, comederos, antenas, montes, pozos de agua, entre otros. En la plataforma se utilizan diversos R-Tree [11] para la indexación de dichos objetos, que en una personalización se definirán cuáles objetos espaciales serán de interés.

R-Tree es el método de acceso estándar para las bases de datos espaciales, utilizado por los Sistemas de Administración de Bases de Datos. Es un árbol balanceado en altura, basado en el B-Tree [3], donde cada nodo almacena mínimos rectángulos envolventes (MBR, *Minimum Bounding Rectangle*). Los nodos hoja contienen los MBR de los objetos espaciales y enlaces a la base de datos espacial donde se encuentra descrita la geometría y referencia geoespacial. Los nodos internos guardan los MBR que contienen a sus hijos. En el procedimiento de búsqueda, se considera una ventana espacial de consulta, por lo que, se desciende el árbol desde la raíz, siguiendo por los hijos cuyo MBR interseca con la ventana de consulta, y así sucesivamente en forma recursiva, hasta llegar a las hojas. Para mayor detalle del funcionamiento de la estructura, ver [11].

### 3.2. Bases de Datos Espacio Temporales

El objeto espacio temporal tiene asociado una forma geométrica y una posición espacial, las cuales pueden cambiar a lo largo del tiempo. En este dominio de aplicación, objetos espacio temporales pueden representar en términos de individuo a un animal o en términos grupales a un rodeo. El rodeo o el individuo son considerados objetos espacio temporales (polígono o punto, respectivamente), que evolucionan de manera continua, por lo que es necesario almacenar su estado en cada instante de tiempo.

Dadas las complejidades que surgen intrínsecamente de la administración de este tipo de dato, la evolución se registra discretamente respecto del tiempo y del espacio. Cuando se produce el cambio de un objeto espacial en forma y/o posición, se almacena en la base de datos el nuevo estado del objeto en el instante de tiempo correspondiente. Para el caso del sistema de monitoreo de la plataforma, un individuo informa con cierta periodicidad cuál es su nueva posición; de esta manera, el sistema mantiene un registro discreto de la trayectoria espacial de cada individuo.

Debido a la componente temporal, se deben manejar grandes cantidades de datos registrados durante un largo periodo de tiempo. Por tanto, se requiere de índices de acceso sobre los datos a fin de responder eficientemente las consultas espacio temporales [5,6,8,9]. En esta plataforma se utiliza el D\*R-Tree [6], el cual es un método de acceso espacio temporal de tipo histórico y actual. Mantiene puntos de referencia para ciertos instantes de tiempo donde se almacena la información espacial de los objetos en un R-Tree. Todos los cambios ocurridos entre dos puntos temporales de referencia consecutivos se almacenan en una lista llamada *bitácora*. Este método de acceso está diseñado para responder eficiente e integralmente las consultas espacio temporales más comunes: *Instante*, *Intervalo*, *Evento* y *Trajectory*.

Las de Instante e Intervalo de tiempo son consultas que recuperan todos los objetos que se encuentran en una región demarcada en el espacio en un instante o en un intervalo de tiempo específico. En el caso del sistema de monitoreo de la plataforma, se recuperan los individuos que se encuentran en determinada parcela o región, en un instante o periodo de tiempo.

La consulta Evento recupera los eventos sucedidos en una región en un instante de tiempo determinado. Los eventos son las apariciones o desapariciones de los objetos espacio temporales en una región en cierto instante de tiempo. Sólo retorna aquellos objetos en los cuales su tiempo inicial o tiempo final de entrada/salida al área coincide con el tiempo de la consulta en cuestión. Por ejemplo, consiste en recuperar los individuos que ingresaron o salieron de una parcela o región determinada en un momento dado. Otro ejemplo, si se establece como región de consulta el área del bebedero, se pueden conocer los individuos que se acercaron o se retiraron del mismo en un instante de tiempo.

La consulta Trayectoria permite recuperar el recorrido de posiciones espaciales a lo largo del tiempo en las que un objeto ha permanecido. Por ejemplo, se puede recuperar la ruta recorrida por un individuo o un rodeo, en un intervalo de tiempo dado. En el estudio del comportamiento natural de las caminatas de los individuos durante un periodo de tiempo, se pueden considerar indicadores de hábitos alimenticios y físicos, que determinen su estatus sanitario.

### 3.3. Geometría Computacional

La Geometría Computacional se ocupa de resolver problemas geométricos de modo constructivo. Se interesa por demostrar la existencia de la solución de un problema y por encontrar los algoritmos y estructuras de datos eficientes, medidos respecto de su complejidad en tiempo y espacio respectivamente [4]. En ocasiones, brinda soluciones eficientes en problemas cuyos datos verifican propiedades geométricas, que permite aplicar técnicas algorítmicas o utilizar estructuras geométricas.

En relación al sistema de monitoreo, se pueden enviar reportes al productor acerca del distanciamiento físico entre los individuos o rodeos, su acercamiento a un monte o zona de pastura, a un bebedero o a una zona de particular atención. También, si siguen a un individuo guía, si se mueven en manada o individualmente, entre otros informes. Esto se obtiene utilizando herramientas que brinda la Geometría Computacional, tales como, diámetros y calibres, grafos de proximidad, visibilidad [4].

Dados los objetos espaciales, tales como bebederos o comederos, es posible que se desee saber dónde ubicar un servicio de atención para los individuos que, dependiendo de las características del servicio, se pretenda minimizar o maximizar la distancia de acercamiento. Por ejemplo, si se desea colocar un servicio de comedero entonces se requiere minimizar las distancias; o en el caso de existir un vertedero incinerador de residuos domésticos o una zona de pastizal perjudicial para la salud de los rodeos, se deben maximizar las distancias de los individuos al vertedero o región.

El cálculo del Cierre Convexo de un conjunto de puntos permite encontrar el polígono convexo más pequeño que contiene a todos los puntos [4]. Si los puntos modelizan los individuos del rodeo, al utilizar el Cierre Convexo se obtiene el diámetro y la anchura de un conjunto, y en consecuencia se pueden calcular medidas de dispersión y trayectoria del rodeo.

El Diagrama de Voronoi de un conjunto de puntos es una subdivisión del plano en un conjunto de regiones de modo que, cada región de Voronoi contiene todas las posiciones más cercanas a esa región que a cualquier otro punto [4]. Si se consideran los bebederos como los puntos (pivotes) del Diagrama de Voronoi, cada región representa la zona de influencia de dicho bebedero y presupone su uso por parte de los individuos de rodeo posicionados en dichas áreas. Si se considera al bebedero como pivote de región, se pueden simular distintos emplazamientos a efectos de optimizar sus ubicaciones en función de homogeneizar las áreas de influencia.

La triangulación de un conjunto de puntos es un conjunto maximal de segmentos cuyos extremos son los puntos, tales que, dos segmentos cualesquiera no se cortan en puntos interiores de ellos [4]. Las triangulaciones de conjuntos de puntos se utilizan para el cálculo del área de parcelas de terreno de forma irregular, en el problema de visibilidad desde el mínimo número de puntos posible, en el cálculo de otras estructuras (árboles y grafos). Las triangulaciones de Delaunay se emplean a menudo para diseñar redes en las que los nodos más próximos estén conectados entre sí. De esta manera se pueden diseñar redes de comunicaciones debido a que es un problema de optimización geométrica.

#### 4. Presentación de la plataforma

En el desarrollo de esta plataforma se separó cada componente en diferentes capas, como se mencionó en la Sección 2 que se trata de múltiples capas, las cuales se implementaron de manera independiente a fin de lograr escalabilidad y encapsulamiento. Luego, para coordinar la recepción de diferentes fuentes de datos y su procesamiento, se trabaja en modo cliente - servidor. El servidor es el encargado de almacenar los datos recibidos a partir de diferentes fuentes en estructuras adecuadas y, además, de realizar el procesamiento de los mismos para generar información relevante para el usuario. El cliente resulta en una aplicación móvil o de escritorio, que en tiempo real o diferido, puede mostrar dicha información. En la Figura 1a se muestra la aplicación de escritorio y en la 1b la aplicación móvil.

El diseño de la capa Presentación considera que la ubicación de los individuos del rodeo pueda visualizarse en tiempo real. La interfaz principal muestra, con simplicidad, las diferentes acciones que un usuario puede realizar, información y navegación sobre el sistema. En la aplicación móvil, se tiene la posibilidad de deslizar el área de visualización, lo cual permite, en tiempo real, visualizar la actividad del establecimiento agropecuario.

En particular, la personalización de la plataforma incluirá el descubrimiento de conocimientos que sean seguros y confiables, para brindar soluciones según los intereses del productor agropecuario. Para ello se prevé el uso de técnicas de Data Mining [13]. La interfaz presenta las siguientes secciones de trabajo: Consultas, Opciones y Estado del campo, las cuales se detallan a continuación.

**Consultas.** En esta sección, el usuario puede realizar consultas, sobre los atributos tradicionales o aquellos de tipo complejo, como los espaciales o espa-

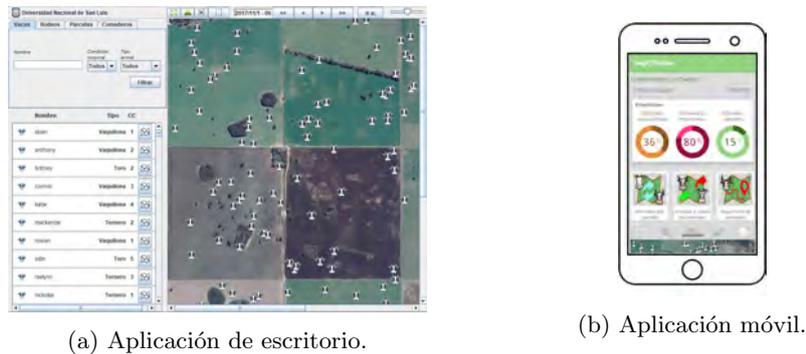


Figura 1: Interfaces de usuario.

cio temporales. También, puede hacer uso de la aplicación de herramientas de Geometría Computacional. Mediante el uso de web services se puede acceder a las bases de datos que integran la plataforma, las cuales están almacenadas en un servidor, enviar y recibir la información solicitada. Respecto de las consultas de tipo complejo mencionadas para las bases de datos espaciales o espacio temporales, se muestran a continuación descripción e imágenes de las mismas.

Los elementos que mantienen una posición fija dentro del establecimiento, tales como parcelas, bebederos y antenas, están almacenados en base de datos espaciales y se utilizan índices R-Tree para realizar las consultas de tipo espacial. Estas consultas se pueden relacionar y obtener información de los objetos espaciales (ensamble espacial, *spatial join*) [13,12]. En particular en la aplicación, se pueden consultar las antenas que están dentro de una parcela indicada por el usuario, como se muestra en la Figura 2.



Figura 2: Consulta espacial. Tipo Ventana.

Para encontrar el polígono más pequeño que encierra a todos los individuos de un rodeo, esta plataforma cuenta con herramientas de Geometría Computacional, tal como el Cierre Convexo. En el ejemplo de la Figura 3a se muestra

el cierre convexo del rodeo en el instante de tiempo  $T1$ . Si se considera un periodo de tiempo ( $T1$ ,  $T2$ ,  $T3$ ), y en cada instante se obtiene el cierre convexo del rodeo, entonces la secuencia de imágenes (Figura 3) muestra cómo el rodeo en conjunto se movilizó en el espacio. Es decir, se puede trazar la trayectoria del rodeo, y de este modo se da respuesta a una consulta espacio temporal de trayectoria sobre objetos espacio temporales que representan los rodeos.

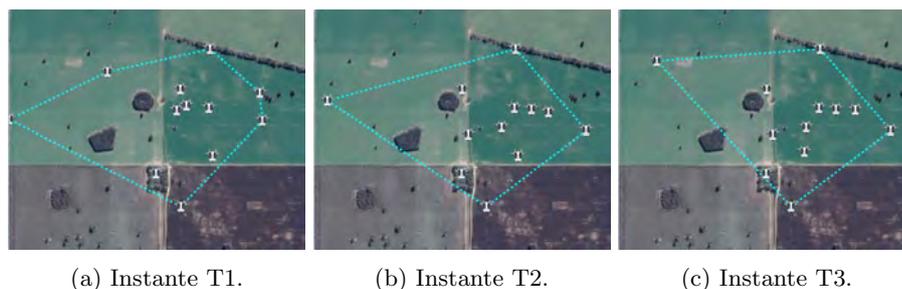


Figura 3: Consulta Espacio Temporal. Cierres Convexos del rodeo en distintos instantes de tiempo.

Dentro de un establecimiento se encuentran objetos de tratamiento ubicados en puntos fijos o con poco movimiento, como es el caso de los bebederos (o comederos). En la plataforma se pueden determinar las zonas más cercanas a estos objetos (modelados como puntos) mediante la utilización de herramientas de Geometría Computacional, en particular, el Diagrama de Voronoi. En la Figura 4a se muestra con línea punteada amarilla las regiones que delimitan las posiciones más cercanas a esa región que a cualquier otro bebedero (o comedero).

La Triangulación de Delaunay puede ser utilizada para trazar redes en la que los nodos más próximos están conectados entre sí. En la plataforma se utiliza para indicar las distancias más próximas entre antenas de telecomunicación ubicadas en el establecimiento. En el ejemplo de la Figura 4b, las antenas son representadas como nodos de la red.

La consulta *Individuos por parcela* en un intervalo o instante de tiempo (fecha y hora inicial, fecha y hora final) retorna todos los individuos que estuvieron en una parcela dada durante un intervalo o instante de tiempo determinado. En el ejemplo de la Figura 5a, se muestran los individuos que estuvieron en la parcela seleccionada en el intervalo indicado (resaltados con círculos amarillos).

La consulta *Entrada y salida de individuos* retorna todos los individuos que ingresaron (resaltados con círculos verdes) y los que salieron (resaltados con círculos rojos) de una parcela seleccionada para un momento determinado (fecha y hora). En el ejemplo de la Figura 5b se muestran dos individuos: uno que ingresó (círculo verde) y otro que salió (círculo rojo) de la parcela seleccionada en el momento elegido.



(a) Zonas de puntos más cercanos a cada comedero. (b) Distancias más cortas entre antenas de telecomunicación.

Figura 4: Herramientas de Geometría Computacional.

La consulta *Trayectoria de individuos* retorna el recorrido de un individuo durante un intervalo de tiempo. En el ejemplo de la Figura 5c se muestra el recorrido de un individuo en un determinado periodo de tiempo (flechas naranjas).



(a) Consulta *Individuos por parcela* (b) Consulta *Entrada y salida de individuos* (c) Consulta *Trayectoria de individuos*

Figura 5: Consultas Espacio Temporales.

**Opciones.** La aplicación cuenta con una barra de opciones, para la actualización de los datos, la búsqueda de un individuo particular, la configuración de opciones y la consulta de alertas y mensajes emitidos por el servidor.

**Estado del campo.** La aplicación muestra el estado actual del campo y sus individuos. Mediante el deslizamiento de la barra que se encuentra en la parte superior de esta sección se pueden reproducir los movimientos históricos de cada individuo.

## 5. Conclusiones y visión de futuro

A través de este desarrollo I+D+i se pudo realizar un proceso de diseño y desarrollo de una plataforma para la integración de TIC orientada al Agro, a fin

de alcanzar los objetivos mencionados. Como parte de la propuesta innovadora, la aplicación de Modelos Avanzados de Bases de Datos y Geometría Computacional, permitió hallar soluciones en los problemas planteados en el dominio de aplicación.

Como visión de futuro se espera detectar, evaluar y analizar indicadores de gestión y para la toma de decisiones, evaluar y analizar los resultados obtenidos, la pertinencia y la transferencia de los proyectos I+D+i en las áreas de intervención seleccionadas, consolidar el ámbito interdisciplinario de formación y fortalecimiento de competencias, como así también considerar líneas de trabajo futuro, con nuevos actores, proyectos y objetivos.

## Referencias

1. AgroTIC. <http://www.agrotic.info> (2016)
2. Campo Conectado. <http://inta.gob.ar/noticias/campo-conectado-un-nuevo-proyecto-interinstitucional> (2017)
3. Bayer, R., McCreight, E.: Organization and maintenance of large ordered indices. In: Proceedings of the 1970 ACM SIGFIDET (Now SIGMOD) Workshop on Data Description, Access and Control. p. 107–141. SIGFIDET '70, Association for Computing Machinery, New York, NY, USA (1970), <https://doi.org/10.1145/1734663.1734671>
4. Berg, M.d., Cheong, O., Kreveld, M.v., Overmars, M.: Computational Geometry: Algorithms and Applications. Springer-Verlag, Santa Clara, CA, USA, 3rd edn. (2008)
5. Carrasco, F.D., Gagliardi, E.O., García Sosa, J.C., Gutierrez, G.: Una propuesta de un método de acceso espacio-temporal: I+ 3 R-Tree. In: Congreso Argentino de Ciencias de la Computación (2006)
6. Dorzán, M., Gagliardi, E., Gómez Barroso, J., Gutiérrez Retamal, G.: Un nuevo índice eficiente para resolver diversas consultas espacio-temporales. In: Conferencia Latinoamericana de Informática (2006)
7. Elmasri, R., Navathe, S.: Fundamentals of Database Systems. Addison-Wesley Publishing Company, USA, 6th edn. (2010)
8. Gagliardi, E.O., Carrasco, F.D., García Sosa, J.C.: I+ 3 R-Tree: un método de acceso espacio-temporal. In: XV Congreso Argentino de Ciencias de la Computación (2009)
9. Gagliardi, E.O., Dorzán, M.G., Gómez Barroso, J.G., Gutiérrez Retamal, G.A.: D\* R-Tree: un método eficiente para responder consultas espacio-temporales. In: XII Congreso Argentino de Ciencias de la Computación (2006)
10. Gagliardi, E.O., Dorzán, M.G., Taranilla, M.T., Palmero, P.R., Casanova, C.: Geoseguimiento de rodeos, hacia una plataforma integral para el agro. In: 47JAIIO - Congreso de AgroInformática (2018)
11. Guttman, A.: R-trees: A dynamic index structure for spatial searching. In: Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data. pp. 47–57. SIGMOD '84 (1984)
12. Manolopoulos, Y., Papadopoulos, A., Vassilakopoulos, M.: Spatial Databases: Technologies, Techniques and Trends. Idea Group (2005)
13. Shekhar, S., Chawla, S.: Spatial databases - a tour. Prentice Hall (2003)

# Conducting a Systematic Review: Trends in Machine Learning and Text Mining <sup>\*</sup>

Mariana Falco<sup>1</sup>[1234–5678–9012] and Ignacio Berdiñas<sup>2</sup>

<sup>1</sup> LIDTUA/CONICET, Engineering School, Austral University; Mariano Acosta 1611, Pilar, Buenos Aires, Argentina [mfalco@austral.edu.ar](mailto:mfalco@austral.edu.ar)

<sup>2</sup> Engineering School, Austral University; Mariano Acosta 1611, Pilar, Buenos Aires, Argentina [ignacio.berdiñas@ing.austral.edu.ar](mailto:ignacio.berdiñas@ing.austral.edu.ar)

**Abstract.** The main goal of a Systematic Review is to identify, evaluate, and summarize the findings of all relevant studies over a topic or an issue, making the evidence accessible to decision makers. But the process of manually conducting a systematic reviews takes a lot of time and researchers often have to limit their procedures. With the recent technological advantages, machine learning (ML) and text mining (TM) became useful to aid the systematic review process. The objective of this study is to detect the main trends of these disciplines by carrying out an analysis of a set of relevant articles, identified with a scientific database search between 2015 and 2020. Our analysis showed that mostly ML and TM techniques were applied to three steps: search, screening and data extraction. Huge progresses have been made over the years, but full automation remains a distant goal at present.

**Keywords:** Systematic Reviews · Literature Reviews · Machine Learning · Text Mining.

## 1 Introduction

A Systematic Review (SR) can be defined as a method that allows to give meaning and identity to a large amount of information with a clear stated purpose, usually in the form of research studies [1] to answer a set of research questions combining evidence found on those studies [2]. The importance of SRs is that their main goal is to identify, evaluate, and summarize the findings of all relevant studies over a topic or issue, making the resulting evidence more accessible to decision makers [3].

Nevertheless, researchers are forced to limit their search procedures due to the time it takes to conduct a proper systematic review [4]. As pointed out by Zachary [5] reviewers regularly identify relevant searches by performing extensive searches and scanning contents, citations and references. Manually conducting a systematic review is no longer sustainable because practitioners and researchers

---

<sup>\*</sup> Supported by a Research Grant from Universidad Austral

use enormous amount of time to perform the tasks of searching, screening, mapping and synthesizing within the process, reducing the hours to explode the creativity.

The process of systematic reviewing includes the following stages: searching, screening, mapping and synthesizing [7]. Text Mining makes possible to analyze collections of textual materials, in order to identify key concepts uncovering also hidden relationships within concepts, allowing the users to efficiently discover, interpret and curate knowledge [6]. Cohen et al. [8] demonstrated that machine learning techniques can reduce the labor required to update systematic reviews. Different authors have produced systematic reviews or reviews, but there are three to mention: Jonnalagadda et al. [9] analyzed methods to automate data extraction, Omara et al. [10] studied the screening phase, and Feng et al. [11] identified and classified text-mining techniques and tools to facilitate conducted a SLR, mostly focused on the SE domain. The biggest difference is that the mentioned articles carried out their analysis considering the studies published from 2014 downward.

The main objective of the present paper is to conduct an analysis of the main contributions of machine learning and text mining for each step within a systematic review process, from 2015 to June 2020 in the point of view of practitioners and researchers, in the industrial and academic context. In order to identify the contributions as well as the existing trends, we performed a comprehensive study based on searches on ACM, IEEE, Springer and Science Direct databases. The goal is to obtain trends over the past five years of technological progresses. The reminder of this paper is as follows: Section 2 presents the followed methodology while Section 3 summarizes the results obtained and theorems defined. Finally, Section 4 provides the conclusions and future work.

## 2 Methodological Process

Our methodology consisted on the following tasks: define the main goal, perform a search on scientific digital libraries, analyze the search results, identify relevant articles and extract data from these relevant articles.

As stated in the Introduction, our goal can be written as the following research question: *What are the main contributions to each step of conducting a systematic review, between 2015 and 2020?* Also, the selected digital databases were ACM, IEEE, Springer and Science Direct. In order to make a feasible search, we considered the term extraction which improve the search strategy by creating metadata that can improve its accuracy [7]. In our case, we used the TerMine service which automatically extracts and ranks technical terms.

In this context, the identification of relevant articles were made with the inclusion criteria. Articles published between 2015 and June 2020 were included if they fulfilled the following topics: (a) the title and/or the keywords should contain "text mining applied to systematic reviews" or "machine learning applied to systematic reviews" or similar phrases, and (b) studies which were directed related to text mining and machine learning techniques, approaches, and imple-

mentations to help the process of a systematic review. Papers on the following topics were excluded: (a) studies on a non-English language; (b) duplicate or updated studies (we selected the most recent one); (c) journals with low impact factors; (d) books; (e) extended abstracts; (f) technical reports; (g) doctoral dissertations; (h) thesis, and (i) non-direct application of TX or ML techniques to improve a step within the systematic review process.

From the searches, we obtained 105310 results while querying on ACM, 3005 results on Science Direct, 10 results on IEEE and 1920 results on Springer. After applying the inclusion and exclusion criteria defined above, we gathered the following studies per each database: eight (8) studies from ACM, five (5) studies from Springer, and three (3) studies from Science Direct summarizing sixteen (16) studies to be analyzed.

### 3 Description of Results and Theorems

As a general overview, we have identified the following publication years: 2015, 2016, 2018, 2019, and 2020; where the years with the highest amount of published studies are 2016 and 2018, followed by 2020, where the latter is still current so new contributions can be added in the following months.

**Theorem 1.** *The studies were mostly published in journals (8 studies) and conferences (6 studies).*

The studies were mainly published in journals and conferences, while only 2 studies were extracted from symposiums. While the conferences showed dispersion counting six different of them, the journals showed a more fixed distribution including three studies in *Systematic reviews*, and three in *Journal of biomedical informatics*. This is due to the fact that systematic reviews are being performed for different domains, and depending with the main goal of the study as well as the application field, the decision for the publication venue is made around this ideas.

**Theorem 2.** *The two journals with published analyzed studies were Systematic reviews, and three in Journal of biomedical informatics.*

**Theorem 3.** *The main steps covered by the studies are search, screening and data extraction.*

**Theorem 4.** *The step with the higher amount of contributions between 2015 and June 2020 is the screening step.*

Theorem 3 and 4 are condensed in Table 1, which describes the amount of contributions per each step while conducting a systematic review.

**Theorem 5. Search** *The studies proposed a set of approaches that includes: iterative methods to build the search string, an automatic query formulation, and an automated approach to extend a search.*

**Table 1.** Summary of findings.

Step	Studies	References
Search	5	(Cairo et al., 2019), (Lanera et al., 2018), (Marcos-Pablos et al., 2018), (Mergel et al., 2015), (Scells et al., 2020)
Screening	10	(Bannach-Brown et al., 2019), (Howard et al., 2016), (Hashimoto et al., 2016), (Kontonatsios et al., 2020), (Lee and Sun, 2018), (Lee et al., 2020), (Ouhbi et al., 2016), (Sellak et al., 2015), (Tsafnat et al., 2018)
Data extraction	3	(Blake and Lucic, 2015), (Bui et al., 2016), (Chatterjee et al., 2017)

**Theorem 6.** *Within the screening step, there are several contributions such as high-performing algorithms, approaches to active learning, new topic detection method, automatic text classification approach, approaches for semi-automating screening, and screening systems like SWIFT-Review [17] and SLR Toolkit [25].*

**Theorem 7.** *In the data extraction step, extraction technologies are still in formative stages.*

**Theorem 8.** *Methods for automating are still far away for current capacities of machine learning and text mining tools [26].*

### 3.1 Understanding the techniques

Table 2 summarizes the analyzed articles with respect to the technique used per author and per step.

Table 2: Extracted Techniques Applied on the Relevant Studies

Begin of Table			
Step	Year	Technique used	Evaluation-Performance
Search	2019	TF-IDF, CBOW and SkipGram	Recall and Workload
	2018	TF-IDF, Support vector machine	Area under the receiver operator characteristic curve (AUC)
	2018	TF-IDF, Multinomial Naive Bayes, Bernouli Naive Bayes, k-Nearest Neighbors, Support Vector Machines	F1 measure
	2015 2020	TF-IDF, heatmap process to create binary questions	Tool analysis with users Recall, F1, F3 and Work Saved Over Sampling (WSS)

Title Suppressed Due to Excessive Length 5

Continuation of Table 2			
Step	Year	Technique used	Evaluation-Performance
Screening	2019	TF-IDF, LDA, SVM	Recall, specificity, precision, accuracy, WSS@95%, positive likelihood ratio
	2016	new topic detection model	yield, burden
	2020	novel neural network-based feature extraction method	Work Saved Over Sampling (WSS) at r% recall (WSS@r%)
	2018	seed-driven document ranking (SDR) model	Average precision, precision, and recall
	2020	multi-modal missing Data aware stacked auto-encoder	Work Saved Over Sampling (WSS)
	2016	TF-IDF, LDA, Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm	Work Saved over Sampling (WSS)
	2016	new algorithm called Rules7-hybrid feature selection (Rules7-HFSRM)	Precision, recall
	2015	novel Hybrid Feature Selection Method (HFSM) within a Class Association Rules (CARs) algorithm	Precision, recall
	2018	The algorithm was developed using the General Architecture for Text Engineering (GATE)	Precision, recall
Search	2015	Endpoint detection	Precision, recall
	2016	multi-pass sieve algorithm	Accuracy, recall, precision, F-measure
	2017	set of heuristics	Precision, recall
End of Table			

As shown in Table 2, we have found different techniques within the studies, where the most applied is TF-IDF (Term Frequency-Inverse Document Frequency) which is a statistical measure used to evaluate how important a word is to a document in a collection or corpus. The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus. By means of TF-IDF, knowledge of the research domain is expanded and improved [12].

**Theorem 9.** *The most applied technique is TF-IDF (Term Frequency-Inverse Document Frequency).*

The following items organize the techniques discovered in the studies.

- TF-IDF: six (6) studies.

6 Falco and Berdiñas

- *Latent Dirichlet allocation (LDA), Support Vector Machines (SVM) and Hybrid Feature Selection Method (HFSM)*: two (2) studies per each.
- *Singular Value Decomposition (SVD), CBOW (Continuous Bag of Words), SkipGram, Multinomial Naive Bayes, Bernouli Naive Bayes, k-Nearest Neighbors, Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm, active learning, Heatmap, Neural network, and Seed-driven document ranking*: one (1) study per each.

With respect to evaluation and performance metrics, the most applied where, on the one hand, Precision, or the positive predictive value, refers to the fraction of relevant instances among the total retrieved instances; and on the other hand, Recall, also known as sensitivity, refers to the fraction of relevant instances retrieved over the total amount of relevant instances. In short, precision and recall are measurements of relevance.

**Theorem 10.** *The most applied evaluation and performance metrics were Precision and Recall.*

- *Recall*: nine (9) studies.
- *Precision*: seven (7) studies.
- *Work Saved over Sampling (WSS)*: four (4) studies.
- *F-measure*: four (4) studies.
- *Accuracy, Workload, WSS@95%, and Specificity*: two (2) studies per each.
- *Average precision, Positive likelihood ratio, Yield, area under the receiver operator characteristic curve (AUC), and Burden*: one (1) study per each.

### 3.2 Search Step

**Techniques** Following Table 2, devising an appropriate search string for a secondary study is not a trivial task and identifying suitable keywords has been reported in the literature as a difficulty faced by researchers. In this context, the ML algorithm TF-IDF was applied to different approaches for search string construction [12–14], and to extend a search on PubMed to clinical trials, as well a cross-validated support-vector machine (SVM) model as the classifier [15]. Also, the search string studies applied CBOW (Continuous Bag of Words) and SkipGram in [12]; Multinomial Naive Bayes, Bernouli Naive Bayes, and k-Nearest Neighbors in [13]. Mergel and others [14] have applied the heatmap for visualizing differences between features.

**Evaluation and Performance Metrics** The most applied performance metric is F1 [13, 22]. Cairo and others [12] used recall and workload, where the latter is used to measure workload in SRs, in the task of search strings. Also, Scells and others [22] measured recall, F1, F3, and Work Saved Over Sampling (WSS), while proposing a five-step approach to automatic query formulation, specific to boolean queries.

### 3.3 Screening Step

**Techniques** TF-IDF was used by Bannach-Brown and others [16] to identify potential errors made during the human screening process (including also Latent Dirichlet allocation (LDA), Support Vector Machines (SVM), and Singular Value Decomposition (SVD)), and by Howard and others [17] while introducing the characteristics of SWIFT (Sciome Workbench for Interactive computer-Facilitated Text-mining) a workbench to assist in the problem formulation and literature prioritization, which also includes LDA and the Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm for document prioritization [17].

Hashimoto and others [18] used a neural network-based vector space model to capture semantic similarities between documents; representing documents within the vector space, and cluster the documents into a predefined number of clusters. Ouhbi and others [19] proposed a new algorithm called Rules7-hybrid feature selection (Rules7-HFSRM) by combining the classical algorithm Rules7 and the Hybrid Feature election measure (HFSRM), for text classification. Also, Sellak and others [20] contributed to this line of work by proposing an alternative approach, not yet tested in this domain based on semantic rule-based classifiers. This approach involved applying a novel Hybrid Feature Selection Method (HFSM) within a Class Association Rules (CARs) algorithm.

**Evaluation and Performance Metrics** Bannach-Brown and others [16] assessed performance using recall (or sensitivity), specificity, precision, accuracy, WSS and the Positive likelihood ratio (LR+). They have obtained that the ML approaches reached 98.7% sensitivity based on learning from a training set of 5749 records, with an inclusion prevalence of 13.2%. The highest level of specificity reached was 86%. Hashimoto and others [18] evaluated performance of the active learning process, over different learning iterations, using two metrics, namely Yield (percentage of eligible studies identified by the active learner), and Burden (percentage of studies that are manually labelled). Kontonatsios and others [23] used WSS, recall and WSS@95%. The proposed method outperforms 10 baseline feature extraction methods by approximately 6% in terms of the WSS@95% metric.

### 3.4 Data Extraction

**Techniques** In the experiments performed by Blake and Lucic [21] they used a collection of more than 2 million sentences from three journals Diabetes, Carcinogenesis and Endocrinology and two machine learning algorithms, support vector machines (SVM) and a general linear model (GLM).

**Evaluation and Performance Metrics** In Blake and Lucic [21], F1 and accuracy measures for the SVM and GLM differed by only 0.01 across all three comparison facets in a randomly selected set of test sentences.

## 4 Conclusions

The amount of published studies have been growing over the years, and this volume of work lead to develop methods that aim to semi-automate different steps while conducting a systematic review, including machine learning and text mining techniques. Even though there is no unified methodology for applying this methods, or more than one method is valid, we believe that the contributions made a huge progress toward the semi-automation of the steps of a SR. The use of text mining as a second screener may also be used cautiously. The use of text mining to eliminate studies automatically should be considered promising, but not yet fully proven [10].

The present article performed a database search in order to obtain a set of relevant studies from 2015 to 2020, to find out the trends on machine learning and text mining when these disciplines are applied to each of the steps while conducting a systematic review. In this context, it was possible to define a set of theorems to show some trends within the studies. For example, Theorem 3 describes that the techniques were mostly applied to the steps of search, screening, and data extraction.

Most of the tools we encountered were written by academic groups involved in research of machine learning and text mining techniques, but very often the produced prototype were not fully maintainable or even thought to used for other practitioners. Nonetheless, for the pioneering systematic review team, many of the methods described can be used now. Users should expect to remain fully involved in each step of the review and to deal with some rough edges of the software. Data extraction tools are designed to assist the manual process, e.g. drawing the user's attention to relevant text or making suggestions to the user that they may validate, or change if needed.

As a conclusion, it is possible to point out that SRs require very high accuracy in their methods, which may be difficult for automation to attain. Yet accuracy is not the only barrier to full automation. In areas with a degree of subjectivity (e.g. determining whether a trial is at risk of bias), readers are more likely to be reassured by the subjective but considered opinion of an expert human versus a machine. As a side comment, peer reviewing is a standard process for assessing the quality of submissions at academic conferences and journals. We have found a generalized framework for fair reviewer assignment [24], which has been proved that it is superior to the current state-of-the-art.

As future work, we will perform a deeper analysis of the relevant studies in order to describe the hidden relationships as well as a bigger study of the techniques mentioned. Also, we will use abstractive summarization to conduct an evaluation of the included abstract.

## References

1. Petticrew, M., and Roberts, H. (2006) *Systematic Reviews in the Social Sciences: A Practical Guide*, Oxford: Blackwell.

2. Millard, L. A., Flach, P. A., and Higgins, J. P. (2015). Machine learning to assist risk-of-bias assessments in systematic reviews. *International journal of epidemiology*, 45(1), 266-277.
3. Gopalakrishnan, S., and Ganeshkumar, P. (2013). Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. *Journal of family medicine and primary care*, 2(1), 9.
4. Boland, A., Cherry, G., and Dickson, R. (Eds.). (2017). *Doing a systematic review: A student's guide*. Sage.
5. Zachary, M. A., Gianiodis, P. T., Payne, G. T., and Markman, G. D. (2015). Entry timing: Enduring lessons and future directions. *Journal of Management*, 41(5), 1388-1415.
6. Ananiadou, S., and McNaught, J. *Text mining for biology and biomedicine*. Boston/London: Artech House; 2006
7. Ananiadou, S., Rea, B., Okazaki, N., Procter, R., and Thomas, J. (2009). Supporting systematic reviews using text mining. *Social Science Computer Review*, 27(4), 509-523.
8. Cohen, A., Hersh, W., Peterson, K., and Yen, P.Y.: Reducing Workload in Systematic Review Preparation Using Automated Citation Classification. *J Am Med Inform Assoc* 2006, 13: 206–219. 10.1197/jamia.M1929
9. Jonnalagadda, S. R., Goyal, P., and Huffman, M. D. (2015). Automating data extraction in systematic reviews: a systematic review. *Systematic reviews*, 4(1), 78.
10. O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., and Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic reviews*, 4(1), 5.
11. Feng, L., Chiam, Y. K., and Lo, S. K. (2017, December). Text-mining techniques and tools for systematic literature reviews: A systematic literature review. In *2017 24th Asia-Pacific Software Engineering Conference (APSEC)* (pp. 41-50). IEEE.
12. Cairo, L., de F. Carneiro, G., Monteiro, M. P., and Abreu, F. B. (2019, September). Towards the Use of Machine Learning Algorithms to Enhance the Effectiveness of Search Strings in Secondary Studies. In *Proceedings of the XXXIII Brazilian Symposium on Software Engineering* (pp. 22-26).
13. Marcos-Pablos, S., and García-Peñalvo, F. J. (2018, October). Decision support tools for SLR search string construction. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'18)*, pp. 660-667. DOI:<https://doi.org/10.1145/3284179.3284292>
14. Mergel, G. D., Silveira, M. S., and da Silva, T. S. (2015, April). A method to support search string building in systematic literature reviews through visual text mining. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing (SAC '15)*, pp. 1594-1601. DOI:<https://doi.org/10.1145/2695664.2695902>
15. Lanera, C., Minto, C., Sharma, A., Gregori, D., Berchialla, P., and Baldi, I. (2018). Extending PubMed searches to ClinicalTrials.gov through a machine learning approach for systematic reviews. *Journal of clinical epidemiology*, 103, 22-30.
16. Bannach-Brown, A., Przybyła, P., Thomas, J., Rice, A. S., Ananiadou, S., Liao, J., and Macleod, M. R. (2019). Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. *Systematic reviews*, 8(1), 1-12.
17. Howard, B. E., Phillips, J., Miller, K., Tandon, A., Mav, D., Shah, M. R., ... and Macleod, M. (2016). SWIFT-Review: a text-mining workbench for systematic review. *Systematic reviews*, 5(1), 87.

18. Hashimoto, K., Kontonatsios, G., Miwa, M., and Ananiadou, S. (2016). Topic detection using paragraph vectors to support active learning in systematic reviews. *Journal of biomedical informatics*, 62, 59-65.
19. Ouhbi, B., Kamoune, M., Frikh, B., Moukhtar Zemmouri, E., and Behja, H. 2016. A hybrid feature selection rule measure and its application to systematic review. In *Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services (iiWAS '16)*. Association for Computing Machinery, New York, NY, USA, 106–114. DOI:<https://doi.org/10.1145/3011141.3011177>
20. Sellak, H., Ouhbi, B., and Frikh, B. (2015, December). Using rule-based classifiers in systematic reviews: a semantic class association rules approach. In *Proceedings of the 17th International Conference on Information Integration and Web-based Applications Services (iiWAS '15)*. Association for Computing Machinery, New York, NY, USA, Article 43, 1–5. DOI:<https://doi.org/10.1145/2837185.2837279>
21. Blake, C., and Lucic, A. (2015). Automatic endpoint detection to support the systematic review process. *Journal of biomedical informatics*, 56, 42-56.
22. Scells, H., Zuccon, G., Koopman, B., and Clark, J. (2020, April). Automatic Boolean Query Formulation for Systematic Review Literature Search. In *Proceedings of The Web Conference 2020, (WWW '20)*, Association for Computing Machinery, New York, NY, USA, 1071–1081. DOI:<https://doi.org/10.1145/3366423.3380185>
23. Kontonatsios, G., Spencer, S., Matthew, P., and Korkontzelos, I. (2020). Using a Neural Network-based Feature Extraction Method to Facilitate Citation Screening for Systematic Reviews. *Expert Systems with Applications: X*, 100030.
24. Kou, N. M., U, L. H., Mamoulis, N., and Gong, Z. (2015, May). Weighted coverage based reviewer assignment. In *Proceedings of the 2015 ACM SIGMOD international conference on management of data* (pp. 2031-2046).
25. Götz, S. (2018, October). Supporting systematic literature reviews in computer science: the systematic literature review toolkit. In *Proceedings of the 21st ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings* (pp. 22-26).
26. Marshall, I. J., and Wallace, B. C. (2019). Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Systematic reviews*, 8(1), 163.

# Una Aplicación de Espacios Métricos en el Ámbito del Comercio Electrónico

**Libertad Speranza, María de los Ángeles de la Torre, Norma Herrera**

Departamento de Informática, Universidad Nacional de San Luis, Argentina,  
libertadsperanza@gmail.com, angelesdlt@gmail.com,  
nherrera@unsl.edu.ar

**Abstract.** El comercio electrónico (e-commerce) es un modelo de negocio que gestiona compra y venta de productos y/o servicios, y utiliza Internet como principal medio de intercambio. En este ámbito, los sistemas de recomendación son herramientas utilizadas para generar sugerencias sobre un determinado producto. En este trabajo abordamos el estudio de espacios métricos para la búsqueda de objetos similares en el entorno de un sistema de recomendación. Utilizamos para ello técnicas de indexación basadas en pivotes para resolver de manera eficiente la búsqueda de productos similares. Presentamos un estudio empírico de la eficiencia de estos algoritmos en este caso real de estudio.

**Palabras claves:** Espacios Métricos, Comercio Electrónico, Pivotes.

## 1 Introducción

Las operaciones de búsquedas en una base de datos requieren de algún soporte y organización especial a nivel físico. En el caso de las bases de datos clásicas, la organización de la información se basa en el concepto de búsqueda exacta sobre datos estructurados. Esto significa que la información se organiza en registros con campos completamente comparables (datos estructurados). Una búsqueda en la base retorna todos aquellos registros cuyos campos coinciden con los aportados en la consulta (búsqueda exacta).

Con el crecimiento de la web el volumen de datos e información que se maneja ha crecido exponencialmente, con lo cual las búsquedas se enfrentan a nuevos desafíos de rendimiento y velocidad de respuesta, dejando en segundo plano la congruencia de datos que ofrece el modelo relacional.

El comercio electrónico (e-commerce) es un modelo de negocio que gestiona compra y venta de productos y/o servicios, y utiliza Internet como principal medio de intercambio. Dentro de los desafíos que implica construir un sitio de e-commerce exitoso, sin duda uno de los más importantes es contar con un catálogo de productos bien organizado, donde el cliente no solo encuentre rápidamente lo que busca, sino que los resultados ofrecidos sean relevantes a lo que necesita.

En el campo del e-commerce, a menudo se necesita seleccionar una de varias alternativas sin tener un conocimiento exacto de cada una de ellas. La decisión final, en estos casos, puede depender de las recomendaciones realizadas por el propio sistema. La recomendación de productos a clientes es un caso de recuperación de información que no se puede abordar con la visión clásica de bases de datos relacionales.

En este trabajo abordamos el problema de buscar productos similares a un producto dado, obteniendo de esa manera recomendaciones para los usuarios de un sitio de e-commerce. Para ello aplicamos la teoría de espacios métricos [5], donde nuestro universo de datos será un conjunto de productos disponibles en la plataforma Mercado Libre, y la función que provee la medida de distancia, será 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 es construir un sistema de recomendación basado en el modelo de espacios métricos.

Este artículo está organizado de la siguiente manera: en la sección 2 y 3 describimos el marco de desarrollo de este trabajo dando un reseña de espacios métricos y de comercio electrónico. En la sección 4 presentamos nuestra propuesta, realizando la descripción y evaluación experimental de la misma en las secciones 5 y 6. Finalizamos en la sección 7 dando las conclusiones y el trabajo futuro.

## 2 El Modelo de Espacios Métricos

Nos interesa estudiar el problema de buscar con tolerancia en conjuntos de objetos. Dicho de otra manera, las búsquedas en donde se puedan recuperar objetos *similares* a uno dado. Para ello, abstraemos el problema a un universo de objetos  $\mathcal{X}$  y una función de distancia  $d$  que modela la similitud entre los objetos del universo. Esta función  $d$  cumple con las propiedades características de una función de distancia: *positividad* ( $d(x, y) \geq 0$ ), *simetría* ( $d(x, y) = d(y, x)$ ) y *desigualdad triangular* ( $d(x, y) \leq d(x, z) + d(z, y)$ ). El par  $(\mathcal{X}, d)$  se denomina *espacio métrico*. La base de datos será cualquier subconjunto finito  $\mathcal{U} \subseteq \mathcal{X}$ .

En este modelo de bases de datos, una de las consultas típicas que implica recuperar objetos similares es la *búsqueda por rango*, que denotaremos con  $(q, r)_d$ . Dado un elemento  $q \in \mathcal{X}$ , al que llamaremos *query* y un radio de tolerancia  $r$ , una búsqueda por rango consiste en recuperar los objetos de la base de datos cuya distancia a  $q$  no sea mayor que  $r$ , es decir,  $(q, r)_d = \{u \in \mathcal{U} : d(q, u) \leq r\}$ .

El tiempo total de resolución de una búsqueda contiene tres términos, a saber:  $T = \#evaluaciones\ de\ d \times complejidad(d) + tiempo\ extra\ de\ CPU + tiempo\ de\ I/O$ . En muchas aplicaciones la evaluación de la función  $d$  es tan costosa que las demás componentes de la fórmula anterior pueden ser despreciadas. Éste es el modelo usado en este trabajo; por consiguiente, nuestra medida de complejidad será la cantidad de evaluaciones de la función de distancia  $d$ .

Claramente una búsqueda por similitud puede resolverse de forma ineficiente en tiempo  $O(n)$  examinando exhaustivamente la base de datos  $\mathcal{U}$ . Para evitar esto, se preprocesa  $\mathcal{U}$  y se construye un índice, diseñado para ahorrar cálculos en el momento de resolver una búsqueda. Básicamente, los algoritmos de indexación pueden clasificarse en dos grupos [5, 3, 7, 9, 8]: *basados en pivotes* y *basados en particiones compactas*.

En este trabajo estamos interesados en los algoritmos basados en pivotes. Este grupo de algoritmos construyen el índice basándose en la distancia de los objetos de la base de datos a un conjunto de elementos preseleccionados que llamaremos *pivotes*. Para ello, se seleccionan  $k$  pivotes  $\{p_1, p_2, \dots, p_k\}$ , y se le asigna a cada elemento  $a$  de la base de datos, el vector o firma  $\Phi(a) = (d(a, p_1), d(a, p_2), \dots, d(a, p_k))$ . Durante la búsqueda

se usa la desigualdad triangular junto con la firma de cada elemento para filtrar objetos de la base de datos sin medir su distancia a la query  $q$ . Dada una búsqueda  $(q, r)_d$ , se computa la firma de la query  $q$ ,  $\Phi(q) = (d(q, p_1), d(q, p_2), \dots, d(q, p_k))$ , y luego se descartan todos aquellos elementos  $a$ , tales que para algún pivote  $p_i$  se cumple que  $|d(q, p_i) - d(a, p_i)| > r$ .

Uno de los principales obstáculos en el diseño de buenas técnicas de indexación es lo que se conoce con el nombre de *maldición de la dimensionalidad*. El concepto de dimensionalidad está relacionado a la dificultad o facilidad de buscar en un determinado espacio métrico. La dimensión intrínseca de un espacio métrico se define en [5] como  $\rho = \frac{\mu^2}{2\sigma^2}$ , siendo  $\mu$  y  $\sigma^2$  la media y la varianza respectivamente de su histograma de distancias. Es decir que, a medida que la dimensionalidad intrínseca crece, la media crece y su varianza se reduce. Esto significa que el histograma de distancia se concentra más alrededor de su media, lo que influye negativamente en el desempeño de los índices.

### 3 Comercio Electrónico: Sistemas de Recomendación

La forma más conocida y popular de almacenar información se basa en las relaciones entre los distintos atributos (bases de datos relacionales). En éste tipo de universo, las consultas vienen dadas por una especificación de criterios que deben cumplirse para que un objeto forme parte del resultado; pero ¿qué pasa cuando en la consulta no podemos definir ese conjunto de criterios en forma determinística?

En el campo del e-commerce, a menudo se necesita seleccionar una de varias alternativas sin tener un conocimiento exacto de cada una de ellas. La decisión final, en estos casos, puede depender de las recomendaciones realizadas por el propio sistema. Los sistemas de recomendaciones [1, 6] son herramientas que generan sugerencias sobre un determinado objeto de estudio, a partir de las preferencias y opiniones dadas por los usuarios. El uso de estos sistemas va en aumento debido a que son muy útiles para evaluar y filtrar la gran cantidad de información disponible en la Web. El objetivo final es asistir a los usuarios en sus procesos de búsqueda y recuperación de información.

Un sistema de recomendación es un software que filtra información de interés para el usuario con el fin de proponerle aquel producto más adecuado a sus necesidades. Estos sistemas evalúan cuál es el grado de interés de un usuario por ciertos productos y buscan productos similares y con una alta probabilidad de atraer su atención.

Las recomendaciones pretenden ser una forma de ayudar al usuario a encontrar productos de su agrado realizando una preselección basándose, por ejemplo, en su historial de búsquedas. Esto proporciona un alivio a los consumidores, ya que les evita tener que recorrer una lista interminable de ofertas poco relevantes. Por otro lado, se espera que estos beneficios para los usuarios se terminen reflejando en aumentos de tráfico y ventas, ya que en el e-commerce, las buenas recomendaciones siempre conducen a incrementos en las compras y, por ende, en los márgenes de ganancia. Es por ello que la efectividad del sistema de recomendación elegido es crucial.

El funcionamiento de un sistema de recomendación está basado en cierta información que es procesada por los algoritmos de filtrado. Dependiendo de la naturaleza de dicha información, podemos clasificar estos algoritmos en: basados en contenido, colaborativos, sensibles al contexto, entre otros.

En este trabajo nos hemos centrado en los algoritmos basados en contenido. Estos algoritmos filtran **objetos o contenidos similares** a los que el usuario ya ha buscado, comprado o calificado positivamente. Por ejemplo, en las plataformas de reproducción de música online, el software evalúa las piezas musicales analizando su estructura interna para encontrar piezas similares que podrían tener una línea de bajo parecida.

## 4 Espacios Métricos para un Sistema de Recomendación

En la plataforma de e-commerce Mercado Libre los usuarios publican productos dentro de un árbol de categorías; dichos productos constan de un título principal, un título secundario y una descripción.

El problema que abordamos en este trabajo es la búsqueda de los productos similares a un producto específico, obteniendo de esa manera recomendaciones para los usuarios. En este caso el universo de datos del espacio métrico está formado el conjunto de productos disponibles y la función de distancia será la distancia de edición (o Levenshtein) aplicada al título principal de los productos. Esta función de distancia es discreta y calcula la mínima cantidad de caracteres que hay que agregar, intercambiar y/o eliminar a una palabra para obtener otra. Este modelo es el más comúnmente usado en recuperación de texto.

El objetivo de este trabajo es, teniendo en mente un sistema de recomendación, encontrar una forma eficiente de resolver la búsqueda por similitud evaluando técnicas conocidas en este caso real de aplicación. Para lograr nuestro objetivo, primero organizamos los productos por categoría con el enfoque clásico de búsqueda exacta y luego dentro de cada categoría aplicamos el enfoque de espacios métricos. En consecuencia, la búsqueda de un producto implica primero filtrar por categoría realizando una búsqueda exacta y luego realizar la búsqueda por similitud sobre los productos que pertenecen a la categoría elegida.

Para el filtrado por categoría armamos un hashing lineal en memoria principal. Para la búsqueda por similitud utilizamos el enfoque basado en pivotes: elegimos  $k$  pivotes y para cada producto  $x$  almacenamos su firma  $\Phi(x) = (d(x, p_1), \dots, d(x, p_k))$  donde  $p_i$  es el  $i$ -ésimo pivote y  $d$  es la función de distancia de edición. Un punto crucial es la técnica de selección de pivotes usados para construir la firma de los productos. La forma en que los pivotes son seleccionados afectan el desempeño de un índice cuando se realiza una búsqueda [2, 5, 4]. En nuestro caso elegimos la técnica de selección incremental de pivotes [2] que ha demostrado ser eficiente en los espacios métricos normalmente usados por la comunidad científica que trabaja en búsqueda por similitud. Esta técnica la comparamos con la selección aleatoria de pivotes, para poder tener una base de comparación sobre su desempeño en este caso de estudio.

## 5 Evaluación Experimental

La evaluación experimental tuvo como objetivo analizar la eficiencia de los algoritmos basados en pivotes en este caso real de aplicación. Cabe señalar que hubo un exhaustivo trabajo previo de seteo de parámetros, dado que los valores que comúnmente figuran

Grupo	Cant. de Categorías	Rango del tamaño	Cant. de Pivotes
1	6	[1.034 - 15.964]	16, 32, 64, 128, 256
2	12	[19.032 - 46.530]	64, 128, 256, 512, 1024
3	8	[57.198 - 136,323]	256, 512, 1024, 2048
4	4	[16.7995- 213.578]	512, 1024, 2048, 4096

**Table 1.** Organización de categorías en grupos.

en los trabajos de espacios métricos no resultaban adecuados en este ámbito. Describimos a continuación cada uno de los aspectos considerados.

**Base de Datos.** Para la evaluación experimental se utilizaron productos ofrecidos en la plataforma de e-commerce Mercado Libre. Estos productos están clasificados dentro de un árbol de categorías, donde cada categoría reúne productos relacionados. Se obtuvieron alrededor de 2 millones de productos, distribuidos en 12.000 categorías hojas y 30 categorías iniciales. Para armar el hashing se utilizó la categoría inicial, dado que utilizar la categoría hoja restringía demasiado el universo de búsqueda produciendo que las respuestas no eran adecuadas a lo que el usuario pretendería como recomendación.

**Elección de pivotes.** El hecho de que la cantidad de elementos de las categorías estuviera distribuido en un rango muy amplio (de 1034 elementos para la categoría más pequeña a 213.578 elementos para la categoría más grande), nos obligó a segmentar dichas categorías en 4 grupos de acuerdo a su tamaño. En base a este mismo criterio, realizamos la elección de la cantidad de pivotes, teniendo en cuenta el porcentaje que los mismos representaban sobre el total de productos de cada categoría y la capacidad de memoria principal, dado que en esta etapa del trabajo el objetivo era diseñar el índice en memoria principal. La distribución resultante se presenta en la Tabla 1.

Otro punto a tener en cuenta era si se utilizaba el mismo grupo de pivotes para todas las categorías elegidos del universo total de búsqueda o cada categoría utilizaba su propio conjunto de pivotes elegidos dentro de esa categoría. Experimentalmente se estableció que la segunda opción era la mas adecuada y es la utilizada en los experimentos que se muestran en la próxima sección.

**Selección del rango de búsqueda.** La mayoría de los trabajos de investigación existentes sobre búsqueda por similitud en textos, utilizan como universo de datos diccionarios de palabras y como radio de búsquedas valores que van de 1 a 5. En nuestro trabajo, 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, donde cada elemento de dicho universo es una combinación de palabras, abreviaciones, números y caracteres especiales. Se realizaron una serie de experimentos previos que nos permitieron concluir que el radio de búsqueda mas adecuado para este caso es  $r = 23$  (los resultados de estos experimentos no se muestran por razones de espacio).

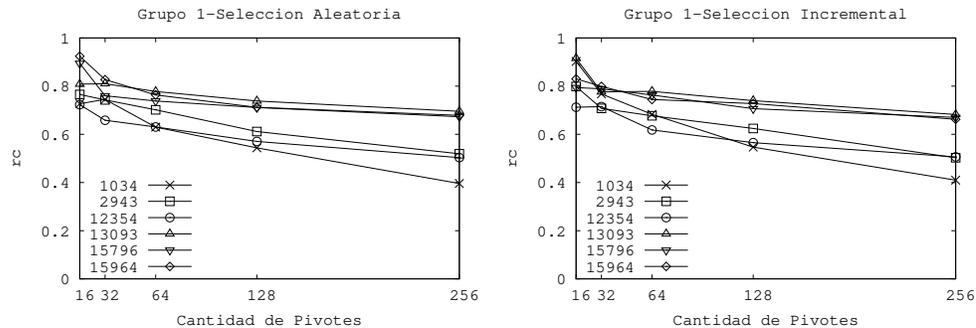


Fig. 1. Grupo 1: Efecto de la cantidad de pivotes sobre  $rc$ .

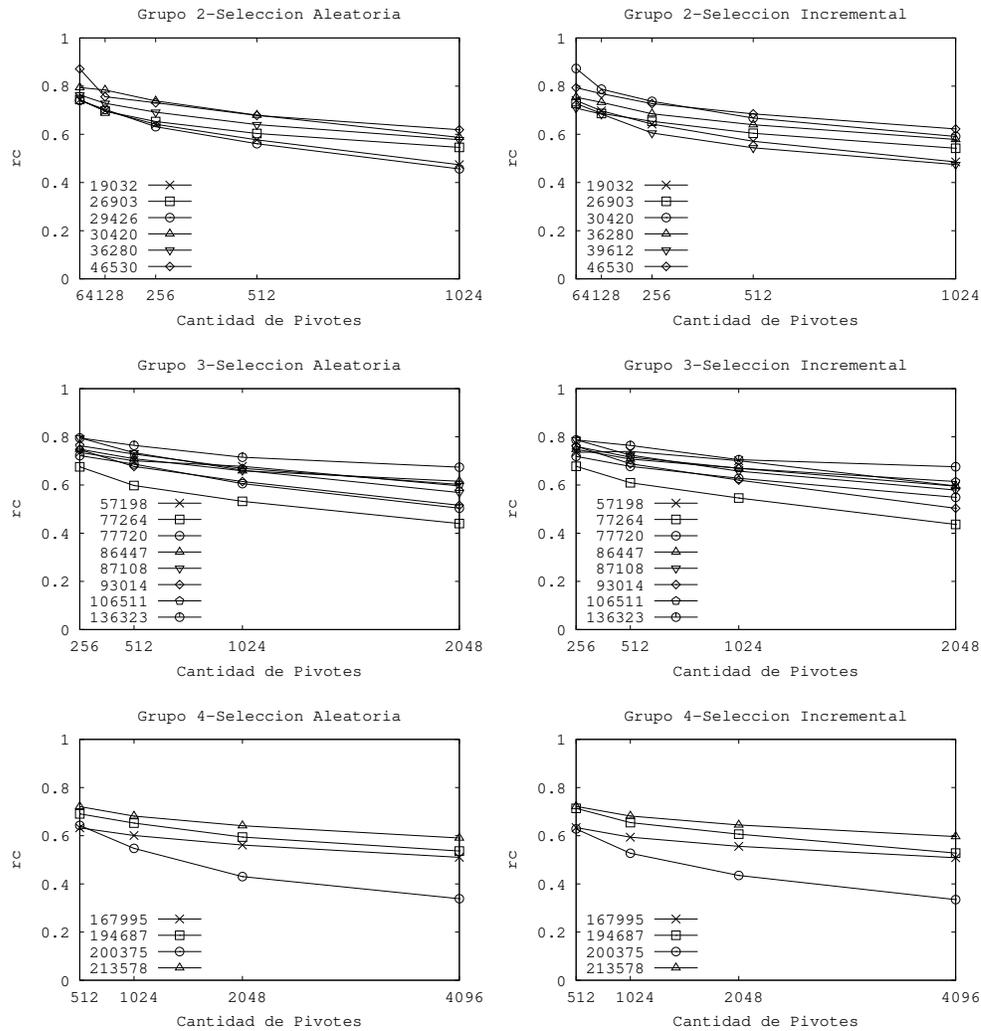
## 6 Análisis de Resultados

### 6.1 Efecto de la Cantidad de Pivotes

Como ya mencionamos, el criterio de eficiencia del algoritmo de búsqueda, para cuando las estructuras se manejan en memoria principal, es la cantidad de evaluaciones de la función de distancia. En nuestro caso, como las categorías tienen diferentes cantidad de productos, para poder realizar una comparación global utilizamos el *ratio de comparaciones* ( $rc$ ) definido como la cantidad de evaluaciones de la función de distancia dividido por la cantidad de elementos de la categoría. Es decir,  $rc$  nos da el porcentaje de elementos de la categoría con la que se comparó un elemento de búsqueda. De esta forma podemos obtener valores comparables independientemente de la cantidad de elementos que tenga cada categoría. El primer paso fue analizar el efecto de la cantidad de pivotes sobre  $rc$  para cada categoría considerada.

Para los experimentos, se seleccionaron al azar el 10% de los elementos de cada una de las categorías para realizar las búsquedas. Los resultados que se muestra son el promedio de las búsquedas realizadas. En las gráficas, cada categoría analizada queda identificada por la cantidad de elementos que contiene. Recordemos que las categorías están organizadas en grupos y para cada grupo hemos usado diferentes cantidades de pivotes. Los resultados que aquí se muestran reflejan esta organización. En el caso del grupo 2, por cuestiones de claridad, solo se muestran los resultados para 6 de las categorías de dicho grupo; las restantes categorías de ese grupo tienen un comportamiento similar a lo aquí mostrado.

En la Figura 1 observamos los resultados obtenidos con las categorías del Grupo 1. Sobre el eje  $x$  están representadas la cantidad de pivotes utilizados y sobre el eje  $y$  el  $rc$ . Como era de esperar, tanto para la selección de pivotes aleatoria (izquierda) como para selección de pivotes incremental (derecha), el  $rc$  disminuye a medida que aumentamos el número de pivotes. Puede observarse que el mejor valor logrado para  $rc$  es cercano a un 70% para las categorías más grandes y 40% para las categorías más chicas.



**Fig. 2.** Grupos 2, 3, y 4: efecto de la cantidad de pivotes sobre  $rc$ .

Este mismo comportamiento se repite con los otros 3 grupos (ver Figura 2). En todos los grupos de categorías el  $rc$  está entre un 40% y un 70% aproximadamente. La diferencia entre los distintos grupos radica en la cantidad de pivotes necesarios para lograr estos resultados: para las categorías más grandes (grupo 4) se necesitan 4096 y para las más chicas con 256 pivotes. Esto nos indica de un claro efecto del tamaño de la base de datos sobre los resultados obtenidos.

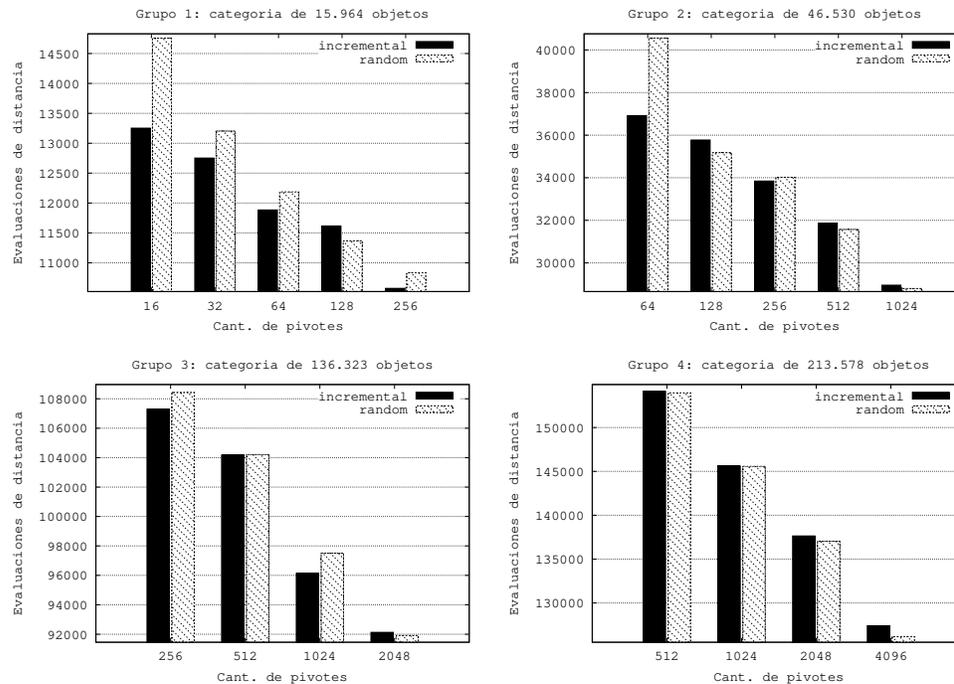


Fig. 3. Selección de pivotes: Incremental vs. aleatoria.

Podemos concluir que para todas las categorías consideradas a medida que aumentamos el número de pivotes mejoramos el  $rc$ . Este resultado es el esperado cuando trabajamos con algoritmos de indexación basados en pivotes. La diferencia que encontramos respecto de lo que se encuentra en trabajos de investigación referidos a esta temática, es que la eficiencia lograda es menor para ambas técnicas de selección de pivotes.

## 6.2 Comparación de las Técnicas de Selección de Pivotes

Otro punto que se evaluó es si la técnica de selección incremental lograba mejorar a la selección random en la misma proporción que ocurre en los espacios métricos de prueba utilizados por la comunidad dedicada al estudio de espacios métricos. En la Figura 3 se muestran las gráficas para las categorías de mayor tamaño de cada grupo porque consideramos que son los más representativos; sobre el eje  $x$  está representada la cantidad de pivotes y sobre el eje  $y$  la cantidad de evaluaciones de la función de distancias hechas en la categoría considerada en cada caso.

Como puede observarse, en las categorías más chicas (grupo 1 y 2), la técnica de selección de pivotes incremental tiene, en la mayoría de los casos, mejor desempeño que la selección aleatoria. Pero a medida que aumentamos el tamaño de las categorías

(grupo 3 y 4), esta diferencia disminuye llegando incluso a superar en eficiencia la selección random a la incremental. Por ejemplo, para la categoría más grande (213.578 elementos) con 4096 pivotes, la selección aleatoria hace aproximadamente un 50% menos de comparaciones que la selección incremental. Esto marca una diferencia de comportamiento de esta técnicas en este caso real de estudio respecto de lo que ocurre en los espacios métricos de pruebas normalmente utilizados en la temática.

### 6.3 Estudio de la Dimensionalidad

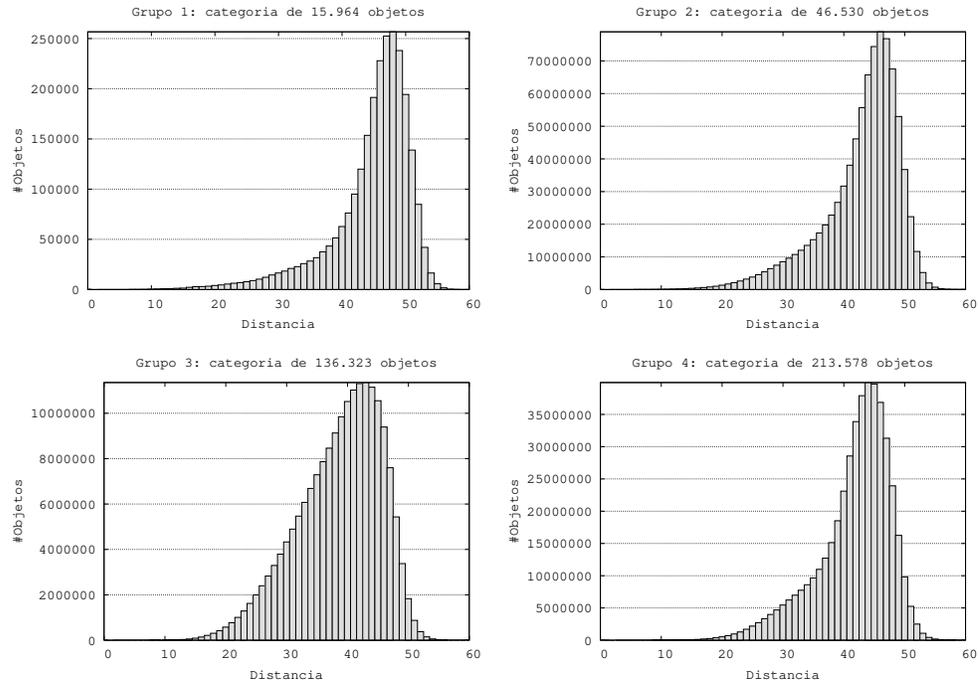
Los resultados obtenidos con la evaluación experimental fueron un indicativo de que podíamos estar ante un espacio de alta dimensionalidad. Dado que cada categoría es un universo completamente distinto, decidimos realizar el cálculo de los histogramas de distancia para todas las categorías consideradas. En la Figura 4 se muestran los histogramas de las categorías con mayor cantidad de elementos de cada grupo porque son las más representativas; sobre el eje  $x$  están representados los distintos valores de la función de distancia  $d$  y sobre el eje  $y$  la cantidad de elementos que están a esa distancia. Puede observar como se concentran los elementos alrededor de la media lo que indica que estamos ante un espacio de alta dimensionalidad, lo que influye negativamente en las búsquedas. En las secciones anteriores vimos que, independientemente de la técnica de selección de pivotes usada, el  $rc$  para la categoría más grande con 4096 pivotes ronda el 60%. Esto significa que ante una búsqueda, el algoritmo debe comparar con mas de la mitad de los elementos de la base de datos para obtener la respuesta a la consulta.

## 7 Conclusiones y Trabajo Futuro

En este trabajo hemos abordado el estudio modelo de espacios métricos aplicado a un caso real de estudio. Específicamente hemos usado el enfoque de indexación basado en pivotes para resolver la búsqueda de productos similares a uno dado, en un entorno de comercio electrónico. Nos propusimos estudiar el comportamiento de las búsquedas por rango utilizando las técnicas de selección de pivotes incremental y aleatoria. Podemos concluir que 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. También se pudo observar que la selección incremental no siempre mejora a la selección aleatoria. Como trabajo futuro nos proponemos estudiar el comportamiento de índices basados en particiones compactas y diseñar además un índice en memoria secundaria, para luego implementar un sistema de recomendación basado en espacios métricos.

## References

1. Asim Ansari and Skander Essegaier. Internet recommendation systems. In *Journal of Marketing Research, ACM E-commerce 2000 Conference*, pages 17–20, 2000.
2. B. Bustos, G. Navarro, and E. Chávez. Pivot selection techniques for proximity searching in metric spaces. In *Proc. of the XXI Conference of the Chilean Computer Science Society (SCCC'01)*, pages 33–40. IEEE CS Press, 2001.



**Fig. 4.** Histogramas de distancias para las categorías mas grandes de cada grupos.

3. Benjamin Bustos and Nelson Morales. On the asymptotic behavior of nearest neighbor search using pivot-based indexes. In *Proc. 3rd International Conference on Similarity Search and Applications (SISAP'10)*, pages 33–39, New York, NY, USA, 2010. ACM.
4. Benjamin Bustos, Oscar Pedreira, and Nieves Brisaboa. A dynamic pivot selection technique for similarity search. In *Proc. 1st International Workshop on Similarity Search and Applications (SISAP'08)*, pages 105–112, 2008.
5. E. Chávez, G. Navarro, R. Baeza-Yates, and J.L. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
6. J. Franganillor. El funware en los sitios de recomendación social: características, claves e inconvenientes. In : *Anuario ThinkEPI: análisis de tendencias en información y documentación, vol. 12*, pages 242–250, 2018.
7. G. Navarro. Searching in metric spaces by spatial approximation. In *Proc. String Processing and Information Retrieval (SPIRE'99)*, pages 141–148. IEEE CS Press, 1999.
8. G. Navarro and N. Reyes. Dynamic spatial approximation trees for massive data. In *Proc. 2nd International Workshop on Similarity Search and Applications (SISAP)*, pages 81–88. IEEE CS Press, 2009.
9. G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 2016. To appear.

## **Análisis del Desempeño de Clustering y Árboles de Decisión en la Evaluación Clínica de Microbiomas de Pacientes con Cáncer Colorrectal.**

Laura Avila\*, Victoria Santa María\*\*, Luis López\*, Marcelo Soria\*\*\*, Cristóbal R. Santa María\*,

\*DIIT-UNLaM, \*\*Instituto Lanari-FMed-UBA, \*\*\*FAUBA

Florencio Varela 1903 San Justo Pcia. de Buenos Aires

54-011-44808952

Laura\_avila75@yahoo.com.ar

vctrstmr@hotmail.com

llopez@ing.unlam.edu.ar soria@agro.uba.ar

csantamaria@unlam.edu.ar

**Abstract.** La metagenómica orientada hacia el uso de genes marcadores como el 16S rRNA permite establecer el perfil taxonómico del microbioma de pacientes con cáncer colorrectal. Cabe entonces explorar el papel del análisis taxonómico del microbioma como herramienta de diagnóstico y evaluación de la enfermedad. En tal sentido debe ajustarse la interrelación bioinformático-médica. Cada algoritmo a utilizar, cada parámetro a ajustar, requieren de una evaluación acerca del grado en que colaboran a mejorar el análisis en términos médicos. El objetivo general del trabajo es entonces caracterizar el microbioma de pacientes del AMBA en cuanto a riqueza, diversidad y distribución estadística, a través de muestras del gen marcador 16S rRNA obtenidas de materia fecal. En particular, se procuró reproducir la pipeline desarrollada anteriormente con muestras extraídas de repositorios internacionales mejorando los aspectos de automatización y ajustando la elección de parámetros. También se validó la metodología de trabajo por medio de comparación con los procesos llevados a cabo en el marco de la Large Bowel Microbiome Disease Network. A su vez, se realizó el análisis estadístico correspondiente para establecer la riqueza, diversidad de los microbiomas autóctonos. Finalmente se evaluó el desempeño de métodos supervisados y no supervisados de clasificación y predicción respecto del diagnóstico

**Palabras Clave:** Microbioma-Cáncer-Secuenciación-Explotación de Datos-Evaluación Médica

### **1. Introducción**

Los métodos de nueva generación para secuenciación de ADN posibilitan el análisis masivo y a bajo costo de las comunidades de microorganismos alojados en el intestino humano. El creciente interés médico que suscitan estos estudios se basa en la probada asociación de estados de riqueza y diversidad del microbioma con patologías importantes como el cáncer colorrectal sobre el cual se focaliza este artículo. En

trabajos anteriores [1] se ha dejado establecida una línea de procedimientos a efectuar sobre las lecturas desde que salen del secuenciador hasta que resultan procesadas en términos de explotación de datos. También se ha probado la potencialidad de estos métodos para caracterizar el microbioma [2]. Sin embargo, la elección de parámetros y algoritmos debe estar guiada por el criterio médico para el cual resulte útil la información aportada en términos clínicos de diagnóstico y evaluación. Este artículo se propone exhibir aspectos de la vinculación bioinformático- médica y ajustar la metodología hasta aquí desarrollada a efecto de hacer evidentes los aspectos clínicos de interés. Por primera vez se realiza el estudio sobre pacientes autóctonos, para los cuales la composición del microbioma varía de acuerdo a factores como tipo de alimentación, edad y localización geográfica. La tarea se realizó en el marco de un convenio firmado entre la Universidad Nacional de La Matanza y el Hospital Italiano de Buenos Aires, Sector de Coloproctología. A través del mismo se cuenta además con la inserción en la Large Bowel Microbiome Disease Network de la Universidad de Leeds, Inglaterra, lo que permite validar los procedimientos que se lleven a cabo.

Las tecnologías de nueva generación para la secuenciación de ADN han potenciado notablemente las posibilidades de los estudios metagenómicos, que involucran el conocimiento simultáneo de los genes de todos los individuos que forman una comunidad, extendiendo sus alcances al análisis de la composición microbiana de suelos, aguas y al microbioma humano. Éste no es otra cosa que la comunidad de microorganismos presentes en el cuerpo humano que contiene diez veces más microorganismos que células propias. Se han presentado entonces probabilidades ciertas de evaluar la interacción entre esta microbiota y el organismo alojante que resulta clave en el mantenimiento de la inmunidad y la protección contra agentes patógenos externos al organismo humano. La composición del microbioma, que se ha considerado como un órgano adicional en las personas [2], varía según el estilo de vida, la dieta y su genotipo, pero es estable dentro de una misma persona. Si se producen modificaciones de tipo permanente esto conlleva una disbiosis que es la alteración de la influencia de la comunidad en los procesos metabólicos y que se asocia con enfermedades tales como la inflamación intestinal, el asma o los desórdenes mentales. En particular la disbiosis puede estar implicada en la carcinogénesis al ser iniciadora de procesos inflamatorios y su presencia da señal de inmunodepresión [3].

Algunos argumentos indirectos sugieren este rol potencial de la microbiota intestinal en la carcinogénesis colorrectal. El cáncer colorrectal es básicamente una enfermedad genética pero el microbioma alojado por el paciente puede explicar la interacción entre los genes del paciente y el entorno de microorganismos presentes que se manifiesta tanto en su diversidad y riqueza taxonómica cuanto en las vías metabólicas que tienen lugar. Frecuentemente aparecen asociados el cáncer colorrectal y las variaciones de las frecuencias con que algunas especies bacterianas se encuentran en el microbioma [4] y [5]. A su vez la disminución en la diversidad total se ha vinculado con distintas patologías que incluyen el cáncer colorrectal, la obesidad, enfermedades autoinmunes y neurológicas [6]. Esta asociación no es clara aún para determinar si la variación del microbioma es una causa o un efecto del cáncer. Incluso recientemente se ha sugerido que el microbioma puede jugar el rol de control sobre la enfermedad. En todo caso existe una perspectiva interesante en los estudios metagenómicos, pues no solo permiten la determinación taxonómica de la comunidad microbiana a través de la utilización de genes marcadores sino que también, al utilizar

la información de todas las secuencias obtenidas del microbioma (WGS), pueden establecer las vías metabólicas que potencialmente sigan los procesos celulares en el paciente [7]. Esto ha motivado un profundo interés en la comunidad médica que ha buscado avanzar en la comprensión, y eventualmente en el diagnóstico y pronóstico de enfermedades, utilizando estos métodos de análisis.

Al respecto hay que señalar no solo la tecnología de secuenciación sino también los desarrollos de algoritmos de aprendizaje automático supervisado y no supervisado. En lo referido al microbioma humano, se ha hecho evidente la necesidad de contar con un esquema seriado de procesos computacionales a aplicar desde que las secuencias salen del secuenciador hasta que resultan transformadas en información útil para la investigación clínica. Esto involucra la confección de software de filtrado de las secuencias, de evaluación de contaminación del conjunto con secuencias humanas, de ensamblado de secuencias, de anotación de las mismas según sus niveles taxonómicos, de identificación de vías metabólicas presentes, de agrupamiento en conglomerados o clusters según taxonomía o metabolismo, y de aprendizaje sobre conjuntos de entrenamiento y testeo para clasificar microbiomas según los mismos principios.

En el proyecto Aplicación de Técnicas de Data Mining para Análisis del Microbioma Humano según Funcionalidades Metabólicas, desarrollado por el grupo en el período 2017-2018, se ha podido establecer una “pipeline”, con varios pasos automatizados, para tratar las secuencias de ADN microbiómico. Comprende el tratamiento de las lecturas desde que salen del secuenciador hasta que resultan datos para explotación por técnicas estadísticas multivariadas y de aprendizaje supervisado y no supervisado, de forma de ponerlos al servicio de la interpretación médica. Estos procesos comienzan con el filtrado de las lecturas para quitar posibles contaminaciones con los reactivos utilizados en la secuenciación, continúan con el ensamblado en contigs, luego con el filtrado de las secuencias humanas que pudieran haber sido obtenidas también en la muestra y finalmente con la anotación taxonómica y funcional. Luego de esto la información debe disponerse de manera adecuada para iniciar el proceso de explotación de los datos que consiste en la aplicación de variadas técnicas estadísticas y de aprendizaje automático a efecto de establecer las características y patrones de comportamiento que puedan asociarse a la condición clínica de los pacientes. Para este trabajo se logró contar con muestras de materia fecal de pacientes autóctonos para iniciar así un estudio sobre las características locales de la enfermedad que se supone presentarán variaciones ligadas a dieta, condiciones de hábitat, etc. [8]

## **2. Materiales y Métodos**

### **2.1 Muestras**

Diseño:

Corte transversal.

1. 20 pacientes (10 con CCR y 10 controles) tratados por la Sección de Coloproctología del Hospital Italiano de Buenos Aires.
2. 15 pacientes (7 con CCR y 8 controles) tratados por la Sección de Coloproctología del Hospital Italiano de Buenos Aires.

Criterio de inclusión:

Casos: - Edad mayor a 18 años. - Adenocarcinoma de colon confirmado con histología.

Controles: - Edad mayor a 18 años - Ausencia de neoplasia colónica (adenocarcinoma y adenoma) confirmada por video colonoscopia completa, con Boston mayor a 6 (al menos 2 puntos por sector).

Criterio de exclusión: - Consumo de antibióticos o probióticos en los últimos 6 meses. - CCR en tratamiento - Antecedentes de cirugía colorrectal, CCR, radioterapia pélvica o quimioterapia. - Antecedentes familiares compatibles con síndromes de CCR hereditario - Enfermedad inflamatoria intestinal o enfermedad intestinal infecciosa. - Incapacidad de dar consentimiento informado.

Muestras empleadas en el estudio:

Muestra 1: materia fecal de 10 pacientes con CCR no tratado, material fecal de 10 voluntarios sanos que se sometieron a una colonoscopia por alguna razón y se haya demostrado que tienen un intestino normal en la colonoscopia.

Muestra 2: materia fecal de 7 pacientes con CCR no tratado, material fecal de 8 voluntarios sanos que se sometieron a una colonoscopia por alguna razón y se haya demostrado que tienen un intestino normal en la colonoscopia.

Mezcla de muestras 1 y 2: Se identificaron 216 géneros comunes entre la Muestra 1 y la Muestra 2. Con ellos y conservando el diagnóstico clínico efectuado se integró la mezcla de muestras con el objetivo de lograr una mayor representatividad y homogeneidad.

## 2.2 Secuenciación

Muestra 1: Se realizó con secuenciador Illumina HiSeq sobre la región V4 del gen 16S rRNA. Cada secuencia representa 150 pares de bases

Muestra 2: Se realizó con secuenciador Illumina MiSeq sobre las regiones V3 y V4 del gen 16S rRNA. Cada secuencia representa 300 pares de bases.

## 2.3 Procesamiento inicial

Ambas muestras fueron tratadas en una cadena de procesos establecida en trabajos anteriores [1]. La metodología empleada en estos procesos iniciales fue validada aquí, por comparación con trabajos similares realizados por grupos dentro de la Large Bowel Microbiome Disease Network, la cual integra el Hospital Italiano de Buenos Aires. Se importaron las lecturas del microbioma de cada paciente al software QIIME2 [9]. Luego se eliminó el ruido. Se filtraron las secuencias y se eliminaron las lecturas ambiguas o de baja calidad. A continuación, las distintas secuencias fueron alineadas contra los alineamientos de referencia para el gen 16S rRNA. Para cada metagenoma intestinal, se generó una tabla de frecuencias de las secuencias agrupadas en Unidades Taxonómicas Operacionales (OTU) y se confeccionó el árbol filogenético. En la Muestra 1, cuyas secuencias comprendieron solo la región V4 del gen, éstas se agruparon en 239 OTUs distintas, correspondientes al nivel taxonómico género. En la Muestra 2, más rica por contener las regiones V3 y V4 del gen, se pudieron identificar 370 taxones género.

## 2.4 Clustering

Se realizaron distintos experimentos de agrupamiento de pacientes a efecto de evaluar las posibilidades de la técnica en la clasificación clínica adecuada de los pacientes de acuerdo a su perfil microbiómico. Se realizaron pruebas de clustering jerárquico, con distancia euclídea, otras con agrupamiento no jerárquico por medio del algoritmo k-means, con distancia euclídea y encadenamiento promedio, variando el número inicial de centroides. Y finalmente se construyó “ad hoc” una distancia entre microbiomas que tiene en cuenta el peso de la diferencia de cada taxón entre pacientes sanos y enfermos [1]. En estos procesos se utilizó software INFOSTAT [10], WEKA [11] y desarrollos propios en lenguaje C para operar entre paquetes cambiando formatos y armar la matriz de distancias pesadas. Los agrupamientos fueron evaluados por el índice Silhouette.

## 2.5 Árboles de decisión

En relación con los métodos de aprendizaje automático, en base a los antecedentes de desempeño [12], se decidió entrenar y testear dos algoritmos de árboles de decisión. Por un lado, el C4.5 [13] disponible en Weka bajo el nombre J48 y por otro, el ensamble Random Forest, también incorporado a WEKA. Desde el punto de vista computacional se utilizaron matrices de confusión y curvas ROC para evaluar tanto el entrenamiento, realizado a partir de la Muestra 1, como el testeo, efectuado sobre la Muestra 2. La consideración comparativa de ambas muestras requirió la identificación de los taxones presentes simultáneamente en ambas. Se identificaron 216 géneros comunes con los cuales se trabajó en los dos tipos de árboles. Además, los mismos algoritmos se probaron con la mezcla de muestras 1 y 2. Así se seleccionó convenientemente un conjunto de entrenamiento de 18 pacientes y otro de testeo de 17. En todos los casos, se estableció como criterio relevante en términos clínicos que la clasificación fuera muy eficiente en la detección de pacientes enfermos y menos importante en cuanto a la verificación de los sanos.

## 3. Resultados obtenidos

Los primeros resultados obtenidos corresponden a los procesos iniciales realizados con QIIME2. Por ejemplo, la distribución estadística de frecuencias de OTUs o taxones se dispuso como exhibe la Tabla 1.

**Tabla 1.** Datos de la muestra 1, cantidad presente de cada taxón u OTU por paciente.

Taxón	D_0_Archae	D_0_Archae	D_0_Bacter												
GCRFNG_AR-CC-A1	0	0	5	0	42	0	0	0	0	0	0	321	0	0	0
GCRFNG_AR-CC-B2	0	0	0	0	22	0	0	0	0	0	0	73	0	4	0
GCRFNG_AR-CC-C3	160	0	2	0	213	0	0	0	0	23	103	0	1	0	0
GCRFNG_AR-CC-D4	0	0	0	0	359	1	0	0	0	0	77	0	0	0	0
GCRFNG_AR-CC-E5	0	0	0	0	143	0	0	0	0	0	0	4	0	0	0
GCRFNG_AR-CC-F6	0	0	10	0	1705	0	0	0	0	0	34	0	0	0	0
GCRFNG_AR-CC-G7	58	0	0	0	3090	0	0	0	0	0	571	0	0	0	0
GCRFNG_AR-CC-H8	0	0	0	0	89	0	0	0	0	0	86	0	0	0	0
GCRFNG_AR-CC-I9	0	0	0	0	40	0	0	0	0	0	284	0	0	0	0
GCRFNG_AR-CC-J10	49	64	141	0	64	0	4	0	0	0	134	0	0	0	0
GCRFNG_AR-HV-A1	1	0	0	0	21	0	0	0	0	0	295	0	0	0	0
GCRFNG_AR-HV-B2	10	0	0	0	95	0	0	0	12	0	55	0	0	0	0
GCRFNG_AR-HV-C3	0	0	7	0	192	0	0	0	0	24	180	0	0	0	0
GCRFNG_AR-HV-D4	0	0	0	0	180	0	0	11	0	0	114	0	10	0	0
GCRFNG_AR-HV-E5	222	99	0	0	112	0	0	23	23	38	130	0	8	0	0
GCRFNG_AR-HV-F6	0	0	18	0	128	0	0	0	0	3	227	0	0	0	0
GCRFNG_AR-HV-G7	0	0	0	0	1105	0	0	0	0	4	95	0	3	0	0
GCRFNG_AR-HV-H8	0	0	0	0	297	0	0	0	0	61	106	0	0	0	0
GCRFNG_AR-HV-I9	0	0	0	0	3711	0	0	0	0	0	537	0	0	0	0
GCRFNG_AR-HV-J10	1	0	13	1	48	0	8	0	0	0	66	0	0	0	0

En la primera columna de cada tabla se anotan los pacientes detallados con un código, y la primera fila nombra cada uno de los taxones u Otus identificados. En la muestra 1, a nivel género, se pudieron identificar 239 Otus, y en la muestra 2, 368. Las últimas tres columnas corresponden a la clasificación, la edad y el sexo.

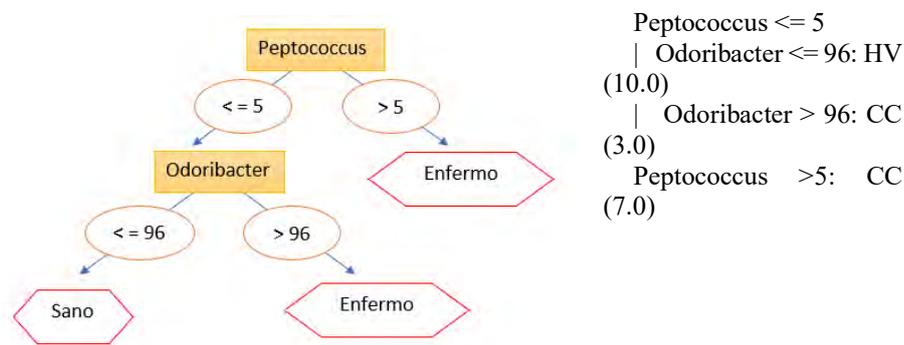
El clustering realizado con la muestra 1 arrojó resultados dispares. El método jerárquico fue poco adecuado para producir agrupamientos que se correlacionaran con la clasificación clínica por sano o enfermo. En cambio, K-means con distancia euclídea y encadenamiento promedio arrojó mejores resultados, aunque insuficientes para asegurar una clasificación adecuadamente correlacionada con el diagnóstico conocido. Esto se logró al establecer una distancia pesada “ad hoc”. Para ello se consideró la media de las frecuencias de cada taxón teniendo en cuenta el diagnóstico clínico. Para cada taxón se calculó el valor absoluto de la diferencia entre la media de los pacientes sanos y la media de los enfermos. A cada resultado se lo dividió por la suma total de las diferencias y se lo multiplicó por un factor positivo constante para constituir el peso de cada taxón. Estos pesos se incorporaron al cálculo de la distancia entre pacientes de

acuerdo a:  $d = \sqrt{\sum_{j=1}^{239} (f_{ij} - f_{kj})^2 P_j}$ . Con la matriz de las nuevas distancias se aplicó k-means para obtener ahora dos clusters. Se observó que los casos enfermos fueron todos bien clasificados, mientras que solo resultaron bien clasificados la mitad de los pacientes sanos. El test chi-cuadrado para evaluar la asociación entre la clasificación clínica y los clusters obtenidos arrojó un valor  $p < 0.009$  lo que indica que puede rechazarse la independencia entre ambas variables cualitativas. Los agrupamientos óptimos alcanzaron índices silueta de 0.49 para el cluster 1 que agrupó todos los casos enfermos y la mitad de los sanos y de 0.12 para el cluster 2. El índice silueta general fue de 0.4 lo que se consideró aceptable habida cuenta de la óptima clasificación de los casos enfermos.

Al realizar sobre la muestra 2 el agrupamiento por medio de k-means, con la distancia pesada y encadenamiento promedio se obtuvo un resultado parecido. El valor  $p$  para la prueba chi-cuadrado resultó  $p < 0.07$  por lo cual puede rechazarse la hipótesis de independencia entre diagnóstico y cluster con ese nivel de significancia. El cluster 1 agrupó el total de los 7 casos enfermos y 5 de los voluntarios sanos, mientras que el cluster 2 se integró con 3 de los 8 casos sanos (37.5 %). El índice silueta del

agrupamiento total resultó de 0.73 y como ocurrió para la otra muestra fue mejor el índice silueta del cluster 1, 0.84 que el del cluster 2, 0.27.

El algoritmo J48 se corrió sobre la muestra 1 dividida en conjuntos de entrenamiento y testeo. Como el desempeño fue pobre, en este caso además se realizó una selección de atributos por medio de un procedimiento que establece un ranking de variables según la información que aportan a la variable de clasificación [14]. Así se seleccionaron solo 20 géneros para entrenar y testear. De ellos lo que mejor rankearon fueron el 119, Peptococcus, y 22, Odoribacter. Ambos le bastaron al modelo predictivo J48 para establecer, podando los otros, la regla de inferencia de la Figura 1.



**Figura 1.** Diagrama de árbol determinado por las reglas de inferencia J48.

Solo el 30% de los casos fue bien clasificado en el testeo, lo que motivo el descarte del algoritmo en este trabajo.

A continuación, sobre las muestras homogeneizadas en los 216 taxones comunes se aplicó el ensamble Random Forest [15]. Se realizaron distintas experiencias. Se tomó como conjunto de entrenamiento, la muestra 1 de 20 pacientes, y se testeó con la muestra 2 de 15 pacientes. El porcentaje de casos de testeo bien clasificados fue del 60% pero lo importante es que el algoritmo detectó bien todos los casos enfermos, aunque solo clasificó adecuadamente a la cuarta parte de los sanos. El área bajo la curva ROC de testeo fue de 0.946 por lo que la diferencia con la de entrenamiento, que había clasificado bien todos los casos, es de 0.054 lo que revela un entrenamiento adecuado.

Se corrió también el algoritmo Random Forest sobre la mezcla de las muestras 1 y 2. En este caso se realizó una selección previa de atributos basada en el criterio de pesos ya utilizado en el clustering para calcular las distancias. Con 9 atributos para entrenar el ensamble el 64 % de los casos resultaron bien clasificados, pero aquí solo el 75 % de los enfermos fue clasificado como tal. La diferencia entre el área bajo las curvas ROC fue de 0.354 lo que revela el sobreentrenamiento a pesar de la poda de atributos efectuada.

Un resumen de los métodos aplicados y sus resultados se muestra en la Tabla 2.

**Tabla 2.** Desempeño de Algoritmos

Método	Algoritmo	M	Selección	%	%E/CC	%S/HV	Sil/DifARoc
Cluster:	Jerárquico	1	Dist Eucl	5	100	10	-----
Cluster	Kmeans	1	Dist Eucl	10	90	10	0.26
Cluster	Kmeans	1	Dist Pes	75	100	50	0.40
Cluster	Kmeans	2	Dist Pes	67	100	37	0.73
Ar. Dec	J48	1	Infogain	30	20	40	0,700
Ar. Dec	R. Forest	1y 2	Sin Selec	60	100	25	0.054
Ar. Dec	R. Forest	Me12	Pesos	64	75	56	0.354

El porcentaje de enfermos bien clasificados como tales se resume en la Tabla 2 como %E/CC y el porcentaje de pacientes sanos correctamente clasificados se simboliza por %S/HV

#### 4. Conclusiones

Se ha logrado realizar toda la cadena de análisis necesaria para la determinación microbiómica por genes marcadores con pacientes autóctonos de la zona del AMBA. Se ha realizado la secuenciación de muestras de ADN de materia fecal, se han completado los procesos de filtrado, alineamiento y reconocimiento taxonómico siguiendo el método validado a nivel internacional. Durante la ejecución de esos procesos se han concretado también todos los enlaces necesarios relativos a cambios de formatos y presentaciones de la información lo cual, detallado parcialmente en trabajos anteriores [1], está aquí implícito. Así la información obtenida ha estado disponible para realizar pruebas de desempeño de algoritmos de explotación de datos en la determinación clínica. Respecto al clustering, se han dado resultados prometedores con la distancia pesada definida. Lo mismo ha ocurrido con la aplicación del ensamble de árboles de decisión Random Forest teniendo en cuenta la alta proporción de clasificación correcta de los pacientes enfermos. Resulta claro que deben realizarse ensayos más amplios utilizando muestras de mayor tamaño para afinar y confirmar la efectividad al utilizar estas técnicas para apoyar el diagnóstico. Sin embargo, tanto los clusters hallados con distancia pesada, como los ensayos con el ensamble de árboles han cumplido con el criterio general de mínimo error en la clasificación de los pacientes enfermos, lo que puede constituir una herramienta no invasiva para determinar la realización de otros estudios.

## 5. Referencias

1. Avila Laura, Santa María Victoria, López Luis, Soria Marcelo y Santa María Cristóbal.: Tratamiento de Secuencias de ADN y Clustering de Pacientes con Cáncer Colorrectal. WICC2020. El Calafate. (2020) <https://wicc2020.unpa.edu.ar/>
2. O'Hara AM, Shanahan F.: The gut flora as a forgotten organ. *EMBO Rep.* 2006 Jul;7(7):688–93. (2006)
3. Lopez, A et al.: Microbiota in digestive cancers: our new partner? *Carcinogenesis*, 1-10. doi:10.1093/carcin/bgx087 (2017)
4. Kosumi K, Hamada T, Koh H, Borowsky J, Bullman S, Twombly TS, et al.: The Amount of Bifidobacterium Genus in Colorectal Carcinoma Tissue in Relation to Tumor Characteristics and Clinical Outcome. *Am J Pathol* [Internet]. 2018 Sep 20; (2018) Available from: <http://dx.doi.org/10.1016/j.ajpath.2018.08.015>
5. Youssef O, Lahti L, Kokkola A, Karla T, Tikkanen M, Ehsan H, et al.: Stool Microbiota Composition Differs in Patients with Stomach, Colon, and Rectal Neoplasms. *Dig Dis Sci* [Internet]. 2018 Jul 11; Available from: <http://dx.doi.org/10.1007/s10620-018-5190-5>
6. Shreiner AB, Kao JY, Young VB.: The gut microbiome in health and in disease. *Curr Opin Gastroenterol*; 31(1):69–75. (2015)
7. Jones, R B. et al.: Inter-niche and inter-individual variation in gut microbial community assessment using stool, rectal swab and mucosal samples. *Scientific Reports* volume 8, Article number: 4139. (2018) [www.nature.com/scientificreports](http://www.nature.com/scientificreports)
8. Taylor M, Wood HM, Halloran SP, Quirke P.: Examining the potential use and longterm stability of guaiac faecal occult blood test cards for microbial DNA 16S rRNA sequencing. *J Clin Pathol.* 2017 Jul;70(7):600–6. (2017)
9. Bolyen E, et al.: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* 37: 852–857. (2019) <https://doi.org/10.1038/s41587-019-0209-9>
10. Di Rienzo J.A., Casanoves F., Balzarini M.G., Gonzalez L., Tablada M., Robledo C.W.: InfoStat versión 2018. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. (2018) URL <http://www.infostat.com.ar>
11. [https://www.cs.waikato.ac.nz/ml/weka/Witten\\_et\\_al\\_2016\\_appendix.pdf](https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf)
12. Statnikov A. Henaff M. Narendra V. Konganti K. Li Z. Yang L. Pei Z. Blaser M. Aliferis C y Alekseyenko A.: A comprehensive evaluation of multcategory classification methods for microbiomic data. *Microbiome* 2013 1:11 (2013)
13. Quinlan, J.R.: C4.5 Programs for Machine Learning, San Mateo, CA: Morgan Kaufmann (1992)
14. Eibe Frank, Mark A. Hall e Ian H. Witten.: El banco de trabajo WEKA. Apéndice en línea para "Minería de datos: herramientas y técnicas prácticas de aprendizaje automático", Morgan Kaufmann, cuarta edición. (2016)
15. Breiman, Leo.: Random Forests. *Machine Learning* 45 : 5–32. doi:10.1023/A:1010933404324. (2001)

# Decisiones y lecciones aprendidas en un proceso ETL aplicado a sistemas con testimonios de delitos de lesa humanidad

David Troncoso<sup>1</sup> \*, Agustina Buccella<sup>1</sup>, and Alejandra Cechich<sup>1</sup>

GIISCO Research Group  
Departamento de Ingeniería de Sistemas - Facultad de Informática  
Universidad Nacional del Comahue  
Neuquen, Argentina  
david.troncoso, agustina.buccella, alejandra.cechich@fi.uncoma.edu.ar

**Resumen** El proceso de selección de las fuentes de datos (extracción-E), el procesado y adaptación de los datos (transformación-T) y la carga de los mismos a un repositorio (carga-L), recibe el nombre de *proceso ETL*. El diseño e implementación de estos procesos son áreas de estudio importantes debido a la proliferación de sistemas similares cuya información necesita ser integrada y reestructurada para ser de utilidad. En este trabajo describimos la ejecución de un proceso ETL aplicado a dos sistemas que almacenan declaraciones y testimonios de delitos de lesa humanidad. Siguiendo dos objetivos necesarios para la integración y explotación de la información, describimos decisiones realizadas y lecciones aprendidas.

**Keywords:** Proceso ETL, Análisis de Datos, Declaraciones y testimonios

## 1. Introducción

Para realizar un desarrollo orientado al análisis de los datos en cualquier organización, se debe primero seleccionar la información fuente que pueda resultar útil, procesarla para que tenga sentido, y luego almacenarla en algún tipo de repositorio, en la forma de depósitos de datos [6] o lagos de datos [3]. La característica principal de estos repositorios es que actúan como centros de información en donde se vuelcan, en algún formato específico, los datos de todas las fuentes que se deseen explotar en un proceso de extracción, transformación y carga (ETL)[2,4,5,7].

Para el diseño de procesos ETL existen en la literatura muchas propuestas diferentes las cuales podemos clasificar entre las que se basan en UML, en BPMN, y en el uso de ontologías [1]. Todas estas propuestas van dirigidas a simplificar y organizar el proceso ETL debido a la complejidad derivada de la

---

\* Este trabajo esta parcialmente soportado por el Proyecto Desarrollo de Software basado en Reuso Parte II

2 Troncoso et al.

forma en que pueden estar representadas las fuentes de información. Es muy común encontrarnos con información no estructurada, incompleta, inconsistente y/o redundante.

En este artículo presentamos una experiencia en el desarrollo de un proceso ETL aplicado sobre dos sistemas informáticos que recaban información de testimonios y juicios realizados por delitos de lesa humanidad. El primero, denominado *Sistema Informático de Procesamiento de Declaraciones en Juicios de DDHH* (SIPDJ) desarrollado por la Facultad de Informática (FAI) de la Universidad Nacional del Comahue (UNCO), gestiona toda la información generada en las declaraciones durante los juicios de la Escuelita I y II. Por otro lado el *Sistema de Análisis Sociológico de Querellas* desarrollado por el Equipo de Asistencia Sociológica a las Querellas (EASQ) a través del Centro de Estudios sobre el Genocidio (CEG)<sup>1</sup>, comprende la Asistencia Sociológica a las Querellas (ASQ) almacenando datos correspondientes a parte del Juicio ABO (Atlético, Banco y Olimpo), Juicio Operativo Independencia de Tucumán y la mega causa de Santiago del Estero. Ambos sistemas gestionan información similar. La diferencia fundamental es que SIPDJ se centra en los juicios de DDHH y las declaraciones que se almacenan son las brindadas en los casos pertenecientes a dichos juicios. En cambio, ASQ se enfoca específicamente en registrar información necesaria para colaborar cualitativamente con las querellas de los juicios.

Como estos sistemas requieren almacenar información textual (de las declaraciones y testimonios) poseen atributos que tienen formatos de texto extensos y sin estructura. A su vez, como las declaraciones y/o testimonios no poseen información precisa, por ejemplo, el declarante no puede especificar lugares precisos donde estuvo o personas que lo retuvieron, hace más difícil la extracción de información para su análisis. La información está desnormalizada y es muchas veces redundante<sup>2</sup>. Con esto en mente, se definieron varios objetivos de manera de crear un depósito de datos consistente que permita un análisis de datos posterior. Entre esos análisis, en este trabajo nos centramos en el proceso ETL para dos objetivos específicos: (1) *Identificar a las personas que se nombraron en los testimonios para definir el recorrido a través de los centros clandestinos de detención (CCD)* y (2) *Obtener información georeferencial de las fuentes, para ser reflejada en un mapa interactivo. De esta forma se habilita la posibilidad de realizar análisis visuales respecto a la distribución geográfica de los datos.*

De esta forma este trabajo contribuye a mostrar una experiencia real de un proceso ETL sobre datos de texto, aportando decisiones y lecciones que pueden ser de interés en situaciones similares. El artículo se organiza de la siguiente manera. En la sección siguiente describimos el proceso realizado para cumplir con los dos objetivos propuestos. Luego en la Sección 3 analizamos las lecciones aprendidas para que sean útiles para la aplicación de procesos ETL con características similares. Finalmente se describen las conclusiones y trabajos futuros.

<sup>1</sup> <https://www.untref.edu.ar/instituto/ceg-centro-de-estudios-sobre-genocidio>

<sup>2</sup> Es importante aclarar que esta redundancia muchas veces es necesaria ya que intenta no obviar información útil de los testimonios y/o declaraciones

## 2. Aplicación de un Proceso ETL para normalizar e integrar información sobre sistemas de delitos de lesa humanidad

Como se mencionó previamente el proceso ETL debe estar orientado a preparar la información para cumplir con los dos objetivos propuestos. Para analizar cada objetivo y el proceso ETL definido, analizaremos el problema, su diseño y su implementación final o solución. Para el diseño utilizamos el enfoque definido en [5] el cual se basa en el uso de clases UML estereotipadas y mecanismos asociados.

### 2.1. Objetivo 1: Identificar las personas que se nombraron en los testimonios para definir el recorrido a través de los centros clandestinos de detención.

Tomamos como base las tablas *cumpa* y *testimoniante* del sistema ASQ, y *victima*, *persona*, *secuestro*, *traslado* y *cautiverio* del sistema SIPDJ. Por razones de espacio describimos algunas de las reglas más importantes aplicadas a la creación de una tabla *recorrido*. Sin embargo para cumplir el objetivo realizamos un cambio sobre la estructura de varias de las tablas mencionadas previamente. Para mostrar el panorama sobre la forma en que los datos están almacenados en la mayoría de estas tablas fuentes, en la Figura 1<sup>3</sup> mostramos algunos datos de la tabla *cumpa* del sistema ASQ. Esta tabla representa a los individuos secuestrados a quienes el testificante vio en los diversos CCD por los que pasó, y compañeros de quienes supo que allí estaban o habían estado<sup>4</sup>.

#### Problemas encontrados

(1) *Valores nulos, repetidos*. En la figura podemos observar que hay algunos valores nulos sobre los atributos como *apellido1* y *apellido2*, *apodo*, etc. A su vez, algunos valores en los atributos se pueden repetir si en distintos testimonios se nombra a la misma persona o si el testificante se describe en diferentes CCD, lo que significa que por cada fila se repiten varios datos como *oriundo*, *ocupacion*, *sexo*, etc.

(2) *Información que puede resultar inconsistente, redundante e incluso incompleta*. También se da el caso que una persona fuese divisada por 2 o más testificantes, entonces aparecerá en la tabla con posibles datos repetidos con algunos detalles descriptos por cada testificante. De esta forma, la información puede resultar inconsistente, redundante e incluso incompleta. Sin embargo, a los principios para lo que fue creado el sistema, esto es útil ya que confirma la presencia de cierta persona en un centro clandestino, y se diferencian los datos aportados por cada testificante.

<sup>3</sup> Hemos difuminado algunos datos debido a la sensibilidad de los mismos.

<sup>4</sup> Esta tabla posee más de 30 atributos

4 Troncoso et al.

id	nombre1	nombre2	apellido1	apellido2	apodo	fecha_nacimiento	sexo	militancia
1	Jorge	Oswaldo	T41			1978-10-02	1978-12-21	M
2	Rebeca					1978-07-21	1978-08-08	F
3	Oscar					1978-07-21	1978-08-08	M
4	Guillermo					1978-08-21	1978-08-16	F
5	Jorge	Caballo Loco				1978-12-22	1978-10-02	M
6	Luis	Salvador				1975-06-01	1975-06-30	M
7	Luis	Salvador				1975-06-01	1975-06-30	M
8	Luis	Salvador				1975-06-11	1975-07-17	M
9	Luis	Salvador				1975-06-17		M
10	Hector	Ricardo				1975-08-19	1975-10-31	M
11	Victor	Hugo				1975-08-19	1975-10-31	M
12	Hector	Ricardo				1975-10-15	1975-11-08	M
13	Hector	Ricardo				1975-10-15	1975-11-08	M
14	Hector	Ricardo				1975-10-15	1975-11-08	M

Figura 1: Formas en que los datos se encuentran almacenados en la tabla *cumpa*

**Diseño ETL.** En la Figura 2 podemos observar la parte del diseño del proceso ETL desde el sistema ASQ que involucra las tablas *cumpa* y *testimoniante* para crear la tabla *recorrido*. Las decisiones involucran el uso de join, agregación y filtros.

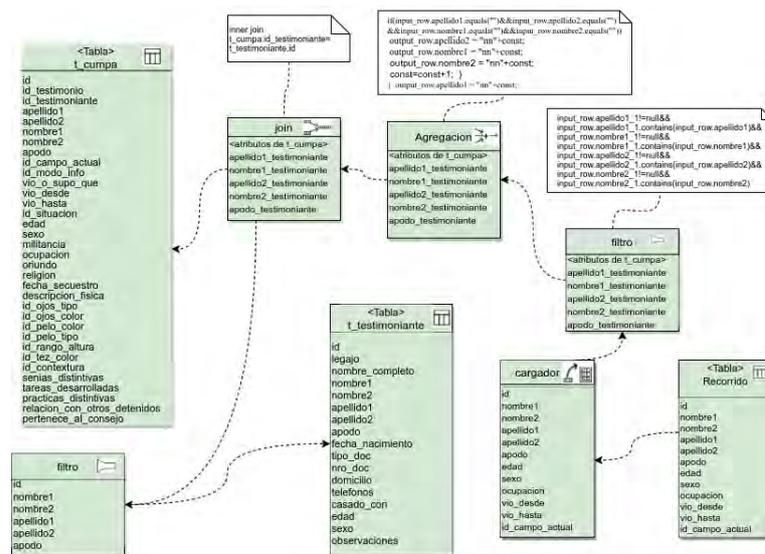


Figura 2: Parte del diseño ETL para crear la tabla *recorrido*

Componente de *Join*. La tabla *cumpa* cuenta con los atributos *nombre1*, *nombre2*, *apellido1*, *apellido2* y *apodo* que corresponden a la persona a la que se logra reconocer en un CCD, pudiendo ser una auto descripción del mismo testimoniante. Para detectar estos casos se busca, a través de este componente, dejar sobre cada registro de la tabla, también los atributos de *nombres*, *ape-*

lidos y apodo pero correspondientes al testimoniante (*nombre1\_testimoniante*, *nombre2\_testimoniante*, etc.).

Componente de *agregación*. Se encarga de asignar a cada uno de los registros que no tengan datos en ninguno de los atributos *nombre1*, *nombre2*, *apellido1*, *apellido2* un valor único. La finalidad de esto es que permanezca el resto de la información asociada a una persona de la cual no se supo nombre ni apellido.

Componente de *filtro*. Se coloca para apartar todos aquellos registros donde la persona que se describe es la misma que el testimoniante. Esto se realiza comparando los atributos *nombres1*, *nombre2*, *apellido1*, *apellido2* con *nombre1\_testimoniante*, *nombre2\_testimoniante*, *apellido1\_testimoniante*, *apellido2\_testimoniante*. El conjunto de datos que se corresponden con la condición son dispuestos sobre un componente *cargador* con los atributos seleccionados para ser ingresados a la nueva tabla *recorrido*.

**Implementación de las decisiones.** Para implementar el proceso ETL visto en la Figura 2 se utilizó la herramienta Talend Open Studio<sup>5</sup> la cual es una herramienta de código abierto y software libre. En la Figura 3 podemos observar la forma en que cada uno de los filtros, agregaciones y joins fueron implementados.

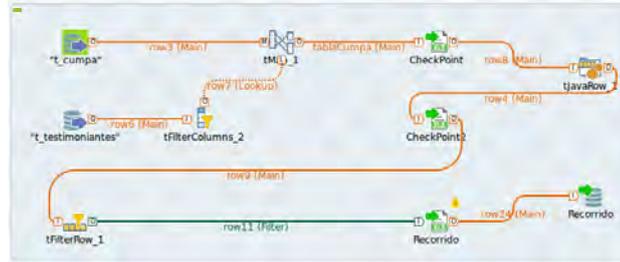
Al igual de como se detalló en el proceso de ETL, se parte de las 2 tablas (*cumpa* y *testimoniante*), para filtrar los atributos deseados del testimoniante mediante el elemento `<< tFilterColumns.2 >>`. Luego, el componente `<< tMap.1 >>` realiza el *join* por *id\_testimoniante* agregando a cada registro de *cumpa* los atributos detallado en el diseño (*nombre1\_testimoniante*, *nombre2\_testimoniante*, etc.) El componente `<< tJavaRow.1 >>` permite exponer cada atributo de cada fila del conjunto a un código en java. Con esto se realiza un chequeo de los valores de los atributos *nombre1*, *nombre2*, *apellido1* y *apellido2* en *cumpa*. Si todos ellos son nulos, se les cambia el valor por las siglas `<< nn >>` concatenado a un valor incremental. Los elementos denominados CheckPoint corresponden a un componente «tFileOutputExcel» de Talend que permite volcar los valores en un XLS con fines de validar las operaciones sobre los datos y detectar errores. El elemento `<< tFilterRow.1 >>` se encarga de realizar el filtrado de aquellos registros donde *nombre1\_testimoniante*, *nombre2\_testimoniante*, *apellido1\_testimoniante*, *apellido2\_testimoniante* se correspondan con *nombre1*, *nombre2*, *apellido1* y *apellido2*. Este conjunto resultante del filtrado es verificado nuevamente para almacenarlo en la nueva tabla *recorrido*.

## 2.2. Objetivo 2: Obtener información georeferencial de las fuentes, para ser reflejada en un mapa interactivo.

Aquí, utilizamos la tabla *secuestros* de ambos sistemas y *victima*, *persona*, *traslado* y *cautiverio* del sistema SIPDJ. Para el caso de las tablas *secuestro* la información almacenada en ambos sistemas es similar con algunas diferencias en cuanto a datos que se pudieron recuperar de los testimonios, como la existencia

<sup>5</sup> <https://es.talend.com/products/talend-open-studio/>

6 Troncoso et al.

Figura 3: Fragmento del trabajo en Talend para crear la tabla *recorrido*

de otras personas secuestradas en el mismo periodo.

### Problemas encontrados.

*Información incompleta, formatos diferentes.* En este caso, como queremos reconocer los lugares para georeferenciarlos en un mapa, analizamos las direcciones de los lugares del secuestro (o domicilios) tal como están almacenados. Como en el caso anterior, vemos en la Figura 4 una de las tablas *secuestros* donde las direcciones se encuentran incompletas y almacenadas en muchas formas diferentes. Por ejemplo se observan direcciones del tipo:

1. San Martín 151, San Miguel de Tucumán
2. Su domicilio, sito en el Pasaje Ecuador 135, barrio El Palomar en La Banda del Río Salí (Cruz Alta, Tucumán)
3. Su domicilio calle Uruguay 4532, San Miguel de Tucumán, Tucumán
4. Me secuestraron el 1ro. de Junio de 1978 en un bar a una cuadra de General Paz y Av. de los Constituyentes.

id	lugar_secuestro	momento_del_dia_secuestro
1	5 lugar	momento
2	13 su domicilio	noche
3	1 "el día 2 de octubre de 1978 lo secuestran en Juan B. Alberdi 5045, lugar donde trabajaba en la ciudad de Buenos A	A las cinco de la tarde
4	40 su domicilio	en la madrugada
5	45 "A mí secuestran el 2 de octubre de 1978 cuando salía de mi trabajo en Juan Bautista Alberdi 5045 de esta ciudad.	a las 15.30 de la tarde
6	47 En la casa de su tía Carmen Aguilar de Lapaco, en Marcelo T. de Alvear 934, entre Carlos Pellegrini y Supacha.	Noche, después de cenar
7	50 Domicilio particular (Lules, Tucumán)	aproximadamente a las dos de la
8	51 El Empalme, Ranchillos (Cruz Alta, Tucumán). ATENCIÓN: no lo explicita en el testimonio pero se infiere que fue allí	
9	52 San Miguel de Tucumán, Tucumán. ATENCIÓN: se infiere, no es especificado en el testimonio.	
10	52 San Miguel de Tucumán (Tucumán). ATENCIÓN: no lo dice el testimonio pero se infiere	
11	54 Vía pública, La Cocha (La Cocha, Tucumán)	*fue detenido, mejor dicho secu
12	55 sin datos, (aparentemente) fue en Ingenio Lules (Lules, Tucumán)	
13	56 su domicilio, Pasaje Vyetes 1462, frente a la plaza del Barrio Victoria a la altura de la Avda. Alem 1450 (San Miguel d	por la mañana, tipo 7.00 hs

Figura 4: Formas en que el *lugar\_secuestro* esta almacenado en la tabla *secuestro* de ASQ

**Diseño ETL.** El diseño ETL para analizar y mejorar las direcciones se puede observar en Figura 5. Las decisiones involucran el uso de filtros y agregación.

Uso de *filtros*. Aquí se decidió en primera instancia, una vez obtenidas las tablas de *secuestro* de ambos sistemas, unificar los mismos datos en una única tabla

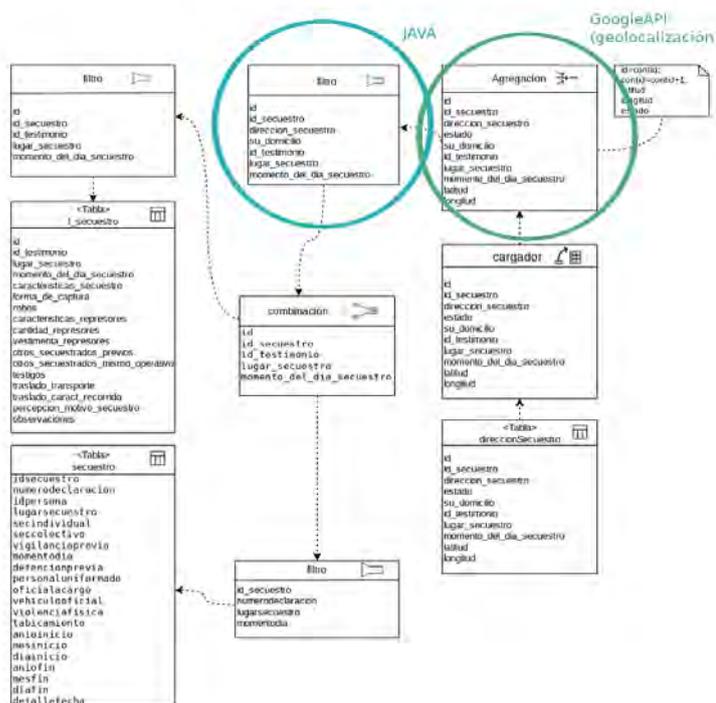


Figura 5: Diseño del proceso ETL para el segundo objetivo

llamada *direccionSecuestro*, con la información referente a la dirección de secuestro y algunos atributos mas. Así para ambas tablas se les aplica un componente *filtro* donde se dejan los atributos *id*, *id\_secuestro*, *id.testimonio*, *lugar\_secuestro* y *momento\_del.dia\_secuestro* para la tabla *sequestro* de ASQ y los atributos *id\_secuestro*, *numerodeclaracion*, *lugarsecuestro* y *momentodia* de SIPDJ. Estos dos conjuntos se combinan en un nuevo *filtro* (en celeste en la Figura 5) encargado de conservar solo los datos que representan una dirección geográfica. Estos datos quedan almacenados en el nuevo atributo *direccion\_secuestro*, y sobre otro nuevo atributo denominado *su.domicilio* se mantiene un valor que indica si sobre la información filtrada se nombró el domicilio de la víctima.

Componente de *agregación* (en verde en la Figura 5). Se presentan los datos filtrados a un sistema de geolocalización que busca convertir esas direcciones en coordenadas geográficas. Los datos que obtienen un resultado en un mapa son almacenados en los nuevos atributos *latitud* y *longitud*. Todos estos nuevos datos resultantes se preparan para ser almacenados mediante el componente cargador sobre la nueva tabla resultante *direccionSecuestro*.

**Implementación de la solución.** Aquí también se utilizaron algunas de las funcionalidades de Talend Studio, como por ejemplo filtrar y combinar ambas

tablas de *secuestro*. Luego, este nuevo conjunto, fue volcado sobre un archivo como entrada del algoritmo de filtrado desarrollado en java<sup>6</sup> y que posee como salida otro archivo normalizado (filtro en celeste en la Figura 5). A partir del filtrado y análisis de estas expresiones, se obtuvo que aproximadamente 162 registros contenían las expresiones «su domicilio» o «su casa», lo que representa un 44% del total de registros en las tablas *secuestro*.

Luego definimos una repetitiva que analiza el archivo de texto como entrada y organiza los datos de salida acorde a los datos que espera un sistema de geo-referenciación. Adicionalmente, el algoritmo crea un nuevo atributo numérico con valores del 1 al 4. Estos valores se generan de acuerdo a la precisión que podríamos obtener en el proceso de geo-referenciación. El valor 1 representa que se disponen con todos los datos para que sea una ubicación específica en un mapa, en cambio, para los casos donde se tienen por ejemplo solo una calle con su altura y una provincia. Así este valor se define de la siguiente forma:

$$\left. \begin{array}{l} \text{calle} + \text{altura} + \text{ciudad} + \text{provincia} = 1 \\ \text{calle} + \text{altura} + [\text{ciudad}][\text{provincia}] = 2 \\ \text{calle} + \text{altura} = 3 \\ \emptyset = 4 \end{array} \right\} \text{precision\_de\_direccion}$$

Una vez finalizado el algoritmo, la dirección es procesada por la API de Google (`<< googleMapsClient.geocode >>`). Aquellos registros cuyo valor fue 4 no se exponen al siguiente proceso de detección, ya que se sabe que no van a poder ser geo-referenciados. A partir del análisis con la API se lograron obtener aproximadamente 300 resultados geo-referenciables que se diferencian por el peso obtenido. Para finalizar se depositan los valores geo-referenciados sobre los atributos *latitud* y *longitud* de la tabla *direccionSecuestro* y se vuelcan en un archivo CSV, mediante Talend. Estos datos nuevos son visibles a través de una Web que se desarrolló para la consulta y visualización de los datos. En la Figura 6<sup>7</sup> podemos observar una pantalla del sistema donde al interactuar con un indicador (punto de geolocalización), se despliega un cuadro de información con un número que corresponde al *Id del registro*, y en la parte inferior la dirección geo-referenciada. A su vez podemos observar los diferentes colores mostrando gráficamente la precisión de la dirección mostrada. Las banderas que se pueden observar en el mapa representan los CCD, que se obtienen a partir de otras fuentes externas a las bases con las que se contaban, y son filtradas, geolocalizadas y almacenadas.

### 3. Lecciones Aprendidas

Luego de haber realizado el proceso ETL, podemos destacar las siguientes lecciones aprendidas:

<sup>6</sup> Aquí se decidió no utilizar la funcionalidad de Talend para expresiones regulares ya que es compleja y posee pocas formas de personalización

<sup>7</sup> Los datos mostrados no representan necesariamente la realidad

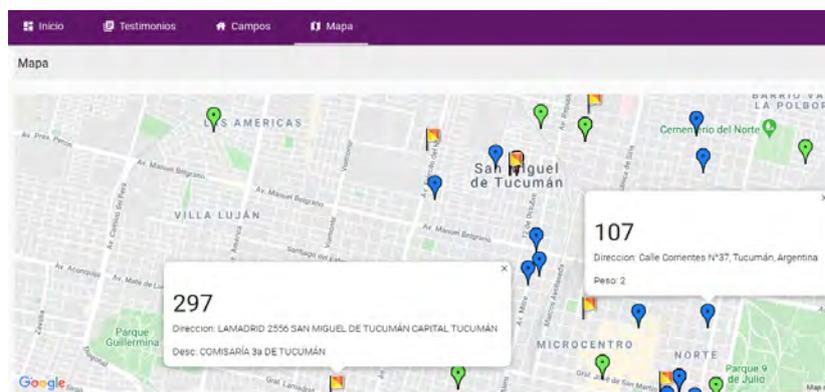


Figura 6: Interfaz del sistema mostrando indicadores geolocalizados

- *El conocimiento del dominio fue fundamental en el proceso:* el trabajo conjunto entre informáticos y expertos del dominio fue fundamental para interiorizarse sobre la forma en que la información estaba almacenada y la forma correcta de normalizarla o reestructurarla. Por la naturaleza del sistema, se almacenan muchos atributos de texto muy largos y con muchas descripciones. Estas descripciones a pesar de que a simple vista parecían redundantes, los expertos hicieron notar la importancia de tenerlas en las declaraciones o testimonios ya que recaban información no precisa, pero sí importante sobre las detenciones. Por lo tanto, se debía realizar una extracción muy cuidadosa de los datos de forma tal de no eliminar dicha información. Es por esto que, por ejemplo, para el caso de los lugares de secuestro, se eligió utilizar colores para indicar la precisión de las direcciones.
- *La aplicación de una metodología de diseño del proceso ETL sirvió para organizar al equipo y reproducir prácticas:* el enfoque de diseño de ETL utilizado basado en UML también facilitó el trabajo debido a que UML era conocido y aplicado por los informáticos que participaron; sólo tuvieron que aprender los mecanismos particulares de ETL. Al mismo tiempo, la metodología permitió que se puedan replicar decisiones de diseño entre los problemas abordados; particularmente cuando se debieron cargar los CCDs en el mapa en donde muchas direcciones estaban también incompletas y se repitieron mecanismos realizados para el primer objetivo.
- *La aplicación de la herramienta open source Talend Studio agilizó también la implementación del proceso:* esta herramienta es muy intuitiva y sencilla, con lo que la implementación pudo realizarse relativamente rápido. A su vez, al tener el diseño ETL ya planteado, el mapeo entre éste y su implementación en Talend fue muchas veces trivial. Igualmente, en algunos casos tuvimos que agregar cierta programación para plasmar justamente las particularidades de este dominio.
- *El resultado del diseño e implementación del proceso ETL generó información útil para ser analizada:* aunque lo definimos fuera del alcance de este

10 Troncoso et al.

trabajo, la información generada se encuentra en un formato ideal para ser analizada. Dentro de los dos objetivos planteados, se pueden efectuar análisis sencillos que involucren recorridos de personas por CCDs, cantidad de CCDs por los que pasó la misma persona, lugares donde fueron secuestrados y recorridos que realizaban una vez detenidos, etc. A su vez se podrían realizar análisis más complejos que ayuden a determinar patrones de comportamiento en cuanto a los secuestros, recorridos recurrentes sobre los CCDs, etc. Esto se plantea como trabajo futuro.

#### 4. Conclusiones y Trabajo Futuro

En este trabajo hemos descrito el diseño e implementación de un proceso ETL aplicado a dos sistemas que almacenan información sobre declaraciones y testimonios de delitos de lesa humanidad. En particular nos hemos centrado en dos objetivos principales y en base a ellos hemos descrito el proceso realizado destacando las lecciones aprendidas durante el mismo.

Como trabajo futuro se plantea continuar registrando las decisiones de diseño e implementación derivadas de nuevos ETLs para ese dominio y a su vez realizar las tareas de análisis de datos propuestas para evaluar la real utilidad de la información resultante para efectuar análisis de datos complejos.

#### Referencias

1. Hamed, I., Ghazzi, F.: A knowledge-based approach for quality-aware etl process. In: 2015 6th International Conference on Information Systems and Economic Intelligence (SIIE). pp. 104–112 (2015)
2. Luján-Mora, S., Vassiliadis, P., Trujillo, J.: Data mapping diagrams for data warehouse design with uml. In: Atzeni, P., Chu, W., Lu, H., Zhou, S., Ling, T.W. (eds.) Conceptual Modeling – ER 2004. pp. 191–204. Springer Berlin Heidelberg, Berlin, Heidelberg (2004)
3. Quix, C., Hai, R.: Data Lake, pp. 1–8. Springer International Publishing, Cham (2018), [https://doi.org/10.1007/978-3-319-63962-8\\_7-1](https://doi.org/10.1007/978-3-319-63962-8_7-1)
4. Simitsis, A., Skoutas, D., Castellanos, M.: Natural language reporting for etl processes. In: Proceedings of the ACM 11th International Workshop on Data Warehousing and OLAP. p. 65–72. DOLAP '08, Association for Computing Machinery, New York, NY, USA (2008), <https://doi.org/10.1145/1458432.1458444>
5. Trujillo, J., Luján-Mora, S.: A uml based approach for modeling etl processes in data warehouses. In: Song, I.Y., Liddle, S.W., Ling, T.W., Scheuermann, P. (eds.) Conceptual Modeling - ER 2003. pp. 307–320. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
6. Vaisman, A., Zimnyi, E.: Data Warehouse Systems: Design and Implementation. Springer Publishing Company, Incorporated, 1st edn. (2016)
7. Vassiliadis, P., Simitsis, A., Skiadopoulos, S.: Conceptual modeling for etl processes. In: Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP. p. 14–21. DOLAP '02, Association for Computing Machinery, New York, NY, USA (2002), <https://doi.org/10.1145/583890.583893>

## SMOTE, Algoritmo para balanceo de clases en un estudio aplicado a la ganadería.

Oswaldo Sposito<sup>1</sup>, Gabriel Blanco<sup>1</sup>, Lorena Matteo<sup>1</sup>, Marcelo Levi<sup>1</sup> y Julio Bossero<sup>1</sup>

<sup>1</sup>Universidad Nacional de La Matanza. Departamento de Ingeniería e Investigaciones Tecnológicas.  
{sposito, g2blanco, lmatteo, mlevi, jbossero}@unlam.edu.ar

**Abstract.** En el estudio de los algoritmos de Minería de Datos del tipo supervisados surge el problema del desbalance de clases, que implica que la información no se encuentre distribuida equitativamente entre todas las clases que la componen, por lo que se generan efectos no deseados en el proceso de clasificación. Este trabajo considera el caso de conjuntos de datos que solamente tiene dos clases y una de ellas cuenta con una mayor cantidad de ejemplos que la otra. El interés principal del trabajo es la aplicación de la técnica de balanceo de clases SMOTE (Synthetic Minority Oversampling Technique), que con algoritmos de interpolación incrementa en forma “sintética” los ejemplos de la clase minoritaria. Los resultados experimentales muestran que algunas técnicas, en el proceso de entrenamiento, obtienen mejores porcentajes de clasificación, cuando se usan estos datos artificiales. El dataset utilizado registra la Diferencia Esperada entre Progenie de animales de la raza Aberdeen Angus.

**Keywords:** Desbalance de clases, SMOTE, Algoritmos Supervisados, Weka, DEP

### 1 Introducción

Se ha observado que algunas de las técnicas de MD del tipo supervisadas, si utilizan un conjunto de datos desequilibrado para la clasificación, presentan un rendimiento de generalización deficiente, debido a un fuerte sesgo hacia las clases mayoritarias [1]. Las clases con el mayor número de instancias se denominan clases mayoritarias y las clases con el menor número de instancias son referidas como las clases minoritarias. Intuitivamente, dado que hay una gran cantidad de ejemplos de clases mayoritarias, un modelo de clasificación tiende a favorecer las clases mayoritarias mientras que clasifica incorrectamente los ejemplos de las clases minoritarias [2].

Las técnicas supervisadas, se expresan mediante algoritmos capaces de tratar y analizar datos de forma automática, con el objeto de extraer cualquier tipo de información subyacente en dichos datos. Como se sabe, en el aprendizaje supervisado, los algoritmos trabajan con datos “*etiquetados*”, intentado encontrar una función que, dadas las variables de entrada, les asigne la etiqueta de salida adecuada.

El algoritmo se entrena con un “histórico” de datos y así “aprende” al asignar la etiqueta de salida a un nuevo valor, es decir, predice el valor de salida [1].

El problema del desbalanceo de las clases, entonces, consiste en la predominancia de ciertos valores en los datos de entrenamiento y la escasez de otros.

Este trabajo, es parte de un proyecto de investigación que se lleva adelante en la Universidad Nacional de La Matanza, denominado “*Uso de Minería de Datos para Mejoramiento Genético en la Raza Aberdeen Angus*”, donde se estudia la aplicación de distintas técnicas de MD con el objeto de encontrar, a partir de los valores genéticos de animales de la raza Aberdeen Angus, patrones o grupos de características que puedan determinar a priori, el peso de los terneros al nacer. En esta investigación se estudiaron tanto algoritmos del tipo No Supervisados como Supervisados. En la aplicación de los primeros, se utilizaron las siguientes técnicas: EM (Expectation Maximization), FarthestFirst, Simple K-Means y Mapas AutoOrganizados (Redes SOM). Ese trabajo titulado “*Peso al Nacer de Terneros Aberdeen Angus mediante Algoritmos No Supervisados*”, fue publicado en CoNaIISI 2019 [3]. En este trabajo se puede leer el procedimiento realizado para obtener los datos usados para confeccionar la vista minable [4]. Estos mismos datos fueron utilizados en los experimentos con métodos supervisados y en particular en el estudio que es motivo de esta presentación.

En cuanto al estudio, sobre el empleo de técnicas Supervisadas, se realizó una comparación entre tres clasificadores: Árbol de Decisión (AD), Red Neuronal Artificial (RNA), del tipo Perceptron Multicapa y una Máquina de Soporte Vectorial (MSV). El trabajo se tituló: “*Clasificación del Peso al Nacer de Terneros Aberdeen Angus mediante Algoritmos Supervisados*”, y fue presentado y aprobado en JAIOO 2019, pero no llegó a publicarse.

En ambos trabajos se utilizó un modelo de datos o vista minable, compuesto con datos provenientes de las evaluaciones genéticas de los toros: conocidos como la Diferencia Esperada entre Progenie (DEP) [5], que permite a los productores tomar decisiones de selección en base a información objetiva. Los DEP’s anticipan, cómo será el comportamiento promedio de las futuras crías de un toro. Esos datos se complementaron, además, con datos de las hembras (edad, cantidad de partos, etc.) y genéticos de los padres de las hembras. Con este conjunto de datos, se alcanzaron valores aceptables de predicción, respecto de la variable a clasificar: Peso al nacer (PN), valor que de acuerdo a los veterinarios es determinante del potencial del desarrollo futuro del animal. Dicha variable se estableció, para los trabajos mencionados anteriormente, en “*Alta*” y “*Baja*”. Un PN promedio es de 38 kilogramos, por tal motivo, a los pesos mayores o iguales a 38 kg., se los clasificó como Alto y al resto como Bajo [3].

Se concluyó que, con los algoritmos utilizados, en este último estudio, se obtuvieron valores aceptables (porcentaje de aciertos superior al 60%), en referencia a las métricas: *precisión, sensibilidad y especificidad*. Y que, para estos datos y con la configuración propuesta, para cada técnica, por el software WEKA<sup>1</sup> (Waikato Environment for Knowledge Analysis, en español “*entorno para análisis del conocimiento de la Universidad de Waikato*”) [7], el algoritmo Árbol de Decisión

<sup>1</sup> Es una plataforma de software para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato. Weka es software libre distribuido bajo la licencia GNU-GPL. [www.cs.waikato.ac.nz/~ml/weka/](http://www.cs.waikato.ac.nz/~ml/weka/)

tuvo la mejor precisión, es decir, la mejor probabilidad de discriminar correctamente, debido a que el valor de su media muestral fue del 72.5%. Cabe resaltar que el indicador precisión es una de las medidas principales, para establecer el desempeño de un algoritmo de clasificación en el área de la MD [1].

El principal interés en este trabajo es mejorar la clasificación obtenida en el trabajo expuesto en el párrafo anterior. Para ello, se modificaron, en la vista minable, los objetos de la clase minoritaria, sin eliminar objetos de la clase mayoritaria, lo cual, según la literatura actual, puede producir pérdida de información importante [2][8], mediante la aplicación del filtro SMOTE (cuya traducción al español es “*técnica de sobre muestreo de minorías sintéticas*”) [9]. Luego, se realizó una nueva comparación, utilizando ese nuevo set de datos, a través del mismo software WEKA, usando siempre la configuración por defecto. A continuación, se hace una reseña de algunos trabajos relacionados con el uso del algoritmo SMOTE. No se han encontrado trabajos en donde se aplique la técnica SMOTE relacionado con la ganadería.

### Antecedentes y Trabajos Relacionados

El problema del desbalanceo de clases es una cuestión que se está abordando en la actualidad de forma activa y son muchos los investigadores que estudian y proponen nuevas técnicas para poder hacerle frente a este problema. La mayoría de ellos solo se han concentrado en resolver situaciones como la nuestra de clasificación que tienen que ver con dos clases. En [8], se encuentra un detalle de varios trabajos relacionados con el desbalanceo de clases ordenados cronológicamente. Algunos otros trabajos encontrados respecto al uso del algoritmo SMOTE son los siguientes:

- De Jesús, Juan. (2016). *Técnicas de muestreo para mejorar el rendimiento del algoritmo back-propagation en problemas de desbalance de clases: Un estudio empírico sobre la clasificación en imágenes de percepción remota*.  
Resumen: En este trabajo se analizaron las diferentes técnicas de re muestreo para tratar el desbalance de clases en dominios de dos clases con ayuda de datos de imágenes de percepción, para determinar que técnica arroja un mejor clasificador. Los resultados mostraron aquellos algoritmos que mejor funcionaron al momento de clasificar, de acuerdo a un análisis estadístico aplicado a los algoritmos.
- J. Monroy de Jesús y otros. (2018). *Algoritmo de aprendizaje eficiente para tratar el problema del desbalance de múltiples clases*.  
Resumen: Los resultados demostraron que la diversidad de versiones de algoritmos y cuan competitivos resultaron en el desempeño de la clasificación con respecto a los métodos de sobre-muestreo y sub-muestreo (ROS, SMOTE y RUS).
- David Municio Duran. (2019). *Técnicas de oversampling aplicadas al análisis de imágenes hiperespectrales*.  
Resumen: En este trabajo se ha estudiado el impacto de diferentes algoritmos de oversampling en el proceso de clasificación de imágenes hiperespectrales. El objetivo del trabajo fue mejorar los resultados de clasificación en imágenes con alta dimensionalidad y gran desbalanceo entre clases.

- Rosa María Valdovinos Rosas. (2006). *Técnicas de Submuestreo, Toma de Decisiones y Análisis de Diversidad en Aprendizaje Supervisado con Sistemas Múltiples de Clasificación*.

Resumen: Este trabajo se encarga del estudio de Sistemas Múltiples de Clasificación (SMC), para el reconocimiento de patrones. El trabajo se centró en la limpieza de un conjunto de datos, se emplearon para la reducción del tamaño del conjunto de entrenamiento, un algoritmo de subconjunto selectivo modificado, y para la generación de patrones sintéticos, se utilizó el algoritmo SMOTE.

## 2. Algoritmo de balanceo de clases

Para este trabajo se hace uso del algoritmo SMOTE como parte de los métodos basados en muestreo (sampling), para balanceo de clases [2]. Se ha elegido SMOTE debido a que es uno de los algoritmos más utilizados para tratar el problema de desbalanceo de clases [8][10]. SMOTE es una técnica basada en sobremuestreo (oversampling), que genera instancias “*sintéticas*” o artificiales en el espacio de atributos con el objetivo de equilibrar la muestra de datos basado en la regla del Vecino más cercano [1]. Esta regla consiste en suponer que instancias próximas entre sí, tienen mayor probabilidad de pertenencia a la misma clase. La generación se realiza interpolando nuevas instancias en lugar de duplicarlas como hacen los algoritmos del tipo re-muestreo (resampling) [1]. Para cada una de las instancias minoritarias se buscan las instancias minoritarias vecinas (más cercanas). Se crean  $N$  o  $\alpha$  (alfa) instancias sobre el segmento que une la instancia original y cada una de las vecinas. Un trabajo de Rodríguez Torres [11], realizó una descripción algorítmica de esta técnica, que puede ser reescrita en estos tres pasos:

- 1) Se determina la cantidad promedio de interpolaciones que debe aportar cada elemento minoritaria:  $\alpha$  (ya sea calculado o elegido).
- 2) Se asigna a cada muestra minoritario el número  $\lfloor \alpha \rfloor$  o  $\lceil \alpha \rceil$ , cantidad de interpolaciones en las cuales intervendrá ya sea con ayuda de azar o determinístico de modo tal de lograr la cantidad de muestras artificiales deseada.
- 3) Luego se procesa cada elemento de la muestra calculando según el número asignado la cantidad de vecinos cercanos interpolando por azar un punto sobre el segmento que los une.

La Figura 1 muestra un ejemplo de procedimiento que utiliza el algoritmo SMOTE, a continuación, se explica el proceso:

- A) Para cada ejemplo de la clase minoritaria  $k$ , se calcula el vecino más cercano ( $i, j, l, n, m$ ).
- B) Se elige aleatoriamente un ejemplo de 5 puntos más cercanos.
- C) Se genera sintéticamente el evento  $k_l$ , de modo que  $k_l$  se encuentra entre  $k$  e  $i$ .
- D) En esta imagen se ve el conjunto de datos después de aplicar SMOTE 3 veces.

<sup>2</sup> En matemática, una forma de considerar un número entero más próximo a un número real dado, se pueden considerar: Al entero inferior como piso ( $\lfloor \cdot \rfloor$ ) y al entero superior como techo ( $\lceil \cdot \rceil$ ).

Estos ejemplos sintéticos ayudan a equilibrar la distribución original de la clase, que generalmente mejora significativamente el aprendizaje. Sin embargo, el algoritmo SMOTE también tiene su desventaja, como por ejemplo generalización del espacio de clase minoritaria [12].

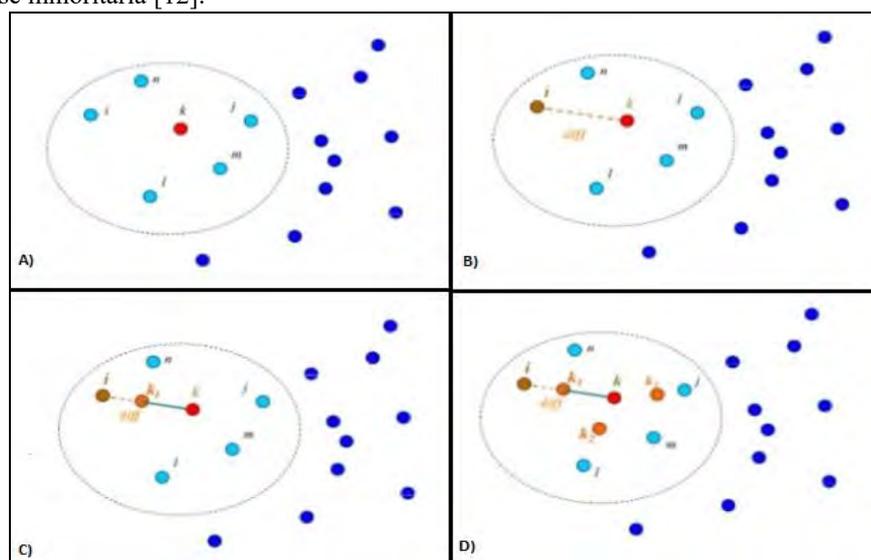


Fig. 1. Ejemplo del algoritmo SMOTE [16].

A partir del algoritmo SMOTE original, se han desarrollado muchos otros algoritmos basados en SMOTE a lo largo de los años y algunos de ellos mejoran efectivamente el rendimiento en el aprendizaje desequilibrado [2][10]. El principal inconveniente del algoritmo es el alto coste computacional que tiene.

### 3 Método utilizado y resultados

Como ya se mencionó, este trabajo representa una continuación de un trabajo ya realizado. Pero, en esta oportunidad, se realizó una comparación entre los resultados con los datos originales del trabajo, y se los comparó contra los resultados obtenidos, al ejecutar los mismos algoritmos con datos sintéticos. Estos datos creados con el filtro SMOTE, a través del software WEKA. El modelo original, descrito en [3], corresponde a los años 2017 y 2018 y contiene un total de 360 ejemplares hembras. Estas cuales fueron inseminadas, por inseminación artificial, por dos reproductores de la cabaña Las Lilas. La nómina de variables utilizadas se muestra en la Tabla 1.

**Tabla 1.** Descripción de las variables del conjunto de datos.

Nomenclatura	Tipo de Dato	Descripción
ID	Númérico	Identificación de la instancia
PAN_Padre	Númérico	Peso al Nacer
PAD_Padre	Númérico	Peso al Destete
PAF_Padre	Númérico	Peso Final
PAdulto_Padre	Númérico	Peso Real
CEsc_Padre	Númérico	Circunferencia Escrotal
Frame_Padre	Númérico	Altura
Certiv_Padre	Númérico	Edad promedio de los vientres primerizos
PNDEP_Padre	Númérico	DEPs del Toro Progenitor (Padre)
PDDEP_Padre	Númérico	
AMDEP_Padre	Númérico	
CMDEP_Padre	Númérico	
PFDEP_Padre	Númérico	
CEDEP_Padre	Númérico	
AOBDEP_Padre	Númérico	
GDDEP_Padre	Númérico	
MARDEP_Padre	Númérico	
EdadMeses_Madre	Númérico	
PAN_Madre	Númérico	
PAD_Madre	Númérico	
UltPeso_Madre	Númérico	
CantNac_Madre	Númérico	
CantAbortos_Madre	Númérico	
CantCesareas_Madre	Númérico	
MuertesAntesDestete_Madre	Númérico	
CEsc_AbueloM	Númérico	DEPs del Toro Progenitor de la Vaca (Abuelo Materno)
Frame_AbueloM	Númérico	
Certiv_AbueloM	Númérico	
PNDEP_AbueloM	Númérico	
PDDEP_AbueloM	Númérico	
AMDEP_AbueloM	Númérico	
CMDEP_AbueloM	Númérico	
PFDEP_AbueloM	Númérico	
CEDEP_AbueloM	Númérico	
AOBDEP_AbueloM	Númérico	
GDDEP_AbueloM	Númérico	
MARDEP_AbueloM	Númérico	
Pnacer_Hijo	Texto	Variable Objetivo en [16], una más en este estudio.

Se describen brevemente las técnicas que se compararon. Para más información, sobre estos algoritmos se pueden consultar en: [1][7][13].

- Árboles de Decisión (AD): como su nombre lo indica es una estructura que se forma por las bifurcaciones en cada una de las decisiones, descubriendo reglas. En WEKA se lo conoce como algoritmo J48, que es una implementación libre en java del algoritmo C4.5, que utiliza el concepto de entropía de la información para la selección de variables que mejor clasifiquen a la variable PN (clase) estudiada.
- Red Neuronal Artificial (RNA): Esta implementación imita el funcionamiento interno de las neuronas humanas [1]. En general, aunque pueden usarse muchos tipos de RNA para clasificación, se han usado las redes multicapa feedforward o perceptron multicapa (MLP) que son los clasificadores basados en redes neuronales más ampliamente estudiados y utilizados. Este algoritmo es entrenado para realizar conexiones entre los valores de entrada y salida, aprendiendo de su error de pronóstico.
- Máquinas de Soporte Vectorial, buscan el límite que separa las clases con el mayor margen posible [1][14], Una de las características de esta técnica es que cuando no se pueden separar correctamente las dos clases, el algoritmo busca el mejor límite posible. Las MVS efectúan esto, sólo con una línea recta (usa un kernel lineal) y gracias a esta aproximación lineal, se puede ejecutar con bastante rapidez.

Como se hizo anteriormente, para la evaluación de los clasificadores se empleó una Matriz de Confusión (MC) [1][13] y el análisis o curvas ROC (acrónimo de Receiver

Operating Characteristic) [1][15], que entrega WEKA, como resultado luego de testear cada uno de los clasificadores. A modo de resumen, una matriz de confusión muestra la clasificación de las instancias. Brinda información muy útil porque no sólo refleja los errores producidos sino también informa del tipo de éstos. Donde:

- VP es la cantidad de positivos que fueron clasificados correctamente como positivos por el modelo.
- VN es la cantidad de negativos que fueron clasificados correctamente como negativos por el modelo.
- FN es la cantidad de positivos que fueron clasificados incorrectamente como negativos.
- FP es la cantidad de negativos que fueron clasificados incorrectamente como positivos.

De estos valores se definen dos métricas asociadas importantes: Sensibilidad y especificidad:

- La sensibilidad nos indica la capacidad de nuestro estimador para dar como casos positivos los casos realmente lo son; proporción de pesos altos correctamente identificados.
- La especificidad nos indica la capacidad de nuestro estimador para dar como casos negativos los casos realmente lo sean; en nuestro caso, proporción de pesos bajos correctamente identificados.

Por último, la curva ROC, es una herramienta estadística utilizada en el análisis de los clasificadores, determinando la capacidad discriminante de una prueba.

Para realizar la comparación se tomaron los 360 datos originales y luego de cargarlos en WEKA, se le aplicó el filtro SMOTE. Weka permite configurar el valor del vecino más cercano ( $\alpha$ ) y por el porcentaje de instancias que se necesita crear para que las clases se balanceen. En este caso no se usó el valor por defecto del programa (100%), sino que se ajustó al 40%, para igualar las clases. En la figura 2 se puede observar la distribución de la variable PN, en los datos originales y en la figura 3, la distribución de valores luego de aplicar el filtro.

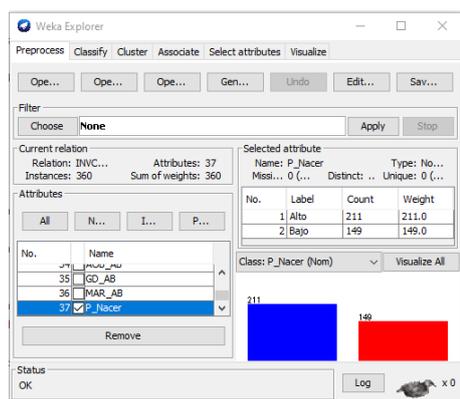


Fig. 2. Distribución con datos originales.

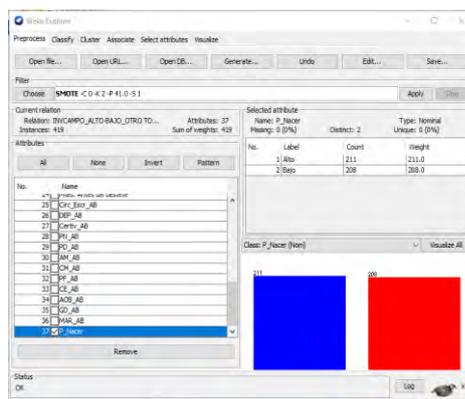


Fig. 3. Distribución con datos sintético.

Como se observa los 360 datos originales, se distribuían en: 211 instancias de PN Altas y 149 del tipo Bajas, luego de aplicar el filtro, las instancias se convirtieron en 419, es decir, 59 instancias más en el peso bajo. Dejando al PN Alta, con la misma cantidad.

Luego de ejecutar cada uno de los algoritmos, con el nuevo set de datos, en la siguiente Tabla es posible comparar los resultados obtenidos. En la solapa Clasificar (Classify), se realiza el entrenamiento de cada algoritmo en WEKA, cada uno de ellos se entrenó usando la opción *Use training set* [7].

**Tabla 2.** Comparación de los resultados obtenidos de las pruebas de clasificación con datos originales y con datos sintéticos.

Porcentaje de Instancias Clasificadas Datos Originales						Porcentaje de Instancias Clasificadas Datos con instancias artificiales					
RNA		MVS		Árbol de Decisión		RNA		MVS		Árbol de Decisión	
Correctas: 63.9 %		Correctas: 60.3 %		Correctas: 72.5 %		Correctas: 66.8 %		Correctas: 52.9 %		Correctas: 65.9 %	
Incorrectas: 36.1 %		Incorrectas: 39.7 %		Incorrectas: 27.5 %		Incorrectas: 33.2 %		Incorrectas: 47.1 %		Incorrectas: 34.1 %	
Matriz de Confusión a-Alto b-Bajo						Matriz de Confusión a-Alto b-Bajo					
a	b	a	b	a	b	a	b	a	b	a	b
146	65	173	38	184	27	186	25	194	17	150	61
65	84	105	44	72	77	114	94	180	28	82	126
Curva Roc						Curva Roc					
0.68		0.558		0.746		0.749		0.527		0.688	

Se observa que solo la RNA mejoró su precisión, respecto a la prueba anterior, pero sin llegar a tener el porcentaje obtenido por el mejor clasificador (AD). En cambio, los otros clasificadores (AR y MVS) mermaron sus porcentajes, respecto a la prueba anterior. En relación con el área bajo la curva (AUC) en las figuras 4 y 5, está la gráfica de ambas curvas. AUC puede interpretarse como la probabilidad de que, ante una instancia nueva de datos, la prueba los clasifique correctamente. Su rango de valores va desde 0, siendo este valor el correspondiente a una prueba sin capacidad discriminante, hasta 1, que es cuando los dos grupos están perfectamente diferenciados por la prueba. Por tanto, podemos decir que cuanto mayor sea el AUC, mejor será la prueba. También la técnica RNA, tuvo el mejor valor AUC, sin llegar al nivel de la técnica AR, de la prueba anterior.

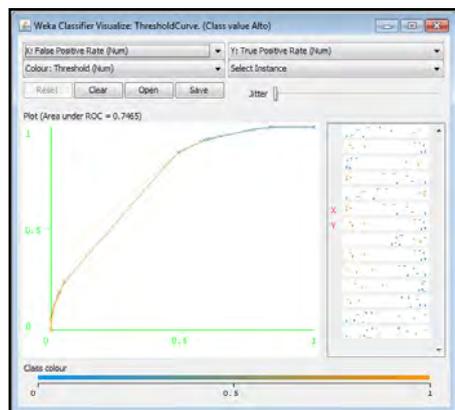


Fig. 4. Curva Roc del AD con los datos originales.

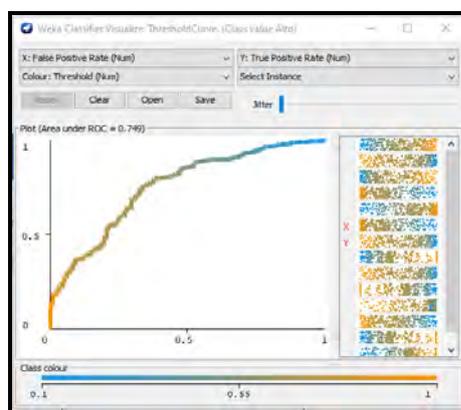


Fig. 5. Curva Roc de la RNA con datos formados por el filtro SMOTE.

#### 4 Conclusiones y Trabajo a Futuro

Este estudio tuvo como objetivo verificar el comportamiento de tres algoritmos supervisados, con un dataset que tiene una porción de los datos construidos en forma artificial, mediante la aplicación del algoritmo de sobre-muestreo SMOTE, provisto por el software WEKA.

Después de haber efectuado todas las pruebas pertinentes sobre los modelos de clasificación propuestos usando estos datos sintéticos, es posible elaborar una serie de conclusiones:

- Es el primer trabajo sobre desbalanceo de clases en el área de la ganadería con estos tipos de datos.
- Dependiendo del tipo de algoritmo, se demuestra que, si bien, con algunas técnicas se puede mejorar los porcentajes de instancias bien clasificadas, en otros casos, los porcentajes disminuyen.
- Si bien, para la comparación se usó el valor alfa propuesto por WEKA ( $\alpha=5$ ), se probó con distintos valores: 2,3,4 y 10, para realizar las pruebas. Dando distintos porcentajes dependiendo del algoritmo.

Para futuras investigaciones se prevé seguir con las siguientes líneas:

- Realizar una comparación con los algoritmos No supervisados, empleados también en la investigación anterior [3].
- Probar con vistas minables de mayor cantidad de muestras y con un mayor porcentaje de diferencia entre las clases mayoritarias y minoritarias.
- Probar con vistas minables que posean más de dos clases.
- Probar con otras variantes del algoritmo SMOTE [8] y con otros algoritmos que tratan el desbalanceo de clase. Por ejemplo: el método de Sobremuestreo por agrupaciones o Clustered Based Oversampling (CBOS).
- Estudiar la relación de los porcentajes de precisión, dependiendo del valor  $\alpha$  propuesto.
- Incursionar en otros programas para el aprendizaje automático y la minería de datos.

## Referencias

- 1 Hernández Orallo, Introducción a la minería de datos, Pearson, 2004.
- 2 Mera, C. & Arrieta Ramos, J.M., «Estudio Comparativo de Técnicas de Balanceo de Datos en el Aprendizaje de Múltiples Instancias», 2015. [En línea]: [https://www.researchgate.net/publication/283642919\\_Estudio\\_Comparativo\\_de\\_Tecnicas\\_de\\_Balanceo\\_de\\_Datos\\_en\\_el\\_Aprendizaje\\_de\\_Multiples\\_Instancias/link/5642151a08aec448fa621f60/download](https://www.researchgate.net/publication/283642919_Estudio_Comparativo_de_Tecnicas_de_Balanceo_de_Datos_en_el_Aprendizaje_de_Multiples_Instancias/link/5642151a08aec448fa621f60/download). [Último acceso: 01/07/2020].
- 3 Sposito, O., «Peso al Nacer de Terneros Aberdeen Angus mediante Algoritmos No Supervisados,» 2019. [En línea]: [https://www.researchgate.net/profile/Lorena\\_Matteo/publication/337445353\\_Peso\\_al\\_Nacer\\_de\\_Terneros\\_Aberdeen\\_Angus\\_mediante\\_Algoritmos\\_No\\_Supervisados/links/5dd7f187458515dc2f439029/Peso-al-Nacer-de-Terneros-Aberdeen-Angus-mediante-Algoritmos-No-Supervisad](https://www.researchgate.net/profile/Lorena_Matteo/publication/337445353_Peso_al_Nacer_de_Terneros_Aberdeen_Angus_mediante_Algoritmos_No_Supervisados/links/5dd7f187458515dc2f439029/Peso-al-Nacer-de-Terneros-Aberdeen-Angus-mediante-Algoritmos-No-Supervisad). [Último acceso: 01/07/2020].
- 4 Quinteros, O. y otros., «Construcción de una vista minable para aplicar minería de datos secuenciales temporales», 2016. [En línea]: <http://sedici.unlp.edu.ar/handle/10915/56747>. [Último acceso: 01/07/2020].
- 5 Monti, A., «Interpretación y uso correcto de los DEPs como herramienta de selección», 1998. [En línea]: <https://es.scribd.com/document/337947981/20-Interpretacion-Deps>. [Último acceso: 01/07/2020].
- 6 Simeone, O., «A Very Brief Introduction to Machine Learning», 2018. [En línea]: <https://arxiv.org/pdf/1808.02342.pdf>. [Último acceso: 01/07/2020].
- 7 Witten I., Data Mining. Practical Machine Learning Tools and Techniques., Morgan Kaufmann. ISBN: 0-12-088407-0., 2005.
- 8 Castro Pérez, N. «Preprocesamiento de datos termográficos por medio de técnicas de balanceo de clases y análisis de cúmulos (Clustering)», 2013. [En línea]: <http://docplayer.es/17472207-Preprocesamiento-de-datos-termograficos-por-medio-de-tecnicas-de-balanceo-de-clases-y-analisis-de-cumulos-clustering.html>. [Último acceso: 2020 07 01].
- 9 Chawla, N.V., «SMOTE: Synthetic Minority Oversampling Technique», 2002. [En línea]: [https://www.jair.org/index.php/jair/\\_article/view/10302](https://www.jair.org/index.php/jair/_article/view/10302). [Último acceso: 01/07/2020].
- 10 Moreno, J. «SMOTE-I: mejora del algoritmo SMOTE para balanceo de clases minoritarias», 2009. [En línea]: [https://www.researchgate.net/publication/229045207\\_SMOTE-I\\_mejora\\_del\\_algoritmo\\_SMOTE\\_para\\_balanceo\\_de\\_clases\\_minoritarias/\\_citation/download](https://www.researchgate.net/publication/229045207_SMOTE-I_mejora_del_algoritmo_SMOTE_para_balanceo_de_clases_minoritarias/_citation/download). [Último acceso: 01/07/2020].
- 11 Rodríguez Torres, F., «SMOTE-D, una versión determinista de smote», 2017. [En línea]: <https://inaoe.repositorioinstitucional.mx/jspui/bitstream/1009/335/1/RodriguezTF.pdf>. [Último acceso: 01/07/2020].
- 12 Huang, P. «Classification of Imbalanced Data Using Synthetic Oversampling Techniques», 2015. [En línea]: <https://escholarship.org/content/qt72w743h7/qt72w743h7.pdf>. [Último acceso: 01/07/2020].
- 13 Jiawei H. y otros, Data Mining: Concepts and Techniques, <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>: 3ra. Edición. (2011). ISBN 978-0-12-381479-1.
- 14 Fariás Concha, M., «Máquinas Vectoriales híbridas para clasificar accidentes de tránsito en la región metropolitana», 2011. [En línea]: [http://opac.pucv.cl/pucv\\_txt/txt-9500/UCF9980\\_01.pdf](http://opac.pucv.cl/pucv_txt/txt-9500/UCF9980_01.pdf). [Último acceso: 01/07/2020].
- 15 Benavides, Ana, «Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones», 2017. [En línea]: <https://idus.us.es/bitstream/handle/11441/63201/Valle%20Benavides%20Ana%20Roc%20del%20del%20TFG.pdf?sequence=1>. [Último acceso: 01/07/2020].
- 16 Dal Pozzo, A. y otros. «Racing for unbalanced methods selection», 2013. [En línea]: <https://www.slideshare.net/dalpozz/racing-for-unbalanced-methods-selection>. [Último acceso: 01/07/2020].

## Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales

Luciano Marrero<sup>1</sup> , Verena Olsowy<sup>1</sup> , Fernando Tesone<sup>1</sup> , Pablo Thomas<sup>1</sup> , Lisandro Delia<sup>1</sup> , Patricia Pesado<sup>1</sup> 

<sup>1</sup> Instituto de Investigación en Informática LIDI  
Facultad de Informática - Universidad Nacional de La Plata – Argentina  
Centro Asociado Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

{lmarrero, volsowy, ftesone, pthomas, ldelia, ppesado}@lidi.info.unlp.edu.ar

**Resumen:** Los Sistemas de Gestión de Bases de Datos (SGBDs) no relacionales (NoSQL) surgen como una alternativa de solución a problemas no resueltos eficientemente por los SGBDs tradicionales. NoSQL, a diferencia del modelo relacional, no responde a un tipo de Base de Datos, sino que representa un conjunto de tipos de Bases de Datos, con diferentes formas y características para representar la información. Este trabajo representa la continuación del estudio presentado en CACIC 2019 [1] y tiene como objetivo comparar y analizar cuatro motores de Bases de Datos NoSQL y un motor de Base de Datos relacional utilizando diferentes esquemas bajo un gran volumen de datos.

**Keywords:** Bases de Datos Relacionales, Bases de Datos NoSQL, Almacenamiento Clave-Valor, Almacenamiento Documental, Almacenamiento de Familia de Columnas, Almacenamiento Orientado a Grafos.

### 1 Introducción

Actualmente, la mayoría de las aplicaciones son multiplataforma. El avance en la comunicación digital, el acceso constante a la información y el aumento en la utilización de tecnología móvil genera que estas aplicaciones sean de uso común por millones de usuarios simultáneamente. El modelo relacional de Codd (1970), que dio origen a los sistemas de Bases de Datos Relacionales, es sin duda, 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, y ha dado lugar a un conjunto de alternativas que plantean almacenar la información de forma no estructurada. 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 adecuadas por su escalabilidad y tienden a utilizar modelos de consistencia relajados con el objetivo de lograr mayor performance y disponibilidad [2, 3, 4, 12]. Existe una gran variedad de motores de Bases de Datos NoSQL que se pueden catalogar en una de estas cuatro categorías:

**Almacenamiento Clave/Valor:** simples en cuanto a su implementación, almacenan datos como un conjunto de pares “clave/valor” (key-value). La clave (key) representa un identificador único que puede retornar un objeto denominado valor (value). Por ejemplo, Redis y Amazon DynamoDB, entre otros, implementan este tipo de almacenamiento [20, 21].

**Almacenamiento Documental:** el concepto central de este tipo de almacenamiento es el documento. Una Base de Datos NoSQL Documental, almacena, recupera y gestiona datos de documentos. Estos documentos encapsulan y codifican datos o información bajo algún formato estándar (XML, YAML, JSON, BSON). Por ejemplo, MongoDB y Apache CouchDB, entre otros, son implementaciones de Bases de Datos Documentales [5, 8, 17, 22].

**Almacenamiento de Familia de Columnas:** en este tipo de almacenamiento los datos se encuentran organizados por columnas, en lugar de por filas. En estas Bases de Datos Columnares por cada entrada, hay una columna, por lo tanto, los datos de cada entrada están dispuestos uno debajo del otro (y no uno al lado del otro, como en la variante orientada a filas). El objetivo de este tipo de almacenamiento es acceder con mayor precisión a la información que se necesita para responder a una consulta. Además, elimina la necesidad de explorar y descartar datos no deseados sobre una fila. Por ejemplo, Apache Cassandra y Apache HBase, entre otros, utilizan este tipo de almacenamiento [8, 9, 19, 23].

**Almacenamiento de Grafos:** en este tipo de almacenamiento se representa a la Base de Datos bajo el concepto de un grafo. Permite almacenar la información como nodos de un grafo y sus respectivas relaciones (aristas) con otros nodos. Aplica la teoría de grafos y son muy útiles para almacenar información en modelos que poseen numerosas relaciones entre sus datos. Por ejemplo, Neo4j y OrientDB, entre otros, permiten este tipo de almacenamiento [11, 24].

Así como las propiedades de ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad) están presentes en las Bases de Datos Relacionales, en las Bases de Datos NoSQL se tienen las propiedades BASE (**B**ase **A**vailability, **S**oft State, **E**ventual Consistency). Estas propiedades tienen su pilar principal en la disponibilidad. El sistema estará básicamente disponible todo el tiempo (**BA**), no necesariamente será consistente (**S**), pero en algún momento lo será, y eventualmente estará en un estado conocido (**E**). Las propiedades BASE difieren de las propiedades ACID, mientras ACID es más rígido y fuerza la consistencia, BASE permite un estado de consistencia flexible. Esto permite grandes niveles de escalabilidad [4, 7, 10, 13].

Este trabajo se centra en un análisis comparativo de rendimiento para cuatro casos de estudio, implementados en 5 motores de Bases de Datos distintos, 4 NoSQL (Cassandra, MongoDB, Neo4j y Redis) y un motor de Base de Datos Relacional (MySQL). Se plantean un conjunto de consultas para cada caso de estudio y para cada motor de Bases de Datos. Se realizaron 10 consultas para MySQL, 7 para MongoDB, 8 para Cassandra, 8 para Neo4j y 2 para Redis. En total se realizaron pruebas con 35 consultas utilizando un gran volumen de información. Cada caso de estudio posee sus propias características con el objetivo de comparar el desempeño de diferentes motores de Bases de Datos en una diversidad de escenarios.

A partir de la sección 2 el trabajo se organiza del siguiente modo: se explican características generales de los motores de Bases de Datos seleccionados, en la sección 3 se presentan los casos de estudio, en la sección 4 se analizan los resultados obtenidos y finalmente se presentan las conclusiones y trabajos a futuro.

## 2 Motores de Bases de Datos utilizados

Para realizar la experimentación presentada en este trabajo, se seleccionaron los siguientes motores de Bases de Datos: MySQL (Almacenamiento Relacional), MongoDB (Almacenamiento Documental), Apache Cassandra (Almacenamiento de Familia de Columnas), Neo4j (Almacenamiento Orientado a Grafos) y Redis (Almacenamiento Clave-Valor). Los motores de Bases de Datos NoSQL elegidos son los más populares en sus categorías [14].

*MySQL* es uno de los sistemas de Base de Datos Relacional más reconocido y popular del mercado con licencia dual (licencia pública general y licencia comercial). Es utilizado por numerosas empresas a nivel mundial, entre ellas se destacan, Facebook, Twitter, YouTube, GitHub, Booking.com, Spotify [15, 16].

*MongoDB* es la Base de Datos NoSQL Documental con mayor popularidad, es de código abierto y multiplataforma. Diferentes empresas como Telefónica, Facebook, Nokia, entre otras, hacen uso de MongoDB [5, 6, 17, 18].

*Apache Cassandra* es la Base de Datos NoSQL con almacenamiento en Familia de Columnas con mayor popularidad. Es multiplataforma, diseñada para trabajar de forma distribuida con grandes volúmenes de datos. Empresas como Instagram, Netflix, GitHub, entre otras, hacen uso de Cassandra [19, 21].

*Neo4j* es la Base de Datos Orientada a Grafos (BDOG) con mayor popularidad, escrita en JAVA y posee licencia dual (comercial y AGPL). Se basa en Grafos para representar datos y las relaciones entre ellos. Empresas como Walmart, Ebay y Cisco hacen uso de Neo4j [11, 24].

Finalmente, *Redis* es la Base de Datos Clave-Valor con mayor popularidad. Redis es una Base de Datos que opera sus estructuras de datos en memoria primaria, ofrece alto rendimiento, posee persistencia opcional y tiene licencia de Software Libre BSD. Es utilizada por diferentes empresas, Stack Overflow, Twitter, GitHub, entre otras [20].

## 3 Casos de estudio

Para cada caso de estudio se realizó previamente el Modelo Conceptual de Datos utilizando un Diagrama de Entidad Relación (DER). Posteriormente, se realizó la derivación al esquema físico correspondiente a cada motor de base de datos utilizado.

En MySQL se generó el modelo relacional. En MongoDB las entidades se representan mediante colecciones de documentos BSON y las relaciones se expresan mediante documentos embebidos [1, 2, 17]. En Cassandra, los datos no se almacenan

en términos de entidades, sino que se representan en términos de consultas, por lo tanto, es importante definir las consultas a realizar [1, 2, 9, 19]. En el caso de Neo4j, los datos se representan mediante nodos y las relaciones son las aristas entre éstos [2, 24]. En el caso de Redis, al ser una Base de Datos en memoria principal, se pensó en el conjunto de claves que serían necesarias tener vigentes de acuerdo con las consultas a realizar [2, 20].

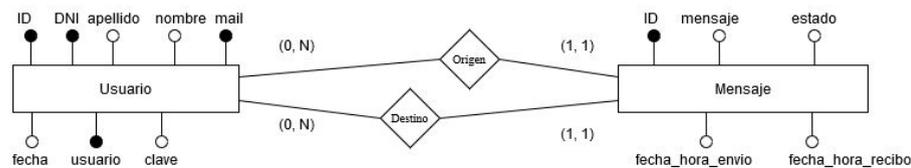
La ejecución de todas las pruebas se realizó utilizando la misma memoria y sistema operativo (4GB de RAM, Ubuntu versión 18.04). En MySQL se utilizó la versión 5.7, en MongoDB la versión 4.0, en Cassandra la versión 3.11, en Neo4j la versión 4.1 y en Redis la versión 4.0. Para cada caso de estudio el volumen de datos utilizado se generó de forma aleatoria.

En los 3 primeros casos de estudios no fue posible evaluar Redis debido a que el gran volumen de datos utilizado no se puede gestionar adecuadamente en memoria principal.

A continuación, se presentan los 4 casos de estudio. A modo de ejemplo y para no exceder la cantidad de páginas, sólo en el primero de ellos se presentan los esquemas físicos y el código de las consultas. No obstante, en todos los casos se expresa el DER y los resultados obtenidos.

### 3.1 Caso de estudio 1

Se presenta un esquema para la bandeja de entrada de un sistema de mensajería interna. Se han generado de forma aleatoria 30.000.000 de mensajes en total correspondientes a 6.000.000 de usuarios distintos. La Figura 1 presenta el DER del problema en cuestión.



**Figura 1.** Esquema Conceptual expresado mediante un DER.

La Figura 2 presenta el modelo relacional (MySQL) resultante de la Figura 1.

```

Usuario=(id, DNI, apellido, nombre, mail, fecha_alta, usuario, clave)
Mensaje=(id, mensaje, estado, fecha_hora_env, fecha_hora_rec,
usuario_orig(FK), usuario_dest(FK))

```

**Figura 2.** Esquema Relacional.

La Figura 3 muestra la estructura del documento para MongoDB.

```
Mensaje = { _id, mensaje, estado, fecha_hora_env, fecha_hora_rec,
  usuario_origen={'dni','nombre','apellido','mail','usuario'},
  usuario_destino={'dni','nombre','apellido','mail','usuario'}}
```

Figura 3. Documento para MongoDB.

En Cassandra, el esquema físico se expresa de acuerdo a las consultas a realizar. La Figura 4 plantea dicho esquema físico.

```
Mensajes_por_usuario =((usuario_orig, usuario_dest,
  fecha_hora_env)PK, mensaje, fecha_hora_rec, estado);
```

Figura 4. Esquema Físico de Cassandra.

La Figura 5, muestra una parte del almacenamiento de Grafos planteado para Neo4j. Los nodos de color verde representan a los usuarios, mientras que los grises representan mensajes.

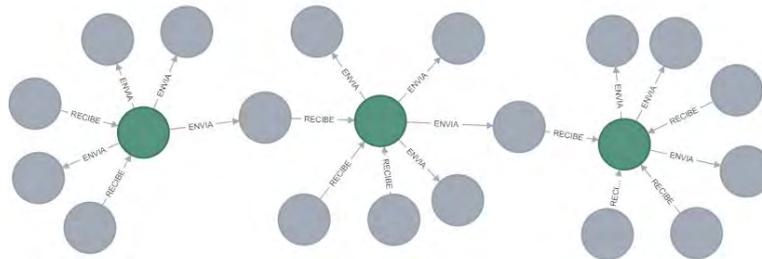


Figura 5. Parte del Grafo generado para el caso de estudio 1.

A continuación, se plantean 3 consultas para evaluar el rendimiento en este caso de estudio.

“C1: Mensajes entre dos usuarios específicos”

“C2: Buscar una cadena de caracteres en los mensajes de un usuario específico”

“C3: Mensajes recibidos y no leídos por un usuario específico”

Las Figuras 6, 7, 8 y 9 muestran la resolución de C1 para 4 motores de Base de Datos (MySQL, Cassandra, MongoDB y Neo4j).

```
SELECT mensaje, fecha_hora_env, estado
FROM mensaje m INNER JOIN usuario o ON (o.id=m.usuario_orig) INNER
JOIN usuario d ON (d.id=m.usuario_dest) WHERE (o.usuario =
'user256' AND d.usuario = 'user1') OR (o.usuario = 'user1' AND
d.usuario = 'user256') ORDER BY m.fecha_hora_env DESC
```

Figura 6. Implementación de C1 en MySQL.

```

SELECT mensaje, fecha_hora_env, estado
FROM mensajes_por_usuario
WHERE usuario_orig IN ('user256','user1') AND usuario_dest IN
('user1', 'user256') ORDER BY fecha_hora_env DESC

```

Figura 7. Implementación de C1 en Cassandra.

```

db.mensajes.find({$or:[ {$and:[{"usuario_origen.usuario": "user1"},
{"usuario_destino.usuario": "user259"}]},
{$and:[{"usuario_origen.usuario": "user259"},
{"usuario_destino.usuario": "user1"}]}
]),{mensaje:1, fecha_hora_env:1, estado:1}).sort({fecha_hora_env: -1})

```

Figura 8. Implementación de C1 en MongoDB.

```

MATCH (u1:Usuario)-[r1]-(m:Mensaje)-[r2]-(u2:Usuario)
WHERE ID(u1)=34000550 AND ID(u2)=34000805
RETURN u1.usuario, r1, m, r2, u2.usuario;

```

Figura 9. Implementación de C1 en Neo4j.

En la Tabla 1 se expresan los tiempos obtenidos en segundos de la ejecución de cada consulta.

Tabla 1. Resultados del caso de estudio 1 expresados en segundos.

Consultas	MySQL	MongoDB	Cassandra	Neo4j
C1	0.33	385	0.35	0.24
C2	259	178	9	0.15
C3	20	166	72	0.24

### 3.2 Caso de estudio 2

Este caso de estudio representa una central de monitoreo que registra y procesa eventos procedentes de distintas fuentes. Se han generado aleatoriamente 40.000.000 de eventos en total. La Figura 10 presenta el DER del problema descrito.

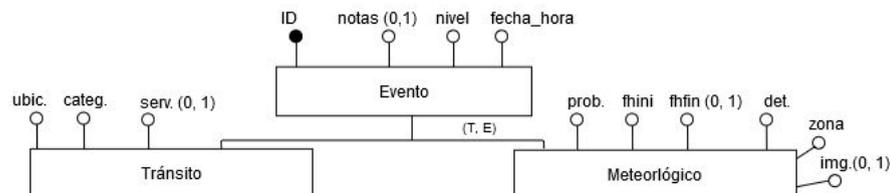


Figura 10. Esquema Conceptual expresado mediante un DER.

A continuación, se plantean 3 consultas para evaluar el rendimiento en este caso de estudio.

“C1: Eventos meteorológicos que superen una determinada probabilidad de ocurrencia”

“C2: Eventos de tránsito con una determinada categoría en un período de tiempo”

“C3: Cantidad de eventos de un determinado nivel de prioridad (Alto, Medio o Bajo)”

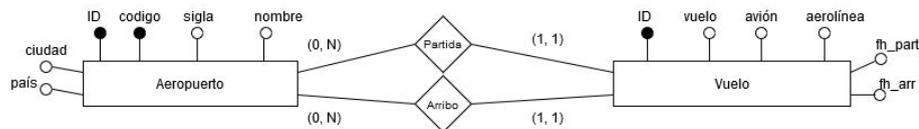
La Tabla 2 expresa los tiempos obtenidos en segundos de la ejecución de cada consulta.

**Tabla 2.** Resultados de este caso de estudio expresado en segundos

Consultas	MySQL	MongoDB	Cassandra	Neo4j
C1	200	119	123	360
C2	233	117	120	134
C3	81	106	200	288

### 3.3 Caso de estudio 3

En este caso, se plantea un esquema de vuelos pertenecientes a distintas aerolíneas. Se han generado aleatoriamente la información de 40.000.000 de vuelos entre 10.000 aeropuertos distintos. La Figura 11 presenta el DER del problema en cuestión.



**Figura 11.** Esquema Conceptual expresado mediante un DER.

A continuación, se plantean 2 consultas para evaluar el rendimiento en este caso de estudio.

“C1: Vuelos directos entre un origen y un destino determinado”

“C2: Vuelos que posean una escala entre un origen y un destino determinado”

La Tabla 3 expresa los tiempos obtenidos en segundos de la ejecución de cada consulta.

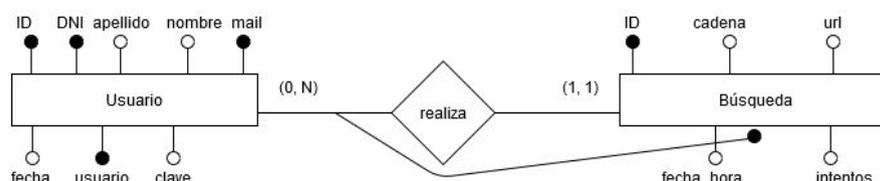
**Tabla 3.** Resultados del caso de estudio 3 expresados en segundos

Consultas	MySQL	MongoDB	Cassandra	Neo4j
C1	18	166.	1	0,1
C2	3720	ver sección 4	1.5	0.4

### 3.4 Caso de estudio 4

El cuarto y último caso de estudio presenta un esquema elegido para evaluar específicamente Redis en comparación con MySQL. Se representa una caché de

búsquedas recientes. Redis, si bien brinda la posibilidad de persistencia, se destaca por disponer sus estructuras de datos en memoria principal. Para este caso de estudio se han generado 100.000 usuarios y 10.000.000 búsquedas en total. La Figura 12 presenta el DER del problema en cuestión. La entidad “Búsqueda” posee un identificador mixto, el cual se compone del identificador del usuario más la fecha y hora de la búsqueda.



**Figura 12.** Esquema Conceptual expresado mediante un DER.

A continuación, se plantean 2 consultas para evaluar el rendimiento en este caso de estudio

“C1: Las 5 búsquedas más recientes para un usuario determinado”

“C2: Las 5 búsquedas más frecuentes para un usuario determinado.”

La Tabla 4 expresa los resultados obtenidos en microsegundos.

**Tabla 4.** Resultados del caso de estudio 4 expresado en microsegundos.

Consultas	MySQL	Redis
C1	580000	0,48
C2	550000	0,53

## 4 Análisis de resultados

### 4.1 Análisis de caso de estudio 1

En este primer caso de estudio, Neo4j obtiene el mejor rendimiento. No obstante, debido a que la cantidad de nodos “mensajes” tiene crecimiento exponencial, el rendimiento puede degradarse.

Para la consulta 1, Cassandra y MySQL obtienen tiempos próximos a Neo4j. Esto se debe a la característica de los campos utilizados en el filtro de la consulta.

MongoDB obtiene tiempos superiores a los otros 3 motores de Bases de Datos. Las consultas que involucran campos que pertenecen a documentos embebidos pueden afectar su performance.

Para la consulta 2, Cassandra obtiene mejor tiempo que MySQL y MongoDB pero necesita de la implementación de un índice para realizar búsquedas relativas. MongoDB obtiene mejor rendimiento que MySQL; las búsquedas relativas y el uso de JOINS afectan el rendimiento en MySQL.

Para la consulta 3, MySQL obtiene mejor rendimiento que MongoDB y Cassandra. Realizar consultas que involucren columnas que no son parte de la Primary Key o Clustering Key en Cassandra, afecta su rendimiento.

#### **4.2 Análisis de caso de estudio 2**

En este caso de estudio, ninguno de los 4 motores de Bases de Datos obtiene el mejor rendimiento para todas las consultas planteadas.

En las consultas 1 y 2, MongoDB obtiene el mejor rendimiento. En estas consultas, el esquema planteado para MongoDB no utiliza documentos embebidos y no requiere almacenar campos opcionales, evitando comparaciones nulas. En MySQL las operaciones de JOINS afectan su performance. En Cassandra, las consultas planteadas involucran campos que no están presentes en la Primary Key o Clustering Key, lo que afecta su tiempo de respuesta. En Neo4j el esquema del caso de estudio no es apropiado para el Almacenamiento de Grafos, dado que no existen relaciones entre los nodos del grafo.

En la consulta 3, MySQL optimiza la función de agregación COUNT, mejorando así, su tiempo de respuesta en comparación a los otros 3 motores de Bases de Datos.

#### **4.3 Análisis de caso de estudio 3**

En este caso de estudio Neo4j obtiene el mejor rendimiento para las consultas planteadas. Esto es debido a que dichas consultas son resueltas eficientemente en el Almacenamiento de Grafos.

Cassandra, obtiene mejores tiempos que MySQL y MongoDB, esto se debe a que el esquema físico en Cassandra se plantea en términos de consultas. En MySQL la baja de performance entre la consulta 1 y la consulta 2 se debe a los JOINS necesarios para resolver cada consulta.

#### **4.4 Análisis de caso de estudio 4**

En este caso de estudio se plantean dos consultas para Redis y MySQL. Redis obtiene el mejor rendimiento. Esto se debe a que Redis define íntegramente sus estructuras en memoria RAM donde opera los datos, mientras que MySQL por su parte distribuye su procesamiento entre memoria RAM y memoria secundaria.

## **5 Conclusiones**

Este trabajo se concentra en un estudio comparativo entre 5 motores de Bases de Datos (4 NoSQL y 1 Relacional) en 4 casos de estudios diferentes. En primer lugar, se

realizó una breve descripción de las diferentes formas de almacenamiento de los motores de Bases de Datos NoSQL, junto a sus propiedades específicas. Los 4 casos de estudios planteados presentan diferentes esquemas de información y para cada uno de ellos se definió un conjunto de consultas específico, resultando 35 consultas en total. Para el primero de ellos se presenta el desarrollo en forma completa, es decir, el esquema DER, el esquema físico y el código de cada una de las consultas. En los 3 casos restantes se presentó el DER.

Los resultados obtenidos en cada uno de los 4 casos de estudio se han analizado con el criterio de justificar porqué en cada caso un motor se comporta mejor que otro.

Para finalizar, se puede concluir que si se dispone un gran volumen de información no hay un motor de Bases de Datos que obtenga el mejor rendimiento para satisfacer todas las consultas planteadas.

En la medida que aumente el volumen de datos, es necesario evaluar qué alternativa de almacenamiento, es decir, que tipo de motor de Bases de Datos, es conveniente utilizar con el objetivo de brindar mejores prestaciones.

Ante la variedad de resultados obtenidos, y la falta de prevalencia de un motor de Bases de Datos con el mejor rendimiento, debe realizarse escalamiento horizontal (aumentar nodos) y plantear una nueva evaluación en ese nuevo contexto.

## 6 Trabajo Futuro

Como trabajo futuro, se prevé incorporar nuevas pruebas y reforzar las existentes con alternativas de motores de Bases de Datos Relacionales y No Relacionales. Además, se buscará explorar la capacidad de escalamiento horizontal de las Bases de Datos NoSQL y de las Bases de Datos Relacionales. Finalmente, se prevé experimentar con Bases de Datos NewSQL y Bases de Datos de Series Temporales.

## 7 Bibliografía

1. 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, Fernandez 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>.
2. Aspectos de la Ingeniería de Software, Bases de Datos Relacionales y Bases de Datos No Relacionales para el desarrollo de Sistemas de Software en Escenarios Híbridos. XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, Universidad Nacional de la Patagonia Austral, El Calafate, Santa Cruz, Argentina).
3. Aspectos de ingeniería de software y bases de datos para el desarrollo de sistemas de software en escenarios híbridos. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan, San Juan, Argentina). ISBN 978-987-3619-27-4. <http://sedici.unlp.edu.ar/handle/10915/77088>.

4. Ajit Singh, Sultan Ahmad. Data Modeling with NoSQL Database. ISBN 978-1072978374 (2019).
5. Jitender Kumar, Varsha Garg. Security analysis of unstructured data in NOSQL MongoDB database. International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN) 2017. <https://ieeexplore.ieee.org/document/8284495>.
6. Benymol Jose, Sajimon Abraham. Exploring the merits of nosql: A study based on MongoDB. International Conference on Networks & Advances in Computational Technologies (NetACT) 2017. <https://ieeexplore.ieee.org/document/8076778>.
7. Replicación de bases de datos NoSQL en dispositivos móviles. <http://sedici.unlp.edu.ar/handle/10915/48085>.
8. Roshni Bajpayee, Sonali Priya Sinha y Vinod Kumar. Big Data: A Brief investigation on NoSQL Databases. International Journal of Innovations & Advancement in Computer Science (IJIACS). ISSN 2347. Volume 4, Issue 1. January 2015.
9. Deepak Vohra. NoSQL Web Development with Apache Cassandra. Cengage Learning, 2015. ISBN 9781305576773.
10. 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>
11. 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>
12. Carlo Batini, Stefano Ceri, Shamkant B. Navathe. Diseño Conceptual de Bases de Datos, un enfoque de entidades-interrelaciones. Addison-Wesley / Díaz de Santos. ISBN 0-201-60120-6 (1994).
13. Flexibilidad en bases de datos NoSQL sobre ambientes web mining. <http://sedici.unlp.edu.ar/handle/10915/59099>
14. Ranking de bases de datos según su popularidad. <https://db-engines.com/en/ranking>. Accedido en julio 2020.
15. MySQL. <https://www.mysql.com/>. Accedido en julio 2020.
16. Clientes de MySQL. <https://www.mysql.com/customers/>. Accedido en julio 2020.
17. MongoDB. <https://www.mongodb.com/>. Accedido en julio 2020.
18. Empresas que utilizan MongoDB. <https://www.mongodb.com/who-uses-mongodb> Accedido en julio 2020.
19. Apache Cassandra. <http://cassandra.apache.org/>. Accedido en agosto 2020.
20. Redis. <https://redis.io/>. Accedido en agosto 2020.
21. Amazon DynamoDB. <https://aws.amazon.com/es/dynamodb/>. Accedido en julio 2020.
22. Apache CouchDB. <http://couchdb.apache.org/>. Accedido en julio 2020.
23. Apache HBase. <http://hbase.apache.org/>. Accedido en julio 2020.
24. Neo4j. <https://neo4j.com/>. Accedido en julio 2020.

# Smart Grids Challenge: A competitive variant for Single Objective Numerical Optimization.

Fabricio Loor<sup>1</sup>, M. Guillermo Leguizamón<sup>1</sup>, Efrén Mezura-Montes<sup>2</sup>

<sup>1</sup> Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)  
Universidad Nacional de San Luis, San Luis D 5700 BPB, Argentina

`faloor,legui@unsl.edu.ar`

<sup>2</sup> Artificial Intelligence Research Center, Universidad Veracruzana, Xalapa, Veracruz,  
91000, Mexico.

`emezura@uv.mx`

**Abstract.** In this work, we present a new algorithm (AJSO) for high-dimensional single objective problems. It is well known that finding high quality solutions is still a challenge for complex problems like those found in the literature as well as in real world concerning Smart Grids scenarios. Our proposal AJSO is an improvement on a state-of-the-art differential Evolution (DE) based algorithm known as SHADE. More specifically, AJSO implements two novel mutation strategies and also incorporates a mechanism for maintaining and taking good solutions from a special archive when a particular condition during the exploration process is detected. To compare the performance of AJSO, the benchmark given in the WCCI/GECCO 2020 [4] is used. This challenge consisted of optimization problems represented in two testbeds of Smart Grids problems. In this paper we adopted the guidelines given in the WCCI/GECCO 2020 competition. Experimental results show that AJSO outperforms SHADE in the two studied testbeds.

**Keywords:** Optimization, Smart Grids, Metaheuristics, Differential Evolution

## 1 Introduction

Optimization problems require minimizing or maximizing a measure according to a given configuration of values. If the possible permutations or combinations in those configurations are really vast, thinking of exhaustive methods or brute force becomes an impractical alternative. Despite this, stochastic algorithms such as metaheuristics have been widely used to solve these situations. In particular, Differential Evolution (DE) [6] is a stochastic algorithm that has proven to be the basis for many improvements and advances, since it is simple and easy to program. For its operation, DE needs to be adjusted by two parameters: a mutation factor ( $F$ ) and a probability of crossover ( $CR$ ). Currently, DE-based algorithms are widely competitive with other efficient and state-of-the-art optimization algorithms [11], [13], [12].

In order to understand the work in this area, we will briefly describe some representative modifications and improvements made throughout the history of DE:

- 2005, FADE [7], Fuzzy Adaptive Differential Evolution. Fuzzy logic controllers are applied for parameters  $F$  and  $CR$ .
- 2005, SaDE [8], Self adaptive differential Evolution. Parameters  $F$  and  $CR$  are updated in each generation by normal distributions.
- 2006, jDE [2]. The values of  $F$  and  $CR$  are given by uniform distributions.
- 2009, JADE [3], introduces a new mutation strategy “*current-to-pbest/1*” by adding an archive of less efficient solutions.
- 2013, SHADE, [4], which is based on JADE and proposes a memory for the satisfactory values of  $F$  and  $CR$ . This algorithm was placed in third place in the CEC 2013 competition.
- 2014, L-SHADE [5], uses SHADE and introduces a population decrease that behaves linearly. It was the winner in the CEC 2014 competition.
- 2016, iL-SHADE [9] is an improvement over L-SHADE. This algorithm introduces changes in initialization and memory update. It also features a new mechanism to update the  $F$  and  $CR$  parameters, varying between the current and maximum generation. It was ranked fourth in the CEC 2016 competition.
- 2017, jSO [10], uses the “*current-to-pbest-w/1*” strategy and introduces modifications according to the search period on the mutation factor. It earned second place in the CEC 2017 restricted limits competition.
- 2019, jDE100 [1] is a modification over jDE that adds two populations, one old and one new. It also adds random restarts when search deadlocks are detected. It was the winner of the CEC 2019 100 Digits competition.

In this work, we use one of these improvements and propose changes to solve a high-dimensional problem applied to intelligent energy distribution.

The rest of the paper is organized as follows. In Section II, we present AJSO by describing its main features and details. In Section III, we describe the experimental design and the framework where the algorithms have been tested. Then, in Section IV, the achieved results are shown and analyzed. Finally, in Section V, the conclusions reached and some future works research lines are summarized.

## 2 Our approach: AJSO (Advanced JSO)

In this section, a new approach (AJSO) for solving single-objective real parameter optimization is presented and its pseudo-code is shown in Algorithm 1. AJSO is an enhanced version of SHADE [4] and jSO [10], two well-known algorithms based on DE. SHADE preserves two population structures, the first one keeps the current population and the other one (an archive) contains the replaced individuals throughout the search process. It also incorporates a recently successful historical parameter memory to guide the generation of parameter control values, for both crossover and mutation operators.

In our proposal, three structures are used to maintain different type of solutions. The first one is the main population  $P_g$ , the second one maintains an

*Archive* of useful recent generated solutions, and the last one preserves the history of best individuals found at each generation ( $A_g$ ). The Population and Archive structures have a fixed  $N$  size, while the third one grows dynamically at each generation. The process to update the control parameters is different than in SHADE and will be described later.

Before entering the main loop, the population is initialized, and the archive equals the population. In this part of the code, the memory is given random values that follow a normal distribution. These values in turn will be the mean for future calculations of the respective values of parameters  $CR$  and  $F$ . After initialization, the algorithm enters the evolutionary stage. The variable  $sF$  is described in [4] This stage continues until the stop criterion is reached, which in our case is given by a maximum number of evaluations completed. In the body of the main loop, a population is evaluated with two different mutation strategies, to subsequently carry out eight evaluations with the best strategy.

---

**Algorithm 1: AJSO**


---

```

1 //Initialization phase
2  $g = 1, N_g = N_{init}$ ;
3 Initialize population  $P_g = (x_{1,g}, \dots, x_{N,g})$  randomly;
4 Set all values in  $M_{CR}, M_F$  following a normal distribution with
    $\mu_{CR} = I_{CR}$  and  $\sigma_{CR} = 0.05$  ;  $\mu_F = I_F$  and  $\sigma_F = 0.1$ ;
5  $Archive = P_g$ ;
6 // Main loop
7 while The termination criterion is not met do
8   Sort and select the best B individuals;
9   Obtain a vector of values with  $CR$  and  $sF$ ;
10  From  $sF$  get  $jF$ ;
11  Generate two new populations with two different strategies;
12  Control limits and evaluate;
13  Select the best strategy S;
14  // Exploitation phase
15  for  $i = 1$  to 8 do
16    Sort and select the best B individuals;
17    Obtain a vector of values with  $CR$  and  $sF$ ;
18    From  $sF$  get  $jF$ ;
19    Generate a new population with the best strategy from 13;
20    Control limits and evaluate;
21    Update Archive;
22    Update Memory Parameters;
23    if  $nsp = 0$  and  $B_g = BSF$  and  $nfes > I_{evalmax} * 0.4$  then
24      | add in the Archive one of the best global past ( $A_g$ );
25    end
26  end
27 end

```

---

4 Loor, Leguizamón, and Mezura-Montes

## 2.1 Mutation and crossing process

The SHADE algorithm [4] builds a mutant vector  $\vec{v}_i$ , according to the strategy “*current-to-pbest/1*” as shown in the next equation:

$$\vec{v}_{i,g+1} = \vec{v}_{i,g} + sF * (\vec{v}_{pbest} - \vec{v}_{i,g} + \vec{v}_{r1} - \vec{v}_{r2}),$$

where individual  $v_{pbest}$  is randomly selected from the top  $N * p$  ( $p \in [0, 1]$ ) members in generation  $g$ . The indices  $i$ ,  $r1$  and  $r2$  are randomly selected from  $[1, N]$  such that they differ from each other.

AJSO (our algorithm) incorporates two ways to mutate solutions by following two different strategies. In order to this, element of the population or elements of the archive are part of the strategies according to the need of the moment. The first strategy uses elements of the population to generate individual  $i$  of generation  $g + 1$ :

$$\vec{v}_{i,g+1} = \begin{cases} \vec{v}_{pr} + sF * (\vec{v}_{r1} - \vec{v}_{i,g}) + jF * (\vec{v}_{r1} - \vec{v}_{r2}), & \text{if } rand(0, 1) < C\vec{R}_i. \\ \vec{v}_{i,g}, & \text{otherwise,} \end{cases}$$

where  $sF$  is a scalar obtained from applying a normal distribution on a random value in memory.  $jF$  is a proportion less than  $sF$  and,  $v_{pr}$  is a random vector among the best  $B$  elements of the previous generation. The second strategy is stated as follows:

$$\vec{v}_{i,g+1} = \begin{cases} \vec{v}_{pr} + sF * (\vec{v}_{r1} - \vec{v}_{i,g}) + jF * (\vec{v}_{r3} - \vec{v}_{r4}), & \text{if } rand(0, 1) < CR_i. \\ v_{i,g}, & \text{otherwise.} \end{cases}$$

Here, vectors  $\vec{v}_{r3}$ ,  $\vec{v}_{r4}$  are element randomly selected from *Archive* and different from each other.

## 2.2 Selecting the best strategy

After checking that the modified solutions are within the allowed limits of the decision variables, the algorithm proceeds to evaluate these solutions. When the evaluation concludes, a selection process is carried out to obtain the best strategy between the two used in the previous step. The selection criterion is that which produced a new population with more improvements with respect to the old one.

## 2.3 Updating the historical memory

In this section we describe the update of the values for the control memories  $M_F$  and  $M_{CR}$ . The elements of historical memory, similarly used in SHADE, are updated in the exploitation stage (lines 14 to 26 of Algorithm 1). For this purpose, the evaluation of each population also counts the numbers of individuals that improve compared with the respective old values. That value is called  $nsp$  for the Number of Successful Parameters, its complement respect to the total number of new individuals generated is called Number of Fail Parameters ( $nfp$ ).

If more than a half of the individuals in a generation improved the previous population ( $nsp > nfp$ ), the following formula is applied to update  $m_i \in Memory$  which are used as explained before for future values of control parameter  $F$ :

$$m_i = \begin{cases} 1.5 * avg(goodF) - 0.5 * m_i & \text{within limits } (0,1]. \\ m_i, & \text{out of limits } (0,1] \end{cases}$$

where  $avg$  calculate the average of a list of numbers and  $goodF$  is a list of  $sF$  values which have improved the individual in the population.

If the number of individuals in a generation that improved does not exceed the half of the population ( $nsp \leq nfp$ ) the following formula is used to update Memory:

$$m_i = \begin{cases} m_i + 0.2 * (0.5 * avg(goodF + BadF) - 0.8 * m_i) & \text{within limits } (0,1]. \\ m_i, & \text{out of limits } (0,1], \end{cases}$$

where  $BadF$  is a listing of  $sF$  values that could not be improved in the population.

The procedure for updating memory elements that store  $CR$  values ( $M_{CR}$ ) is analogous to that described in this subsection.

### 3 The problem and results

#### 3.1 Problem definition: Optimization applied to energy distribution

The distribution of energy has many stages from its generation to its distribution for consumption. In smart grids, the distribution of energy is optimized for both consumers and suppliers, so it is highly desirable to find a good balance between demand and supply. This balance can be modeled as an objective function where the input variables to the function have certain restrictions. Due to space issues in this link<sup>3</sup> the reader may find a more detailed description of the objective functions.

#### 3.2 The benchmark

We have evaluated the performance of the AJSO with the IEEE WCCI / GECCO 2020 benchmark “*CEC-C4 Evolutionary Computation in the Energy Domain: Smart Grid Applications*”. This benchmark has two testbeds. The first one aims at optimizing energy resources management for day-to-day use in smart grids under uncertain environments. 500 scenarios with a high degree of uncertainty are used. The second testbed is concerned with a two-tier optimization of end-user bidding strategies in local energy markets (LM). These two levels represent a

<sup>3</sup> [http://www.gecad.isep.ipp.pt/ERM-competitions/wp-content/uploads/2019/12/WCCI2020\\_Guidelines.pdf](http://www.gecad.isep.ipp.pt/ERM-competitions/wp-content/uploads/2019/12/WCCI2020_Guidelines.pdf)

6 Loor, Leguizamón, and Mezura-Montes

complex problem where competitive agents at the one level try to maximize their profits by modifying a price, on the other hand, an agent tries to minimize costs. The competition proposes 50,000 as a limit for the total number of evaluations for each testbed.

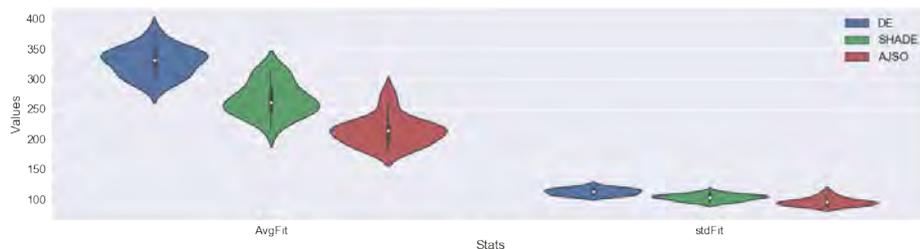
Table 1 and Table 2 contain statistical values and positions for the classification following the criteria of WCCI 2020 for three algorithms studied, DE (a basic version), SHADE, and AJSO. The value that summarizes the respective performances is represented by *RankingIndex*. The calculation of this value can be found in the Guideline of the WCCI Competition.

Alg	RankingIndex	PAvgFit	PstdFit	PminFit	PmaxFit	PvarFit
DE	<b>442.64</b>	329.09	23.38	286.53	376.89	546.70
SHADE	<b>370.81</b>	267.03	27.63	217.93	316.31	763.87
AJSO	<b>311.11</b>	214.68	24.07	182.50	277.26	579.45

**Table 1.** Results obtained by DE, SHADE and AJSO in the *TestBed 1*

Alg	RankingIndex	AvgFit	StdFit	VarFit	minFit	maxFit	AvgProfit
DE	<b>3.03</b>	3.03	0.07	0.0059	2.86	3.20	-4.98
SHADE	<b>2.49</b>	2.49	0.09	0.0084	2.35	2.64	-4.04
AJSO	<b>2.28</b>	2.28	0.04	0.0020	2.20	2.35	-3.60

**Table 2.** Results obtained by DE, SHADE and AJSO in the *TestBed 2*



**Fig. 1.** Comparison between 20 runs of testbed 1 for the three algorithms

Figure 1 and 2 show violin plots to compare the performances of the three algorithms taking into account 20 runs, showing respectively *AvgFit* and *stdFit* in the testbed 1 and *Fit* and *Profit* in the testbed 2. The figure shows the efficiency of AJSO having a better evaluation and converging much faster than DE and SHADE for the WCCI proficiency test suite. Based on this figure, we

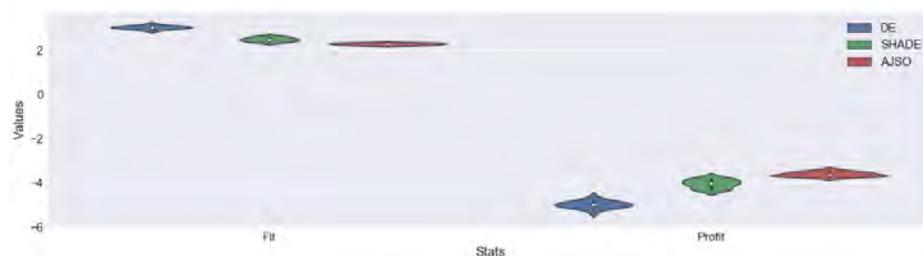


Fig. 2. Comparison between 20 runs of testbed 2 for the three algorithms

observe that AJSO is quicker in finding better values, avoiding the waste of evaluations of the cost function. Therefore, this implies that the exploration phase correctly discards various sectors of search spaces.

## 4 Conclusions and Future Work

In this article we have proposed AJSO, a new algorithm based on SHADE to solve a Smart Grid application presented at WCCI 2020. AJSO saves the best of each generation ( $A_g$ ) and makes more efficient use of the Archive to better explore the search space. Both improvements have been introduced as part of the mutation operator and in the updating of the historical memories of the controls parameters. If stagnation is detected, AJSO inserts good solutions in the Archive from  $A_g$  structure. In turn, the combination of two mutation strategies allows the algorithm to vary between two abilities, one concerning the amplitude and the other one the depth.

The incorporation of all these characteristics allowed our proposal to obtain the best scores compared to DE and SHADE. In a future work, a deeper analysis can be performed by varying Population and Archive sizes. Currently, AJSO has very few parameters to calibrate, but it may be possible to incorporate values that modify the updating of the memories.

## References

1. Brest, J., Maučec, M. S., and Bošković, B. (2019, June). The 100-Digit Challenge: Algorithm jDE100. In 2019 IEEE Congress on Evolutionary Computation (CEC) (pp. 19-26). IEEE.
2. Brest J., Greiner S., Bošković B., Mernik M., Žumer V. (2006). Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems. IEEE Transactions on evolutionary computation, Vol. 10, No. 6.
3. Zhang, J., and Sanderson, A. C. (2009). JADE: adaptive differential evolution with optional external archive. Evolutionary Computation, IEEE Transactions on, 13(5), (pp. 945-958).

8 Loor, Leguizamón, and Mezura-Montes

4. R. Tanabe and A. Fukunaga, Success-History Based Parameter Adaptation for Differential Evolution, in IEEE CEC, 2013, pp. 71–78.
5. R. Tanabe and A. S. Fukunaga, "Improving the search performance of SHADE using linear population size reduction," 2014 IEEE Congress on Evolutionary Computation (CEC), Beijing, 2014, pp. 1658-1665, doi: 10.1109/CEC.2014.6900380.
6. Storn, R., and Price, K. (1995). Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces (Vol. 3). Berkeley: ICSI.
7. J. Liu and J. Lampinen, A fuzzy adaptive differential evolution algorithm, *Soft Computing—A Fusion of Foundations, Methodologies and Applications*, vol. 9, no. 6, pp. 448–462, 2005 [Online]. Available: <http://springerlink.metapress.com/index/10.1007/s00500-004-0363-x>
8. Brest J., Greiner S., Bošković B., Mernik M. and Žumer V., Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems. *IEEE Transactions on evolutionary computation*, Vol. 10, no. 6, 2006.
9. J. Brest, M. Sepesy Maučec and B. Bošković, iL-SHADE: Improved L-SHADE algorithm for single objective real-parameter optimization. *IEEE CEC*, 2016.
10. J. Brest, M. Sepesy Maučec and B. Bošković, Single objective real-parameter optimization: Algorithm jSO. *IEEE CEC*, 2017.
11. R. Mallipeddi, P. N. Suganthan, Q. K. Pan, and M. F. Tasgetiren, *Differential evolution algorithm with ensemble of parameters and mutation strategies*, *Appl. Soft Comput.*, vol. 11, no. 2, pp. 1679–1696, 2011.
12. J. Brest and M. S. Mau Cec, *Population size reduction for the differential evolution algorithm*, *Appl. Intell.*, Vol: 29, no: 3, pp. 228–247, 2008.
13. G. Wu, X. Shen, H. Li, H. Chen, A. Lin and P. N. Suganthan, *Ensemble of differential evolution variants*, *Information Sciences* Vol: 423 pp: 172-186, 2018.

# Representación del Trie de Sufijo: Una Evaluación Empírica

Darío Ruano, Jesica Cornejo, Paola Azar, Norma Herrera

Departamento de Informática, Universidad Nacional de San Luis, Argentina,  
dmruano@unsl.edu.ar, jncornej@gmail.com, pazar18@gmail.com,  
nherrera@unsl.edu.ar

**Abstract.** Un *trie de sufijos* es un índice para bases de datos de texto que permite resolver eficientemente las operaciones de búsqueda pero que necesita en espacio 10 veces el tamaño del texto indexado. En [12] se propone una nueva representación compacta del trie de sufijos que resulta eficiente en espacio y que permite un posterior paginado del índice. Posteriormente, en [13, 3] se presentan mejoras en espacio y en tiempo respectivamente a esta representación secuencial del trie de sufijos. En este artículo presentamos una evaluación experimental global de estas tres versiones del trie de sufijos a fin de analizar la eficiencia de cada una de ellas tanto en tiempo como en espacio.

**Palabras claves:** Bases de Datos de Texto, Índices, Trie de sufijos.

## 1 Introducción

Gran parte de esta información disponible en formato digital se representa en forma de texto, es decir, secuencias de símbolos que pueden representar no sólo lenguaje natural, sino también música, secuencias de ADN, secuencias de proteínas, etc. Las tecnologías tradicionales de bases de datos (modelo relacional) para almacenamiento y búsqueda de información no son adecuadas en este ámbito, dado que no es posible organizar una base de datos de texto en registros y campos.

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 este tipo de búsqueda podemos trabajar directamente sobre el texto sin preprocesarlo [1] o podemos preprocesar el texto para construir un índice. Construir un índice tiene sentido cuando el texto es grande, cuando las búsquedas son más frecuentes que las modificaciones (de manera tal que los costos de construcción se vean amortizados) y cuando hay suficiente espacio como para contener el índice. Un índice debe dar soporte a dos operaciones básicas: *count*, que consiste en contar el número de ocurrencias de  $P$  en  $T$ , y *locate*, que consiste en ubicar todas las posiciones de  $T$  donde  $P$  ocurre.

En bases de datos de texto el índice generalmente ocupa más espacio que el texto pudiendo necesitar de 4 a 20 veces el tamaño del mismo [5, 8]. Algunos índices reducen

el espacio ocupado restringiendo el tipo de búsqueda que se pueden resolver. Así, un **índice orientado a palabra**, permite solamente buscar palabras completas en el texto. Algunos índices orientados a palabras permiten también buscar patrones que sean inicios de palabras. Los índices que son capaces de encontrar todas las ocurrencias de un patrón dentro del texto se denominan índices orientados a caracteres o **full text indexes**.

Otra alternativa para reducir el espacio ocupado por el índice es buscar una representación compacta del mismo, manteniendo las facilidades de navegación sobre la estructura. En los últimos años se han realizado grandes avances en este ámbito. Las nuevas técnicas de compresión de índices no sólo permiten reducir su tamaño, sino que procesan búsquedas más rápido que la versión sin comprimir [6, 9, 11].

Un **trie de sufijos** es un índice que permite resolver eficientemente las operaciones *count* y *locate* pero que necesita en espacio 10 veces el tamaño del texto indexado. En [12] se propone una representación secuencial del trie de sufijos que es adecuada para un posterior proceso de paginación. En [13] se propone una mejora en espacio del trie de sufijos representado de manera secuencial y en [3] se propone una mejora en tiempo del mismo índice.

En este trabajo presentamos una evaluación experimental de estas tres versiones del trie de sufijos que nos permitirá concluir cuál de todas es más adecuada en tiempo y en espacio.

Lo que resta del artículo está organizado de la siguiente manera. En la sección 2 presentamos el marco teórico, en la sección 3 damos el trabajo relacionado explicando las tres representaciones que hemos considerado. La evaluación experimental de las mismas se desarrolla en la sección 4, para finalizar en la sección 5 dando las conclusiones y el trabajo futuro.

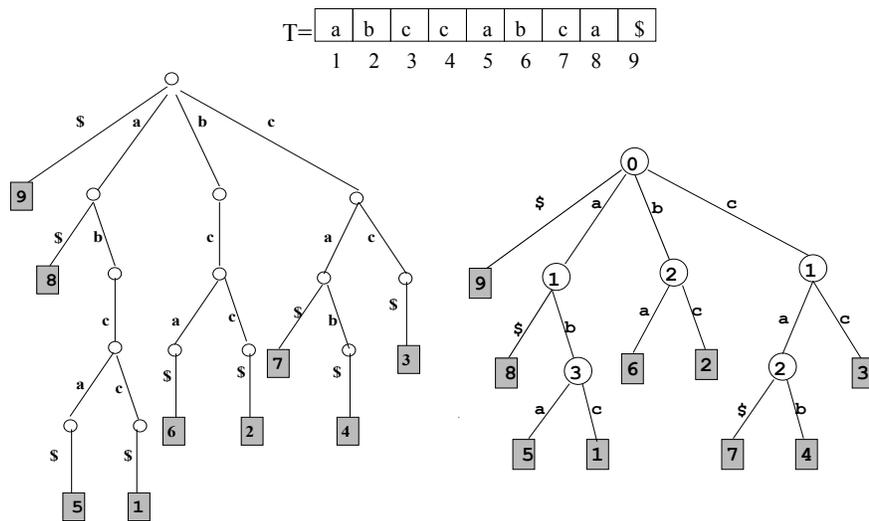
## 2 Marco Teórico

Dado un texto  $T = t_1, \dots, t_n$  sobre un alfabeto  $\Sigma$  de tamaño  $\sigma$ , donde  $t_n = \$ \notin \Sigma$  es un símbolo menor en orden lexicográfico que cualquier otro símbolo de  $\Sigma$ , un sufijo de  $T$  es cualquier string de la forma  $T_{i,n} = t_i, \dots, t_n$  y un prefijo de  $T$  es cualquier string de la forma  $T_{1,i} = t_1, \dots, t_i$  con  $i = 1..n$ . Cada sufijo  $T_{i,n}$  se identifica unívocamente por  $i$ ; llamaremos al valor  $i$  **índice del sufijo**  $T_{i,n}$ . Un patrón de búsqueda  $P = p_1 \dots p_m$  es cualquier string sobre el alfabeto  $\Sigma$ .

Entre los índices más populares para búsqueda de patrones encontramos el arreglo de sufijos [8], el árbol de sufijos [15] y el trie de sufijos [5]. Estos índices se construyen basándose en la observación de que un *patrón*  $P$  ocurre en el texto si es prefijo de algún sufijo del texto.

Un **trie de sufijos** [5] es un trie[4] construido sobre el conjunto de todos los sufijos de  $T$ . Cada rama está rotulada por un símbolo de  $\Sigma$  y cada hoja representa un sufijo de  $T$ . El conjunto de sufijos se obtiene recorriendo todos los caminos posibles desde la raíz hasta una hoja y concatenando los rótulos de las ramas que forman cada uno de esos caminos. En cada nodo hoja se mantiene el índice del sufijo que esa hoja representa.

Una forma de reducir el espacio utilizado es reemplazar las ramas del árbol que han degenerado en una lista, por una única rama y agregar a cada nodo interno un valor que depende de la longitud de la rama que se ha eliminado. Esta longitud se conoce con el



**Fig. 1.** Un texto, su correspondiente trie de sufijos (izquierda) y el trie de sufijos con valores de salto (derecha).

nombre de *valor de salto*. Si la rama eliminada es una rama terminal, no es necesario el valor de salto, solo se elimina y se deja el índice del sufijo. La Figura 1 muestra un ejemplo de un texto, su correspondiente trie de sufijos y la versión del trie que utiliza valores de salto.

Para encontrar todas las ocurrencias de  $P$  en  $T$ , se busca en el trie utilizando los caracteres de  $P$  para direccionar la búsqueda. La búsqueda comienza por la raíz y en cada paso, estando en un nodo  $x$  con valor de salto  $j$ , avanzamos siguiendo la rama rotulada con el  $j$ -ésimo carácter de  $P$  [12].

### 3 Representaciones de un Trie de Sufijos

La representación habitual de un trie consiste en mantener en cada nodo los punteros a sus hijos, junto con el rótulo correspondiente a cada uno de ellos. Existen distintas variantes de representación que consisten en organizar estos punteros a los hijos sobre una lista secuencial, sobre una lista vinculada o sobre una tabla de hashing [7]. Una de las propuestas de representación que mejor desempeño tiene en memoria principal es la de Kurtz, quien propuso que cada nodo mantenga un apuntador al primer hijo y almacenar los nodos hermanos en posiciones consecutivas de memoria. Esto permite durante una búsqueda, realizar una búsqueda binaria sobre los rótulos para decidir por cual hijo seguir.

La mayoría de las propuestas existentes mantienen explícitamente la forma del árbol con punteros, los que pueden ser punteros físicos (direcciones de memoria principal) o punteros lógicos (posiciones de un arreglo). Las representaciones que describimos a continuación tienen la ventaja de evitar mantener explícitamente los punteros físicos,

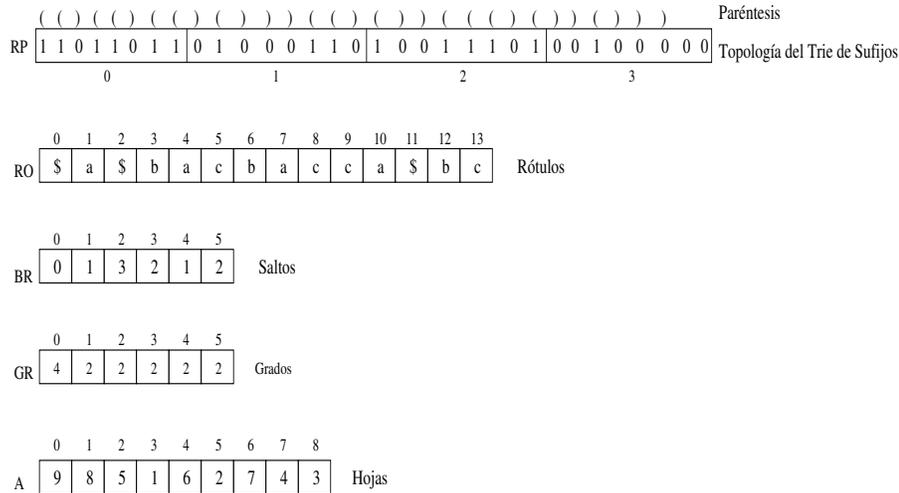


Fig. 2. Representación secuencial del trie de sufijos de la Figura 1

lo que permite ahorrar espacio y permitir un posterior proceso de paginado del índice [14].

### 3.1 Una Representación Secuencial

En [12] se presenta una nueva representación de un trie de sufijos que permite reducir el espacio necesario para almacenar el índice, eliminando la necesidad de mantener los punteros explícitos a los hijos. Los autores proponen representar de manera secuencial cada una de las componentes que conforman un trie de sufijos, manteniendo la posibilidad de navegar eficientemente sobre el trie:

**Topología del árbol.** La forma del árbol se representa utilizando la técnica de representación de paréntesis [10]. Esta representación consiste en realizar un barrido preorden sobre el árbol colocando un paréntesis que abre cuando se visita por primera vez un nodo y un paréntesis que cierra cuando se termina de visitar todo el subárbol de ese nodo. Esta representación utiliza un total de  $2n$  bits para un árbol de  $n$  nodos, manteniendo las facilidades de navegación sobre el mismo [10].

**Rótulos, saltos y grados.** Para la representación de los rótulos de cada rama, de los valores de salto de cada nodo y del grado de cada nodo, se utilizan arreglos colocando los elementos que forman cada arreglo en el orden indicado por un barrido preorden del árbol. Esto permite durante la navegación del árbol moverse coordinadamente sobre todas las secuencias que conforman la representación del mismo.

**Hojas.** Para la hojas también se mantiene un arreglo de índices de sufijos, tomados en orden de izquierda a derecha del árbol.

La figura 2 muestra esta representación para el trie de sufijos de la Figura 1.

Para poder navegar sobre esta representación los autores se basan en el algoritmo que permite navegar sobre una representación de paréntesis [10], adaptándolo para moverse coordinadamente sobre los 5 arreglos. Estos algoritmos necesitan realizar operaciones *findclose*, *excess* y *enclose* sobre secuencias binarias [10].

### 3.2 Una Mejora en Espacio

En [13] se presenta una mejora a la representación anterior que consiste en comprimir las secuencias RO, BR y GR de la Figura 2 usando códigos DAC (Directly Addressable Variable-Length Code) [2]. Estos códigos son una técnica de compresión que permite acceso aleatorio y eficiente a cada código en una secuencia de códigos de longitud variable. Para cada una de las secuencia mencionadas se utilizan códigos DAC para su representación.

Sea  $C = C_1, C_2, \dots, C_k$  una secuencia de  $k$  códigos, se divide cada código  $C_i$  en bloques de  $b$  bits. Luego se crea un arreglo  $A_1$  conteniendo la concatenación de los primeros bloques de cada símbolo, y un mapa de bits (*bitmap*)  $B_1$  de  $k$  bits, donde el  $i$ -ésimo bit está en 1 si el código  $C_i$  está formado por más de un bloque. Se continua con la creación de un arreglo  $A_2$  conteniendo la concatenación de los segundos bloques de cada símbolo y un bitmap  $B_2$  con 1 en aquellos bits correspondientes a los códigos con más de dos bloques. Se continua así hasta alcanzar la máxima cantidad de bloques.

Para acceder al código  $C_i$  de la secuencia original primero buscamos su primer bloque en  $A_1[i]$ . Si el  $i$ -ésimo bit de  $B_1$  está en 0, se finaliza retornando  $C_i = A_1[i]$ . Caso contrario, continuamos buscando el segundo bloque del código  $C_i$  esto es en  $A_2[\text{rank}_1(B_1, i)]$ , donde  $\text{rank}_1(B_1, i)$  es la cantidad de unos en  $B_1[1..i]$ . Este proceso continúa hasta llegar al último nivel de arreglos o hasta encontrar el bit del correspondiente bitmap en cero.

### 3.3 Una Mejora en Tiempo

En [3] se presenta una mejora en tiempo de la representación secuencial del trie de sufijos, que consiste en mantener junto con el trie una estructura llamada *ParentClose*. Esta estructura mejora substancialmente el tiempo de ejecución de la operaciones *findclose* que se realizan sobre la representación de paréntesis del  $TS_S$  (arreglo RP de la Figura 2).

Sea  $x$  la raíz, *ParentClose* en el nivel 1 almacena para cada hijo  $y$  de  $x$  tres valores: la posición del parentesis que cierra de  $y$ , la cantidad de nodos del subárbol con raíz  $y$  y la cantidad de hojas del subárbol con raíz  $y$ . Para poder mantener la navegación dentro de  $TS_S$ , esta información se ordena tomando los hijos de  $x$  de izquierda a derecha.

Esta misma información se puede mantener para los restantes niveles del trie. El problema es que si *ParentClose* se mantiene en todos los niveles del trie, el espacio que se necesita aumenta considerablemente. Además, a medida que se baja en el trie los

subárboles son cada vez mas pequeños por lo que en algún punto será más conveniente (tanto en tiempo como en espacio) trabajar directamente sobre *RP* sin usar *ParentClose*.

## 4 Evaluación Experimental

Presentamos aquí la evaluación experimental de las tres versiones del trie de sufijos secuencial explicadas en la sección anterior. Denotaremos con  $TS_S$  al trie de sufijos secuencial, con  $TS_{DAC}$  al trie de sufijos secuencial que utiliza códigos DAC y con  $TS_{DACN}$  al que utiliza códigos DAC y la estructura de niveles *ParentClose*.

Para realizar esta evaluación hemos tomado tres tipos de texto:

**DNA** que contiene secuencias de ADN.

**PROTEINS** que contiene secuencias de proteínas.

**SOURCE** que contiene códigos fuente de programas escritos en Java y C.

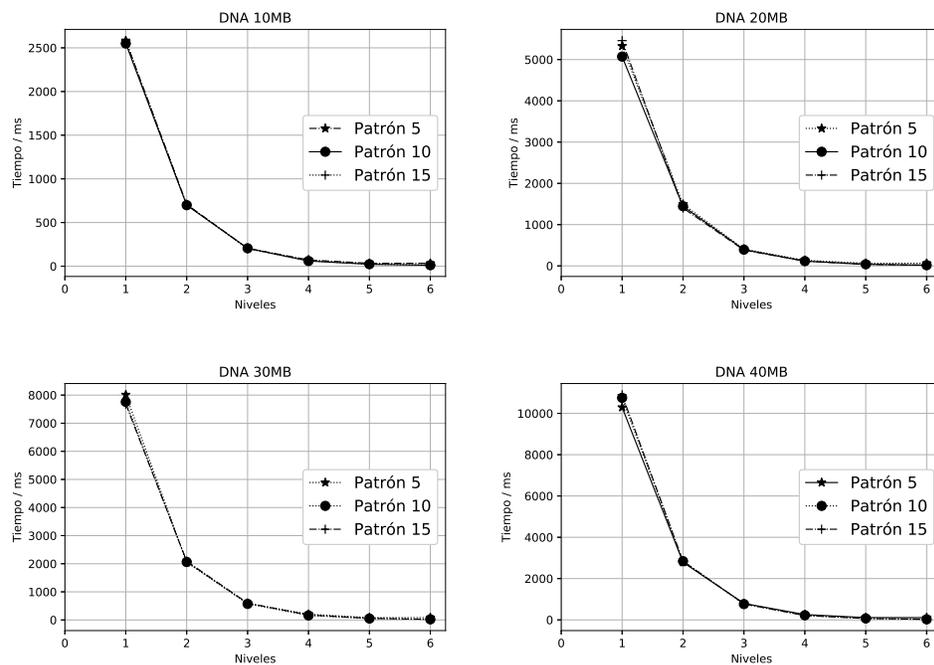
Estos textos han sido tomados del sitio <http://pizzachili.dcc.uchile.cl>. Para cada tipo de texto se han tomado partes de tamaño 10, 20, 30 y 40 MB a fin de poder evaluar la incidencia del tamaño del texto sobre los resultados obtenidos. Los tamaños han sido establecidos tomando en cuenta que si paginamos el índice las búsquedas se realizarán sobre parte pequeñas del mismo.

Para las tres representaciones consideradas del trie de sufijos, hemos evaluado el tamaño del índice y los tiempos de las operaciones *count*. Los resultados aquí mostrados fueron obtenidos realizando búsquedas con patrones de longitud 5, 10 y 15. Para cada longitud de patrón se generaron lotes de 500 patrones, promediándose los resultados obtenidos.

Presentamos a continuación la evaluación de estos resultados mostrando las gráficas que hemos considerado más relevantes. Evaluaremos primero el nivel mas adecuado para *ParentClose* en  $TS_{DACN}$ , para luego comparar la mejor versión de  $TS_{DACN}$  con las demás representaciones.

### 4.1 Efecto de los Niveles sobre $TS_{DACN}$

La Figura 3 muestra los resultados obtenidos para texto DNA realizando búsquedas con patrones de longitud 5, 10 y 15. Sobre el eje  $x$  se han representado los distintos niveles de *ParentClose* y sobre el eje  $y$  está representado el tiempo medio de realizar la operación *count* sobre un patrón, expresado en milisegundos. Como puede observarse el tiempo medio de *count* disminuye a medida que aumentamos la cantidad de niveles. En los primeros niveles se obtiene una mejora notable logrando reducir el tiempo en un 70% entre el nivel 1 y el nivel 2. Si comparamos el nivel 1 con el nivel 4 la mejora es de aproximadamente el 90%. A partir del nivel 4 los tiempos tienden a estabilizarse. Con respecto a la longitud del patrón de búsqueda, puede observarse que no tiene efectos significativos sobre el comportamiento de  $TS_{DACN}$ . Esta misma situación se produce con los otros textos analizados. La Figura 4 muestra los resultados para **PROTEINS** y **SOURCE** de tamaños 10MB y 40MB. En todos los casos los mejores resultados se obtienen a partir del nivel 4 logrando reducciones cercanas al 90%.



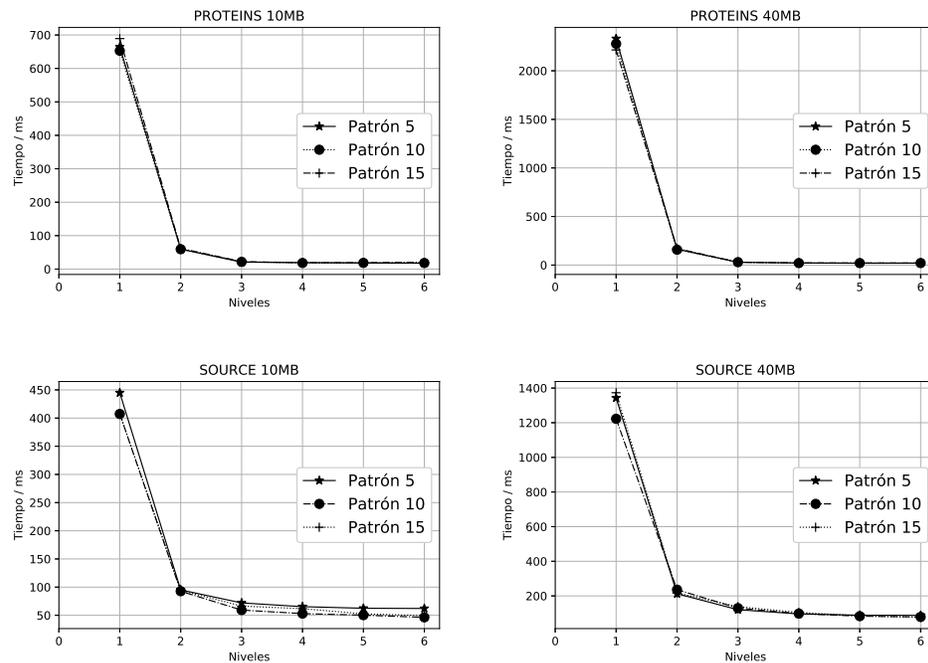
**Fig. 3.** Tiempo medio de *count* para distintos tamaños de patrones de búsqueda con texto DNA.

Tam. del Texto	Tam. de $TS_{DACN}$	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
10	69.21	0.00008	0.0004	0.001	0.004	0.01	0.06
20	139.38	0.0001	0.0006	0.002	0.005	0.01	0.06
30	209.05	0.0001	0.001	0.004	0.01	0.02	0.07
40	278.91	0.0001	0.001	0.004	0.01	0.02	0.07

**Table 1.** Espacio ocupado por  $TS_{DACN}$  para un texto DNA. Todos los valores están expresados en MB.

Con respecto al espacio, la Tabla 1 muestra los valores obtenidos para texto DNA: para cada tamaño de texto se indica el tamaño del trie de sufijos y de *ParentClose* con los diferentes niveles analizado. Se puede observar que es muy poco el espacio adicional que se necesita para obtener las reducciones de tiempo antes señaladas: en nivel 4 y 5 se necesita solo un 0.01% aproximadamente de espacio extra. Esta misma situación se produce con todos los tipos y tamaños de texto que hemos considerado en este trabajo.

Podemos concluir que el nivel 4 y nivel 5 son un punto de compromiso entre espacio y tiempo para este índice. Estos valores serán los utilizados para la comparación global de la siguiente sección.



**Fig. 4.** Tiempo medio de *count* de  $TS_S$  para distintos tamaños de patrones de búsqueda con textos PROTEINS y SOURCES.

## 4.2 Comparación de $TS_S$ , $TS_{DAC}$ y $TS_{DACN}$

En esta sección comparamos en espacio y tiempos las tres representaciones del trie de sufijos consideradas en este trabajo.

La figura 5 muestra los tiempos medios obtenidos con todos los textos considerados para patrones de longitud 5, en escala logarítmica. Se puede observar una clara ventaja de  $TS_{DACN}$  respecto de las otras dos versiones, logrando reducir el tiempo entre un 80% y un 90% aproximadamente. La diferencia entre nivel 4 y nivel 5 no es tan significativa, lo que implica que nivel 4 es también una buena opción si queremos bajar los requerimientos de espacio de  $TS_{DACN}$ . También se observa que  $TS_S$  y  $TS_{DAC}$  tienen un comportamiento similar, habiendo una pequeña ventaja en tiempo para  $TS_S$ ; esto refleja las operaciones adicionales que se hacen sobre los códigos DAC en  $TS_{DAC}$  para recuperar rótulos, grados y saltos. El mismo comportamiento se observó con las demás longitudes de patrones.

La figura 6 muestra el espacio ocupado por  $TS_S$ ,  $TS_{DAC}$  y  $TS_{DACN}$  para los tres textos considerados. Claramente las dos versiones que usan DAC ( $TS_{DAC}$  y  $TS_{DACN}$ ) superan en eficiencia a  $TS_S$ , ocupando un 50% menos de espacio. Entre  $TS_{DAC}$  y  $TS_{DACN}$  la diferencia de espacio es realmente baja:  $TS_{DAC}$  ocupa un 0.01% que  $TS_{DACN}$  en el caso de DNA y aproximadamente un 10% menos para PROTEINS y SOURCES.

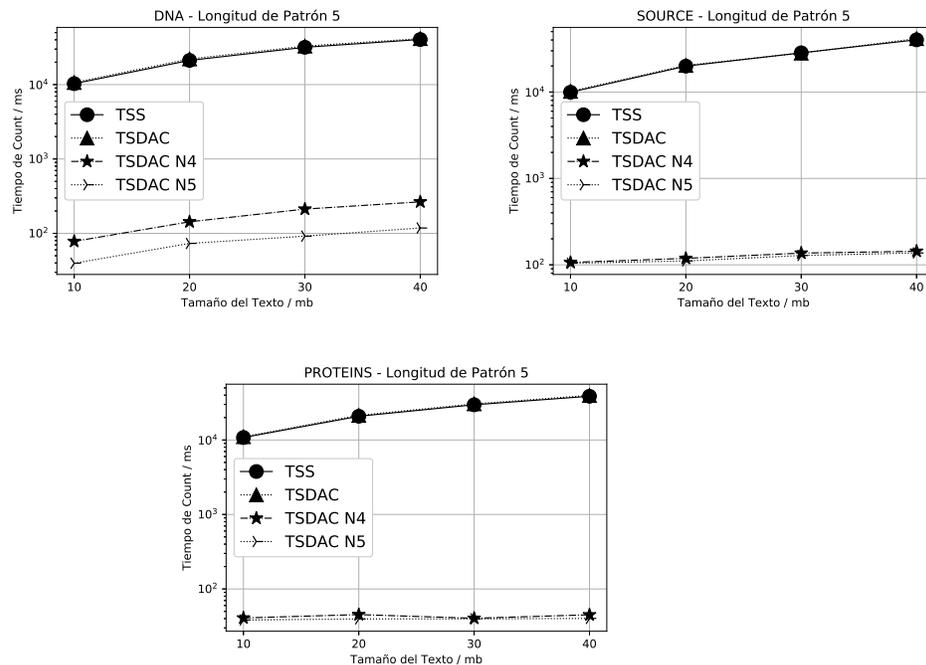


Fig. 5. Tiempo medio de *count* para patrones de longitud 5.

Podemos concluir, que en espacio  $TS_{DAC}$  es la mas eficiente de las 3 y en tiempo  $TS_{DACN}$  es la que mejor desempeño tiene, con un bajo espacio adicional respecto de  $TS_{DAC}$ .

## 5 Conclusiones y Trabajo Futuro

En este trabajo presentamos una evaluación experimental de tres versiones del trie de sufijos. De los experimentos realizados podemos concluir que  $TS_{DAC}$  es la de mejor desempeño en espacio y  $TS_{DACN}$  la mas competitiva en tiempo. Con respecto al trabajo futuro nos proponemos integrar estas dos versiones con técnicas de paginado para memoria secundaria.

## References

1. R. S. Boyer and J. S. Moore. A fast string searching algorithm. *Communications of the ACM*, 20(10):762–772, 1977.
2. Nieves R. Brisaboa, Susana Ladra, and Gonzalo Navarro. Directly addressable variable-length codes. In *SPIRE*, pages 122–130, 2009.
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.

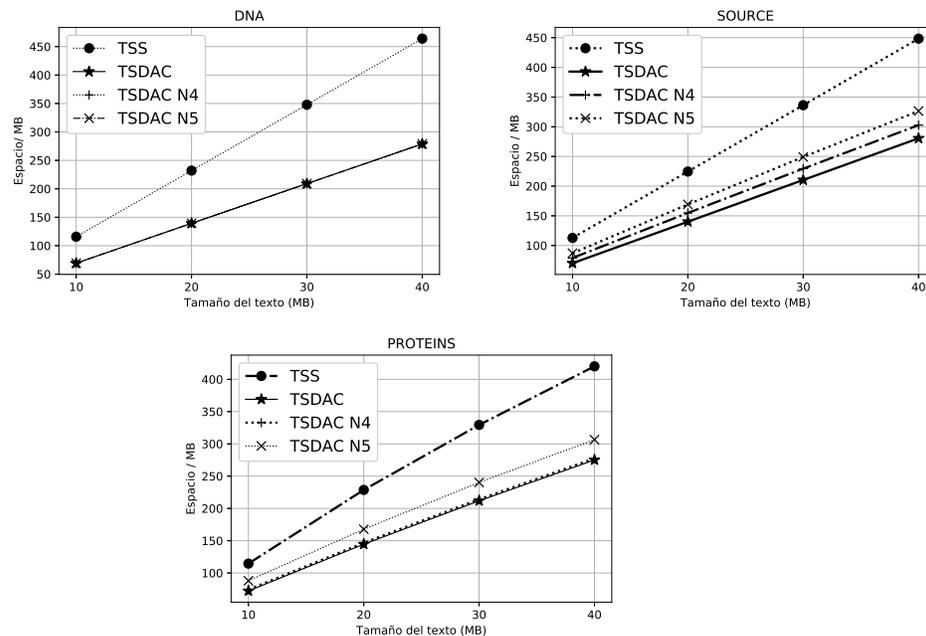


Fig. 6. Espacio ocupado por el índice.

4. G. H. Gonnet and R. Baeza-Yates. *Handbook of Algorithms and Data Structures*. Addison-Wesley, 1991.
5. G. H. Gonnet, R. Baeza-Yates, and T. Snider. *New indices for text: PAT trees and PAT arrays*, pages 66–82. Prentice Hall, New Jersey, 1992.
6. 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.
7. A. Thomo M. Barsky \*, U. Stege. A survey of practical algorithms for suffix tree construction in external memory. In *Software: Practice and Experience*, 2010.
8. U. Manber and G. Myers. Suffix arrays: A new method for on-line string searches. *SIAM Journal of Computing*, 22(5):935–948, 1993.
9. E. Moura, G. Navarro, N. Ziviani, and R. Baeza-Yates. Fast and flexible word searching on compressed text. *ACM Transactions on Information Systems (TOIS)*, 18(2):113–139, 2000.
10. J. Ian Munro and Venkatesh Raman. Succinct representation of balanced parentheses and static trees. *SIAM J. Comput.*, 31(3):762–776, 2001.
11. G. Navarro and V. Mäkinen. Compressed full-text indexes. *ACM Computing Surveys*, 39(1):article 2, 2007.
12. 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.
13. 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.
14. Jeffrey Scott Vitter. Publishers Inc, 2006.
15. P. Weiner. Linear pattern matching algorithm. In *Proc. 14th IEEE Symposium Switching Theory and Automata Theory*, pages 1–11, 1973.

# Workshop Ingeniería de Software

Coordinadores

Patricia Pesado (UNLP)  
Elsa Estevez (UNS)  
Alejandra Cechich (UNCOMA)  
Horacio Kuna (UNaM)

## Transferencia de Conocimiento en la Gestión de Calidad de la Ingeniería de Software

Mauricio Rozo<sup>1 2</sup>, Ines Casanovas<sup>1 2</sup>, Ma. Florencia Pollo Cattaneo<sup>1 2</sup>,

<sup>1</sup>Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS)  
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.  
Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires

<sup>2</sup>Programa de Maestría en Ingeniería de Sistemas de Información  
Universidad Tecnológica Nacional, Facultad Regional Buenos Aires.  
Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires

{mauricio.rozo.rodriguez, inescasanovas, flo.pollo}@gmail.com

**Resumen.** Los procesos y actividades en la gestión de calidad de la ingeniería de software generan un gran volumen de conocimiento, con lo cual es considerado como un factor crítico para la calidad de producto software, por lo tanto, exige una creciente demanda en la mejora de la efectividad y cumplimiento de las tareas que la componen, es ahí donde el uso de métodos y principios de gestión de conocimiento se convierte en la base para gestionarla. Con base en este argumento, se indagan los desafíos de los modelos de gestión de conocimiento existentes en el dominio de la calidad de software para identificar las falencias de los modelos existentes y a partir de estas, proponer una solución para la ingeniería de software en el ámbito de la gestión de conocimiento, basado en la utilización de metodologías ontológicas en el dominio de la fase de pruebas.

**Palabras Clave:** Gestión de Conocimiento, Gestión de Calidad, Fase de Pruebas, Ingeniería de Software, Ontologías.

### 1 Introducción

Los procesos y actividades de la gestión de calidad tienen gran importancia en la ingeniería de software debido que permiten evaluar las soluciones informáticas y determinar el nivel de calidad que poseen. Por tanto, esta fase es una de las más importantes en el desarrollo de software, sin ellas, el producto informático contendría multitud de defectos provocando alto impacto y costos en la solución de los defectos, deteriorando la gestión integral en las organizaciones [1-2]. Por otro lado, en [3], entienden que las pruebas de software son un proceso intensivo en conocimiento, y se hace necesario proporcionar soporte computarizado para las tareas de adquisición, procesamiento, análisis y diseminación de conocimiento para su reutilización.

Es indudable que las pruebas de software (PS) es uno de los aspectos que hacen que la ingeniería de software trascienda en los proyectos informáticos con calidad de

producto. Sin embargo, las dificultades que atraviesan para ser efectivas tienen diferentes ámbitos, en ese sentido, el principal problema en las organizaciones de software relacionadas con esta etapa es la baja tasa de reutilización del conocimiento y las barreras en la transferencia de este [4].

A pesar de la importancia de las PS, existen organizaciones que no se les presta la suficiente atención o son, simplemente, omitidas en la planificación del desarrollo de productos software debido a la presión por tener un producto terminado en una fecha establecida, con lo cual provoca que haya actividades, consideradas erróneamente como prescindibles, que son eliminadas o reducidas en la planificación, afectando la calidad de software [3,5].

De acuerdo con lo anteriormente descrito, en [6], se identificaron varios desafíos para la gestión de la calidad de los productos de software, entre los que destacan: baja tasa de reutilización del conocimiento de PS, barreras en la transferencia de conocimientos, baja posibilidad de lograr rápidamente la distribución más eficiente de recursos humanos durante las pruebas. Además, se enuncia que las debilidades en las PS se deben a que existe una pérdida significativa de capital intelectual debido a la rotación del personal y conocimiento limitado. La obtención, difusión, adquisición, evolución y empaquetamiento del conocimiento reciben poca atención debido a que el conocimiento se maneja principalmente durante las pruebas dentro de un proyecto u organización y se dedica menos atención al intercambio de conocimientos [7-8].

Basado en estos desafíos, la gestión de conocimiento (GC) provee, de manera estructurada y sistemática, las herramientas para facilitar que el conocimiento generado permita alcanzar los objetivos estratégicos de las organizaciones y optimizar las decisiones que se tomen para mejorar los procesos, convirtiéndose en un activo [9].

Ahora bien, entendiendo que, en la ingeniería de software, la GC está desarrollándose en la medida de las necesidades que expresan las organizaciones de este rubro, es necesario expandir el uso de la GC dentro de las diferentes fases de la ingeniería de software, debido a la especificidad que exige la evolución de los procesos de cada una. En consecuencia, surgen como interrogantes cuál es el estado de situación de la GC aplicado a la gestión de calidad de la ingeniería de software y qué técnicas existentes y adaptadas en otros ámbitos informáticos, son aplicables en la GC en la fase de PS y surgen como elementos para potencializar el desarrollo y masificación en el contexto de la ingeniería de software. En vista de lo anterior, es necesario resolver estos cuestionamientos para determinar qué técnica es viable utilizar para potencializar el desarrollo de la GC aplicado a la PS en la ingeniería de software. Es por ello que, aunque reconocida como un instrumento importante por la comunidad de GC, las ontologías no se utilizan ampliamente en estas iniciativas, entendiendo que una ontología es una especificación explícita de una conceptualización, en ésta se modela el vocabulario del dominio, básicamente usando los conceptos, características y sus relaciones [3,10,11]. En ese sentido en [12,20] se propone estudiar cómo las ontologías se pueden utilizar en la GC en el dominio de la gestión de calidad en el ámbito de la ingeniería de software.

Para cumplir con el objetivo propuesto, el presente trabajo se organiza de la siguiente manera: en la Sección 2 se exponen los materiales y métodos, en la Sección 3 se presenta la propuesta metodológica para la transferencia de conocimiento en la gestión de calidad, en la Sección 4 se presentan los resultados de la evaluación de la

propuesta. Finalmente, en la Sección 5 se exponen las conclusiones y trabajos futuros.

## 2 Materiales y Métodos

El presente trabajo se basa metodológicamente en la visión del constructivismo, donde el interés se centra en cómo se crean las significaciones y cómo se construye la realidad. De acuerdo con lo anterior, el enfoque metodológico que sustenta el presente trabajo es el cualitativo, donde el investigador hace cuestionamientos basados principalmente en perspectivas constructivistas que demanda estrategias de investigación [13-14].

El diseño utilizado en esta investigación es el descriptivo, y en relación con los métodos, se utiliza la revisión sistemática documental exploratoria [15-16] aplicándolo en el conocimiento existente en el dominio de la GC, en el dominio de las PS y en el dominio de los modelos ontológicos para determinar el estado de la cuestión y fundamentación de los desafíos en la gestión de calidad de software.

A su vez, se realiza un estudio analítico y comparativo de las metodologías existentes para determinar su aplicabilidad a la solución de la problemática, concluyendo en la elaboración de una propuesta metodológica que solucione o minimice la brecha de conocimiento detectada como desafíos en la gestión de calidad dentro de la ingeniería de software. Como solución se desarrolla un modelo de aplicación de GC a la PS de la ingeniería de software.

## 3 Propuesta Metodológica

El modelo de GC propuesto reúne los descriptores: motor de la GC, uso de las herramientas tecnológicas, contexto, actores, tipo de conocimiento y clima organizacional, a su vez incorpora en su estructura una propuesta ontológica que adopta el método planteado en [17] como pasos para su construcción y define el uso de patrones de diseño, basado en los desafíos de las ontologías y aplicable a la gestión de la calidad de producto de la ingeniería de software [21].

El modelo está basado en los procesos propuestos por los modelos Riesco [9], Paniagua y López [18], Angulo y Negrón [9] y el Modelo ISECO [9], teniendo en cuenta que cumple con lo definido en [18], que plantean como elemento clave en la GC la necesidad de asumir los aspectos relacionados con las personas, los procesos y la tecnología como un todo.

Como lo propone [19], los procesos de GC ocurren cíclicamente y acompañan la evolución de las PS, maximizando la tasa de reutilización, la transferencia de conocimiento, a su vez, minimizan los desafíos que tiene la fase de pruebas en la ingeniería de software. El modelo propuesto incorpora una capa ontológica, que adopta mejoras en la formalización y distribución del conocimiento, a través de la propuesta de una ontología de dominio en el contexto de las PS (ver Fig. 1).



**Fig. 1.** Propuesta de un modelo de aplicación de Gestión de Conocimiento a la fase de pruebas de la ingeniería de software.

### 3.1 Estructura

El modelo de GC propuesto está dividido en instancias, entendiendo que una instancia, como lo propone [22], es una entidad que pertenece a un determinado contexto. Las instancias que componen el modelo son: organización, ejecución y conservación o mantenimiento, se describen a continuación. A su vez, cada instancia posee una serie de fases que permiten el desarrollo de las PS en el dominio de la GC.

**Instancia de Organización.** En esta instancia, lo que se busca es la asignación de roles específicos, para que, en la GC desde la perspectiva de la organización, garantice que las habilidades, experiencias y pericias del equipo del proyecto se utilicen antes, durante y después del proyecto [19,23].

Desde las perspectivas antes descritas, se busca definir en el equipo de pruebas, el rol de gestor de conocimiento y a su vez definir el dominio de conocimiento bajo el enfoque de las PS que va a ser incorporado en la GC.

Como lo proponen en [23], para las entradas de esta parte del modelo se tiene el expertise de los integrantes del equipo de pruebas, que proporciona información sobre competencias y experiencias disponibles del proyecto de software, además ayuda a comprender qué conocimiento está o no disponible; también se requiere la documentación del proyecto como registro de lecciones aprendidas, documentación de requisitos. Las actividades principales consisten en seleccionar el gestor de conocimiento, determinar qué tipos de pruebas se van a utilizar en el proyecto, definir la estrategia para la obtención de elementos de conocimiento básico que exista en el proyecto [23]. Como salida se tiene el gestor y dominio de conocimiento seleccionado que involucra la validación del plan de gestión de calidad y sus métricas.

**Instancia de Ejecución.** En la instancia de ejecución, se busca cumplir las etapas cíclicas del modelo para que las PS dentro de la gestión de calidad queden inmersas en el contexto de la GC. En [18], se expone que estas acciones disminuyen la redundancia de tareas y el número de errores en su ejecución, al aprovechar la experiencia existente. De acuerdo con lo anterior, esta instancia posee cuatro etapas que involucran: Identificación del conocimiento, Formalización del Conocimiento, Distribución del Conocimiento y Validación e Incorporación del Conocimiento.

*Identificación del conocimiento.* Se inicia cuando el gestor identifica el conocimiento requerido para que el equipo de pruebas del proyecto realice las PS de acuerdo con el control de gestión de calidad seleccionado. Se establece qué información se tiene disponible y sus fuentes. Se basa en convertir el conocimiento explícito en conocimiento implícito y agregarle experiencia [19,24]. Como entradas, se identifica la necesidad de datos, de información y de conocimiento explícito de las pruebas no formalizado en el dominio del proyecto, se requiere la documentación como el registro de lecciones aprendidas y documentación de requisitos. Se propone como actividades la recopilación de datos, el juicio de expertos y el análisis de datos, para identificar la necesidad de conocimiento en el ámbito de la calidad del proyecto. Como salida se tiene la información y el conocimiento identificado para ser formalizado en la GC [23].

*Formalización del conocimiento.* Toma la salida de la fase previa y a través de la capa ontológica, identifica las características y/o atributos para categorizar, evaluar, seleccionar y formalizar la información que ingresa en la GC. Como entradas se tiene la información y/o conocimiento no formalizado. Como actividades se propone utilizar la metodología ontológica del modelo Stuart [17], es decir, seguir sus procesos para definir las características y/o atributos para categorizar, evaluar, seleccionar y formalizar la información que ingresa en la GC a través de la creación de la ontología. En [21], indican que la practicidad de este modelo es, ser iterativo y permite la reutilización de las ontologías, lo cual reduce costos y facilita la interoperabilidad de los datos. La ontología de dominio que el modelo de GC aplicado a la fase de PS en la ingeniería de software propone, está inmersa en la capa ontológica. Se busca que esta ontología permita ser más eficiente en resolver la identificación de un concepto o conjuntos de conceptos. Como última actividad se tiene el almacenamiento, dando crecimiento a la memoria corporativa y los sistemas basados en conocimiento [18]. Como salida se tiene el conocimiento implícito formalizado y almacenado, disponible para ser presentado.

*Distribución de conocimiento.* Comparte el conocimiento entre los miembros del equipo de PS, con lo cual, el conocimiento implícito es convertido a explícito y combina este conocimiento a partir de las experiencias adquiridas en la fase de PS del proyecto, el énfasis es adquirir los mecanismos necesarios para dejar disponible la transmisión de la información adecuada para aplicar y crear conocimiento [9]. Como entradas se tiene el conocimiento implícito formalizado. Dentro de las actividades se propone compartir el conocimiento entre los miembros del equipo, exteriorización del conocimiento a través del intercambio de experiencias, generar conocimiento

explícito a través de las experiencias adquiridas en las PS realizadas en el proyecto, transmitir la información para crear conocimiento, aquí cobra participación de la capa ontológica, para mejorar la eficiencia en la identificación y procesamiento en la GC para las PS que se estén realizando. Como salida se tiene el conocimiento explícito.

*Validación e Incorporación del Conocimiento.* En esta etapa, el conocimiento explícito desarrollado en la fase de pruebas es validado e incorporado como activo de conocimiento en el dominio de GC para el proyecto de software. Los integrantes del equipo generan conocimientos explícitos al reunir experiencias provenientes de sus aportes y de otras fuentes, es aquí donde se valida y combina el conocimiento, dando como resultado una fácil comprensión del conocimiento para ser utilizado en la generación de nuevas experiencias [19]. Es importante que en la validación participen los miembros del equipo de PS, que junto al gestor de GC, le dan el valor de aceptación, o no, del conocimiento (En el caso de una validación no aprobada, se reinicia el proceso de GC). Tal como se establece en [9], el conocimiento tendrá un valor apreciado cuando su utilización sea direccionada a mejoras en los procesos, toma de decisiones, innovación, resolución de problemas o cualquier otro aspecto que beneficie la fase de PS del proyecto. Por otro lado, [23] propone que se debe realizar el proceso de control de la calidad, en el cual se monitorea y registra los resultados de la ejecución del plan de pruebas, de esta manera se evalúa el desempeño y se asegura que la fase PS del proyecto se encuentre completa, sea correcta y satisfaga las necesidades la calidad de producto del proyecto de software. Como entrada se tiene el conocimiento explícito. Como actividades se define la realización de reuniones de validación. Como salida se tiene el conocimiento explícito incorporado como activo dentro de la PS para el proyecto de ingeniería de software.

**Instancia de Conservación o Mantenimiento.** Extiende el ciclo de vida del modelo realizando los ajustes en los aspectos que lo requiera, buscando la evolución de este según las necesidades. Es necesario revisar el estado de aplicación del modelo, de esta manera se sigue y controla los resultados, para adelantar en el proyecto la implantación, o no, de mejoras en determinadas etapas [19]. Para [23], este proceso se lleva a cabo a lo largo del proyecto y se realiza a través del control de la calidad, el beneficio es verificar que el trabajo del proyecto cumpla con los requisitos especificados. En esta instancia como entrada se identifican los procesos del modelo de GC propuesto. Como actividades se verifica la aplicabilidad del modelo y se realiza el seguimiento a los procesos, corrección a posibles desvíos en la gestión de los procesos. Como salida se toman las lecciones aprendidas y se reincorporan en la evolución del modelo de GC.

**Capa Ontológica del modelo.** Para el desarrollo de la ontología de dominio que da sustento a la capa ontológica del modelo de GC, se adopta el método propuesto en [17], constituido por doce pasos que incorpora la identificación del software apropiado y la sustentabilidad de la ontología [21]. Además, se complementa la ontología propuesta, con el fortalecimiento de la reutilización con ontologías en la dimensión de cobertura conceptual, fortaleciendo su evaluación e incorporando la implementación de relaciones no taxonómicas en la dimensión de estructura no

propuesta en la ontología ROoST [8]. Es aquí, donde el presente trabajo avanza en fortalecer e incorporar mejoras en la identificación y procesamiento en la GC para las PS que se estén realizando. Se proponen para cada uno de los pasos, una serie de preguntas de competencia, a su vez equivale a la elicitación de términos, que la ontología debe ser capaz de contestar con todos los conceptos, propiedades entre conceptos y datos, los axiomas y reglas definidas en la ontología [17,24].

#### 4 Resultados obtenidos

La validación de este modelo se divide en dos componentes, el primero evalúa el fortalecimiento de la PS dentro de la gestión del proyecto de software en el cual se definen las siguientes dimensiones: gestión de la información y la documentación, gestión de la comunicación, gestión de la innovación y el cambio, gestión del aprendizaje organizacional [25,26]. El segundo componente define la efectividad de operación al aplicar las fases del modelo de GC en el dominio de PS, lo compone: identificación de conocimiento, formalización de conocimiento, distribución del conocimiento, retención de conocimiento [25-27]. A partir de las dimensiones, se definen los indicadores, lo que permite establecer los ítems que componen el instrumento de investigación. La tabla 1 delimita los componentes, dimensiones e indicadores a fines de determinar la estructura de validación.

**Tabla 1.** Componentes, Dimensiones e Indicadores de validación del modelo de GC aplicado a la PS de la ingeniería de software.

Componente	Dimensión	Indicadores
<b>Fortalecimiento de la fase de pruebas dentro de la gestión del proyecto de software.</b>	Gestión de la información y la documentación	1.Análisis de requerimientos de información 2.Gestión documental
	Gestión de la comunicación	1.Estructura de entorno adecuado para comunicar la información 2.Facilidad de comunicar a través de materiales digitales documentación de casos de pruebas funcionales, de regresión, de aceptación.
	Gestión de la innovación y el cambio	1.Sistemas de evaluación del desempeño 2.Gestión del cambio de cultura en la mejora de las actividades en las pruebas
	Gestión del aprendizaje organizacional	1.Incorporación de mejores prácticas en el desarrollo de las actividades de pruebas 2.Mejora de gestión de calidad en la fase de pruebas
<b>Efectividad de operación al aplicar las fases del modelo de GC</b>	Identificación de conocimiento	1.Espacio para el fortalecimiento de lecciones aprendidas 2.Incorporación de mapas de conocimiento
	Formalización de conocimiento	1.Sistematizar las tareas de las pruebas 2.Sistema de gestión de la formación
	Distribución del conocimiento	1.Facilidad al acceso a los escenarios de pruebas 2.Efectividad al acceso a los escenarios de pruebas
	Retención de conocimiento	1.Incorpora centros digitales de información 2.Incorpora cuadernos digitales explicativos sobre las actividades de la fase de pruebas

Por otro lado, tabla 2 resume el resultado obtenido para cada uno de los componentes, producto de aplicar el instrumento de investigación a los Expertos.

El cálculo del valor promedio para cada indicador está basado en la escala Likert. El valor referencial, describe la tendencia de valoración de la escala de acuerdo con la puntuación que el Experto aplica sobre las preguntas del instrumento de investigación [15], para valores menores de 2,5 representa estar en desacuerdo, para valores mayores a 3,5 indica estar de acuerdo con lo que propone el modelo de GC para ese indicador. Para los valores que sean mayores de 2,5 y menores de 3,5 no se está ni de acuerdo ni en desacuerdo. Ahora bien, analizando los resultados en un nivel mayor, se tiene que los Expertos coinciden como mayor aporte del modelo de GC a la dimensión “Gestión del aprendizaje organizacional”, el cual obtuvo una valoración de 4,56 puntos. De esta manera, la propuesta del presente trabajo de investigación aporta crecimiento en términos de aprendizaje dentro de las organizaciones. A su vez, los Expertos concuerdan que la dimensión “Formalización de conocimiento” es la segunda en importancia para el modelo propuesto, debido que obtuvo una valoración de 4,39 puntos.

**Tabla 2.** Resultados para los componentes del modelo de GC producto de la valoración de los Expertos.

Componente	Dimensión	Valor Prom. x Dimen.	Valor Prom. x Componen.	Valor Ref.	En Desacuerdo / De Acuerdo
Fortalecimiento de la fase de pruebas dentro de la gestión del proyecto de software.	Gestión de la información y la documentación	3,94	4,14	2,5 < 3,5	De Acuerdo
	Gestión de la comunicación	3,89			
	Gestión de la innovación y el cambio	4,17			
	Gestión del aprendizaje organizacional	4,56			
Efectividad de operación al aplicar las fases del modelo de GC	Identificación de conocimiento	4,28	4,18	2,5 < 3,5	De Acuerdo
	Formalización de conocimiento	4,39			
	Distribución del conocimiento	4,28			
	Retención de conocimiento	3,78			

## 5 Conclusiones y Trabajo Futuro

El modelo de aplicación de GC a la fase de pruebas de la ingeniería de software propuesto, se validó y los resultados denotan que, en el primer nivel (que refiere a los indicadores), los Expertos estuvieron de acuerdo con la propuesta, destacando la mejora de gestión de calidad en la fase de pruebas. Para el segundo nivel (que define a las dimensiones), los Expertos destacaron que la gestión del aprendizaje organizacional fue el que más se acentuó en el modelo de GC propuesto. En relación con el tercer nivel (que describe los componentes que hacen a la utilización del

modelo de GC aplicado a las PS), los Expertos indicaron que se expresa efectividad de operación al aplicar las fases del modelo. A su vez están de acuerdo ya que advierten el fortalecimiento de la fase de pruebas dentro de la gestión del proyecto de software.

Como trabajo futuro se continuará en mejorar los aspectos relacionados a la incorporación de cuadernos (“Guías”) digitales explicativas sobre las actividades de la PS, el cual fue uno de los indicadores con menor valoración. A su vez, se fortalecerá la dimensión “Retención de conocimiento”. Por otro lado, es necesario continuar incorporando aspectos en el desempeño del modelo GC desde el enfoque de la gestión del recurso humano.

## Referencias

1. Jústiz-Núñez, Dalila, Gómez-Suárez, Darlene, Delgado-Dapena, Marta Dunia: Proceso de pruebas para productos de software en un laboratorio de calidad. *Ingeniería Industrial*, 35(2), 131-145. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1815-59362014000200003&lng=es&tlng=pt](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1815-59362014000200003&lng=es&tlng=pt) (2014).
2. Fernández, F., Ángel, M.: Aplicación de técnicas de pruebas automáticas basadas en propiedades a los diferentes niveles de prueba del software. Recuperado de <http://hdl.handle.net/2183/14814> (2015).
3. Souza, E. F., Falbo, R. A., Vijaykumar, N. L.: Knowledge management applied to software testing: A systematic mapping. In *The 25th International Conference on Software Engineering and Knowledge Engineering (SEKE 2013)*, Boston, USA (pp. 562-567). <https://www.researchgate.net/publication/289638583> (2013).
4. Souza, É. F., Falbo, R. A., Vijaykumar, N. L.: Using the findings of a mapping study to conduct a research project: a case in knowledge management in software testing. In *Software Engineering and Advanced Applications (SEAA), 2015 41st Euromicro Conference on* (pp. 208-215). IEEE. <http://dx.doi.org/10.1109/SEAA.2015.10> (2015b).
5. Dávila, Abraham, García, Cecilia, Córdor, Sandra: Análisis exploratorio en la adopción de prácticas de pruebas de software de la ISO/IEC 29119-2 en organizaciones de Lima, Perú. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Información*, (21), 1-17. <https://dx.doi.org/10.17013/risti.21.1-17> (2017).
6. Wnuk, K., Garrepalli, T.: Knowledge Management in Software Testing: A Systematic Snowball Literature Review. *e-Informatica Software Engineering Journal*, 12(1), 51-78. <http://io.pwr.edu.pl/eprints/id/eprint/81> (2018).
7. Durán, D. E. S., Gamboa, A. X. R., Builes, J. J.: Aplicación de la Gestión de Conocimiento al proceso de pruebas de software. *Ingenierías Revistas USBMed*, 8(2), 6-13. <http://dx.doi.org/10.21500/20275846.2836> (2017).
8. Souza, É. F. D., Falbo, R. D. A., Vijaykumar, N. L.: ROoST: reference ontology on software testing. *Applied Ontology*, 12(1), 59-90. <https://www.researchgate.net/publication/314268288> (2017).
9. Avendaño Pérez, V., Flores Urbáez, M.: Modelos teóricos de gestión del conocimiento: descriptores, conceptualizaciones y enfoques. *Entreciencias: diálogos en la Sociedad del Conocimiento*, 4(10). <http://www.redalyc.org/articulo.oa?id=457646537004> (2016).
10. Flores, V., Hadfeg, Y.: Un método para generar explicaciones de resultados de un Sistema Experto, usando Patrones de discurso y Ontología. *RISTI-Revista Ibérica de Sistemas e Tecnologías de Información*, (21), 99-114. [http://www.scielo.mec.pt/scielo.php?script=sci\\_arttext&pid=S1646-98952017000100008&lng=pt&nrm=iso](http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S1646-98952017000100008&lng=pt&nrm=iso) (2017).

11. González Gola, F., Sánchez, A., Montejano, G. A.: Asistencia dirigida por ontologías al diseño arquitectónico de videojuegos. In XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires). <http://sedici.unlp.edu.ar/handle/10915/61343> (2017).
12. Rozo Rodríguez M., Casanovas I.: La Gestión de Conocimiento Aplicado a la Fase de Pruebas de la Ingeniería de Software - Revisión Sistemática. 6to Congreso Nacional de Ingeniería Informática / Sistemas de Información. CoNaIISI. Publicación online - ISSN 2347-0372. <https://www.conaiksi2018mdp.org/memorias/memorias.html#> (2018).
13. Creswell, J. W.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. University of Nebraska, Lincoln. 4a. ed. Sage Publications, Inc. (2014).
14. Labra, O.: Positivismo y Constructivismo: Un análisis para la investigación social. Rumbos TS. Un espacio crítico para la reflexión en Ciencias Sociales, 0(7), 12-21. <http://revistafacso.ucentral.cl/index.php/rumbos/article/view/55> (2016).
15. Hernández Sampieri, R., Fernández Collado, C., Baptista Lucio, P.: Metodología de la investigación. México: Editorial Mc Graw Hill, 15-40. ISBN 978-92-75-32913-9. (2010).
16. Souza, E. F., de Almeida Falbo, R., Vijaykumar, N. L.: Knowledge management initiatives in software testing: A mapping study. Information and Software Technology, 57, 378-391. <https://doi.org/10.1016/j.infsof.2014.05.016> (2015).
17. Stuart, D.: Practical Ontologies for Information Professionals. Facet. <https://doi.org/10.29085/9781783301522> (2016).
18. Medina Nogueira, Daylin, Medina León, Alberto, Nogueira Rivera, Dianelys.: Procesos y Factores claves de la Gestión del Conocimiento. Revista Universidad y Sociedad, 9(3), 16-23. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2218-36202017000300002&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2218-36202017000300002&lng=es&tlng=es) (2017).
19. Linares Pons, N., Piñero Pérez, Y., Rodríguez Stiven, E., Pérez Quintero, L.: Diseño de un modelo de Gestión del Conocimiento para mejorar el desarrollo de equipos de proyectos informáticos. Revista española de Documentación Científica, 37(2), e044. <http://dx.doi.org/10.3989/redc.2014.2.1036> (2014).
20. Rozo Rodríguez M., Casanovas I.: Aplicación de Gestión de Conocimiento a la Fase de Pruebas de la Ingeniería de Software XXII Workshop de Investigadores en Ciencias de la Computación. WICC. <https://drive.google.com/file/d/1vGiNX9ogumSBjnAnH-4d6N5hn5omJPM4/view> (2020).
21. Barber, E., Pisano, S., Romagnoli, S., de Pedro, G., Gregui, C., Blanco, N., Mostaccio, M.: Metodologías para el diseño de ontologías Web. Información, Cultura Y Sociedad, 0(39), 13-36. <http://ppct.caicyt.gov.ar/index.php/ics/article/view/14453/45454575759139> (2018).
22. Fernández Hernández, A.: Modelo ontológico de recuperación de información para la toma de decisiones en gestión de proyectos. <http://hdl.handle.net/10481/43409> (2016).
23. PMI.: Guía de los Fundamentos para la dirección de Proyectos (Guía del PMBOX®). Sexta Edición. Pennsylvania, EEUU.: PEARSON. (2017).
24. Contreras, M., Gutiérrez, F., Ortiz, J., Ramírez, R.: Ingeniería Ontológica Aplicada en el Diseño de un Sistema de Ontologías para la Gestión de Horarios. Pistas Educativas, 39(128). <http://www.itcelaya.edu.mx/ojs/index.php/pistas/article/viewFile/1173/949> (2018).
25. Arias, I. M.: Diseño y validación de un cuestionario de escala formativa para valorar las competencias transversales de los estudiantes universitarios. Una propuesta para dispositivos móviles basada en Android. Universitas Tarraconensis. Revista de Ciències de l'Educació, 1(1), 84-87. <https://revistes.urv.cat/index.php/ute/article/view/661> (2015).
26. Marulanda, Carlos E, Giraldo, Jaime A, López, Marcelo.: Evaluación de la Gestión del Conocimiento en las Organizaciones de la Red de Tecnologías de Información y Comunicaciones del Eje Cafetero en Colombia. Información tecnológica, 24(4), 105-116. <https://dx.doi.org/10.4067/S0718-07642013000400012> (2013).
27. Romero Rodríguez, J. M.: Representación del conocimiento experto de un tutor e-learning a través del mapa conceptual: Un modelo de buenas prácticas docentes. <http://digibug.ugr.es/handle/10481/46413> (2017).

## ***ScrumGame*: A Serious Game to Initiate Software Trainees in Scrum**

Guillermo Rodriguez<sup>1,3</sup>, Matías Glessi<sup>2</sup>, Alfredo Teyseyre<sup>2,3</sup>, Pablo Gonzalez<sup>1</sup>

<sup>1</sup> UADE-INTEC Business School, Buenos Aires, Argentina

<sup>2</sup>Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Argentina

<sup>3</sup>ISISTAN (UNICEN-CONICET) Research Institute, Tandil, Argentina

guirodriguez@uade.edu.ar, matiasglessi@gmail.com,

alfredo.teyseyre@isistan.unicen.edu.ar, pablogonzalez@uade.edu.ar

**Abstract.** Serious games have emerged to improve the commitment and performance of users, since they accomplish a specific goal, integrating with the mechanics of the game, thus creating a very attractive mission. The use of serious games in Software Engineering to increase developer engagement has been investigated with great interest to train future professionals to experience situations that they might face in software development. This work presents *ScrumGame*, a serious game to train both Software Engineering students and software practitioners in Scrum. The game was assessed through pre-test-post-test design with 10 users who use Scrum in their daily work. For this, the SIMS and MSLQ tests were used, which both were completed by the users before and after playing the game. We aimed to measure how the use of the game impacts on learning strategies and motivation. Backed up with statistical significance support, results show that *ScrumGame* generated a positive impact on the users, indicating an increase in the variables studied, key issues in the development of self-managed education.

**Keywords:** Scrum; Serious Game; Software Engineering Education; m-learning; SIMS; MSLQ

### **1 Introduction**

In recent years, the use of agile methodologies such as Scrum has grown significantly. According to a 2019 survey conducted by Forbes Insights and Scrum Alliance [1], 81% of executives consider agility as the most important characteristic of a successful organization, due to its ability to respond to changing priorities and the visibility of your projects. Furthermore, 76% of surveyed companies use Scrum or hybrid approaches involving Scrum [2]. These results provide several insights about Agile are nowadays the de-facto methodologies in the software industry. This issue implies a clear need for workers prepared to adapt to this development methodology.

Along this line, universities and educational entities are including the teaching of agile methodology within their study plans [13], for which they employ different

learning strategies. Thus, universities need to effectively provide students with the skills needed to succeed in current software organizations. In particular, experience-based learning techniques, such as games, have been widely used. Among these, the use of a LEGO-based simulation game to teaching Scrum [3]; Scrum-X, a spreadsheet-based simulation game for teaching Scrum [4]; Scrumi, an electronic board serious game for teaching concepts inherent to the SCRUM framework [5]; and a gamified system oriented to mobile platforms that allows applying the Scrum process from the Scrum Master's viewpoint [6], among others.

In this work, we have built *ScrumGame*, a serious game developed for the iOS platform. The game aims to train both software engineering students and software practitioners in Scrum by exploiting gamification design and ubiquity. Learning with *ScrumGame* could be fun and addictive; user can earn points for correct answers, for racing against the clock and for leveling up. To validate the game, user tests were carried out with 10 employees of a software development company focused on the development of mobile games that use Scrum in their daily work. We aimed to analyze psychological variables, such as motivation and learning-oriented strategies. For this analysis, two validated questionnaires were used: the SIMS (Situational Motivation Scale) [7] and the MSLQ (Motivated Strategies for Learning Questionnaire) [8], both carried out before and after using the application, with the intention of observing how the use of the game impacts on the variables under study. In addition, we aimed to obtain feedback on important aspects related to games, such as engagement, level of enjoyment and possible future improvements, through interviews. As main results, it was obtained that most of the variables under study in both tests showed positive changes. These insights showed that the use of our game generated a positive impact on the users, both in the learning strategies as motivation, which are crucial issues in the development of self-managed education.

The remainder of this paper is structured as follows. Section 2 describes the background. Section 3 introduces our serious game approach. Section 4 reports the assessment of the approach. Finally, section 5 concludes our work and identifies future line of research.

## 2 Background

Scrum is a framework by which people can tackle complex adaptive problems, where requirements are changing or poorly defined, while delivering incremental value products. According to the last annual survey carried out by the State of Agile in 2020 [2], Scrum is the most used agile methodology with 66%, which, considering also the hybrid combination Scrum and other frameworks, give a total 76%. This methodology consists of teams and their associated roles, events, artifacts, and rules. Each component within the framework serves a specific purpose and is essential to the success of a software application. Scrum ensures that knowledge comes from experience and from making decisions based on what is known. From there, its iterative and incremental approach optimizes predictability and risk control.

The use of games and mobile platforms has boomed as a constant source of learning. In particular, improving the teaching of Scrum to both students and

practitioners has been the subject of research to replace or complement traditional teaching methods. In this context, m-learning learning strategies will allow users to learn Scrum anywhere, anytime.

Gamification is a term that refers to the use of techniques, elements and dynamics typical of games in non-recreational systems in order to enhance motivation, to reinforce behavior to solve a problem, and to improve productivity [9]. Furthermore, gamification is used to improve user experience and the level of commitment and participation of users [10]. More specifically, according to [11] the importance of its application has several reasons, such as activation of the motivation for learning, constant feedback, facilitation of more meaningful learning, commitment to learning, more measurable results such as levels, points and medals, generation of adequate skills, digital alphabetization, generation of autonomous learners, and development of competitiveness as well as collaboration, among others.

Serious games were conceived within the variety of games for a primary purpose beyond mere fun [14]. They aim at a wide variety of audiences, from primary and secondary school students to professionals and consumers. Serious games can be of any genre, use any game technology and be developed for any platform. They are made to provide a context of entertainment and self-empowerment with which to motivate, educate and train users. These games seek to improve learning; they aim to strike a balance between studio content and the gameplay and ability of the player to retain and apply that content in the real world. They are also widely used in the business world to improve employee skills.

### 3 Our Serious Game Approach

This section presents the approach used to develop the *ScrumGame* application. This game aims to support Scrum teaching and training both in academic fields and professional contexts. To deal with this issue, a mobile application was implemented, using gamification, serious games and game-based learning concepts. Fig. 1 depicts a snapshot of *ScrumGame* in action.

The structure of the game content will be presented as follows. *ScrumGame* has a series of levels, which will divide the information into the main topics. Each topic has a series of sub-levels, which divide the content of the level into sub-themes within the main theme. Each sub-level is divided, in turn, into two main parts. On the one hand, the theory shows the theoretical information of the topic within a sub-level. The theory is presented in pages. Each page, like a page of a book, shows the data that the user will need to face the game. On the other hand, a sub-level is divided into games, which are presented in different formats to enrich users' learning experience.

A level has certain information to provide the user with a context. The level highlights the title of the topic, the number of sub-levels necessary to complete to advance to the next one, and the percentage of completeness of the level. This value is obtained by analyzing the percentage of completeness of the contained sub-levels. In addition, the level contains a state, which shows its current situation, namely locked, started and finished. *Locked* means that the level cannot be started because the previous level was not completed. *Started* means that the level is started or ready for

the user to start it. Finally, *finished* means that the level was successfully completed by the user.



**Fig. 1** Snapshot of *Scrum Game* in action.

Each sub-level also has a title, a status and a percentage of completion. Each sub-level is made up of theoretical information on the sub-topic discussed, which is presented to the user in the first instance. Furthermore, a sub-level also contains a series of games where the user can put into practice the concepts learned in the theoretical section, as well as give rise to possible more interpretative theoretical derivations. The completeness of the theoretical concepts and the number of successfully completed games of a sub-level will be used to compute the percentage of completeness that will be presented to the user.

Each sub-level has a theoretical section, where the user can find out about the concepts required to complete that level. For this, the information is presented in pages, which will be made up of different components, as required. A page consists of a title, paragraph, image and video.

The theory module has a control to advance or go back the pages of the game, showing different options depending on the page where the user is. This movement can also be achieved by performing a lateral slide, also known as swipe left/right. In addition, the concept of “Jump” is introduced, which allows users to go directly to the Games module, if they prefer. This definition stems from the need to avoid being pushed to revisit concepts by experienced users. The theoretical section will be considered complete if a user presses the play button or the skip button at the end of the section. The sub-level games section consists of a set of games prepared to give a user a way to validate the concepts learned in the theoretical section, presented in an attractive and motivating way.

Two main types of games are defined in *ScrumGame*: static and dynamic. The former is a non-modifiable game. Its implementation was thought from the beginning to meet a game objective and a particular theme; it belongs to a specific level. The latter is a game that can be reused, since the value of it is in its content. In this case, this is a game where the gameplay module will obtain the necessary information to

display it and the correct values to complete it. Fig 2 illustrates 2 games developed in *ScrumGame*. Some games are based on multiple-choice (Fig. 2a) or drag-and-drop (Fig. 2b) strategies, among others. All the games aim to assess Scrum concepts according to the knowledge level gained by the user along the game. Particularly, games depicted in Fig. 2 propose to train users in the life-cycle of Scrum and the artifacts and events used along the methodology. The goal in the game (b) is to drag the components (Scrum artifacts or event) and drop them in the correct place of the Scrum framework.



**Fig. 2** Example game strategies developed in *ScrumGame*: (a) Multiple-choice (b) Drag-and-drop.

Once the user submits a correct response, the game notifies the main controller so that it can stop the game timer. Meanwhile, it shows the user a message that the game was completed successfully and waits for their interaction. When the user agrees to continue, the main controller is notified to prepare for the next game, and also sends the “correct answer” and “total game time” events to the server. Finally, the main controller removes controller for the recently completed game, leaving room for the next game. It should also be noted that when completing each game within the sublevel, the main controller sends the progress data for storage. For its local save, *ScrumGame* uses Realm (<https://realm.io/>), while for its persistence in the cloud, the game uses Firebase (<https://firebase.google.com/?hl=en>).

## 4 Experimental Results

The goal of the evaluation is to measure the impact of *ScrumGame* on learning experience and motivation of users. To meet our objective, we used a pre-test-post-test design. To do this, two different tests were utilized: the SIMS (Situational Motivation Scale) [7] and the MSLQ (Motivated Strategies for Learning Questionnaire) [8]. Before playing, the users participating in the evaluation completed the tests and, after experiencing a given scenario with *ScrumGame*, the users completed the same tests again.

### 4.1 Setup

For the development of the experiment, 10 users took part in the experience, of whom 7 were men and 3 women. All participants are employees of a software development company focused on the development of mobile games. They are in the age range of 25-40 years old and the distribution of roles is described as follows: Project Manager (1), UX/ UI Designers (2), Technical Leader (T1), Developers (2), Product Owner (2), DevOps Leader (1). Particularly, in this company, knowledge of Scrum is not a mandatory requirement, since each team is organized with the methodology that best suits their projects and members. This allowed the participants to have different initial background of Scrum when conducting the experiment.

### 4.2 Pre-test and Post-test

Two of the important psychological variables to consider in self-managed education are motivation and the learning strategies. In this work, 2 instruments were used: the MSLQ and SIMS tests. The former is a self-report instrument that has been used in research to assess students' motivation and beliefs, among other aspects. This information is useful for evaluating strengths and difficulties of a student and, according to these, optimizing learning. The latter is a situational questionnaire, used to measure motivation. It seeks to judge the constructions of intrinsic motivation, identified regulation, external regulation and amotivation.

To answer the questionnaires, a Likert scale was used, which is the most widely used psychometric scale in research surveys, mainly in the social sciences. When answering a question from a questionnaire developed with the Likert scale, the level of agreement or disagreement with a statement is specified [12]. In our experiment, participants were asked to agree with a statement giving a value from 1 (does not correspond at all) up to 7 (totally corresponds) for SIMS.

In our case, the Motivation and Learning Strategies Questionnaire Short Form – CMEA FC [7] was used, which is the result of the translation and adaptation into Spanish of the MSLQ, in its reduced and validated version. The short form consists of 40 items with a Likert scale of 5 points, which goes from 1 (Never) up to 5 (Always). The short form is structured as follows with the corresponding dimensions (i.e. items): Motivation variable is determined by three items that evaluate the assessment of the task and four items that assess the anxiety under examination situations.

Regarding the Learning Strategies variable, the cognitive and metacognitive strategies, resource management strategies and intrinsic orientation were organized in the following way: five items to evaluate the elaboration strategies, four items to consider organization strategies, three items to assess critical thinking, and seven items to take into account self-regulation to metacognition. The Resource Management Strategies are defined as follows: six items to evaluate study time and habits, six items to talk about self-regulation of effort. And finally two items evaluate the goals of intrinsic orientation.

The SIMS test seeks to judge the constructions of intrinsic motivation, identified regulation, external regulation and amotivation, in a short and versatile way, both in field and laboratory studies [8]. It corresponds to 16 statements that the user must assess, according to the Likert scale, as detailed in the following section. These 16 statements are structured as follows along with the corresponding dimensions: Intrinsic motivation (4 items), Identified regulation (4 items), External regulation (4 items) and Amotivation (4 items).

### 4.3 Psychological variables under study

For our analysis, six psychological variables were taken into account during the experiment: two variables from MSLQ (Learning Strategies and Motivation) and four variables from SIMS (Intrinsic Motivation, External Regulation, Identified Regulation, and Amotivation).

- Learning Strategies: it talks about the imposition of own goals, distribution of time and effort, repetition, organization, critical thinking, and other cognitive processes that determine the acquisition of information, its processing and recovery. This variable is ranged from 33 to 165.
- Motivation: this variable consists of all the internal determinants that stimulate action. Motivation activates, directs and maintains a behavior. In this factor, micro variables such as the assessment of the task to be carried out and the management of anxiety are evaluated. This variable ranges from 7 to 35.
- Intrinsic Motivation: it is a motivation originated within the individual, and is directed by the needs of exploration, experimentation, curiosity and manipulation, which are considered motivating behaviors in themselves. That is, those behaviors that are carried out in the absence of any apparent external contingency are considered intrinsically motivated. The range of this variable is [4...28].
- External regulation: is a prototype of motivation not generated by one. Externally regulated behaviors are carried out to obtain a reward or satisfy external demand. The range of this variable is [4...28].
- Identified Regulation: represents an extrinsic motivation that is internalized and autonomous. They are behaviors regulated by personally important aspects. The range of this variable is [4...28].
- Amotivation: lack of motivation causes individuals to experience a lack of contingency between their behaviors and their results. Unmotivated behaviors are the least self-determined because there is no sense of purpose

and there are no expectations of reward or possibility of changing the course of events. The range of this variable is [4...28].

#### 4.4 Hypothesis Statement

For all variables under study, we defined null hypothesis and alternative hypothesis as follows. Null hypothesis ( $H_0$ ) states that there is no significant difference between the pre-test and post-test values for the variable under study ( $p\text{-value} \geq 0.05$ ). The alternative hypothesis ( $H_1$ ) states that there is a significant difference between the pre-test and post-test values for the variable under study ( $p\text{-value} < 0.05$ ). The hypotheses were corroborated by utilizing the two-tailed Student t-Test with SPSS tool.

#### 4.5 Results

This section reports the results obtained after conducting both tests and analyzing the means for each pair of psychology variables (before-after). All the null hypotheses were rejected and consequently the alternative hypotheses accepted ( $p\text{-value} < 0.0001$ ).



**Fig. 3** Comparison of means of External Regulation and Amotivation in pre- and post-test.

Fig. 3 depicts that External Regulation showed a slow decay (from 17.6 to 16.4). In other words, all those motivational behaviors driven by strictly external stimuli decreased in the post-test, after using *ScrumGame*. Regarding the Amotivation, Fig. 3 shows decay after the post-test (from 12.4 to 7.5). Positively, the lack of motivation fell in the post-test. This variation could mean that the users found a sense of purpose within the game. As can be seen in Fig. 4, Intrinsic Motivation has risen after using the game (from 20.1 to 23.8). Similarly, the Identified Regulation, understood as those external motivations that users understand and accept as their own, increased slightly in the post-test (from 23.8 to 22.8). Fig. 5 illustrates that Motivation values slightly decreased after using *ScrumGame* (from 28.0 to 27.7). In this case, the variables related to the assessment of the task and anxiety management decreased in the post-test. This unfavorable result could be attributed to the lack of psychological

preparation in games and in the application in general, to improve key concepts that affect this variable, such as competition and anxiety management.



**Fig. 4** Comparison of means of Intrinsic Motivation and Identified Regulation in pre- and post-test



**Fig. 5** Comparison of means of Motivation in pre- and post-test.

As for Learning Strategies, Fig. 6 shows that this variable has increased in the post-test (from 163.4 to 169.2). The micro variables related to this aspect, such as the imposition of own goals, distribution of time and effort, among others, improved after playing *ScrumGame*.

## 5 Conclusions

In this work, we presented a serious game for teaching Scrum. After the evaluation carried out through a pre-test-post-test design experiment with 10 users, we concluded that most of the psychological variables under study showed positive changes after the post-test. We can affirm that the use of *ScrumGame* significantly showed a positive impact on users, both in learning-related strategies and motivation.



**Fig. 6** Comparison of means of Learning Strategies in pre- and post-test.

As future work, firstly, we propose to generate new games with a psychopedagogical perspective, in order to optimize the learning experience and obtain better results. Secondly, we plan to incorporate more games to simulate real complex situations, serving as practice for specific scenarios in professional contexts. Thirdly, by using the users' feedback, we will work on improving techniques and components related to gamification issues, in order to generate higher commitment, fun and motivation from users. Finally, we aim to apply data science to the data generated along the game sessions to discover knowledge and assist users in their learning experience.

## References

1. The Elusive Agile Enterprise, <https://www.scrumalliance.org/forbes/the-report>
2. 14th Annual State of Agile Survey, State of Agile, <https://explore.digital.ai/state-of-agile/14th-annual-state-of-agile-report>
3. Paasivaara, M., Heikkilä, V., Lassenius, C., & Toivola, T.: Teaching students scrum using LEGO blocks. In Companion Proceedings of the 36th International Conference on Software Engineering, 382-391 (2014)
4. Lee, W. L. SCRUM-X: An interactive and experiential learning platform for teaching scrum (2016)
5. De Souza, A. D., Seabra, R. D., Ribeiro, J. M., & Rodrigues, L. E. D.: SCRUMI: a board serious virtual game for teaching the SCRUM framework. International Conference on Software Engineering Companion (ICSE-C), 319-321 (2017)
6. Angarita, L. B., & Hernández, J. A. G.: Sistema gamificado para el aprendizaje del proceso de desarrollo Scrum. Iberian Conference on Information Systems and Technologies (CISTI), Coimbra, Portugal (2017)
7. Sabogal Tinoco, L. F., Barraza Heras, E., Hernandez Castellar, A., Zapata, L.: Validación del Cuestionario de Motivación y Estrategias de Aprendizaje Forma Corta -MSLQ SF, en estudiantes universitarios de una institución Pública-Santa Marta. Psicogente, 14(25), 36-50 (2011)
8. Martín-Albo, J., Núñez, J., & Navarro, J.: Validation of the Spanish Version of the Situational Motivation Scale (EMSÍ) in the Educational Context. The Spanish Journal of Psychology, 12(2), 799-807 (2009)

9. Zichermann, G., Cunningham, C.: *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. O'Reilly Media (2011)
10. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., Dixon, D.: Gamification. Using game-design elements in non-gaming contexts. *Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 2425–2428 (2011)
11. Borrás Gené, O.: *Fundamentos de la gamificación*. Madrid: GATE-Universidad Politécnica de Madrid (2015)
12. Burns, A., Burns, R.: *Basic Marketing Research (Second ed.)*. New Jersey: Pearson Education. pp. 245 (2008)
13. Rodríguez, G., Soria, Á., & Campo, M.: Measuring the impact of agile coaching on students' performance. *IEEE Transactions on Education*, 59(3), 202-209 (2016)
14. Alvarez, J., & Djaouti, D.: An introduction to Serious game Definitions and concepts. *Serious Games & Simulation for Risks Management*, 11(1), 11-15 (2011)

# Revisión sistemática de la literatura sobre implementación de arquitecturas software para sistemas críticos

Joaquín Acevedo<sup>1</sup>, Andrea Lezcano, Emanuel Irrazábal<sup>1</sup>

<sup>1</sup> Grupo de Investigación en Innovación de Software y Sistemas Computacionales, FaCENA – UNNE, Corrientes, Argentina  
{jacevedo, alezcano, eirrazabal}@exa.unne.edu.ar

**Abstract.** Contexto: los sistemas críticos presentan funcionalidades específicas y un conjunto de buenas prácticas normativas que buscan asegurar niveles de seguridad mínimos en cada etapa en su ciclo de vida. Esto define características en el software presente en estos sistemas que requieren una integración conjunta con el hardware, particularidad presente en los sistemas embebidos. Por ello, es posible encontrar técnicas constructivas que pueden cumplir los niveles de seguridad requeridos, pero utilizando diferentes estrategias y recursos. Objetivo: realizar 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. Método: se utilizó una revisión sistemática de la literatura para identificar estudios publicados desde enero de 1999 a diciembre de 2019 sobre arquitecturas software para sistemas críticos. Resultados: se lograron identificar 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.

**Keywords:** ISO 61508, Arquitectura, Software, Estudio Secundario.

## 1 Introducción

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 [1]. 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. Además, la norma IEC 61508 detalla el concepto de nivel de integridad de seguridad o SIL por sus siglas en inglés. El SIL ofrece una escala contra la cual medir y cuantificar el nivel de seguridad de un sistema, desde el SIL 1, el nivel más bajo posible hasta el valor SIL 4.

A partir de la estandarización de las medidas de seguridad funcional en el ámbito de los sistemas críticos con la publicación de la norma IEC 61508 [1] en 1998, el tratamiento apropiado de las características del software encargado de controlar procesos industriales es una problemática que fue creciendo con la modernización y diversificación de las aplicaciones industriales. Así, por ejemplo, la industria automotriz y su normativa ISO 26262 [2] establece los niveles de integridad de seguridad automotriz (llamado ASIL por sus siglas en inglés) o la industria aviónica y su normativa DO-178B [3] que establece los niveles de seguridad de diseño (o DAL por sus siglas en inglés).

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. Desde el punto de vista de esta norma la arquitectura software consiste la definición de los subsistemas o módulos junto con sus interconexiones y, especialmente, la manera en la cual el nivel SIL es logrado. También se define el comportamiento general del software, sus interfaces y las decisiones que sostendrán las técnicas detalladas de diseño de componentes.

En este sentido, las elecciones de los tipos de arquitectura software que cumplen con las buenas prácticas antes mencionadas pueden ser complejas. 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 todo ello, en este artículo se ha llevado adelante una Revisión Sistemática de la Literatura (RSL) para identificar las arquitecturas software habituales en el desarrollo de software para sistemas críticos que han demostrado ser validadas para determinados niveles de seguridad. Asimismo, se han analizado los dominios de aplicación y las tecnologías relacionadas.

El trabajo se encuentra organizado de la siguiente manera además de esta introducción. La sección 2 describe los estudios secundarios relacionados con la temática. En la sección 3 se detalla la metodología empleada para el estudio correspondiente a la fase de planificación de una RSL y se presentan las preguntas de investigación en la tabla 1. En la sección 4 se reportan las actividades correspondientes a la etapa de conducción de la RSL. En la sección 5 se proveen los resultados obtenidos; para su presentación se sintetizaron los resultados de 23 estudios y se responden las preguntas de investigación. Finalmente, en la sección 6 se incluye la conclusión de la RSL.

## 2 Trabajos Relacionados

Existen trabajos relacionados con el objetivo de esta revisión y que proveen un estado del arte acerca de diferentes cuestiones relacionadas a las arquitecturas para sistemas críticos. En [4] los autores presentan un mapeo sistemático de la literatura enfocado en las pruebas basado en modelos para seguridad del software, en este estudio se incluye un análisis sobre diferentes publicaciones que presentan un desarrollo software sobre un dominio determinado de la seguridad en software. Sin embargo, no se incluyen preguntas relacionadas con los niveles de seguridad. En [5] se presenta el estado del arte de la combinación de técnicas de ingeniería conducida por modelos e ingeniería de línea de productos para el desarrollo de arquitecturas de software para sistemas críticos. En este caso las preguntas están enfocadas a identificar estudios que traten sistemas embebidos desde la dimensión de la ingeniería dirigida por modelos. Esta aproximación compromete la validez interna de las técnicas evidenciadas y no se tratan en profundidad los requerimientos que aseguran la criticidad de los sistemas resultantes de las técnicas registradas en el estudio secundario.

Asimismo, en [6] se presenta una caracterización de diferentes técnicas empleadas para representar arquitecturas software para sistemas embebidos. Se discute sobre los riesgos asociados de cumplir para arquitecturas de sistemas críticos sin incluir un análisis directo del nivel de seguridad. Por todo ello, esta revisión difiere de las revisiones mencionadas de la siguiente manera:

- Se identifican las normas y estándares relacionados a los sistemas críticos.
- Se incluyen estudios publicados desde 1999 a 2019.
- Se caracteriza la relación entre las arquitecturas software y el nivel de seguridad que presentan al considerar su aplicación en un sistema crítico.
- Se describe la tecnología empleada para el desarrollo de arquitecturas.
- Se provee un análisis de cada estudio incluido en términos de rigor, validez y aplicabilidad.

## 3 Metodología

En esta sección se describe el método empleado para la realización del estudio secundario. En este caso se ha seguido el enfoque identificado por las pautas para realizar Revisiones Sistemáticas de Literatura en Ingeniería de Software [7]. De acuerdo con lo descrito en la introducción de este trabajo, el objetivo de la presente RSL es identificar las arquitecturas software empleadas para aplicaciones críticas de sistemas embebidos, la evidencia del cumplimiento de las normativas internacionales y las tecnologías utilizadas para ello. El análisis se basa en preguntas de investigación que se encuentran en la tabla 1.

**Tabla 1.** Preguntas de investigación.

Preguntas de investigación	Motivación
RQ 1. ¿Qué arquitecturas software existen para software crítico?	Identificar las diferentes arquitecturas software en el campo de aplicación de los sistemas críticos y la frecuencia de su uso y reporte en estudios académicos.
RQ 2. ¿Qué normativa cumplen?	Identificar las normativas que se presentaron como requerimientos a cumplir por las arquitecturas presentadas y el grado de cumplimiento de la misma.
RQ 3. ¿Qué nivel de seguridad verifican?	Determinar la evaluación realizada en cuanto al cumplimiento del nivel de seguridad que la arquitectura software presenta.
RQ 4.1 ¿Qué tecnologías utilizan?	Determinar las tecnologías empleadas para el desarrollo de las arquitecturas software y su capacidad para lograr el cumplimiento de los requisitos incluidos en las normativas.
RQ 4.2 ¿Cuáles son las plataformas utilizadas?	Identificar las plataformas sobre las cuales se desarrollan arquitecturas software, si se utilizan herramienta propietarias o de fuente abierta, propias o de terceros.
RQ 5 ¿Cuáles son los dominios de aplicación?	Determinar la evaluación realizada en cuanto al cumplimiento del nivel de seguridad que la arquitectura software presenta.
RQ 6 ¿Cuáles son las plataformas utilizadas?	Determinar las tecnologías empleadas para el desarrollo de las arquitecturas software y su capacidad para lograr el cumplimiento de los requisitos incluidos en las normativas.

### 3.2 Estrategia de búsqueda

En esta sección se describe la estrategia de búsqueda explicando el alcance en cuanto a fuentes, el método empleado y la cadena de búsqueda.

Se decidió realizar la búsqueda en bases de datos que contengan artículos relacionados a ciencias de la computación e ingeniería. Se emplearon las bases de datos generalmente utilizadas en otras revisiones de la misma temática: Scopus, IEEE Xplore, ACM Digital Library y Springer Link. El artículo se realizó en base a la revisión de publicaciones en revistas, conferencias y artículos. En la tabla 2 se presenta cada término junto con sus palabras claves<sup>1</sup>.

**Tabla 2.** Términos clave y sinónimos empleados para crear la cadena de búsqueda.

Términos	Palabras clave
safety	"safety*" OR "functional safety" OR critical OR "mission-critical" OR "fault tolerance" OR "fault tolerant" OR "fault-tolerance" OR "fault-tolerant" OR sil OR ssil OR rams
architecture	architect* OR "architectural pattern" OR "architectural design" OR design
software	software OR application OR "source code" OR firmware
61508	61508 OR 61508-3 OR 50128 OR MISRA OR "MISRA C" OR 62279 OR do-178* OR 26262

### 3.3 Criterio de selección de los estudios

Se incluyen los artículos referidos al desarrollo de arquitecturas software para aplicaciones en sistemas críticos, que presenten un nivel de seguridad y publicados a partir de 1999 en revistas indexadas y en conferencias. Se excluyen los artículos de descripción o uso de herramientas, duplicados y presentaciones de conferencias. En cuanto al contenido, se excluyen los estudios que describan, implementen o evalúen lenguajes de descripción de requerimientos.

Para seleccionar los estudios primarios se aplican los criterios de inclusión/exclusión leyendo los resúmenes de los artículos encontrados. En caso de que se registren dudas acerca de su pertinencia se leerá el artículo completo. El proceso de revisión será acompañado por el criterio de los investigadores para consensuar sobre la correcta aplicación de los criterios de inclusión/exclusión<sup>2</sup>.

### 3.4 Estrategia para la extracción y síntesis de datos

El procedimiento empleado para la extracción de datos consiste en un primer ordenamiento de todas las publicaciones resultantes de la consulta a la base de datos de investigación en una planilla respetando los identificadores de exportación de la fuente sumado a un campo de control para indicar el estado de inclusión o exclusión del artículo. Cada base de datos tiene sus campos propios para indexar publicaciones, por consiguiente, luego de llegar al conjunto de todos los estudios a incluir en el estudio, se procedió a una normalización de todos sus identificadores.

Para realizar la síntesis de los datos se organizó cada publicación aceptada en un directorio de acuerdo a la fuente de acceso. Luego se procedió a la creación de etiquetas para identificar cada dimensión de análisis correspondiente a los criterios sobre las preguntas de investigación. El resultado de este procedimiento fue presentar la información de cada publicación que signifiquen descripciones concisas relacionadas a las preguntas de investigación.

## 4 Realización de la Revisión

La RSL fue realizada siguiendo todos los pasos del protocolo definido en la sección previa y completada en un año, en este periodo se incluye el tiempo requerido para cada una de las tres fases, es decir, planeamiento, realización y reporte. Inicialmente fueron encontradas 7550 publicaciones. Cada fase de búsqueda se detalla en la tabla 3 y la tabla 4 con mayor precisión.

<sup>1</sup> Anexo: bit.ly/rslcacic2020ASC, en la tabla A1 se describen las cadenas de búsqueda para cada base de datos.

<sup>2</sup> Anexo: bit.ly/rslcacic2020ASC, en la tabla A2 se describen los criterios de inclusión y exclusión.

#### 4.1 Selección de los estudios primarios

En esta etapa de selección de los estudios primarios se encuentran los criterios de aseguramiento, aplicación de los criterios de aseguramiento y extracción de los datos de los artículos. Se encontraron en total de 7550 artículos aplicando la estrategia de búsqueda definida el protocolo. La búsqueda se efectuó empleando título, resumen y palabras claves indexadas (ver tabla 3).

**Tabla 3.** Detalles de la búsqueda automatizada.

Base de datos	Resultados de búsqueda
IEEE Xplore	295
SCOPUS	696
ACM Digital Library	2000
Springer Link	4559
TOTAL	7550

De los primeros 7550 resultados se realizó el filtrado empleando la lista de comprobación establecida en el protocolo, se realizaron revisiones reiteradas hasta llegar a un número final de 23 artículos con fechas de publicación que varían desde del 2004 al 2017 (ver tabla 4). Además del criterio de inclusión/exclusión, se considera también el aseguramiento de la calidad de los estudios primarios para proveer un criterio más detallado y con el sentido de evaluar la importancia de los estudios individuales cuando los resultados están siendo sintetizados. En la presente sección se trata la calidad en términos de minimizar el sesgo y maximizar la validez interna y externa como se indica en el manual Cochrane para revisiones sistemáticas de intervenciones [8] desde las prácticas adoptadas por las buenas prácticas para RSL [7]. Para esta etapa se emplearon una serie de preguntas orientadas a evaluar el modo en que los estudios primarios presentan la información relacionada al objetivo, así como también el rigor de los tópicos desarrollados<sup>3</sup>.

A continuación, se extrajeron los datos cuantitativos y cualitativos de cada uno de los 23 artículos para contar con información válida y objetiva. Se organizó cada estudio en una disposición para enfocar los contenidos de cada publicación, bajo los fines de la investigación. La importancia de esta actividad radica en la presentación apropiada del reporte de la revisión al momento de la realización de la discusión, presentada en la sección 5. Para la síntesis de los datos representativos de cada publicación se emplearon los métodos establecidos en la sección 3. La realización de esta actividad permitió caracterizar y evaluar cada publicación empleando en conjunto el estudio completo y la información extendida de la tabla de clasificación y el ordenamiento de las herramientas de apoyo. Al momento de recolectar la información relacionada a cada pregunta de investigación se registraron datos puntuales en cada campo del formulario de extracción.

**Tabla 4.** Detalles de cada fase de selección de estudios primarios.

Fase	Descripción	Incluidos	Excluidos
Fase 1	Resultados de búsqueda	7550	0
Fase 2	Selección por criterios de inclusión	2899	4651
Fase 3	Selección por criterios de exclusión	57	2842
Fase 4	Selección por título y abstract	26	31
Fase 5	Validación conjunta	23	3

## 5 Reporte de la Revisión

A continuación, se describen los resultados organizados por pregunta de investigación. En la discusión de los resultados se hace referencia a los artículos con la forma Px siendo “x” el número de artículo tal y como está descrito en la sección 2 de la documentación anexa a este trabajo<sup>4</sup>.

<sup>3</sup> Anexo: [bit.ly/rslcacic2020ASC](http://bit.ly/rslcacic2020ASC), en la sección 3 se describe el análisis de calidad de los estudios.

<sup>4</sup> Anexo: [bit.ly/rslcacic2020ASC](http://bit.ly/rslcacic2020ASC), sección 2.

### 5.1 RQ1. ¿Qué arquitecturas software existen para sistemas críticos

En la tabla 5 se resumen las arquitecturas software encontradas en el análisis de los artículos y el número total encontrado. El 39% de los trabajos bajo estudio aplican una implementación de arquitecturas de capas. Esta decisión de diseño de los sistemas bajo estudio requiere pruebas exhaustivas y una clara separación de funciones que pueden introducir más posibilidades de fallo; por tanto, al seleccionar una arquitectura en capas se asegura que cada función tenga acotada su probabilidad de fallo. En P3 se evidencia un caso de seguridad modular para una red de trabajo que requiera cumplir con la normativa IEC 61508 y describe un protocolo para transferir datos de forma segura cumpliendo con el nivel 3 SIL.

El 30% de las publicaciones emplean una arquitectura con un esquema de votación del tipo *M-out-of-N* y un 70% de este subconjunto emplean específicamente el esquema 1oo2. Esta distribución entre los estudios se justifica ya que ciertos dominios requieren que el sistema se mantenga funcional bajo condiciones extremas y cumplimentando el nivel de SIL requerido. En el trabajo P20 los autores presentan una solución que implementa hardware redundante y dado el caso de implementación de la solución sobre un entorno de alta presión alta temperatura, no recomiendan el uso de su esquema de transmisores de doble presión porque podrían agregar una falla en el sensor completo.

Cerca de un 26% de los trabajos mencionan y aplican un esquema software en conjunto con redundancia de hardware como medida para mitigación de riesgos. En P13 se menciona la redundancia de hardware como un método de monitoreo seguro bajo pruebas de comparación empleando un votador.

Asimismo, el 13% de las publicaciones emplean una arquitectura maestro esclavo como solución propuesta. Así, en P9 se recomienda un microcontrolador maestro y diferentes microcontroladores esclavos, con su tratamiento respectivo de fallas posibles.

Solo un 8% de los trabajos proponen una arquitectura que contemple redundancia de software y las mismas no cubren todo el procedimiento de certificación necesario para la implantación de un sistema crítico dirigido por software.

**Tabla 5.** Tipos de arquitectura software utilizadas.

Identificador	Arquitectura	Artículo que implementa la solución	Total
MooN	m-out-of-n	P5 P6 P8 P9 P13 P14 P15 P20	8
MS	Maestro esclavo	P9 P18 P19	3
RHW	Redundancia hardware	P7 P13 P14 P16 P20 P21	6
AC	Arquitectura en capas	P3 P10 P11 P16 P17 P19 P20 P21 P22	9
RSW	Redundancia software	P1 P2	2
FSM	Máquina de estados finitos	P12	1
SCP	Safety channel pattern	P23	1

### 5.2 RQ2. ¿Qué normativa intentan cumplir?

Cerca del 78% de los estudios tratan expresamente la normativa IEC 61508 mientras que el 22% restante indican normativas específicas del dominio y del ámbito de aplicación de la arquitectura del software en un sistema crítico refiriendo a las características y restricciones de diseño demandados: automotriz [P7] [P25] [P15] aviónica [P22] y normativas específicas de telecomunicación [P3].

Un resultado notable es el creciente desarrollo de las tecnologías de conducción automática tratada en las prácticas recomendadas para vehículos de superficie [9] aumentando la rigurosidad requerida para los componentes software y los circuitos programables, con tiempos de respuesta acotado. Se menciona, entonces, a la normativa ISO/IEC/IEEE 42010 que regula esta nueva área sin excluir a la norma ISO 26262 existente.

El trabajo P19 no hace referencia directa la norma IEC 61158 porque el ámbito del estudio en cuestión es la implementación de un protocolo de comunicación segura. Técnicamente la norma IEC 61784 define protocolos basados en la norma mencionada anteriormente e incluye otras definiciones extendidas.

En P21 se presentan resultados experimentales sobre la auditoría del sistema bajo una prueba de inyección de fallas para verificar la implementación de los conceptos de seguridad, además se indica la posibilidad de realizar pruebas automatizadas para validación del sistema en conformidad con el estándar ISO7637[10]. En la tabla 6 se resumen las normativas que se intentan cumplir y los artículos que la mencionan.

**Tabla 6.** Tipos de arquitectura software utilizadas.

Normativa	Artículo
IEC 61508	P1 P2 P3 P4 P5 P6 P7 P9 P10 P11 P13 P14 P15 P16 P18 P19 P20 P21
DO 178	P12
EN 14908	P5
EN 50126	P13
EN 50128	P13
EN 50129	P13
ISO 26262	P8 P9 P11 P17 P21 P22 P23
IEC 61131	P7
EN 954-1	P6
IEC 61800	P6
IEC 60204	P6
IEC 61511	P16
API RP 14C	P16
IEC 61784	P19 P22
IEC 61158	P22
ISO/IEC/IEEE 42010	P23
ISO/IEC 25010	P23

### 5.3 RQ3. ¿Qué nivel de seguridad verifican?

Un 60% de los trabajos incluidos en el estudio tienen como objetivo alcanzar el nivel SIL 3 o similar de acuerdo con la normativa específica. La distribución de publicaciones referidas a arquitecturas software para sistemas críticos que abordan este nivel de seguridad integral permite visualizar la presencia de soluciones software en este ámbito que presenta conformidad con las normas vigentes en cada dominio tratado.

Alrededor de un 39% de los trabajos revisados tienen como objetivo certificar hasta un nivel SIL4 o equivalente en ASIL y DAL. Finalmente, el 30% de las publicaciones tienen como objetivo alcanzar un nivel SIL2 o inferior, es importante aclarar que el nivel de seguridad requerido varía de acuerdo al dominio del sistema crítico y que la presencia de publicaciones que mencionan estos niveles de SIL1 y SIL2.

**Tabla 7.** Nivel SIL identificado en cada artículo.

SSIL	Artículo
SIL1 / ASIL A	P2 P15
SIL2	P2 P6 P10 P11 P16 P20
SIL3 / ASIL B / DAL B	P1 P2 P3 P4 P5 P6 P7 P9 P12 P14 P16 P18 P19 P22
SIL4 / ASIL D / DAL A	P1 P2 P4 P8 P12 P13 P17 P21 P23

### 5.4 RQ4.1 ¿Qué tecnologías utilizan?

El apartado de tecnología elegida como plataforma de implementación es la dimensión de análisis que presenta más diversidad de resultados. El total de las publicaciones emplea un esquema software de alto nivel propio para disminuir la complejidad de los esquemas de hardware específicos. El 30% de los estudios incluidos emplea un *framework* como tecnología para el desarrollo de software. El 48% de los estudios incluidos presenta un desarrollo de software empleando librerías. En menor proporción, cerca del 22% de los estudios incluidos solo describe el desarrollo de software a nivel de código fuente.

En algunos casos del dominio automotriz se describe el uso de un bus CAN [P8] [P11] [P17] [P23] que a nivel de tecnología software emplean librerías o un *framework* para una mayor integración de funciones. La implementación del bus es específica para cada caso tratado. En la tabla 8 se indican los artículos por tecnología.

**Tabla 8.** Tecnologías empleadas para arquitecturas software.

Tecnologías	Artículo	Total
Framework	P3 P10 P13 P14 P19 P21 P22	7
Librerías	P2 P4 P5 P6 P7 P8 P11 P15 P17 P20 P23	11
Código fuente	P1 P9 P12 P16 P18	5

### 5.5 RQ4.2 ¿Cuáles son las plataformas utilizadas?

La distribución de plataformas evidenciadas para implementar una arquitectura software se visualiza en la tabla 9. Cerca del 65% de los estudios emplea una solución propia, el 21% de los estudios describe el uso de un microcontrolador disponible en el mercado. Un resultado notable es el descrito en el artículo [P18] en el que se emplea una plataforma que cuenta con certificación IEC61508 SIL 3 e ISO26262 ASIL D. En menor proporción se encuentra el uso de arquitecturas software sobre controladores lógicos programables (PLC) significando un 13% del total de estudios incluidos. En los casos que contemplan el uso de controladores lógicos programables, su implementación fue acompañada por un entorno operativo empleando un sistema operativo de tiempo real para validar el comportamiento esperado del sistema bajo prueba. Y por último se encuentra la elección de un sistema operativo de tiempo real (RTOS) como plataforma, significando cerca del 8% de los estudios incluidos. En P6 se menciona el uso de un sistema operativo en tiempo real para su implementación en los sistemas de control.

**Tabla 9.** Plataformas empleadas en cada tipo de arquitectura software desarrollada.

Plataforma	Artículo	Total
Solución específica	P1 P2 P3 P4 P5 P8 P10 P11 P14 P15 P16 P17 P19 P22 P23	15
MCU	P9 P12 P13 P18 P21	5
PLC	P6 P7 P20	3
RTOS	P6 P7	2

### 5.6 RQ5. ¿Cuáles son los dominios de aplicación?

Para responder este apartado se toma en consideración el dominio al que pertenece el sistema crítico trabajado en la publicación. Un 43% de las publicaciones revisadas emplean una arquitectura software para sistema crítico bajo el dominio de la seguridad en equipamiento industrial.

Alrededor del 26% de los estudios describen sistemas que se emplean bajo el dominio de la industria automotriz, incluyendo las actividades de validación, análisis y desarrollo de la arquitectura software. Cerca del 8% de los estudios revisados presentan un sistema que se emplea bajo el dominio de trenes [P23] [P9]. En P20 se trata la implementación de sistemas críticos en el dominio de seguridad en instalaciones *offshore*. Los dominios de aplicación para sistemas críticos que incluyen una arquitectura software (ver tabla 10).

**Tabla 10.** Dominio del sistema crítico tratado en cada artículo.

Dominio	Artículo	Total
Seguridad en equipos industriales	P1 P2 P3 P4 P5 P6 P7 P14 P15 P16	10
Automotriz	P8 P9 P10 P11 P17 P21 P23	7
Aviónica	P12	1
Redes de comunicación segura	P19 P22	2
Ferrovionario	P13 P18	2
Seguridad en instalaciones <i>offshore</i>	P20	1

### 5.7 RQ6 ¿Cuáles son las actividades reportadas?

Se requieren actividades que describan diseño de la arquitectura incluyendo casos de seguridad, y actividades que aseguren la aplicación del sistema empleando componentes electrónicos certificados y los roles de prueba respectivos siguiendo los casos de seguridad desarrollados de forma particular. Las actividades cubiertas en cada publicación difieren porque cada nivel de integridad de seguridad tiene mayor rigurosidad en los requerimientos. El 50% de este conjunto de estudios mencionados describe la etapa de pruebas de sistema. Una de las características notables en este dominio es la evaluación integral de todas las restricciones del entorno en el cual se va a implementar el sistema crítico, este hecho requiere una correcta documentación del software y del hardware que se va a emplear.

Finalmente, P23 y P9 incluyen actividades de validación a nivel de pruebas de sistema, esto se relaciona con el nivel de SIL que se pretende lograr en estos dos artículos. La identificación por actividad y artículo se encuentra en la tabla 11.

Tabla 11. Actividades reportadas en el desarrollo de arquitecturas software para sistemas críticos.

Actividad	Artículo	Total
Análisis	Todos los artículos	23
Investigación	P4 P5 P7 P9 P10 P11 P17	7
Desarrollo	P1 P2 P3 P5 P6 P8 P9 P10 P12 P13 P14 P15 P16 P17 P18 P19 P20 P21 P23	19
Validación	P1 P2 P3 P4 P6 P7 P8 P9 P11 P12 P13 P14 P15 P16 P17 P18 P19 P20 P21 P22 P23	22
Implantación	P13 P16	2

### 5.8 Resultados adicionales

Como se muestra en la Fig. 1, la distribución de publicaciones que cumplen con los criterios de investigación propuesto presenta una fluctuación que alcanza sus máximos entre los años 2009 a 2012.

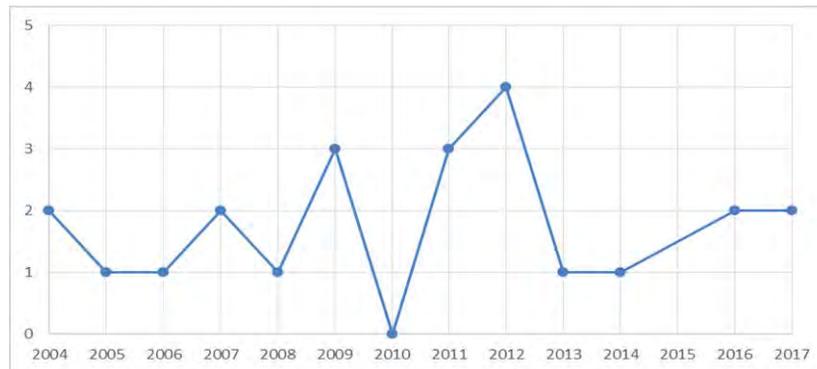


Fig. 1. Distribución de los estudios primarios por dominio.

Una vez extraídos los datos relevantes del conjunto de estudios seleccionados para conducir la revisión se llegó a un esquema que lo describe en función de los parámetros que se relacionan de forma directa con las preguntas de investigación, significando información con granularidad basada en la unicidad de cada publicación y en las características excluyentes y no excluyentes indicadas en la metodología. Este cambio en los esquemas presentados respetando el criterio de unicidad de las publicaciones y presentando vinculaciones ha sido estudiado formalmente en trabajos como Hashemi et al. [11]. Implementados sobre el conjunto de estudios se observa un agrupamiento importante de publicaciones que tratan una arquitectura en capas con SIL 2 y SIL 3. Y otra agrupación presentando una arquitectura tomando un esquema base de M-out-of-N agrupada en los niveles SIL 3 y SIL 4, esto es presentado en la Fig. 2. En cuanto a la proporción en la forma de presentar desarrollos de software relacionado a las arquitecturas incluidas teniendo presente la dimensión temporal, no se presentaron agrupaciones mayores a dos publicaciones y su distribución se visualiza en la Fig. 3.

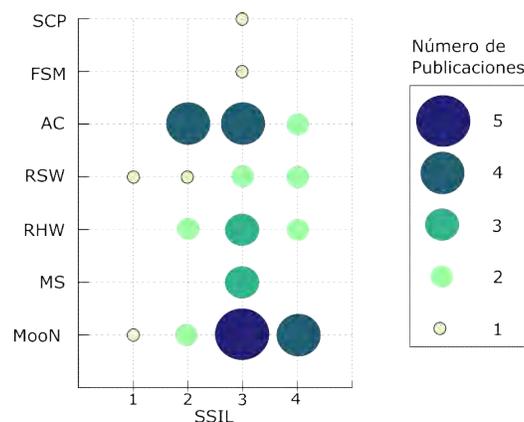


Fig. 2. Agrupación de estudios primarios por arquitectura y por SIL tratado.

### 5.9 Amenazas a la validez

A continuación, se indican las amenazas a la validez identificadas durante el desarrollo del trabajo. En cuanto a la búsqueda realizada: una revisión sistemática de la literatura empleando diversos motores de búsqueda presenta la necesidad de controlar los estudios que se encuentren indexados en varias fuentes y la selección de la versión de última revisión si esto ocurre. Este aspecto fue tratado al emplear un control por índice luego del resultado de la cadena de búsqueda al finalizar la fase 1 y hasta la fase 4 inclusive cuando se realizó la selección por título y resumen.

En cuanto a los estudios incluidos: se ha buscado asegurar la validez de los datos presentados en la presente revisión a partir de la calidad del conjunto de artículos incluidos. Según el enfoque presentado en las buenas prácticas para realizar revisiones los estudios de observación suelen ser más susceptibles al sesgo que los estudios experimentales; las conclusiones que se pueden extraer de ellos son necesariamente más tentativas y a menudo generan hipótesis, destacando áreas para futuras investigaciones [12]. Este enfoque fue adoptado para el criterio de aseguramiento de calidad por Kitchenham y Charters para minimizar el sesgo y maximizar la validez interna y externa. Empleando estas herramientas de aseguramiento de calidad para cada estudio se logró obtener una medida de calidad del conjunto de estudios para tratar este aspecto de validez.

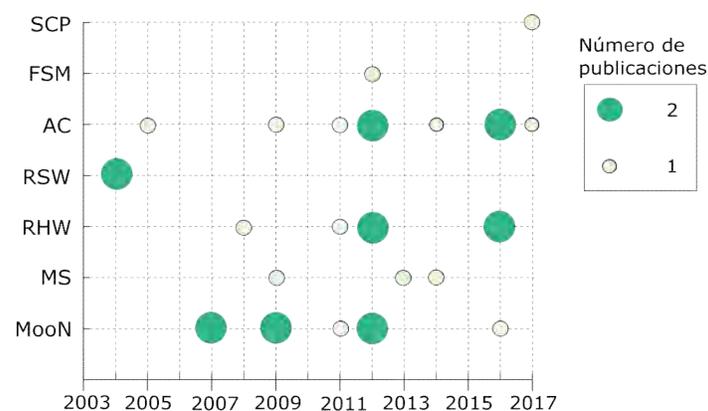


Fig. 3. Agrupación de estudios primarios por arquitectura y por año de publicación.

## 5 Conclusiones

En este trabajo se presentó una RSL enfocada en las arquitecturas software y su aplicación en sistemas críticos. Por lo tanto, se definieron los métodos de búsqueda y selección de los estudios primarios, con el propósito de obtener el estado del arte de los enfoques. Se definieron los criterios de investigación, se presentó el método utilizado para realizar la búsqueda, los criterios de selección de los estudios primarios y se realizó el reporte.

Los resultados indican que hay un amplio desarrollo de arquitecturas software en varios dominios industriales, se evidencia una mayoría dominio automotriz y el dominio de seguridad en equipamiento industrial. Se llegó a una distribución de estudio por año, plataforma, tecnología y arquitectura en el diseño. Se buscó además determinar el nivel de seguridad integral tratado en cada arquitectura, en este sentido, se presenta una distribución importante sobre niveles SIL 3 y SIL 4 en el conjunto de estudios analizados. También se encontraron referencias a normas que regulan cada proceso implementado como función de seguridad delegado al software.

Del presente estudio, podemos concluir que la implementación de arquitecturas software en sistemas críticos requiere una evaluación integral que incluye las tecnologías a emplear, las normas existentes que brindan un marco de trabajo sobre el cual debe adecuarse el ciclo de desarrollo, el nivel de seguridad que deben cumplimentar las funciones implementadas por el software, la tecnología y la plataforma de desarrollo elegidas que permitan acotar el índice de fallos, dentro de lo permitido por el nivel de seguridad requerido.

En términos generales, se evidenciaron arquitecturas software que pueden cumplir altos niveles de SIL, en este aspecto se encontraron evidencias de que predominan las arquitecturas software con un diseño en capas. En cuanto a la normativa que intentan cumplir los desarrollos presentes en los estudios, se evidenció que la utilización del estándar IEC 61508 se presenta en la mayoría de los estudios, pero su conformidad debe presentarse en conjunto con normas específicas del dominio.

Sobre los apartados de tecnología empleadas y plataformas de desarrollo, se identificó una tendencia por emplear soluciones específicas, si bien existen soluciones de hardware certificado que permite un mayor control

durante todo el ciclo de vida del proyecto, gran parte de los estudios emplean definiciones de hardware propias y definiciones de software a nivel de código fuente o librerías que hacen uso del soporte físico establecido, con *frameworks* destinados a la validación de las funciones implementadas por software.

En cuanto al apartado de actividades reportadas, no hay una definición única a seguir para el desarrollo de las arquitecturas, pero existen procesos claros y normalizados para establecer módulos, probar cada elemento software y controlar por simulación su interacción con el hardware, que, en muchos casos evidenciados, significa presentar el nivel de seguridad requerido.

Para resumir, los resultados de la RSL presentados en las secciones anteriores nos permiten afirmar que existe una fuerte tendencia de implementar funciones seguras mediante software, esta tendencia es creciente, y abarca varios dominios de la industria.

**Agradecimientos.** Este desarrollo se realizó a partir del proyecto de investigación PI 17F017 y 17F018 de la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste.

## Referencias

1. International Electro-technical Commission IEC, “IEC 61508: Functional safety of electrical/electronic/programmable electronic safety-related systems.” 1998.
2. International Organization for Standardization (ISO), “ISO26262 Road vehicles – Functional safety.” 2011.
3. Radio Technical Commission for Aeronautics (RTCA), “DO-178B, Software Considerations in Airborne Systems and Equipment Certification.” 1992.
4. H. G. Gurbuz and B. Tekinerdogan, “Model-based testing for software safety: a systematic mapping study,” *Softw. Qual. J.*, vol. 26, no. 4, pp. 1327–1372, 2018, doi: 10.1007/s11219-017-9386-2.
5. P. G. G. Queiroz and R. T. V Braga, “Development of Critical Embedded Systems Using Model-Driven and Product Lines Techniques: A Systematic Review,” in 2014 Eighth Brazilian Symposium on Software Components, Architectures and Reuse, 2014, pp. 74–83.
6. E. A. Antonio, F. C. Ferrari, and S. C. P. F. Fabbri, “A Systematic Mapping of Architectures for Embedded Software,” in 2012 Second Brazilian Conference on Critical Embedded Systems, 2012, pp. 18–23, doi: 10.1109/CBSEC.2012.22.
7. B. A. Kitchenham and S. M. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering | Request PDF[1] B. A. Kitchenham and S. M. Charters, ‘Guidelines for performing Systematic Literature Reviews in Software Engineering | Request PDF.’ [https://www.researchgate.net/publication/302924724\\_Guidelines\\_for\\_performing\\_Systematic\\_Literature\\_Reviews\\_in\\_Software\\_Engineering](https://www.researchgate.net/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering) (accessed May 23, 2020).
8. The Cochrane Collaboration, *Cochrane Handbook for Systematic Reviews of Interventions*. 2019.
9. SAE, “Surface vehicle recommended practice—Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles,” 2018.
10. ISO/DIS 7637-2, “SO7637-2 Road vehicles - Electrical disturbances from conduction and coupling.” International Organization for Standardization, 2011.
11. R. R. Hashemi, S. De Agostino, B. Westgeest, and J. R. Talburt, “Data granulation and formal concept analysis,” in Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS, 2004, vol. 1, pp. 79–83, doi: 10.1109/nafigs.2004.1336253.
12. Centre for Reviews and Dissemination, *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. York Publishing Services Ltd, 2009.

# A Parallel Tableau Algorithm for BIG DATA Verification

Fernando Asteasuain<sup>1,2</sup> and Luciana Rodriguez Caldeira<sup>2</sup>

<sup>1</sup> Universidad Nacional de Avellaneda, Argentina  
fasteasuain@undav.edu.ar

<sup>2</sup> Universidad Abierta Interamericana - Centro de Altos Estudios  
CAETI, Argentina  
luciana.rodriguezcaldeira@alumnos.uai.edu.ar

**Abstract.** BIG DATA systems are becoming more and more present in our everyday life generating data and information that needs to be explored and analyzed. In this sense, formal verification tools and techniques must provide solutions to face with these new challenges since they been pointed out as one of the most needed software engineering activities to consolidate BIG DATA modern systems. In this work we present a parallel implementation of a tableau algorithm aiming to improve the performance of our formal verification scheme. The pursued objective behind this transformation is to adapt our framework to deal with BIG DATA systems.

**Keywords:** BIG DATA, Formal Verification, Parallel Programming

## 1 Introduction

Modern systems reside in a world where everything is connected and information is generated, consumed and exchanged at a surprisingly increasing growth rate. This enormous amount of data and information needs to be explored and analyzed, activities that originated new disciplines such as BIG DATA [34, 16] or Data Science [32] whereas other important areas such as Artificial Intelligence turned their attention into this topic applying techniques such as Automated Reasoning [27], Machine Learning [2] or Neural Networks [22].

In the last years the Software Engineering community did also make an effort to leverage on this emerging topic. Several works like [24, 15, 26, 23, 30, 25] thoroughly present a complete state of the art stating how software engineering tools, methodologies and techniques are applied and adapted to deal with BIG DATA implications, covering all the very well known software developing phases, from requirements gathering to product release, including modeling, design, testing, validation and verification of Big Data Software Systems. In particular, all the mentioned approaches pinpoint the urgent need of a new arising of more research lines focusing on the formal verification of BIG DATA systems. This is a very challenging path to take since it involves dealing with rigorous performance

requirements in a context where data and information is highly informal and unstructured [25, 24, 20].

In this sense, approaches like [10] try to expand current formal verification tools such as model checking [17] to cope with new architectures for BIG DATA systems such as Cloud Computing [19] and distributed environments. Other approaches involves a classic migration from sequential to parallel model checking [11, 13, 7, 8, 14]. However, a prior step in the formal verification road has been somehow neglected, which is the way the behavioral properties to be verified in the model checker are built and specified [31, 18].

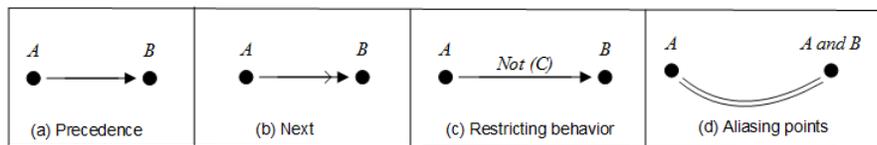
Given this context in this work we present a parallel implementation of a tableau algorithm which translates graphical scenarios specified in the FVS formal specification language [3, 4] describing behavioral properties into Büchi Automata. The automata built by the tableau can latter be used to feed a model checker in order to perform verification tasks [3]. FVS (Feather weight Visual Scenarios) is a formal and graphical specification language which can be used to specify, verify and synthesize behavior [4, 5]. It denotes a very rich and expressive notation (being for example more expressive than Linear Temporal Logics) and linear as well as branching-time type properties can be specified [3]. In order to adapt FVS for formal verification in BIG DATA system in this work we parallelized the tableau algorithm. **This new parallel implementation of the tableau enables the possibility for FVS to make a solid contribution for the formal verification phase applied to BIG DATA systems.** We developed three different implementations of the tableau algorithm, one using *Java* threads and other two employing two well known libraries for parallel programming: *Open MPI* [33] and *MPJ Express* [29]. Since our framework is developed using the *Java* programming language we relied on *Java* oriented tools. We compared the three versions against each other and against the sequential version of the tableau taking as a case study a complex and industrial relevant protocol verification: the *MS-NNS* protocol [1], a lightweight option to provide authenticated and confidential communication between a server and a client over a TCP connection. Although this early results are preliminaries, we believe are encouraging enough to continue exploring FVS's contributions for formal verification in BIG DATA systems.

The rest of this paper is structured as follows. Section 2 briefly presents the FVS specification language and Section 3 introduces the sequential version of the tableau. Section 4 exhibits the details behind the parallelization of the tableau, considering three different versions, a performance comparison between them and some final observations. Finally, Section 5 comments some related and future work while Section 6 concludes this work highlighting its conclusions.

## 2 Feather weight Visual Scenarios

In this section we will informally describe the standing features of FVS. The reader is referred to [3] for a formal characterization of the language. FVS is a graphical language based on scenarios. Scenarios are partial order of events, con-

sisting of points, which are labeled with a logic formula expressing the possible events occurring at that point, and arrows connecting them. An arrow between two points indicates precedence. For instance, in figure 1-(a)  $A$ -event precedes  $B$ -event. In figure 1-b the scenario captures the very next  $B$ -event following an  $A$ -event, and not any other  $B$ -event. Events labeling an arrow are interpreted as forbidden events between both points. In figure 1-c  $A$ -event precedes  $B$ -event such that  $C$ -event does not occur between them. Finally, FVS features aliasing between points. Scenario in 1-d indicates that a point labeled with  $A$  is also labeled with  $A \wedge B$ . It is worth noticing that  $A$ -event is repeated on the labeling of the second point just because of FVS formal syntaxis.

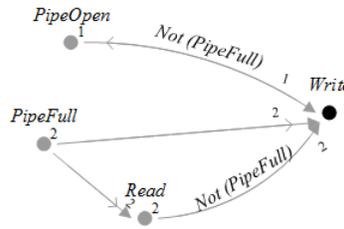


**Fig. 1.** Basic Elements in FVS

We now introduce the concept of FVS rules, a core concept in the language. Roughly speaking, a rule is divided into two parts: a scenario playing the role of an antecedent and at least one scenario playing the role of a consequent. The intuition is that whenever a trace “matches” a given antecedent scenario, then it must also match at least one of the consequents. In other words, rules take the form of an implication: an antecedent scenario and one or more consequent scenarios. Graphically, the antecedent is shown in black, and consequents in grey. Since a rule can feature more than one consequent, elements which do not belong to the antecedent scenario are numbered to identify the consequent they belong to. An example is shown in figure 2. The rule describes requirements for a valid writing pipe operation. For each write event, then it must be the case that either the pipe did not reach its maximum capacity since it was ready to perform (Consequent 1) or the pipe did reach its capacity, but another component performed a read over the pipe (making the pipe available again) afterwards and the pipe capacity did not reach again its maximum (Consequent 2).

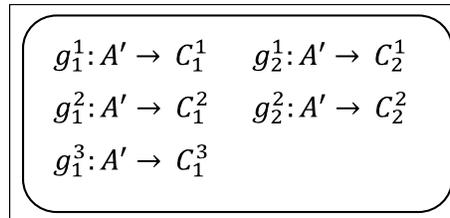
### 3 Tableau Algorithm: From FVS Scenarios to Büchi Automata

We now present some basic concepts to understand the tableau algorithm while the reader is referred to [3] for a more detailed version of it. From a formal point of view, FVS scenarios can be defined as morphisms from the antecedent to the consequent. The algorithm relies on the notion of situations [3]. In few words, a situation represents for a given rule possible combinations of partial matches from the antecedent to the consequent. Consider the following example



**Fig. 2.** An FVS rule example

in figure 3. In this case, a rule with two consequents is shown. Furthermore, there are three partial matches for consequent one, and two for consequent two. Therefore,  $\eta_1$  consists of the three morphisms in the first column ( $g_1^1, g_1^2, g_1^3$ ), whereas  $\eta_2$  consists of the two morphisms in the second column ( $g_2^1, g_2^2$ ).



**Fig. 3.** A situation example

Given a rule  $R$ , the tableau builds a Büchi Automaton  $\mathcal{B} = \langle \Sigma, S, S^0, \Delta, F \rangle$  such that  $\Sigma$  constitutes minterms over  $\Sigma_R$  and the set of states  $S$  are triples  $(\mathcal{Y}_R \times \text{bool} \times \mathcal{P}\mathcal{L}(\Sigma_R))$ , where  $\mathcal{P}\mathcal{L}(\Sigma)$  is a function that labels each point with a given formula. The set  $\mathcal{Y}_R$  associated to a state (a set of situations  $\eta$ ), denoted  $\text{situations}(S)$ , symbolically represents all the possible combination of partial matches obtained up to that state from the antecedent to each consequent. The second term of the triple identify accepting states. This boolean variable is set to **true** when the pattern is completely matched and will make the state transient. Finally, a third element is needed to maintain future obligations of the trace. These formulas are needed when rules predicate about conditions that must hold until the end of the trace.

The pseudo-code sketched in Algorithm 1 computes the successor states for transition relation  $\Delta$ . Starting from the initial state  $(\langle \emptyset, \text{false}, \text{true} \rangle)$ , the automata will try to incrementally “construct” the pattern as events, represented by minterms, occurs. For every minterm, algorithm 1 computes all possible matchings considering matchings in the antecedent and also in each consequent. This is obtained through two auxiliary algorithms, *advanceAntecedent* (line 5) and *advanceConsequent* (line 6). Line 7 analyzes if any successor reaches a *trap sit-*

uation, a situation where the antecedent has been matched (a morphism such that  $A' = A$ ), but matching for all consequents is known unfeasible. Lines 8 and 9 check if any consequent has been matched by the last move. This is,  $goalmatched[i] = true$  if and only if *consequent*  $C_i$  is matched. Line 10 analyzes if the next state is an accepting state: a consequent has been matched and it is not a trap situation. Finally, line 11 returns the expected output.

```

1 Algorithm Succ( $S : State, m : minterm$ ) : set of states;
2 Precondition :  $m \wedge obligations(S)$  is satisfiable;
3  $newSits := \emptyset$ ;
4 foreach  $\eta \in Situations(S)$  do
5    $newSits := add(newSits, advanceAntecedent(\eta, m));$ 
6    $newSits := add(newSits, advanceConsequent(\eta, m));$ 
7  $trapSituation : \exists \eta \in newSits \forall i \forall j \in [1..n] g_j^i : A' \rightarrow C_j^i \in \eta \wedge A' = A \wedge C_j^i$ 
   it is not a configuration of  $C_j$ ;
8 foreach  $j \in [1..n]$  do
9    $goalmatched[j] := \exists \eta \in situations(S) \wedge$ 
    $g_j^i : A' \rightarrow C_j^i \in \eta \wedge C_j^i \xrightarrow{m} C_j \cup \mathcal{F}_j^i \wedge m \in (R_F(C_j^i)) \wedge C_j^i \cup \mathcal{F}_j^i = C_j$ ;
10  $goalMatched := (\exists j (goalmatched[j])) \wedge (\neg trapSituation)$  ;
11 return {  $\langle newSits, GM, Obligations \rangle$  such that
    $GM \rightarrow goalMatched \wedge GM = true \rightarrow \exists j (goalmatched[j]) \wedge Obligations =$ 
    $Obligations(S) \wedge \bigwedge_{j \in I} \mathcal{R}(C_j) \wedge GM = false \rightarrow Obligations = Obligations(S)$ 

```

**Algorithm 1:** Successor states

As an example of the application of the algorithm consider the FVS rule in Figure 4, which represents a classic instantiation of the Response pattern [21]. The automaton built by the tableau is depicted in Figure 5.

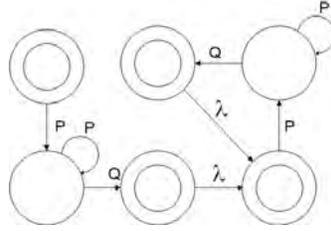


**Fig. 4.** An FVS Rule for the Response Pattern

## 4 Parallel Tableau Implementation

We now describe the main features of the parallel algorithm. We developed three different versions of it, whose implementations details are presented in Section 4.1. Finally, Section 4.2 presents some final remarks.

After a rigorous analysis of Algorithm 1 we detected two natural points suitable for parallelization. Those two points are: the computation of all the possible



**Fig. 5.** The automaton built by the tableau for the Response Pattern

antecedents and consequents and tagging all the possible matches (the *For Each* constructor from lines 4 to 6) and checking whether any consequent has been matched by the last move (line 9). These are individual, orthogonal and repetitive tasks that can be easily divided into different nodes to be realized and then the main algorithm can continue once every one is finished. For example, the calculation of all the possible matches for every situation  $\eta$  can be done in parallel where each node calculates the possible matching in antecedent and consequents for every  $\eta$ . Once all the nodes are finished the results are sent to the main algorithm which simply merges the results. Similarly, for the calculation in line 9 the detection of any matching in the consequent can be done in parallel where a node contemplates one situation. When the nodes are finished the main algorithm can obtain the final result. The parallel pseudo-code for the parallel implementation of the tableau is depicted in Algorithm 2.

```

1 Algorithm Parallel Succ( $S : State, m : minterm$ ) : set of states;
2 Precondition :  $m \wedge obligations(S)$  is satisfiable;
3  $newSits := \emptyset$ ;
4 PrepareNodes( $N_1, N_2, \dots, N_m$ );
5 DistributeAdvancesCalculation( $N_1, N_2, \dots, N_m, Situations(S)$ );
6 JoinNodes( $N_1, N_2, \dots, N_m, newSits$ ) ;
7 trapSituation :  $\exists \eta \in newSits \forall i \forall j \in [1..n] g_j^i : A' \rightarrow C_j^i \in \eta \wedge A' = A \wedge C_j^i$  it
  is not a configuration of  $C_j$ ;
8 PrepareNodes( $N_1, N_2, \dots, N_m$ );
9 DistributeGoalMatched( $N_1, N_2, \dots, N_m, Situations(S)$ );
10 JoinNodes( $N_1, N_2, \dots, N_m, goalMatched$ ) ;
11  $goalMatched := (\exists j (goalmatched[j])) \wedge (\neg trapSituation)$  ;
12 return {  $\langle newSits, GM, Obligations \rangle$  such that
   $GM \rightarrow goalMatched \wedge GM = true \rightarrow \exists j(goalmatched[j]) \wedge Obligations =$ 
   $Obligations(S) \wedge \bigwedge_{j \in I} \mathcal{R}(C_j) \wedge GM = false \rightarrow Obligations = Obligations(S)$ 

```

**Algorithm 2:** Parallel Successor states Calculation

Line 4 in Algorithm 2 deals with the nodes preparation and setup, a typical task in parallel systems. Line 5 is in charge of distributing the task of obtaining

the advances of antecedent and consequents in each situation  $\eta$  among the nodes. Finally, in Line 6 all the tasks done by the nodes is united and the new situations set represented by the variable *newSits* is obtained. Similarly, lines 8 to 10 deal with the parallelization of goal calculation and verify if any antecedent has been satisfied.

#### 4.1 Algorithm Implementation and Evaluation

We developed three different implementation for the parallel algorithm delineated in Algorithm 2. We choose *Java* related tools since our algorithm is implemented in that programming language. In the first one we simply use *Java* predefined constructors to deal with parallelism: *threads*. The others two version handle different parallel libraries for *Java*: *Open MPI* [33] and *MPJ Express* [29].

*Open MPI* is one of the most popular implementations of *MPI*, the Message-Passing Interface, which is one of the predominant programming paradigm for parallel applications on distributed memory computers [33]. This work enables *Java* *MPI* bindings which have been included in the *Open MPI* distribution, exposing *MPI* functionality to *Java* programmers. It can be easily downloaded and installed from its website<sup>3</sup>. The implementation of the algorithm was very straightforward since parallel constructors application is declarative and intuitive.

The third version was implemented using *MPJ Express* [29], an open source *Java* message passing library that enables the possibility of introducing parallel instructions in *Java* programs. As in the previous case, it can be easily downloaded and installed from its website<sup>4</sup>. The setup and integration did not result in an straightforward task since several difficulties arose when trying to integrate it to our *Java* framework (for example, library versions incompatibility). However, once these not that unexpected situations when integrating software tools were solved, the codification of the algorithm was achieved without major problems.

We conducted a typical performance comparison including the three versions of the parallel tableau and also its sequential version. We employ as case of study the verification of the *MS-NNS* protocol specified in [3]. In few words, this protocol was introduced as a lightweight option to provide authenticated and confidential communication between a server and a client over a *TCP* connection protocol. In [3] both the server and the client behavior is specified and verified. For this performance evaluation we consider four cases. In all of them we employ one server whereas the amount of clients was different in each one. In other words, we consider these four cases: *Case 1*: One server and two clients ; *Case 2*: One server and four clients; *Case 3*: One server and eight clients and *Case 4*: One server and sixteen clients. The results are shown in Table 1 .

<sup>3</sup> <https://www.open-mpi.org/>

<sup>4</sup> <http://mpj-express.org/>

Case Number	Sequential	Threads	MPJ	Open MPI
1	60 sec	48 sec	53 sec	50 sec
2	109 sec	80 sec	60 sec	70 sec
3	360 sec	115 sec	70 sec	83 sec
4	600 sec	143 sec	80 sec	98 sec

**Table 1.** Performance Evaluation

## 4.2 Some Observations

It can be noted from the results in Table 1 that there is a considerable gain when the parallel version of the tableau is employed. The threads version had the best performance in the first case. We believe that the node preparation and other necessary parallel settings was too much overload for a simple case with only two clients. In the following cases the *Open MPI* and the *MPJ Express* versions dethrone the threads implementation and the difference increases when more complex case studies are presented. Between the *Open MPI* and the *MPJ Express* the latter version turns out to be slightly more competitive regarding performance.

We are aware there are several threads to the validity of this initial and exploratory results. First of all, we assume there are always lazy nodes available to receive new tasks. It would be interesting to analyze scenarios where this condition is not necessarily meet since more parallel overhead is naturally expected. Secondly, in all the analyzed cases a similar load balance was assigned to each node. It might occur a different setting in other contexts and examples making necessary to introduce some load balancer strategies who will certainly impact in the performance of the algorithm. Finally, more examples are also needed to further validate this initial experimentation.

However, taking all these facts into consideration we believe the results are promising enough to continue exploring this line of research.

## 5 Related and Future Work

Several approaches aim to adapt current formal verification techniques to BIG DATA systems.

In [10, 15] a interesting framework for distributed CTL (computation tree logic) model checker is presented. They present a novel architecture employing *HADOOP MAPREDUCE* as its computational engine. They provide a very solid empiric evaluation with several case of studies employing *Amazon Elastic MapReduce* [15] and the *GRID5000* cloud infrastructure [6]. For generating and building distributed state space exploration they rely on a framework called *Mardigras* [9]. We would definitely like to explore in future work the combination of these advanced tools with our specification language FVS.

Other approaches like [11, 13, 7, 8, 14] provide some tools implementing different versions of parallel model checking algorithms for both linear and branching-

time properties. We believe that a natural continuation of this work is to provide the automata build by the parallel tableau as the behavioral properties to be checked in any of the mentioned approaches. In particular, we would like to explore the integration of FVS with the parallel implementation of MTSA (The Modal Transition System Analyser) [12].

In a different direction, work like [20, 28] employ metamorphic testing as an alternative to validate BIG DATA results. We would like to extend this notion to formally model check behavior pursuing the notion of “metamorphic” properties.

## 6 Conclusions

In this work we present a parallel implementation of the tableau algorithm which translates FVS scenarios into Büchi automata. Actually, three different implementations were developed and analyzed taking as a case of study an industrial relevant protocol with complex behavior. This new version of the algorithm allows FVS specifications to be constructed in a much efficient way, easing the adoption of our language to model and verify behavior in BIG DATA systems. The next step in this direction is to combine FVS specification with parallel and distributed architectures which are the most frequent ones in BIG DATA systems.

## References

1. [ms-nns]: .net negotiatestream protocol specification v2.0. <http://msdn.microsoft.com/en-us/library/cc236723.aspx>, July 2008.
2. E. Alpaydin. *Introduction to machine learning*. MIT press, 2020.
3. F. Asteasuain and V. Braberman. Declaratively building behavior by means of scenario clauses. *Requirements Engineering*, 22(2):239–274, 2017.
4. F. Asteasuain, F. Calonge, and M. Dubinsky. Exploring specification pattern based behavioral synthesis with scenario clauses. In *CACIC*, 2018.
5. F. Asteasuain, F. Calonge, and P. Gamboa. Behavioral synthesis with branching graphical scenarios. In *CONAIIISI*, 2019.
6. D. Balouek, A. C. Amarie, G. Charrier, F. Desprez, E. Jeannot, E. Jeanvoine, A. Lèbre, D. Margery, N. Niclaussé, L. Nussbaum, et al. Adding virtualization capabilities to the grid5000 testbed. In *International Conference on Cloud Computing and Services Science*, pages 3–20. Springer, 2012.
7. J. Barnat, L. Brim, M. Česka, and P. Ročkal. Divine: Parallel distributed model checker. In *2010 ninth PDMC*, pages 4–7. IEEE, 2010.
8. A. Bell and B. R. Haverkort. Sequential and distributed model checking of petri nets. *STTT journal*, 7(1):43–60, 2005.
9. C. Bellettini, M. Camilli, L. Capra, and M. Monga. Mardigras: Simplified building of reachability graphs on large clusters. In *RP workshop*, pages 83–95, 2013.
10. C. Bellettini, M. Camilli, L. Capra, and M. Monga. Distributed ctl model checking using mapreduce: theory and practice. *CCPE*, 28(11):3025–3041, 2016.
11. M. C. Boukala and L. Petrucci. Distributed model-checking and counterexample search for ctl logic. *IJSR* 3, 3(1-2):44–59, 2012.

12. M. V. Brassesco. Síntesis concurrente de controladores para juegos definidos con objetivos de generalized reactivity(1). *Tesis de Licenciatura.*, [http://dc.sigedep.exactas.uba.ar/media/academic/grade/thesis/tesis\\_18.pdf](http://dc.sigedep.exactas.uba.ar/media/academic/grade/thesis/tesis_18.pdf) UBA FCEyN Dpto Computacion 2017.
13. L. Brim, I. Černá, P. Moravec, and J. Šimša. Accepting predecessors are better than back edges in distributed ltl model-checking. In *International Conference on Formal Methods in Computer-Aided Design*, pages 352–366. Springer, 2004.
14. L. Brim, K. Yorav, and J. Žídková. Assumption-based distribution of ctl model checking. *STTT*, 7(1):61–73, 2005.
15. M. Camilli. Formal verification problems in a big data world: towards a mighty synergy. In *ICSE*, pages 638–641, 2014.
16. M. Chen, S. Mao, and Y. Liu. Big data: A survey. *Mobile networks and applications*, 19(2):171–209, 2014.
17. E. M. Clarke, O. Grumberg, and D. Peled. *Model checking*. MIT press, 1999.
18. E. M. Clarke, W. Klieber, M. Nováček, and P. Zuliani. Model checking and the state explosion problem. In *LASER School*, pages 1–30. Springer, 2011.
19. T. Dillon, C. Wu, and E. Chang. Cloud computing: issues and challenges. In *2010 AINA*, pages 27–33. Ieee, 2010.
20. J. Ding, D. Zhang, and X.-H. Hu. A framework for ensuring the quality of a big data service. In *2016 SCC*, pages 82–89. IEEE, 2016.
21. M. Dwyer, M. Avrunin, and M. Corbett. Patterns in property specifications for finite-state verification. In *ICSE*, pages 411–420, 1999.
22. M. H. Hassoun et al. *Fundamentals of artificial neural networks*. MIT press, 1995.
23. O. Hummel, H. Eichelberger, A. Giloj, D. Werle, and K. Schmid. A collection of software engineering challenges for big data system development. In *SEAA*, pages 362–369. IEEE, 2018.
24. V. D. Kumar and P. Alencar. Software engineering for big data projects: Domains, methodologies and gaps. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 2886–2895. IEEE, 2016.
25. R. Laigner, M. Kalinowski, S. Lifschitz, R. S. Monteiro, and D. de Oliveira. A systematic mapping of software engineering approaches to develop big data systems. In *SEAA*, pages 446–453. IEEE, 2018.
26. C. E. Otero and A. Peter. Research directions for engineering big data analytics software. *IEEE Intelligent Systems*, 30(1):13–19, 2014.
27. A. J. Robinson and A. Voronkov. *Handbook of automated reasoning*, volume 1. Gulf Professional Publishing, 2001.
28. S. Segura, G. Fraser, A. B. Sanchez, and A. Ruiz-Cortés. A survey on metamorphic testing. *IEEE Transactions on software engineering*, 42(9):805–824, 2016.
29. A. Shafi, B. Carpenter, and M. Baker. Nested parallelism for multi-core HPC systems using java. *J. Parallel Distributed Comput.*, 69(6):532–545, 2009.
30. P. A. Sri and M. Anusha. Big data-survey. *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, 4(1):74–80, 2016.
31. A. Valmari. The state explosion problem. In *Advanced Course on Petri Nets*, pages 429–528. Springer, 1996.
32. W. Van Der Aalst. Data science in action. In *Process mining*, pages 3–23. Springer, 2016.
33. O. Vega-Gisbert, J. E. Roman, and J. M. Squyres. Design and implementation of java bindings in open mpi. *Parallel Computing*, 59:1–20, 2016.
34. P. Zikopoulos, C. Eaton, et al. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media, 2011.

# Evaluación de un Modelo Ontológico basado en la Adecuación Funcional de la Norma ISO 25010 para la Elicitación de Requisitos de Software

Rebeca Yuan<sup>1</sup>, Carlos Salgado<sup>2</sup>, Mario Peralta<sup>2</sup>, Alberto Sánchez<sup>2</sup>

<sup>1</sup> Ingeniería de Software, Departamento en Ingeniería en Sistemas de Información,  
Facultad Regional San Francisco – Universidad Tecnológica Nacional  
San Francisco, Córdoba, Argentina  
ryuan@sanfrancisco.utn.edu.ar

<sup>2</sup> 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, mperalta, alfanego}@unsl.edu.ar

**Resumen.** En la era de la transformación digital, los productos de software buscan evolucionar más allá del simple hecho de resolver un problema; deben ser capaces de apoyar con la información que generan, las decisiones de las organizaciones. Esta exigencia no proclama cambios de paradigmas para el proceso de desarrollo del software; sino una construcción “consciente” del mismo. Para lograr un juicio común de lo que se busca de un producto de software, se desarrolló un modelo ontológico como herramienta que de soporte a la elicitación de requerimientos basado en la completitud funcional de la Norma ISO 25010. El presente trabajo busca validar el modelo ontológico desarrollado, con el objetivo de establecer si el modelo resulta ser una herramienta capaz de identificar y visualizar, funciones y datos de gran impacto para las organizaciones.

**Palabras Claves:** Calidad de Software, Modelo Ontológico, Elicitación de Requerimientos, Adecuación Funcional, Completitud Funcional.

## 1 Introducción

Bajo el marco de la Cuarta Revolución Industrial, se van evidenciando cambios profundos (como toda revolución), los datos comienzan a jugar un papel importante y primordial en esta transformación digital [1].

Los sistemas de información no son ajenos a esta revolución, es de ellos mismos, donde surgen los datos para futuras tomas de decisiones. Este avance no proclama nuevos ciclos de vida de desarrollo, ni nuevas técnicas que acompañen al mismo; pero llama a hacer un alto y observar qué funciones del dominio generan información que produzca un valor agregado en la organización. Una vez más, el foco está en el análisis de requerimientos del software, etapa inicial en el proceso de desarrollo del mismo.

“Obtener los requisitos del sistema por medio de la observación de los sistemas existentes, discusión con los usuarios potenciales y proveedores, el análisis de tareas, etcétera, puede implicar el desarrollo de uno o más modelos y prototipos del sistema que ayuden al analista a comprender el sistema a especificar” [2]. El análisis de requerimientos basa su desarrollo en la actividad de elicitación de requerimientos, esta última, es la etapa de mayor interacción con el usuario. La extracción de requerimientos necesita distintas herramientas para construir y perfilar el alcance de un sistema [3]. Sobre esta tarea, se sientan los problemas y objetivos, necesidades y restricciones devenidos a requerimientos, que el analista debe saber interpretar y comunicar [4].

El tipo de lenguaje utilizado, el uso de homónimos en contextos poco definidos, son algunos de los peligros que se presentan en la etapa de elicitación de requerimientos; una herramienta que proclama y proporciona un vocabulario común son las ontologías [5]. El fin de la ontología es facilitar un entendimiento común del conocimiento a los miembros de un equipo de desarrollo de software, “una especificación formal, explícita de una conceptualización compartida” [6, 7]. Por las características enunciadas, se propone un modelo ontológico que sirva de base para la elicitación de requerimientos.

Para que el modelo sea capaz de garantizar una descripción completa del dominio que se analiza, se contempló la incorporación en el desarrollo del mismo, a la norma de calidad ISO/IEC 25010 [8]. El modelo de calidad definido por la norma se encuentra compuesto por 8 características: Adecuación Funcional, Eficiencia de Desempeño, Compatibilidad, Usabilidad, Fiabilidad, Seguridad, Mantenibilidad y Portabilidad. Una sub característica de la Adecuación Funcional, es la Completitud Funcional, grado en el cual el conjunto de funcionalidades cubre todas las tareas y los objetivos del usuario especificados [9]. El modelo bajo estudio, busca garantizar que el modelo ontológico cumpla la Completitud Funcional adaptándose a cualquier dominio que se aplique.

El desarrollo del modelo se realizó a través del Método 101, propuesto por Natalya F. Noy y Deborah L. McGuinness, [10,11]. El método establece siete pasos que guían al desarrollo de la ontología. Determinar el dominio y alcance de la ontología, reutilizar ontologías existentes, enumerar términos importantes de la ontología, definir clases y su jerarquía, definir las propiedades de esas clases, definir las facetas de ranuras y crear instancias para las mismas.

En base a lo establecido por el Método 101, el dominio y alcance de la herramienta desarrollada se corresponde a la etapa de elicitación de requerimientos de software.

Contemplar a las normas ISO en el desarrollo del modelo ontológico, permitió definir clases y relaciones representativas y adaptables a cualquier dominio. Las clases, propiedades y relaciones se establecieron para poder lograr información y un entendimiento general sobre cualquier dominio. Bajo la mirada de la norma, se gestó el establecimiento de las clases pensando, ¿para quién se desarrolla el software?, ¿quién consume la información que genera la solución?, ¿quiénes son los interesados o afectados a futuro a la solución que se propone?. El desarrollo completo del modelo se encuentra en [12].

La herramienta que permitió el desarrollo del modelo fue Protégé [13], el mismo es un editor para construir sistemas inteligentes, permite introducir el modelo ontológico, visualizar el resultado y ejecutar sobre el mismo razonadores y consultas que mejoren su estructura y entendimiento por parte de los interesados. En la Figura 1, se presenta la vista OWLViz, donde se observan las clases utilizadas en el modelo ontológico y en la Figura 2, las clases y sus relaciones, ambas vistas son generadas a través de Protégé.

Para determinar la eficiencia de la herramienta, se sometió al modelo ontológico a una evaluación triangular.

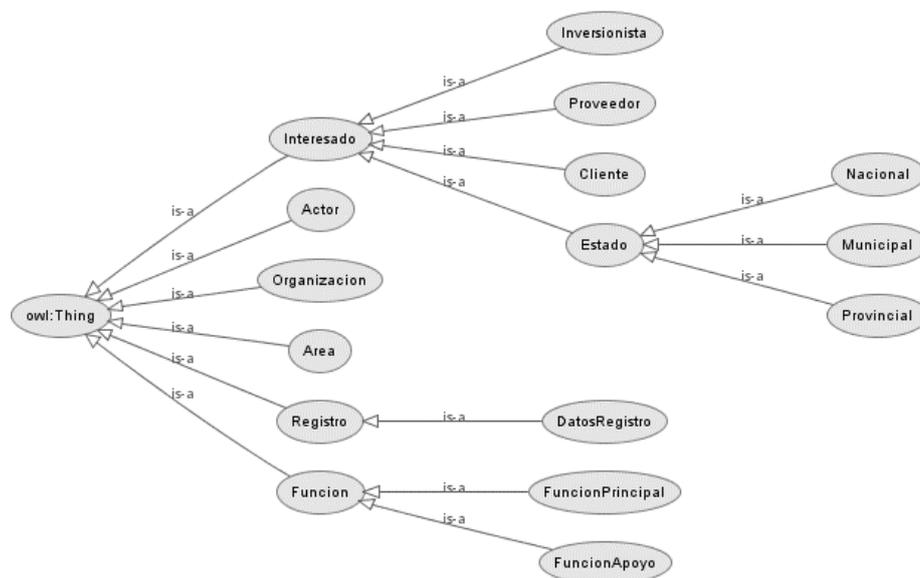


Fig. 1. Vista OWLViz Modelo Ontológico para la Elicitación de Requerimientos.

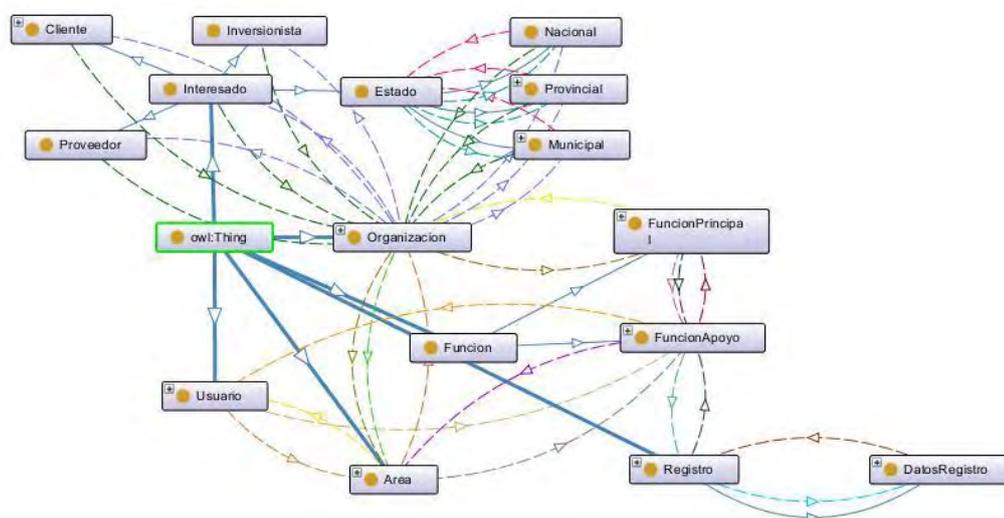


Fig. 2. Vista de Clases y Relaciones del Modelo Ontológico para la Elicitación de Requerimientos.

## 2 Evaluación de la Propuesta

La construcción de modelos ontológicos puede realizarse de diversas maneras. Sin embargo, a pesar de que no existe un único camino para el desarrollo de ontologías, el resultado final debe ser consecuente a los principios que lo propiciaron [14].

Algunas metodologías de desarrollo de modelos ontológicos, exponen junto a la formalización de las ontologías, el camino para evaluar las mismas [15]; otras no indican métodos de evaluación, pero plantean la necesidad de someter al modelo, a una valoración por parte de los usuarios que trabajan sobre el mismo.

La validación se refiere a que las definiciones de la ontología modelen lo más exactamente posible el dominio para el cual fueron creadas [16]. En el presente trabajo de investigación, el marco de evaluación del modelo se gestó bajo la Norma ISO/IEC 25010 [8][17] con el objetivo de que el modelo ontológico asegure la completitud funcional en la elicitación de requerimientos del software.

Para la evaluación del modelo se empleó el método de evaluación triangular propuesto en [14], el cual considera, tanto aspectos internos del modelo, como aquellos relacionados con el usuario. El método propone utilizar casos de pruebas, razonadores y una lista de chequeos con los errores más comunes detectados en el diseño de las ontologías, como se observa en la Figura 3 - “Método de evaluación triangular”. Se utilizará el razonador HermiT para evaluar las propiedades lógico-formales de las ontologías durante el ciclo de vida de las mismas. Las preguntas de competencia se realizarán a través de SPARQL. El razonador y el lenguaje SPARQL son propios de la herramienta Protégé [18], se encuentran integrados en la misma, lo que facilita la etapa de evaluación.

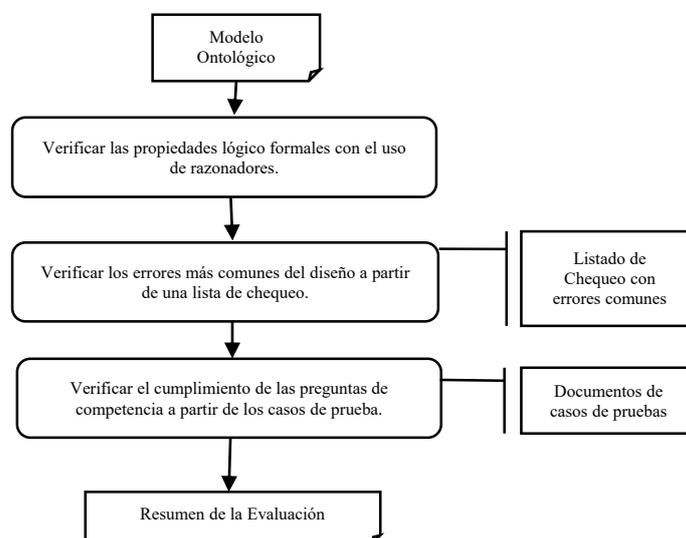


Fig. 3. Modelo de Evaluación Triangular

### 3 Metodología de Evaluación Triangular

#### 3.1 Verificar las propiedades lógico formales con el uso de razonadores

La herramienta Protégé trabaja con razonadores, la función de los mismos es controlar que en el modelo establecido, no existan contradicciones e inferir en nuevo conocimiento, estos razonadores pueden estar agrupados en dos categorías: razonadores de lógica descriptiva y razonadores de programación lógica [19].

La lógica descriptiva permite representar bases de conocimiento que describen un dominio en particular, la misma se representa a través de clases, individuos, roles. Los razonadores de lógica descriptiva (LD) [19] ofrecen los siguientes servicios de inferencia: Validación de la consistencia, Validación de conceptos, Precisiones sobre los conceptos de la jerarquía.

Protégé presenta varios razonadores ELK, Ontop, Pellet y otros más, para el presente proyecto trabajamos con Hermit [20], el cual es un razonador para ontologías escritas empleando OWL (Web Ontology Language, lenguaje de marcado para publicar y compartir datos en la web) y construido empleando cálculo hypertextableau a fin de proveer razonamiento más eficiente. Pese a encontrarse integrado en Protégé, Hermit también es open-source y se anuncia como un razonador veloz y capaz de resolver ontologías complejas.

Una vez realizada la carga completa de instancias al modelo ontológico, se ejecuta el razonador. El resultado guía al analista y a los distintos actores participantes de la elicitación a visualizar aquellas características ausentes; permitiendo lograr un entendimiento compartido de las funciones que el producto de software debe realizar.

#### 3.2 Verificar los errores más comunes del diseño a partir de una lista de chequeo.

Para verificar los errores comunes en el diseño, se toma como referencia el framework de evaluación para una bio-ontología basado en la norma de calidad ISO 9126 establecida por [21].

Los atributos a evaluar de nuestro modelo, quedan establecidos en la Tabla 1.

**Tabla 1.** Marco de Evaluación para el Modelo Ontológico para la Elicitación de Requerimientos de Software.

Marco de Evaluación para el Modelo Ontológico para la Elicitación de Requerimientos	
Funcionalidad	<ul style="list-style-type: none"> <li>– Referencia Ontológica</li> <li>– Vocabulario Controlado</li> <li>– Búsquedas y Consultas Consistentes.</li> <li>– Adquisición de Conocimiento.</li> <li>– Agrupamiento y similitudes.</li> <li>– Representación de resultados.</li> <li>– Reutilización.</li> <li>– Inferencia.</li> </ul>

En este contexto, la *referencia ontológica*, se corresponde con el uso de otras ontologías sobre el dominio con el que se está trabajando. Por ejemplo, si el dominio de trabajo pertenece a una empresa de Seguros, podría enlazar modelos ontológicos que

correspondan a accidentes, seguros de vida, modelos de autos, etc. El *vocabulario controlado* se establece en función de si el modelo logró extraer el vocabulario propio del dominio bajo estudio, si logra plasmar los modismos y palabras de uso común del dominio. *Adquisición de conocimiento*, se corresponde a si el modelo logra captar el entendimiento y conocimiento completo del dominio, ya sea por parte del analista, como de los usuarios y clientes; se busca determinar si el modelo sirve como herramienta que abra el debate sobre un entendimiento compartido. La característica *Representación de Resultados*, busca establecer si la herramienta presenta una interface simple para observar las instancias establecidas, bajo el mismo marco se evalúa si corresponde el *agrupamiento y similitud de las clases*, relaciones y distintos elementos del modelo. La *reutilización del modelo* para la elicitación de requerimientos en dominios diferentes va a establecer el nivel de generalidad del mismo. La *inferencia, la búsqueda y consultas de consistencia*, se trabajaron con el razonador Hermit.

### 3.3 Verificar el cumplimiento de las preguntas de competencia a partir de los casos de prueba

El último punto a evaluar corresponde a las preguntas de competencia que debe responder el modelo. Nuestro objetivo se centra en maximizar el reconocimiento de funciones del dominio; a través de la visualización por parte de los distintos actores.

Las preguntas de competencia hacen referencia a aquellos interrogantes que los actores aplican al modelo y que éste debe responder. La estructura del modelo ontológico debe encontrarse bien definida, para que permita dar respuestas verdaderas a estas preguntas.

Para realizar esta etapa de evaluación, se utilizó el Lenguaje SPARQL (Protocol and RDF Query Language/Lenguaje de Protocolo y Consulta RDF) integrado a Protégé. RDF (Por su sigla del Inglés *Resource Description Framework* - Marco de Descripción de Recursos) es un método para expresar el conocimiento en un entorno descentralizado y es el fundamento de la web semántica, en el que las aplicaciones informáticas utilizan información estructurada distribuida por toda la red [22]. RDF permite descomponer cualquier tipo de conocimiento en trozos pequeños, con reglas semánticas. SPARQL representa las RDF como un conjunto de triples que consisten en un sujeto, un predicado y un objeto como expresión básica de los datos almacenados en la base de conocimientos [18]. SPARQL brinda la posibilidad de identificar las entidades o clases que intervienen en la relación y las formas mediante las cuales se establece la relación [23]. La siguiente estructura de consulta, cubre los requisitos anteriores.

```
SELECT ?x ?y ?z
WHERE {
  ?x rdf:subClassOf ?y.
  ?y owl:onProperty ?z.
}
```

## 4 Casos de Estudio - Evaluación

Para realizar la evaluación del modelo ontológico; se solicitó a cuatro profesionales

en Sistemas de Información, que se desempeñan como analistas, desarrolladores y tester en distintas empresas de software, que utilicen y califiquen la herramienta.

Para calificar las variables bajo estudio, se estableció una calificación numérica que corresponde: de 1 al 3, “Regular”, de 4 a 6 calificación “Buena”; de 7 a 8 “Muy Buena” y 9 o 10 “Excelente”. Esta evaluación se realizó sobre 10 proyectos de software, de las tres empresas de software entrevistadas. Los proyectos se corresponden a aplicaciones de control, sitios web y sistemas de gestión. Los resultados obtenidos se muestran en la Tabla 2.

**Tabla 2.** Valoración de los distintos atributos de Funcionalidad.

	Proyecto 1	Proyecto 2	Proyecto 3	Proyecto 4	Proyecto 5	Proyecto 6	Proyecto 7	Proyecto 8	Proyecto 9	Proyecto 10	
Funcionalidad	Referencia Ontológica	9	8	2	9	6	8	4	5	7	5
	Vocabulario Controlado	8	7	5	7	4	7	4	4	4	3
	Búsquedas y Consultas Consistentes.	9	7	5	8	5	6	6	6	7	6
	Adquisición de Conocimiento.	8	7	7	8	6	6	7	7	7	7
	Agrupamiento y similitudes.	7	7	6	7	7	7	7	6	5	6
	Representación de resultados.	7	7	6	7	8	8	7	8	7	8
	Reutilización.	9	8	6	7	7	6	7	7	7	7
	Inferencia.	8	8	8	6	7	7	7	6	7	7

Sobre los resultados arrojados se calculó la media y desviación estándar para analizar cada una de las características bajo estudio. La tabla 3 muestra un resumen de los resultados obtenidos en este análisis.

El análisis busca determinar si los valores establecidos por los evaluadores se corresponden a una aprobación de la herramienta, pero también es necesario establecer si existe una variación en las calificaciones, y establecer si corresponden a la herramienta o al tipo de proyecto al que se aplica.

**Tabla 3.** Análisis Descriptivo

	Promedio	Desviación Estándar Muestral	Máximo	Mínimo
Referencia Ontológica	6,3	2,31	9	2
Vocabulario Controlado	5,3	1,77	8	3
Búsquedas y Consultas Consistentes.	6,5	1,27	9	5
Adquisición de Conocimiento.	7	0,67	8	6
Agrupamiento y similitudes.	6,5	0,71	7	5
Representación de resultados.	7,3	0,67	8	6
Reutilización.	7,1	0,88	9	6
Inferencia.	7,1	0,74	8	6

Como se observa en la tabla 3 en los atributos Búsquedas y consultas consistentes, Adquisición de Conocimiento, Agrupamiento y similitudes, Representación de Resultados, Reutilización e Inferencia la desviación con respecto a la media es baja, su lectura se corresponde a decir que: **“de 10 proyectos analizados, representación de resultados, obtuvo una calificación media, Muy Buena, con una desviación que varía entre Bueno y Muy Bueno, por lo tanto la herramienta cumple con la expectativa de representar en forma correcta los datos que se están modelando”**.

Analizamos también aquellos puntos donde la desviación es un poco más representativa. Se realizó un análisis más exhaustivo sobre los datos de la tabla 2, para conocer qué motivó a valores tan extremos en la puntuación.

La distribución normal de la *referencia ontológica*, se corresponde a la gráfica de la figura 4.

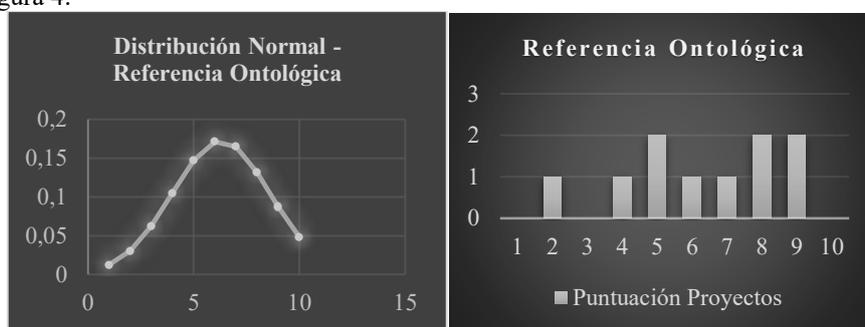


Fig. 4. Distribución Normal – Referencia Ontológica.

La gráfica muestra una distribución con una leve asimetría negativa, esto se debe a que existen en la calificación valores por debajo de la media. Analizando la distribución de los datos, se observan valores muy distantes, entre el máximo y mínimo valor, que hacen que la desviación estándar sea importante. Analizando los tipos de proyectos en los que se estableció este valor, resultó que los mismos se corresponden a un dominio muy cerrado, donde no se encontró en la web referencias ontológicas para ampliar el conocimiento del mismo. Los dominios cerrados, se corresponden a proyectos que, en su desarrollo, apuntan a solucionar problemas puntuales correspondientes a políticas de la organización.

El atributo *vocabulario controlado*, presenta una distribución asociada a la normal (Figura 5), su desviación establece que el modelo ontológico propuesto no logra extraer todo el vocabulario del dominio bajo estudio. Al analizar los proyectos asociados a valores bajos, los mismos se corresponden con proyectos de amplio alcance, es decir, dominios muy grandes.

Pese a evaluarse la herramienta en 10 proyectos, se pudo establecer dos características importantes: el *tipo* de dominio con el que se va a trabajar y el *tamaño* del mismo, infieren sobre los resultados de los atributos vocabulario y referencias, evaluados en el modelo ontológico.

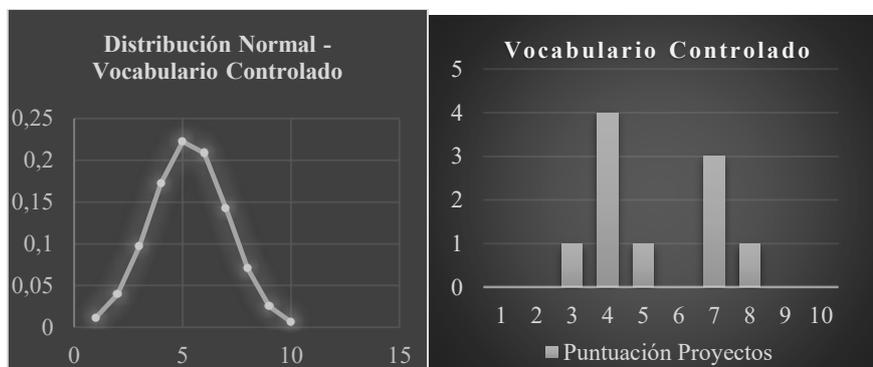


Fig. 5. Distribución Normal – Vocabulario Controlado

## 5 Conclusiones

La Elicitación de Requerimientos es el fundamento o base primordial en el desarrollo de proyectos software y es la fase que proporciona el impacto más alto en el diseño y en las demás fases del ciclo de vida del producto. Es muy importante asegurar la validez de los requisitos antes de comenzar el desarrollo del software.

El modelo propuesto, establece una herramienta para la elicitación de requerimientos que contribuye a solucionar el problema de completitud funcional, brindando un marco de trabajo que permite el entendimiento compartido por parte de los actores interesados en el dominio. Contemplar a la familia de normas ISO 2500 en el desarrollo y evaluación del modelo permitió visualizar la funcionalidad solicitada por los distintos actores y establecer el vocabulario del dominio bajo estudio. Contemplar a los “interesados” como clase del modelo; permitió entender el dominio como “parte de” y no atomizado al problema a resolver. Esta mirada brinda distintas perspectivas al problema y busca preguntarnos qué datos necesitamos incluir para ser competitivos a futuro. Las distinciones entre funciones principal y de apoyo, fue otro punto favorable observado, ya que indican un camino para lograr el objetivo final, y se pueden establecer planificaciones sobre las distintas funciones.

Siguiendo esta línea de investigación, sería posible aplicar herramientas de aprendizaje automático sobre la base de conocimiento que genera Protégé. Como mejoras al modelo ontológico, se observa la necesidad de incluir restricciones que hacen a la realidad de los dominios. Se concluye que evaluar al modelo ontológico bajo el marco de la norma ISO 25010 sobre adecuación funcional contribuye a la disminución de los costos de mantenimiento de software, al reducir esfuerzos y optimizar sus procesos; ya que logra obtener la comprensión y la totalidad de las funciones que los actores solicitan en la etapa de elicitación de requerimientos.

## Referencias

1. Schwab, K.: La Cuarta Revolución Industrial. Penguin Random House. ISBN: 978-84-9992-

- 699-5 2016.
2. Sommerville, I.: INGENIERÍA DE SOFTWARE. Novena edición. Pearson Education. ISBN: 9786073206037. 2011.
  3. M. A. Chaves, “La ingeniería de requerimientos y su importancia en el desarrollo de proyectos de software,” *InterSedes Rev. las Sedes Reg.*, vol. VI, no. 10, pp. 1–13, 2005.
  4. M. Griselda Báez and S. I. B. Brunner, “Metodología DoRCU para la Ingeniería de Requerimientos,” pp. 210–222.
  5. C. M. Z. Jaramillo, G. L. Giraldo, and G. A. U. Giraldo, “Las ontologías de la ingeniería de software: un acercamiento de dos grandes áreas del conocimiento,” *Rev. Ing. Univ. Medellin*, vol. 9, no. 16, pp. 91–99, 2010, doi: 10.1017/CBO9781107415324.004.
  6. G. N. Aranda and F. Ruiz, “Clasificación y ejemplos del uso de ontologías en Ingeniería del Software,” *XI Congr. Argentino Ciencias la Comput.*, 2005.
  7. T. R. Gruber, “A Translation Approach to Portable Ontology Specifications,” 1993.
  8. ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation - System and software quality models. 2011.
  9. J. A. M. López, “Adaptación del proceso de desarrollo software para cumplimiento de la adecuación funcional según ISO/IEC 25000,” 2017.
  10. Natalya F. Noy and Deborah L. McGuinness, “Ontology Development 101: A Guide to Creating Your First Ontology,” *Sustain.*, vol. 9, no. 12, pp. 1–25, 2017, doi: 10.3390/su9122317.
  11. O. Lassila and D. McGuinness, “The Role of Frame-Based Representation on the Semantic Web Introduction : Frame-based Representation Systems,” 1992.
  12. R. Yuan, C. Salgado, A. Sánchez, and M. Peralta, “Modelo Ontológico para la Educación de Requerimientos basado en la Adecuación Funcional de la Norma ISO 25010.” 2019 - UTN San Francisco - Universidad Nacional de San Luis.
  13. S. University, “Protégé.” Available: <https://protege.stanford.edu/>. [Accessed: 10-Mar-2020].
  14. Y. A. López Rodríguez, Y. Hidalgo-delgado, and N. Silega, “Un método práctico para la evaluación de ontologías,” no. May, 2018.
  15. M. Fernandez, A. Gómez-Pérez, and N. Juristo, “Methontology: from ontological art towards ontological engineering,” *Proc. AAAI97 Spring Symp. Ser. Ontol. Eng.*, no. May 2014, pp. 33–40, 1997.
  16. R. C. Esmeralda Ramos, Haydemar Núñez, “Esquema para evaluar ontologías únicas para un dominio de conocimiento,” *Art War*, vol. 6, no. 3, p. 32 ، ص 117, 2007, doi: 10.23943/9781400889877.
  17. E. De, S. D. De, M. J. Blas, S. Gonnet, and H. Leone, “Esquema de Calidad basado en el Estándar ISO / IEC 25010 Especificación de la Calidad en Software-as-a-Service : Definición de un Esquema de Calidad basado en el Estándar ISO / IEC 25010,” no. March 2018, 2016.
  18. H. Dong, F. K. Hussain, and E. Chang, “Application of Protégé and SPARQL in the field of project knowledge management,” *Second Int. Conf. Syst. Networks Commun. ICSNC 2007*, no. Table 1, 2007, doi: 10.1109/ICSNC.2007.22.
  19. N. N. A. Helguero, “Un agente basado en un razonador de ontologías,” Facultad de Ingeniería - Universidad de Buenos Aires, 2011.
  20. D. A. K. G. O. C. S. O. OXFORD, “Hermit OWL Reasoner.” [Online]. Available: <http://www.hermit-reasoner.com/>.
  21. J. T. Fernandez-Breis, M. Egaña Aranguren, and R. Stevens, “A Quality Evaluation Framework for Bio-Ontologies,” *Nat. Preced.*, no. July, 2009, doi: 10.1038/npre.2009.3479.1.
  22. R. Menor, “Ricard Menor,” *¿Qué es RDF?*, 2018. [Online]. Available: <https://www.seofreelance.es/que-es-rdf-introduccion-a-rdf/>.
  23. A. Vega-Ramírez, I. Grangel-González, I. Sáez-Mosquera, and R. García-Castro, “Procedimiento para la obtención de un modelo ontológico para representar la información contenida en bases de datos,” *CEUR Workshop Proc.*, vol. 1219, pp. 46–59, 2014.

# Impacto de una debilidad de ciberseguridad en la arquitectura de un sistema electromédico

Diego Coulombie<sup>1</sup>, Agustín Reyes<sup>1</sup>, Alberto Miguens<sup>1</sup>

<sup>1</sup> Dto de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza  
San Justo, Buenos Aires, Argentina  
{dcoulombie, aireyes, amiguens}@unlam.edu.ar

**Resumen.** El diseño de un equipamiento médico dedicado al monitoreo de profundidad de anestesia, que utiliza conectividad en su arquitectura con el fin de hacer el monitoreo de forma remota, se ve perjudicado cuando se enfrenta a las amenazas de ciberseguridad en la etapa de análisis de riesgos. La imposibilidad de validar de manera objetiva las medidas de control por parte de la organización que lo diseña, obligó a mitigar el elevado riesgo ante ciberataques con la drástica solución de hacer un cambio de arquitectura y de modelo de negocio abandonando la posibilidad de conectividad remota.

**Palabras clave:** Aplicaciones de uso médico, Internet de las Cosas, Telemedicina, Análisis de riesgo, Requisitos regulatorios de seguridad y protección.

## 1 Introducción

A veces las cosas se complican. Aquello que en un boceto parece una idea maravillosa y revolucionaria se desarma cuando se enfrenta con la realidad. Incluso cuando se tomaron precauciones previas puede ocurrir que algo inesperado, que no se imaginó al momento del diseño, se revele ante nosotros como un problema. La envergadura del problema puede tener gran importancia según el campo de aplicación de esa idea. En áreas que se saben sensibles, los contextos en los que se practican los diseños están ordenados por regulaciones y por autoridades de control que se encargan de verificar su cumplimiento.

En el área de la tecnología médica las regulaciones no solo plantean requisitos sobre el producto final, sino que también exigen, controlan y evalúan las estructuras y metodologías destinadas al diseño. La mejor forma de satisfacer esas regulaciones es logrando el cumplimiento de normas específicas, definidas por la industria y aceptadas por los organismos de control.

El conjunto de normas aplicables depende de las características del producto. En particular nos enfocamos en el marco regulatorio que cubre al producto electromédico para monitorizar la profundidad de anestesia. El desarrollo hecho en el ámbito de la universidad en los proyectos “PICTO 0087 Monitor de profundidad de anestesia” y “C2-ING-061 Comunicación inalámbrica de baja energía para aplicaciones electromédicas” consta de un sistema de adquisición de neuroseñales inalámbrico (Fig.1, Cabezal) con una pantalla de monitorización (Fig.1, Monitor) dentro del área

de uso y otra interfaz remota (Fig.1, Aplicación). Se trata de un dispositivo médico de aplicación en quirófano y terapia intensiva [1].



**Fig. 1.** Arquitectura del Monitor de profundidad de anestesia

El modelo de negocio al que responde esta arquitectura es el de telemedicina aplicada a un subsector de la salud en el que hay una dependencia absoluta del profesional de manera presencial. La anestesia en las intervenciones quirúrgicas y la sedación en terapia intensiva se hicieron durante muchas décadas mediante un profesional médico especializado en anestesiología. Este ajustaba la dosis para obtener el estado anestésico del paciente deseado, midiendo variables hemodinámicas como el ritmo cardíaco y la presión arterial; y parámetros clínicos como la sudoración, la piloerección, la dilatación de las pupilas, la sialorrea, y los movimientos espontáneos [2]. Hace unos cuantos años vienen usándose monitores de profundidad anestésica basados en parámetros objetivos que complementan la observación del profesional. Existen pocos fabricantes en el mundo. La mayoría usa señales electrofisiológicas que procesadas por un algoritmo propietario, dan un indicador dentro de una escala y registran los valores durante toda la maniobra anestésica con fines legales y de historia clínica [3].

La arquitectura propuesta con un “Cabezal” que registra y adecúa la señal junto a un “Monitor” con conectividad (que podría ser un Smartphone u otra plataforma con software embebido) permitirían sacar del área de quirófano al profesional de anestesia, para que actúe, verifique, avale, revise y registre de manera remota el efecto de la dosis suministrada al paciente. En un caso superador de telemedicina podría también ajustar la dosis de manera remota, si es que la mesa de anestesia lo permite

Considerando el producto electromédico descrito, el conjunto de normas aplicables para gestionar su seguridad y eficacia, abarca desde el sistema de gestión de calidad (ISO13485) [4], la gestión de riesgos (ISO 14971) [5], la de usabilidad (IEC 62366) [6], la del proceso del ciclo de vida del software (IEC 62304) [7], la de seguridad básica y funcionamiento esencial (IEC 60601-1, cuyo punto 14 se encarga del software del producto electro medico o PEMS) [8] y la de seguridad del software que puede funcionar autónomamente (IEC 82304-1) [9]. La relación entre estas normas y el producto se puede ver en la Fig. 2.

La gestión de riesgos es un estudio sistemático y predefinido hecho por el fabricante que busca conocer de antemano la mayor cantidad posible de los peligros que puedan afectar al usuario o paciente al entrar en contacto con el dispositivo médico. Atraviesa todo el ciclo de vida del producto, desde la concepción de la idea, las diferentes etapas del diseño, la producción, la posproducción, el uso, el mantenimiento y finaliza con la retirada del producto del mercado. Busca también que se valoren los peligros ponderándolos como riesgos para así poder justificar cual será la acción que se tomará con cada uno de ellos, con el fin de eliminarlos, mitigarlos o aceptarlos. La ISO 14971 brinda un marco de referencia y propone metodologías para abordar ese estudio sistemático para encontrar peligros, evaluar su gravedad y su probabilidad de ocurrencia, encontrando así riesgos. Requiere que el fabricante defina

critérios de aceptación o rechazo de esos riesgos, para definir cuales de ellos serán mitigados con soluciones técnicas que este proponga. La norma también establece requisitos de trazabilidad entre los peligros y entradas al diseño. La IEC 62304-1 establece el marco de responsabilidades, registro y verificación para cada una de las etapas del proceso del ciclo de vida del software, incluyendo obviamente las entradas de diseño que surgen de la gestión de riesgo y especialmente sobre como se trata la realimentación que brinda el proceso de resolución de problemas (bugs). Así como la norma de gestión de riesgos alimenta las entradas del proceso de diseño regido por la de ciclo de vida de software, la ISO 14971 se alimenta también de requisitos de seguridad y eficacia planteados por la IEC 60601-1 y la IEC 82604-1 que actúan como disparadores para el análisis de los riesgos. Estas normas son específicas de producto y proponen requisitos destinados a controlar peligros que están reconocidos por la industria.

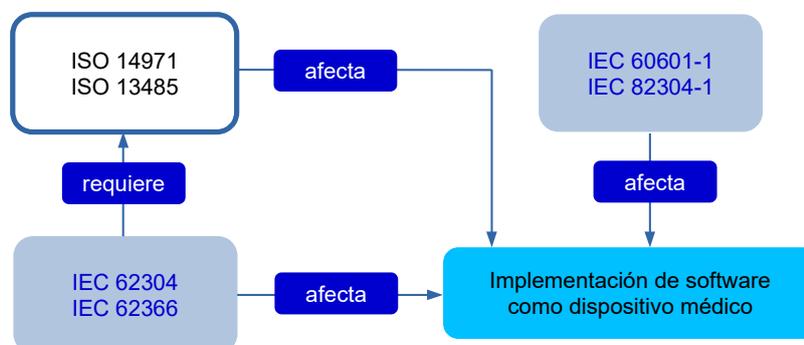


Fig. 2. Relaciones entre normas aplicables al producto, extraído y adaptado de IEC 62304-1 AMD:2015

Uno de esos peligros conocidos es la vulnerabilidad ante ataques a la ciberseguridad (Fig.3). Si bien es un elemento no definido en la norma de equipamiento médico [8] que se encarga estrictamente de los peligros relacionados con la seguridad (safety: prevención de accidentes sin intencionalidad), si se hace mención específica en la norma de software autónomo [9] considerando también como fuente de peligros a la amenaza proveniente de la falta de protección del dispositivo (security: prevención de actividades maliciosas intencionadas).



Fig. 3. Impacto de las medidas de protección en la seguridad, adaptado de TIR57 [10]

Esta diferencia que no era considerada hace unos años, hoy con el avance de la conectividad de los productos médicos hace que el problema de la ciberseguridad esté tomando cada vez mas relevancia [11]. Esto significa un cambio de paradigma en el enfoque de la gestión de riesgos, en donde transversalmente al diseño se debe considerar la protección del sistema como una entrada más al diseño.

El objetivo de este trabajo es mostrar cual fue el impacto que causó en la arquitectura de un sistema electromédico cuando en la gestión de riesgos se incorporaron herramientas de análisis de la ciberseguridad.

## 2 Métodos

Para hacer la gestión de riesgos de todo el producto médico se usó la herramienta Análisis Modal de Fallos y Efectos (AMFE) enunciándose aquí únicamente aquellos ítems relacionados con la ciberseguridad.

Las siguientes tablas establecen criterios para la aceptabilidad del riesgo. Las Tablas 1 y 2 dan una guía la valoración de la gravedad y la ocurrencia. La Tabla 3 establece los límites que definen cuando un riesgo es aceptable o cuando es inaceptable

**Tabla 1.** Guía para valoración de la Gravedad:

Gravedad	Consecuencias asociadas	Valor
Catastrófico	Origina la muerte del paciente	5
Crítico	Origina un deterioro permanente o una lesión que pone en peligro la vida	4
Serio	Origina una lesión o un deterioro que requiere intervención médica profesional	3
Pequeño	Origina una lesión temporal o un deterioro que no requiere intervención médica profesional	2
Insignificante	Inconveniente o molestia transitoria	1

**Tabla 2.** Guía para valoración de la Ocurrencia:

Probabilidad	Frecuencia del fallo	Valor
Frecuente	Fallo casi inevitable. Es seguro que el fallo se producirá frecuentemente.	5
Probable	El fallo se ha presentado con cierta frecuencia en el pasado en procesos similares o previos procesos que han fallado.	4
Ocasional	Defecto aparecido ocasionalmente en procesos similares o previos al actual. Probablemente aparecerá algunas veces en la vida del componente/sistema.	3
Remoto	Fallos aislados en procesos similares o casi idénticos. Es	2

	razonablemente esperable en la vida del sistema, aunque es poco probable que suceda.	
Improbable	Ningún fallo se asocia a procesos casi idénticos, ni se ha dado nunca en el pasado, pero es concebible.	1

**Tabla 3.** Guía para aceptar el Riesgo A = ADMISIBLE; I = INADMISIBLE:

Ocurrencia	Gravedad	Improbable	Remoto	Ocasional	Probable	Frecuente
		1	2	3	4	5
Catastrófico	5	A	I	I	I	I
Crítico	4	A	A	I	I	I
Serio	3	A	A	I	I	I
Pequeño	2	A	A	A	I	I
Insignificante	1	A	A	A	A	I

Para la detección de los peligros se usó la técnica de análisis de situaciones relacionando los aspectos de uso de cada uno de los componentes del sistema y las amenazas de ciberseguridad a las que están expuestos. Estas amenazas pueden ser:

- Interrupción de la atención / servicio
- Engaño del personal con correo electrónico falso o sitios web falsos para obtener credenciales de inicio de sesión o instalar malware
- Amenaza interna, involuntaria o intencional, que puede representar una amenaza significativa debido a la posición de confianza dentro de la organización
- Pérdida de información del paciente
- Violación de datos, filtración de información y pérdida de privacidad
- Chantaje, extorsión y coacción a través de la explotación de datos filtrados.

El análisis permitió descubrir los peligros, y estos se registraron y sistematizaron como ítems en el AMFE de la sección resultados. De esta manera se identificó cual

es la secuencia de sucesos que desencadena una situación peligrosa causando un daño específico (columnas: Secuencia de sucesos, Situación peligrosa, Daño). También se ponderó el riesgo mediante las tablas 1, 2 y 3 antes mencionadas (columnas: O, G, R) y se propuso la medida de control que busca que el riesgo baje a niveles aceptables (columna: Mitigación). Como cada situación peligrosa debe ser trazable se la identificó alfa numericamente (columna: ID).

### 3 Resultados

Los resultados del análisis de riesgos y soluciones propuestas para mitigar los riesgos se enuncian en la Tabla 4 en el formato tradicional de un AMFE, aceptado por la industria y por los organismos regulatorios.

Se analizaron las secuencias de sucesos para ataques que logran modificar parámetros de funcionamiento, suplantando dispositivos y usuarios, detienen la acción del sistema, corrompen su funcionamiento y/o generan nuevas vulnerabilidades.

Los daños que responden a cada uno de los peligros detectados se circunscriben a la falta de monitorización tanto presencial como remota, a una posible dosificación incorrecta del agente anestésico y a la pérdida de privacidad de los pacientes.

Se proponen las medidas de mitigación que buscan disminuir el riesgo mediante un blindaje del sistema minimizando las vías de conectividad lo que implica suprimir la aplicación remota. Las propuestas incluyen mejorar la protección de partes vitales del software impidiendo el acceso remoto, permitiendo solo el resguardo de datos de manera física y la ejecución de actualizaciones solo por personal autorizado. Las medidas de control también impactan en la selección del hardware donde ya no se permite la selección de una plataforma de uso general y se exige una plataforma de uso exclusivo para cumplir con la función de monitor.

**Tabla 4.** Extracto del AMFE con los ítems relacionados a la ciberseguridad con riesgo Inadmisibles (I). MPA = Monitoreo de Profundidad Anestésica

ID	Secuencia de sucesos	Situación peligrosa	Daño	O	G	R	Mitigación
CS01	El monitor sufre un ataque que modifica parámetros de funcionamiento	El cabezal interrumpe su transmisión	MPA no disponible	3	3	I	Desactivar conectividad monitor red externa. Protección contra escrituras y CRC de partes vitales.
CS02	El cabezal del cabezal	El cabezal envía datos erróneos	Dosificación incorrecta de anestesia	2	5	I	Resguardo de datos presencial (USB). Monitor en

CS03	El monitor sufre un ataque que modifica sus propios parámetros de funcionamiento	El monitor genera e informa valores espurios al usuario local y/o al usuario remoto de la aplicación	Dosificación incorrecta de anestesia	2	5	I	plataforma exclusiva.
CS04	Suplantación de dispositivo cabezal por comunicación Bluetooth	El monitor recibe datos espurios interrumpiendo su funcionamiento	MPA no disponible	3	3	I	Identificación de cabezal encriptada periódica
CS05		El monitor recibe datos espurios y continúa funcionando	Dosificación incorrecta de anestesia	2	5	I	
CS06	Suplantación de dispositivo monitor	La aplicación recibe valores espurios	Dosificación incorrecta de anestesia	2	5	I	
CS07	Suplantación de usuario en aplicación remota	Un tercero toma control de la maniobra de dosificación	Dosificación incorrecta de anestesia	2	5	I	
CS08		Un tercero se apodera de la base de datos de pacientes.	Perdida de privacidad de pacientes	4	3	I	Suprimir la aplicación remota
CS09	La aplicación sufre un ataque	La aplicación no inicia, se desinstaló o muestra error.	MPA remoto disponible	4	3	I	
CS10		La aplicación muestra valores espurios	Dosificación incorrecta de anestesia	2	5	I	
CS11	Una actualización de alguna parte del sistema no controlada corrompe el funcionamiento o genera vulnerabilidades	El sistema queda fuera de funcionamiento	MPA no disponible	3	3	I	Permitir actualizaciones validadas del sistema solo de forma presencial en fábrica o por personal autorizado
CS12		El sistema queda vulnerable a ataques	MPA no disponible Dosificación incorrecta de anestesia	2	5	I	

## 4 Conclusiones

Las medidas para mitigar los peligros que surgen de los ciber-ataques, pueden parecer extremas a primera vista (Fig. 4). Todas las medidas de control no solo deben implementarse sino que también deben validarse. La validación de una solución muchas veces implica un esfuerzo mayor al de la implementación misma. Este motivo justifica lo radical de las soluciones planteadas que obligaron a abandonar el modelo de negocio de la telemedicina y blindaron al dispositivo de monitoreo de toda conexión externa.

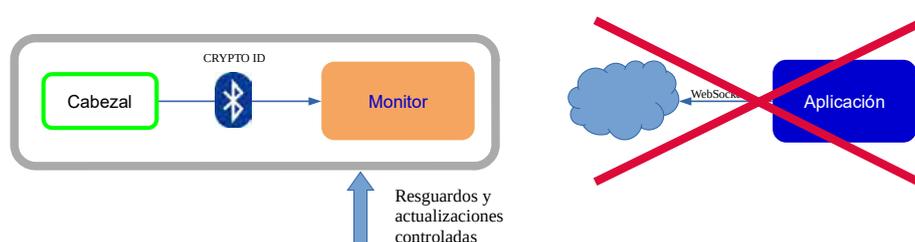


Fig. 4. Cambios en la Arquitectura del Monitor de profundidad de anestesia

Es importante destacar que no es imposible implementar y validar soluciones mejores. En este caso y en particular para la organización encargada de hacer el diseño, resultó estar muy por encima de sus posibilidades y recursos el hecho de proporcionar de manera objetiva evidencias que permitieran llevar adelante estrategias superadoras. Esta situación puede extenderse a muchos de los fabricantes de equipamiento médico tecnológico local que son PyMES [12]. Sus limitados recursos destinados al I+D+i están más inclinados a satisfacer los requisitos de funcionamiento y a mejorar las soluciones clínicas que a la protección de los sistemas. La incorporación de la ciberseguridad implica un cambio de paradigma donde más temprano que tarde los expertos en equipamiento médico deberán volverse también expertos en protección de sistemas médicos para hacer productos sostenibles en un mercado con cada vez más avidez de dispositivos conectados.

## Referencias

1. D. Coulombie, F. Orthusteguy, A. Reyes, F. Ortalda. Monitor de Profundidad Anestésica , proyecto. Argentina. Buenos Aires. 2017. Libro. Artículo Breve. Workshop. WICC 2017 XIX Workshop de Investigadores en Ciencias de la Computación. Instituto Tecnológico de Buenos Aires.
2. Luis Miguel Torres Morera, “Tratado de anestesia y reanimación”, Madrid Arán cop. (2001)
3. J.Bruhn etal; “Depth of anaesthesia monitoring: what's available, what's validated and what's next?” British Journal of Anaesthesia,(2006), 97(1):85-94
4. ISO 13485:2016 Medical devices Quality management systems Requirements for regulatory purposes (2016)
5. ISO 14971:2019 Medical devices Application of risk management to medical devices

6. IEC 62366-1:2015 Medical devices Part 1: Application of usability engineering to medical devices.
7. IEC 62304:2006 Medical device software Software life cycle processes
8. IEC 60601-1:2005+AMD1:2012 CSV Consolidated version Medical electrical equipment Part 1: General requirements for basic safety and essential performance
9. IEC 82304-1:2016 Health software Part 1: General requirements for product safety
10. AAMI TIR57: 2016 Principles For Medical Device Security - Risk Management
11. Williams, Patricia & Woodward, Andrew. (2015). Cybersecurity vulnerabilities in medical devices: A complex environment and multifaceted problem. Medical devices (Auckland, N.Z.). 8. 305-16. 10.2147/MDER.S50048.
12. F. Porta, G. Baruj, Nucleo Socio Productivo Estratégico Equipamiento Médico- Documento de Referencia, Argentina Innovadora 2020, 2012 Ministerio de Ciencia Tecnología e innovación productiva,

# TAPIR: An Object-Oriented Programming Testing Framework based on Message Sequence Specification with Aspect-Oriented Programming

Martín L. Larrea and Dana K. Urribarri

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}@cs.uns.edu.ar  
<http://vyglab.cs.uns.edu.ar>

**Abstract.** Today we can see a significant increase in testing tools available for both the developer and the tester. Those tools aim at the testing of system implementation and vary according to the implementation paradigm, the programming language or the type of errors they seek to detect. In this paper, we present TAPIR, a white-box testing framework for Object-Oriented Programming. It was designed and implemented to detect failures in the sequence of calls that objects make. In that sense, we rely on Message Sequence Specification and Aspect-Oriented Programming. Hence, TAPIR can be used in any Java project without the need to modify the implementation of such a project. Our framework is open source and is freely available.

**Keywords:** Verification and Validation, Testing, Message Sequence Specification, Aspect-Oriented Programming

## 1 Introduction

In 1994, Kirani and Tsai [2] presented a technique called Message Sequence Specification (MSS) that, in the context of an Object-Oriented program, describes the correct order in which the methods of a class should be invoked. The MSS associated with an object specifies all sequences of messages that the object can receive while still providing correct behavior. In this paper, we present TAPIR a framework that implements the concept of MSS as a dynamic testing tool. The framework was designed to detect failures in the sequence of calls made by objects. In that sense, we rely on the concept of Message Sequence Specification. Its implementation was done using Aspect-Oriented Programming [3] (AOP) and, therefore, it can be used in any Java project without the need to modify the implementation of the project. The first version, 1.0, of the framework was

introduced in [7] and then 2.0 was described in [4]. Its latest version, 3.0, is presented in this paper. The change track of each version can be summarize here; Version 1.0 was limited to test only one instance of one class in the system. It was suitable only for usages where the singleton pattern was appropriate. Version 2.0 could test multiple instances of a single class in the system. Version 3.0, the one presented here, can test multiple instances of multiples classes in the system.

## 2 A review of Message Sequence Specification

The MSS [2] associated with an object specifies all sequences of messages that the object can receive while still providing correct behavior. Their strategy used regular expressions to model the constraints over the correct order of the method invocation. Method names were used as the alphabet of the expression which was then used to statically verify the program's implementation for improper method sequences. A runtime verification system identifies incorrect method invocations by checking for sequence consistency with respect to the sequencing constraints. According to Kirani's specification, if a class  $C$  has a method  $M_1$ , this is noted as  $C_{M_1}$ . Sequence relationships between two methods were classified into three categories, sequential, optional, and repeated. If the method  $M_1$  of  $C$  should be invoked before the method  $M_2$  of the same class, then this relationship is sequential and is represented as  $C_{M_1} \bullet C_{M_2}$ . If only one of the methods  $M_1$  and  $M_2$  can be invoked, then this relationship is optional and is represented as  $C_{M_1} | C_{M_2}$ . Finally, if the method  $M_1$  can be invoked many times in a row then this is a repeated relationship and is represented as  $(C_{M_1})^*$ . For example, if a class  $X$  has three methods called *create*, *process*, and *close*, a possible sequencing constraint based on MSS could look like  $X_{create} \bullet (X_{process})^* \bullet X_{close}$ . If class  $X$  is part of a larger system  $S$ , then we could statically check the source code of  $S$  to see if all calls to  $X$ 's methods follow the defined expression. If a static analysis is not enough, we could implement a runtime verification system that tracks all calls to  $X$ 's methods and dynamically checks the sequence of calls against its regular expression. This technique can also be used to test the robustness of a system. Continuing with class  $X$  as an example, we can use the defined regular expression to create method sequences that are not a derivation from it, i.e. incorrect method sequences. These new sequences can be used to test how the class handles a misuse. For example, how does class  $X$  respond to the following sequence of calls?:  $X_{create} \bullet X_{close} \bullet X_{process}$

## 3 Previous Applications of Message Sequence Specification

Daniels & Tsai [1] extended the work of Kirani et al. [2] by testing with some sequences generated by the expression and others not generated by the expression. Also in 1999, Tsai et al. [9] presented Message Framework Sequence Specifications (MfSS), for generating scenario templates that can be used to generate

test cases to test applications developed using extensible design patterns and an object-oriented framework. This framework aim to test the dynamic typing and dynamic binding of an Object-Oriented program. MfSS involves with multiple objects, thus a sequence expression must use the object name together with its method name. The framework is suitable for the application of several testing technique such as positive testing, negative testing, test slicing, partition testing, boundary testing, random testing and stress testing based on scenario templates. The MSS presented in that paper is far more expressive than the one we are presenting but, as far as we know, there is no tool or software that implement this methodology. The application of the MfSS shown in the original paper was done by hand and not with a software. In 2003, Tsai [8] used MSS as a verification mechanism to the UDDI servers in the context of Web Services (WS). WS are particularly interoperative between each other. This kind of complex relationship was expressed using the MfSS presented in [9]. This approach was part of a larger set of testing mechanism for WS presented in that work. Unfortunately, only a partial part of the approach was implemented and it did not include the MfSS. In 2014, a Java-based tool for monitoring sequences of method calls was introduced [6], it had similar objectives as our work but they used annotations instead of AOP. In their work, they used annotations to specify method-call sequences in terms of regular expressions of method-call signatures. They included an implementation of the proposal, in the form of a tool called JMSeq. This tool runs alongside the application under test but in a different java virtual machine, hence there are two virtual machine running at the same time: one for the program under test and the other for the JMSeq execution. Since the annotations must be included in the source code to test, the source code must be modified, and, potentially, new errors may arise. In our approach, AOP avoids the need of modifying the source code under test, hence reducing the possibility of new bugs. We introduced MSS as a black-box technique for testing visualizations interactions [5] in 2018. The technique is built on constraints imposed over the sequences of low-level interactions available in the visualization with User Action Notation and MSS. Instead of specifying a sequence constraint on the methods of a class, in this work we specify a Sequence Constraint on the Interactions (SCI) available in the visualization. This research also included the definition of coverage criteria for the technique, both for valid and invalid sequences. The work aims to generate two types of testing tools: one to dynamically test the correct usage of a visualization tool, i.e., while the user is using the visualization our method checks that interactions are being used accordingly to the SCI; and the other, to generate the test cases based on the SCI. So far, this research is only available as theory without a proper implementation. Turner used sequence specification for GUI testing in her Ph.D. thesis in 2019, with previous publications about this topic [10], [11]. In these articles, interaction sequences are used as an abstraction of the interactive system to inform a model-based testing approach using lightweight formal methods. Interaction sequences provide an abstract view of the point at which the functional and interactive components

intersect. The formalization of interaction sequences and the modelling of those sequences was done using Finite State Automata.

## 4 TAPIR - An Object-Oriented Programming Testing Framework

TAPIR is a testing framework for object-oriented source code based on MSS using AOP. AOP allows us to create test cases without modifying the source code, and those test cases run automatically for every execution of the program under test. The use of MSS allows the developer to describe a regular expression for each class which represents its correct behavior. The framework takes each of these expressions, runs the program and checks that the methods are used according to its class specification. TAPIR can be classified as a dynamic analysis testing tool. A major feature of our framework is to be easy to use, with an easy to read and understand representation of the correct usage of each classes methods. Particularly, the framework was designed to be used by the developer, without the need of a testing specialist. The first thing the developer must do to use the framework is to create the regular expressions associated with the classes under test. These regular expressions must specify the correct behavior or order in which the methods of the classes should be called. In order to express this in a simple way, the developer must use symbols (i.e. characters) to represent each method. This means that the actual names of the methods are not used in the expressions. But, to be able to interpret it at some point the developer must create a map between the actual methods' names and their corresponding symbol. Any method not included in the class's regular expression is ignore by the framework. Hence, the developer is not required to use all the class's methods in the regular expression. The developer must also specify how he/she wants the framework to behave in the event of an error. When the framework detects a sequence of calls that is not derived from its associated regular expression, it reports the error and can abort the execution or allow it to continue. This decision is in the hands of the developer and it can be specified independently for each defined regular expression. The regular expressions and the maps between methods and symbols are set in the *TestingSetup.java* class. The framework consists of two main components, an aspect, and a java class. The aspect is named *TestingCore.aj* and it contains the implementation of the framework's core. Listing 1.1 shows an example with two classes: CA and CB. In this case, the correct order to use the CA class is: first, the object must be created. Then, there should be a call to *f()* followed by a call to *g()*. After that, there can be as many call as desired to either *g()* or *h()*. The final call of the sequence must be to *h()*. For the CB class, the correct use is: first, there should be a call to *alpha()* followed by a call to *gamma()* or, a call to *gamma()* follow by a call to *beta()*. Afterward, any method between alpha, beta or gamma can be called. The listing 1.1 shows how this information is input into TAPIR in the *TestingSetup.java* class.

Listing 1.1. TAPIR configuration

```

//Testing setup for CA class
//Definition of the methods and their corresponding symbols
mapObjectsToCallSequence = new HashMap<>();
mapMethodsToSymbols = new HashMap<String, String>();
mapMethodsToSymbols.put(" main.CA.<init>", "c"); mapMethodsToSymbols.put(" main
.CA.f", "f"); mapMethodsToSymbols.put(" main.CA.g", "g");
mapMethodsToSymbols.put(" main.CA.h", "h");
//Definition of the regular expression
regularExpression = Pattern.compile(" cfg(g|h)*h");
//Initializing the regular expressions controller
matcher = regularExpression.matcher("");
// All information related to how the class is tested is store in a TestingInformation
instance
TestingInformation ti = new TestingInformation(CA.class.toString(),
mapObjectsToCallSequence, mapMethodsToSymbols, regularExpression, matcher,
true);
TestingCore.mapClassToTestingInformation.put(CA.class.toString(), ti);
//Testing setup for CB class
//Definition of the methods and their corresponding symbols
mapObjectsToCallSequence = new HashMap<>();
mapMethodsToSymbols = new HashMap<String, String>(); mapMethodsToSymbols.
put(" main.CB.alpha", "a"); mapMethodsToSymbols.put(" main.CB.gamma", "g");
mapMethodsToSymbols.put(" main.CB.beta", "b");
//Definition of the regular expression
regularExpression = Pattern.compile(" (ag|gb)(a|g|b)*");
//Initializing the regular expressions controller
matcher = regularExpression.matcher("");
// All information related to how the class is tested is store in a TestingInformation
instance
ti = new TestingInformation(CB.class.toString(), mapObjectsToCallSequence,
mapMethodsToSymbols, regularExpression, matcher, false);
TestingCore.mapClassToTestingInformation.put(CB.class.toString(), ti);

```

In Listing 1.2, we can see the framework output when the code portion of Listing 1.1 corresponding to the CA class is executed. In this case, the last call to *f()* does not follow the MSS specified for the CA class. As mentioned above, when an error is detected, TAPIR informs this by console indicating the class and object that produced the error. The method that violated the MSS, the MSS in question and the actual sequence of calls are also shown in the console. Finally, the system abort the execution because this is what the last true parameter of method *TestingInformation* call indicates.

**Listing 1.2.** Error example for the CA class. The execution is aborted when the error is found.

```

CA ca1 = new CA();
ca1.f();
ca1.g();
ca1.h();

```

```

ca1.f());

---- ERROR FOUND ----
Class: class main.CA
Object Code: 977993101
Method Executed: main.CA.f
Regular Expression: cfg(g|h)*h
Execution Sequence: cfghf
----- SYSTEM ABORTING... -----

```

**Listing 1.3.** Error example for the CB class. The execution is allow to continue when the error is found.

```

CB cb1 = new CB();
cb1.alpha();
cb1.alpha();
cb1.gamma();
cb1.gamma();

---- ERROR FOUND ----
Class: class main.CB
Object Code: 859417998
Method Executed: main.CB.alpha
Regular Expression: (ag|gb)(a|g|b)*
Execution Sequence: aa
-- CONTINUING EXECUTION... ----
---- ERROR FOUND ----
Class: class main.CB
Object Code: 859417998
Method Executed: main.CB.gamma
Regular Expression: (ag|gb)(a|g|b)*
Execution Sequence: aag
-- CONTINUING EXECUTION... ----

```

Listing 1.2 shows the framework output when the code portion of Listing 1.1 corresponding to the CB class is executed. In this case, the second call to *alpha()* does not follow the MSS specification for class *CB*. As configured in Listing 1.1, the last false parameter in the call to method *TestingInformation* indicates that the execution must continue despite the existing errors. This is why, Listing 1.3 shows multiple errors. The next section shows how the framework can be useful in a more complex real-life situation.

## 5 Case Study. Earth Defender

Earth Defender is a video game developed in one of our programming courses at the university. It's a classic vertical shooter as Space Invader but with more modern features. It has 28000 lines of code distributed in 217 classes. It was developed in Java 1.8 with Eclipse IDE 4.4. There are three classes that are important for our case study; the *PlayerInteractionMananger* class is responsible

for capturing the user interactions that control the space ship's movements and shots. The *Enemy* class's responsibility is to manage each enemy ship in the game; in particular, how each ship return a power up when destroyed. Finally, the *GULGame* class deals with all the graphical user interfaces of the game. Because of the limitation of the number of pages in this congress, we are not showing how the *TestingSetup* class should be configure for this case. We are only focusing on the regular expressions and the errors founds.

For the *PlayerInteractionMananger* class, the three main methods to test are *playerStartMove*, *playerStopMove* and *playerShoot*. Its MSS is  $((a|b)^*x)^*$ , where *a* stands for *playerStartMove*, *b* for *playerStopMove* and *x* for *playerShoot*. This means that, under a correct behavior, the player can move several times before shooting. This can be repeated multiple times during a game session. Two methods are considered for the *Enemy* class's MSS, those are *takeDamage* represented with a *t* and *dropPowerUp*, with a *d*. The MSS for this class is  $t^*d$ ; this means that each enemy in the game can take a lot of damage until it drops its power up. Power up are dropped when enemies are destroyed. Finally, the *GULGame* class has the more complex MSS. Six methods are used: *inicializar* starts the GUI game level, *initLifeBar* draws the life bar on screen, *initScore* draws the initial score of the player on screen, *shoot* draws the shooting effects on screen, *changeLevel* starts the next level once the player completes the current one and, finally, *stopGame* deals with the player's death. These methods are represented by the symbols *a*, *b*, *c*, *h*, *i*, and *j*, respectively. The MSS for the *GULGame* class is  $((bca(h)^*)i)^*(bca(h)^*)j$ . The  $(bca(h)^*)$  part of the MSS represents the user playing a level. If the user wins the level, then he/she moves to the next level. This is shown in  $((bca(h)^*)i)^*$ . The game continues until the players dies, which is the second part of the MSS  $(bca(h)^*)j$ .

### 5.1 Testing Earth Defender

We configured TAPIR to stop the execution of the game only when an error is encounter on the *GULGame* class. If errors are founds on the other two classes, TAPIR will inform them but allow the execution to continue. On the first test run of the game, we played the first level and lost. When we finished playing we found that TAPIR had reported three errors. The output of the framework can be seen in Listing 1.4. The three errors correspond to three different objects of the *Enemy* class. We can affirm that they are different objects because the object code that is reported in each error is different. In all three cases, by observing the execution sequence, we can detect that once a ship delivers its power-up, meaning it was destroyed, it still takes new damage. This should not happen since the ship was already destroyed. An analysis of the code involved with this situation showed that the problem was how the enemies were removed from the game once destroyed. Each time an enemy's life reached 0, the power-up was first delivered, then the score and other features of the game were updated and, finally, the enemy was removed from the level. Since the instance is eliminated last, any other shot that was on the way could still impact it. Consequently, the code was modified to remove the enemy immediately after its life reached zero.

Listing 1.4. TAPIR output of the first test run of Earth Defender

```

---- ERROR FOUND ----
Class: class Entity.Enemy
Object Code: 1256786520
Method Executed: Entity.Enemy.takeDamage
Regular Expression: t*d
Execution Sequence: ttttdt
-- CONTINUING EXECUTION... ----
---- ERROR FOUND ----
Class: class Entity.Enemy
Object Code: 1337065869
Method Executed: Entity.Enemy.takeDamage
Regular Expression: t*d
Execution Sequence: ttttdt
-- CONTINUING EXECUTION... ----
---- ERROR FOUND ----
Class: class Entity.Enemy
Object Code: 1128333601
Method Executed: Entity.Enemy.takeDamage
Regular Expression: t*d
Execution Sequence: ttttttdt
-- CONTINUING EXECUTION... ----

```

We successfully completed the first level, however, since TAPIR detected errors in the *GUI.class* and it was configured to abort execution in that case, the game closed abruptly at the end of the level. The output of TAPIR can be seen in Listing 1.5. Two errors were found for the same instance of the *GUI.Game* class. In this case we can see that the object code reported are the same. We believe that TAPIR reported two errors that should abort the execution because while processing the first error, the second error was generated. In the first error, the method that causes the error is *initLifeBar*, represented by the symbol *b*. The last shot of the player can lead to two possible actions later, if with that last shot he won the level then the *changeLevel* method must be executed; on the other hand if after that last shot the player dies then the *stopLevel* method must be executed. As can be seen in the first error report detected by TAPIR, after the last shot that is represented in the sequence with *h* the method *initLifeBar* was executed. As we know that we successfully completed the level, the correct thing would have been the execution of the method *changeLevel*, symbol *i*. The second error reported is a consequence of the first, so correcting the first should fix the second. After inspecting the code, it was discovered that the method *initLifeBar* was called twice each time a level ended. This error was not noticeable from the graphical interface. One of the calls was made after changing levels, this call was correct; the second call was made as part of the logic that prepares the level change. This line of code should have been deleted in a game update but it was not. By removing this line, both errors detected by TAPIR were corrected. Three more executions of the game were made after this and the framework reported no errors.

**Listing 1.5.** TAPIR output on the first test run of Earth Defender

```
---- ERROR FOUND ----
Class: class GUI.GUI_Game
Object Code: 103210007
Method Executed: GUI.GUI_Game.initLifeBar
Regular Expression: ((bca(h)*i)*(bca(h)*j)
Execution Sequence:
bcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
----- SYSTEM ABORTING... -----
---- ERROR FOUND ----
Class: class GUI.GUI_Game
Object Code: 103210007
Method Executed: GUI.GUI_Game.shoot
Regular Expression: ((bca(h)*i)*(bca(h)*j)
Execution Sequence:
bcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
----- SYSTEM ABORTING... -----
```

## 6 Conclusions & Future Work

As previously mentioned, our goal was to create an easy-to-use framework for developers to test their Java source code without any modification to it, hence reducing the possibilities of introducing new errors and minimizing preparation time. The objective of this framework is to find those vulnerabilities that allow breaking the expected order of execution of the methods of a class. As shown in the case study, the framework helps detect errors that would otherwise be difficult to find. Future work will consider a more expressive framework. For the moment, the framework can only test the order in which methods of a class are being called. However, the actual values of the parameters of a method or the inner state of the instance can also be relevant in the execution order. For example, a class could require a method  $x$  to be called after method  $y$  if the value of a particular attribute is equal to 0. In addition, we will look for the possibility of defining a MSS that involves methods of different classes; i.e., being able to combine methods of different classes in a common MSS. We will also begin to translate TAPIR into other programming languages, such as C# and Python. The framework is available for downloading<sup>1</sup>. The source code is available and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

### Acknowledgment

This work was partially supported by the following research projects: PGI 24/N050 and PGI 24/ZN35 from the Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur, Argentina.

<sup>1</sup> <http://cs.uns.edu.ar/~mll/lapaz/>

## References

1. FJ Daniels and KC Tai. Measuring the effectiveness of method test sequences derived from sequencing constraints. In *Proceedings of Technology of Object-Oriented Languages and Systems-TOOLS 30 (Cat. No. PR00278)*, pages 74–83. IEEE, 1999. doi:10.1109/TOOLS.1999.787537.
2. Shekhar Kirani and W. T. Tsai. Specification and verification of object-oriented programs. Technical report, Computer Science Department, University of Minnesota, 1994.
3. Ramnivas Laddad. *AspectJ in Action: Practical Aspect-Oriented Programming*. Manning Publications Co., Greenwich, CT, USA, 2003.
4. Martín L Larrea, Juan Ignacio Rodríguez Silva, Matías N Selzer, and Dana K Urribarri. White-box testing framework for object-oriented programming. an approach based on message sequence specification and aspect oriented programming. In *Argentine Congress of Computer Science*, pages 143–156. Springer, 2018.
5. Martín Leonardo Larrea. Black-box testing technique for information visualization. sequencing constraints with low-level interactions. *Journal of Computer Science & Technology*, 17, 2017.
6. Behrooz Nobakht, Frank S de Boer, Marcello M Bonsangue, Stijn de Gouw, and Mohammad Mahdi Jaghoori. Monitoring method call sequences using annotations. *Science of Computer Programming*, 94:362–378, 2014. doi:10.1016/j.scico.2013.11.030.
7. Juan Ignacio Rodríguez Silva and Martín Larrea. White-box testing framework for object-oriented programming based on message sequence specification. In *XXIV Congreso Argentino de Ciencias de la Computación (Tandil, 2018)*., pages 532–541, 2018.
8. Wei-Tek Tsai, Ray Paul, Zhibin Cao, Lian Yu, and Akihiro Saimi. Verification of web services using an enhanced uddi server. In *Proceedings of the Eighth International Workshop on Object-Oriented Real-Time Dependable Systems, 2003.(WORDS 2003)*., pages 131–138. IEEE, 2003. doi:10.1109/WORDS.2003.1218075.
9. Weik-Tek Tsai, Yongzhong Tu, Weiguang Shao, and Ezra Ebner. Testing extensible design patterns in object-oriented frameworks through scenario templates. In *Proceedings. Twenty-Third Annual International Computer Software and Applications Conference (Cat. No. 99CB37032)*, pages 166–171. IEEE, 1999. doi:10.1109/CMPSAC.1999.812695.
10. Jessica Turner, Judy Bowen, and Steve Reeves. Simulating interaction sequences. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, page 8. ACM, 2018.
11. Jessica Turner, Judy Bowen, and Steve Reeves. Using abstraction with interaction sequences for interactive system modelling. In *Federation of International Conferences on Software Technologies: Applications and Foundations*, pages 257–273. Springer, 2018.

## Risk refinement in the deployment process of software systems: a case study

Felipe Ortiz<sup>1</sup>, Marisa Panizzi<sup>1,2</sup> Rodolfo Bertone<sup>3</sup>

<sup>1</sup> Master's Program in Information Systems Engineering. Graduate School. Universidad Tecnológica Nacional. Regional Buenos Aires. Castro Barros 91. (C1178AAA). CABA. Argentina.

<sup>2</sup> Department of Information Systems Engineering. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires. Medrano 951 (C1179AAQ), C.A.B.A, Argentina.

<sup>3</sup> Instituto de Investigación en Informática [Information Systems Research Institute]- III-LIDI. School of Information Systems. Universidad Nacional de La Plata (UNLP) 50 y 120 – La Plata, Argentina.

[ortizfd@gmail.com](mailto:ortizfd@gmail.com); [marisapanizzi@outlook.com](mailto:marisapanizzi@outlook.com); [pbertone@lidi.unlp.edu.ar](mailto:pbertone@lidi.unlp.edu.ar)

**Abstract.** Deployment is the process by which a software system is transferred to a business client. A risk is defined as the likelihood for a loss to occur. In a software project, a risk might imply decreased quality of the software product, increased costs, a delay in project completion or a flaw, among others. A case study is developed with the aim to refine the set of risks. Furthermore, procedures are proposed for their prevention, mitigation and/or transfer for the software system deployment process. This article presents the results of a case study which analyzed the documentation related to deployment of functionalities in a bank's Human Resources Portal conducted by an Argentina-based software Small and Medium Enterprise (SME<sup>1</sup>).

**Key words:** software system deployment process, risk management, case study.

### 1 Introduction

There are various factors that can affect software projects, such as modifications in priorities and inadequate planning [1]. One of the most important factors might be unmanaged risks. A risk is the probability for a loss to occur. In a software project, such loss might take the form of decreased quality of the software product, increased development costs, a delay in project completion or a flaw [2].

A large number of projects lack formal approaches for risk management. The identification thereof usually depends, at an informal level, on the abilities and level of experience of software managers [3]. Although software risk management plays a key role in successful project management, it is usually not properly implemented in real world software projects, particularly in SMEs in Argentina [4].

<sup>1</sup> Presidencia de la Nación. (2020). <https://www.argentina.gob.ar/noticias/nuevas-categorias-para-ser-pyme>. Last updated on 07/06/2020.

Software system deployment is the phase of the development life cycle in which the software product is transferred to the client. The deployment process entails practices which tend to pose problems, such as the lack of components (generally external), incomplete downloads and faulty installations [5].

The problems that might arise in the deployment phase are transferred and they are eventually resolved during the maintenance phase. For this reason, an efficient software deployment process should save resources in terms of costs and effort [6].

Software deployment is usually conducted in distributed and heterogeneous environments, which add complexity, thus causing time consumption and additional costs [7]. Deployment entails a series of changes at several levels: processes, working methods, technology and organizational structure [8].

According to Reascos Paredes et al. [9], the main causes of technological risks include heterogeneous and incompatible infrastructure, SMEs' poor technological capabilities and competences, the complexity of these systems, and bad data quality and safety.

Forbes et al. [10] argue that the results of non-standardized and inadequate deployment practices are reflected in the information systems, which are difficult to maintain and operate.

This work presents the results of a case study aimed at refining (if necessary) the set of risks, as well as the procedures for their prevention, mitigation and/or transfer defined for the deployment process of software systems.

This article is organized as follows: related works are described in section 2; section 3 presents the set of risks for the deployment process; section 4 addresses the case study; and finally, section 5 presents the conclusions and future works.

## 2 Related works

A Systematic Mapping Study (SMS) was performed to build the state of the art on risk management for the deployment process of software systems [11]. After analyzing 100 primary studies, it was found that the most commonly used methodologies, methods and standards addressing risk management are CMMI [12], PMBOK [13] and SOFTWARE RISK EVALUATION [14].

To complement the SMS, a comparative analysis of the previously mentioned methodologies, methods and standards was conducted based on the DESMET method characteristics [15]. MAGERIT [16] was added to the comparison since it is one of the pioneering risk management methodologies [17].

The comparative analysis for the deployment addressed three dimensions: "Process", "Person" and "Product" [18]. After this comparative analysis, it was concluded that in the "Process" dimension all the methodologies, methods and standards analyzed address the risks for the deployment process. In the "Product" dimension, SOFTWARE RISK EVALUATION as well as PMBOK and MAGERIT include the risks of the deployment process while CMMI does not. Finally, in the "Person" dimension, none of the methodologies, methods or standards evaluated address the risks of the deployment process.

### 3 Risks of the deployment process

The activities and tasks considered for the definition of the risks of the deployment process are those stated in the technical process called “Transition” of the ISO/IEC/IEEE 12207:2017 standard [19]. This standard was chosen because it is internationally recognized. The activities and tasks are detailed in [20].

The risk classification used is the one proposed in [3], with adjustments made considering the evolution of software engineering in the last few decades and the deployment process of software systems. For risk weighting, the proposal established in the ISO/IEC 31010:2009 standard [21] is adopted, since it is one of the main international references in terms of risk management for the software industry.

The definition of risks was established considering a three-dimensional approach, given by the “Process” dimension, the “Person” dimension and the “Product” dimension [18]. The risks proposed for these three dimensions are described in [20].

### 4 Description of the case study

This section presents a detailed account of the case study following the guidelines proposed in [22].

#### 4.1 Case study design

The main objective is to examine the feasibility of the application of a set of risks, as well as the procedures for their prevention, mitigation and/or transfer in the deployment process of software systems in a real environment with the aim to refine them (if necessary). According to Robson's classification [23], case studies fall under the scope of exploratory studies. We worked with documentation related to the deployment of capability deliverables for a bank's Human Resources Portal performed by an Argentina-based software SME.

#### 4.2 Research questions

In order to address the objective of this study, the following research questions (RQ) are posed:

*RQ1: How were risks managed during the activities of the software system deployment process (identification, analysis and severity)?*

This question is intended to provide information about the risks encountered during the execution of the deployment process and the treatment provided by the consulting company in order to compare them with the proposal made.

*RQ2: How can the software system deployment process be strengthened in this company?*

This question is intended to determine the way in which the consulting company can enhance its deployment process. For this purpose, the identification of a set of

risks is proposed, along with the procedures for their prevention, mitigation and/or transfer.

### 4.3 Case and unit of analysis

This section describes the context, the case and the unit of analysis of the case study. According to Yin's classification [24], it is a holistic single-case study.

Context: the case study was conducted in a software SME located in the Autonomous City of Buenos Aires, with a total of 430 employees. This company develops customized information systems for clients of different industry sectors, including finance, automotive, pharmaceutical and banking. Its software projects combine agile practices with iterative life cycle development methodologies. Access was granted to the documentation of the project subject to an agreement not to disclose the name of the company and a commitment to inform about any findings and recommendations to be considered for deployment process risk management.

Case: deployment of deliverables for a Human Resources Portal conducted at a bank based in Argentina. It consisted in adding new capabilities, using a modular strategy. These were: integration with a new data source, publication of Application Programming Interfaces (APIs), integration with a distance learning portal, modification of the final user interface, new employee management alerts and notifications, appearance modifications to the application organigram, and modification to approval flows.

Unit of analysis: documentation related to the deployment of deliverables for a Human Resources Portal.

### 4.4 Preparation for data collection

A third-degree technique was used combined with an independent method according to the classification proposed in [25]. A template with a coding scheme made up of 3 groups was used. Each group coincides with the 3 activities of the technical process called "Transition" of the ISO / IEC / IEEE 12207: 2017 Standard [19] (A1 Preparation for deployment, A2 Deployment Execution and A3 Deployment Results Management).

Table 1 shows the traceability of the documents analyzed and the risks associated with each of the dimensions. The calculated risk weight is found in [20].

**Table 1:** Traceability of the documents analyzed for the case study (the defined risk coding scheme is detailed in [20]).

Documents/ Activities	A1	A2	A3
Risk monitoring spreadsheet	RProc6, RPers3 and RProd1	RProc10	RProd15
Progress Report	RPers4	RProc7 and RProd9	RPers13
Deliverable 1 - Closing report	RProd4	RProc8 and RPers9	RProc14, RPers15 and RProd13
Deliverable 1 -		RProd8	RProc15 and

Deployment report			RPers12
Deliverable 1 - Deployment Summary		RPers8	RProc11 and RProd12
Deliverable 1 - Deployment Tests Guide	RProc4, RPers2 and RProd5	RProd10	
Deliverable 1 - Deployment Test cases	RProc4, RPers2 and RProd3		
Deliverable 1 – installation scripts	RPers1 and RProd2		RProc12
Deliverable 1 – Work Plan	RProc5 and RPers5	RProd7 and RPers10	
Deliverable 1 – Installation Requirements	RProc1 and RProd6	RProc9 and RPers7	
Deliverable 1 - Deployment Completion report	RProc2 and RPers6	RProd9	RProc13, RPers14 and RProd14
Deliverable 2 - Closing Report	RProd4	RProc8 and RPers9	RProc14, RPers15 and RProd13
Deliverable 2 - Deployment Report		RProd8	RPers12
Deliverable 2 – Deployment Summary		RPers8	RProc11, RProc15 and RProd12
Deliverable 2 - Deployment Tests Guide	RProc4, RPers2 and RProd5	RProd10	
Deliverable 2 - Deployment Test cases	RProc4, RPers2 and RProd3		
Deliverable 2 – installation scripts	RPers1 and RProd2		RProc12
Deliverable 2 – Work Plan	RProc5 and RPers5	RPers10 and RProd7	
Deliverable 2 – Installation Requirements	RProc1 and RProd6	RProc9 and RPers7	
Delivery 2 - Deployment Completion Report	RProc2 and RPers6	RProd9	RProc13, RPers14 and RProd14
General Documentation	RProc3		RPers11 and RProd11

### 4.3 Analysis and Interpretation of Results

The results of the research questions defined for the case study are presented below:

*RQ1: How were risks managed during the activities of the software system deployment process (identification, analysis and severity)?*

Based on the documentation analyzed, it was possible to find flaws in the risk management proposed for the activities of the deployment process:

- Activity 1 (A1) – Preparation for Deployment: The deployment progress reports showed that, due to the few investments in technology made in recent years, the resources (hardware and basic software) assigned to the production environment did not comply with the minimum requirements requested by the

consulting company to carry out the deployment in accordance with the established work plan.

According to the deployment reports analyzed, the technicians (bank employees) did not have the knowledge and skills necessary for the correct deployment of scripts and monitoring of the guides sent by the consulting company. This is because the technicians who participated in the original deployment left the organization and were replaced by personnel with little technical or functional experience.

The general documentation of the project shows that the bank does not have an adequate personnel retention policy, which generates frequent rotation.

- Activity 2 (A2) – Deployment Execution: according to the progress reports of the deployment project, the technical flaws mentioned in the previous stage (separation of technical personnel with experience in the technologies involved and greater complexity of the product) generated friction between the consulting company and the managers of the bank. This was due to non-compliance with the deadlines established in the work plan, which ended up activating a penalty clause against the consulting company.

During the documentary analysis, incomplete test plans and inadequate deployment metrics were found. According to the deployment completion reports, the consulting company had to face cost overruns for not having the document management procedures required by the bank in the contract and in corporate policy. In addition, it was necessary to add technical resources from the consulting company to address the lack of technical expertise of the bank's employees, who had to be trained to carry out future deployments.

These technical drawbacks, added to a very demanding work schedule for internal reasons and needs of the bank (shown in the closing reports), were some of the causes that produced very important delays and friction between different sectors of the organization that even considered the cancellation of the deployment project on several occasions.

- Activity 3 (A3) - Deployment Results Management: problems with the software repositories (lack of necessary permissions, previous versions, lack of components, etc.), in addition to the low commitment and inexperience of the bank's technicians, generated multiple drawbacks during the deployment. These technical drawbacks strongly impacted on the quality of the final product and the satisfaction of the users who saw their productivity affected due to failures in the application's capabilities once the deployment was complete.

In the deployment completion reports, it was also evidenced that there was a wrong dimensioning of the deliverables and that the necessary security tests were not carried out. This gave end users access to sensitive human resource information.

*RQ2: How can the software systems deployment process be strengthened in this company?*

Proper risk management minimizes drawbacks in the deployment process. In [20], the recommended procedures are presented to the software consulting company in order to prevent, mitigate and / or transfer each of the risks associated with the "Process", "Person" and "Product" dimensions.

#### 4.4 Threats to validity

To analyze the validity of the study, the factors proposed in [25] were considered:

**Construct validity.** The results were obtained based on the documentary analysis of a set of risks for the process of deployment of software systems in a real context. This allowed us to answer the defined research questions, determining their relevance and suitability for the case.

**Internal validity.** The documentation used refers to a real case, a deployment of new deliverables for a Human Resources Portal performed in a bank in Argentina. In order to achieve greater precision and validity of the studied process, the need to combine the data source (project documentation) with other types of sources, such as interviews and / or focus groups to guarantee "data triangulation (source)", is recognized. Furthermore, the qualitative data collected and analyzed could be combined with quantitative data resulting from the project, thus ensuring a "Methodological Triangulation".

**External validity.** Carrying out a single case study may limit the generalizability of the results. However, a preliminary case study was conducted in [18]. These two experiences allow us to present results, which can be used by other researchers to carry out more studies with the same principles.

**Reliability.** The study data was collected and analyzed by the research group.

#### 4.5 Lessons learned

- **Method selection:** a validation of a set of risks, as well as the procedures for their prevention, mitigation and / or transfer, for the process of deployment of software systems, was needed in a real environment, in order to refine them (if required). The results obtained allowed us to analyze the application of the set of risks defined in a real environment. Therefore, the method used is considered to have yielded the expected results.
- **Data collection:** although the documentation of the software system deployment process has been reviewed in order to analyze how the risks were managed, it is considered that the case could be strengthened if the data collected were complemented by another source or by quantitative data.
- **Selected coding.** The coding scheme selected for the design of the data collection and analysis template was adequate and allowed the systematic recording of risk information.
- **Results report:** Although the case is made up of two research questions, it is considered that the work carried out took into account an adequate level of detail for understanding the phenomenon under study.

### 5 Conclusions and future work

The results of a case study were presented to determine the feasibility of applying a set of risks, as well as the procedures for their prevention, mitigation and / or transfer

for the process of deploying software systems in a real environment. It consisted of the risk analysis of the deployment of new deliverables for a Human Resources Portal carried out by a software SME in a bank in Argentina. After conducting the case study, it is concluded that:

- The first question allowed us to identify shortcomings in risk management through documentary analysis. These shortcomings include the lack of specialization of project personnel, mixed interests between the intervening areas and non-compliance with requirements of the installation environment.

- The second question allowed us to design a set of recommended procedures (presented in section 4.3) for the company to improve its deployment process and to introduce good risk management practices for future software system deployments.

The lessons learned from the case showed that the research method was adequate to validate the proposal.

The following are identified as future works: (a) to validate the risk proposal for the software deployment process in different case studies in order to refine it. (b) To propose the use of the risks defined for the deployment of software systems, as well as the procedures for the prevention, mitigation and / or transfer thereof, by other professionals in the industry.

## References

1. Charette R. Why software fails [software failure]. *IEEE spectrum*, 42(9), 42-49. (2005).
2. Dhlamini, J. & Nhamu, I. y Kaihepa, A. Intelligent risk management tools for software development. 33-40 (2009)
3. Jones C., *Assessment and control of software risk*. Yourdon Press (1994).
4. Liu D., Wang Q., Xiao J. The role of software process simulation modeling in software risk management: A systematic review. In *Proceedings of the 3rd International Symposium on Empirical Software Engineering and Measurement*. Empirical Software Engineering and Measurement, pp. 302-311 (2009).
5. Jansen S., Brinkkemper S. Definition and validation of the key process of release, delivery and deployment for product software vendors: Turning the ugly duckling into a swan *IEEE International Conference on Software Maintenance, ICSM*, art. no. 4021334, pp. 166-175. (2006).
6. Subramanian, N. The software deployment process and automation. *CrossTalk*, 30 (2), pp. 28-34 (2017).
7. Tyndall J. Building an effective software deployment process. In *Proceedings of the 40th annual ACM SIGUCCS conference on User services*, pp. 109-114 (2012).
8. Reascos I., Carvalho J., Bossano S. Implanting IT Applications in Government Institutions: A Process Model Emerging from a Case Study in a Medium-Sized Municipality. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance*, pp. 80-85 (2019).
9. Paredes I., Carvalho J. Research in Progress: Understanding the process of implantation IT Enterprise Applications in Small and Medium Enterprises (SMEs). In *Atas da Conferência da Associação Portuguesa de Sistemas de Informação*, Vol. 17, No. 17, pp. 270-283. (2017).

10. Forbes J., Baker E. Improving Hardware, Software, and Training Deployment Processes. In: Proceedings of 19th International Conference on Software Maintenance, pp. 377-380. IEEE, The Netherlands. (2003).
11. Ortiz F, Davila M., Panizzi M. y Bertone R. State of the art determination of risk management in the implantation process of computing systems. En las Actas del I Congreso Internacional sobre Avances en Nuevas Tendencias y Tecnologías (ICAETT 2019). Ecuador, Guayaquil Ecuador, 29 al 31 de mayo, pp- 23-32 (2019). ISBN 978-3-030-32022-5.
12. CMMI Institute, «Capability Maturity Model Integration,». <https://cmmiinstitute.com> Página vigente al 24/06/2020.
13. Project Management Institute. <https://www.pmi.org/pmbok-guide-standards>. Página vigente al 24/06/2020.
14. Software Engineering Institute, «Software Risk Evaluation Method» (1999). [https://resources.sei.cmu.edu/asset\\_files/TechnicalReport/1999\\_005\\_001\\_16799.pdf](https://resources.sei.cmu.edu/asset_files/TechnicalReport/1999_005_001_16799.pdf)
15. Kitchenham B., Linkman S., Law D.T. DESMET: A method for evaluating software engineering methods and tools. Keele University (1996).
16. Portal de administración electrónica, «MAGERIT v.3: Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información» 2012. Página vigente al 24/06/2020.
17. Ortiz F., Panizzi M., y Bertone R., Risk determination for the implantation process of software systems. En las Actas del XXV Congreso Argentino de Ciencias de la Computación - CACIC 2019. Universidad Nacional de Río Cuarto, 14 al 18 de octubre, pp- 817- 825 (2019). ISBN 978-987-688-377-1
18. Panizzi M., Davila M., Hodes A., Vázquez P., Ortiz F., Arana F., Bertone R. Desafíos para la implantación de sistemas de software. En las Actas del XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020), El Calafate, Argentina 7 y 8 de Mayo de 2020. ISBN en trámite.
19. ISO/IEC/IEEE 12207:2017. Systems and software engineering — Software life cycle processes (2017).
20. Felipe Ortiz, Marisa Panizzi, Rodolfo Bertone. Appendix – Risk refinement in the deployment process of software systems: a case study. <https://doi.org/10.6084/m9.figshare.12670967.v1>.
21. International Organization for Standardization, «ISO/IEC 31010:2009». <https://www.iso.org/standard/51073.html>. Página vigente al 24/06/2020.
22. Runeson P, Höst M, Rainer A, Regnell B. Case study research in software engineering: guidelines and examples. Wiley Publishing, Hoboken (2012).
23. C. Robson. Real world research 2nd edition. Blackwell (2002)
24. Yin, R., Case study research: design and methods. 5th Edition. Sage Publications. (2014).
25. Lethbridge T., Sim S., Singer J., Studying software engineers: data collection techniques for software field studies. Empir Softw Eng 10(3):311–341 (2005).

## Aplicaciones Móviles 3D: un estudio comparativo de performance y consumo de energía

Federico Cristina<sup>1</sup>, Sebastián Dapoto<sup>1</sup>, Pablo Thomas<sup>1</sup>, Patricia Pesado<sup>1</sup>,  
Jefferson Perez Altamirano, Martin De la Canal Erbetta

<sup>1</sup> Instituto de Investigación en Informática LIDI,  
Universidad Nacional de La Plata – Argentina  
Centro Asociado Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

{fcristina, sdapoto, pthomas, ppesado}@lidi.info.unlp.edu.ar  
jefferson556@gmail.com  
martin\_delacanal@hotmail.com

**Resumen.** En los últimos años la tecnología ha evolucionado exponencialmente, con avances en todos los ámbitos. Particularmente, la tecnología en dispositivos móviles se ha vuelto cada vez más sofisticada, robusta y con gran capacidad de cómputo. Debido a estos avances, actualmente es posible ejecutar aplicaciones móviles complejas con exigentes requisitos de hardware y energía, como por ejemplo las aplicaciones 3D. Existen, sin embargo, dos factores primordiales a considerar al momento de desarrollar aplicaciones móviles 3D: la performance y el consumo de energía, los cuales tienen un impacto directo en la experiencia de uso. El presente trabajo evalúa y compara estos aspectos en los dos frameworks de desarrollo 3D más utilizados actualmente: Unity y Unreal Engine.

**Palabras Clave:** dispositivos móviles, aplicaciones 3D, Unity, Unreal Engine, performance, consumo de energía

### 1 Introducción

El continuo avance en la capacidad de procesamiento de los dispositivos móviles ha permitido el desarrollo de aplicaciones cada vez más complejas, las cuales incluyen características tales como visualización 3D o realidad aumentada, características computacionalmente exigentes.

Este tipo de aplicaciones demandantes de alto requerimiento de hardware, presentan básicamente dos factores a considerar al momento del desarrollo:

- La performance, la cual tiene un impacto directo en la fluidez de la visualización y por ende en la experiencia de uso.
- El consumo de energía, que puede llegar a definir la utilización - o no - de una aplicación.

Es posible aseverar que en cierta medida ambos aspectos se encuentran en estrecha relación, dado que mayor performance demanda mayor potencia de cómputo y por

ende mayor consumo de energía. Sin embargo, es necesario considerar a estos dos factores por separado a fin de poder realizar un estudio independiente sobre las herramientas de desarrollo de aplicaciones 3D presentadas en este trabajo.

Existen diversos frameworks de desarrollo de aplicaciones 3D. Dos de los frameworks más populares actualmente son Unity [1] y Unreal Engine [2].

Unity se destaca por la cantidad de documentación disponible, una enorme y muy activa comunidad de usuarios, gran variedad de componentes pre-desarrollados (assets) y plugins que facilitan la integración con otras herramientas. Además, tiene una elevada curva de aprendizaje, dado la sencillez de su editor y lenguaje de programación requerido. Haciendo uso de los manuales y tutoriales existentes es posible simplificar el desarrollo de una aplicación móvil 3D. Por otro lado, Unity ofrece una gran cantidad de plataformas de publicación.

Unreal Engine ofrece una gran calidad y potencia de gráficos. Aunque no tan extensa como la de Unity, Unreal cuenta con una comunidad muy activa y colaborativa. Dispone además de una buena cantidad de documentación y tutoriales. En términos generales, el desarrollo de aplicaciones 3D en Unreal es más complejo que en Unity. Sin embargo, en las últimas versiones se ha incluido el sistema de Blueprints, una opción de programación basada en nodos y componentes que facilita el trabajo de codificación.

El presente estudio tiene como finalidad analizar y comparar la performance y el consumo de energía independientemente sobre estos dos frameworks de desarrollo de aplicaciones 3D. El trabajo se organiza del siguiente modo, en el capítulo 2 se presenta la motivación, el capítulo 3 muestra una propuesta de evaluación, el capítulo 4 detalla el desarrollo de los prototipos, el capítulo 5 muestra las pruebas efectuadas y presenta los resultados obtenidos, en el capítulo 6 se presentan las conclusiones y trabajo a futuro.

## 2 Motivación

El continuo avance en la capacidad de procesamiento de los dispositivos móviles ha permitido el desarrollo de aplicaciones cada vez más complejas, las cuales incluyen características tales como visualización 3D o realidad aumentada, características computacionalmente exigentes.

Este tipo de aplicaciones demandantes de alto requerimiento de hardware, presentan básicamente dos factores a considerar al momento del desarrollo:

El análisis de performance se inició durante el proceso de desarrollo de aplicaciones móviles 3D inmersivas en Unity las cuales presentan un costo computacional elevado, principalmente en lo que refiere a visualización. Tanto en el proceso de implementación como en el de ejecución, se encontraron limitaciones en cuanto a la performance obtenida. Se procedió a reconocer los puntos críticos que incidían en la fluidez visual y se determinaron ciertos umbrales que no debían ser superados. Este análisis permitió modificar las aplicaciones desarrolladas y lograr un mejor funcionamiento en los distintos tipos de dispositivos móviles.

Por otra parte, la batería, componente indispensable en los dispositivos móviles, no ha sufrido grandes cambios en los últimos años. Los fabricantes han logrado

paulatinamente aumentar su capacidad y autonomía, pero sin avances notables. Actualmente se pueden conseguir celulares con baterías de hasta 5500 miliamperios hora (mAh) y autonomía de dos días. Dicha autonomía se estima a partir de la necesidad de suministro eléctrico requerido por los componentes internos del dispositivo, creando así un compromiso entre el poder de procesamiento, la visualización y la duración de la carga de la batería. Como consecuencia de esto, si una aplicación realiza un procesamiento intensivo, limitará el tiempo que el usuario puede usar el dispositivo sin necesitar una recarga.

Este mayor consumo plantea problemas para la evolución de la computación móvil, ya que los desarrolladores no pueden utilizar todo el potencial de la tecnología actual sin sacrificar la autonomía de la batería hasta una nueva recarga.

Por lo tanto, al momento de desarrollar una aplicación móvil, es de suma importancia tener en cuenta el consumo energético que pueda llegar a generar [3][4]. Esto toma mayor relevancia en las aplicaciones móviles 3D, dado que éstas son implícitamente demandantes de poder de cómputo y por consiguiente de consumo de energía.

A fin de ayudar a conseguir un equilibrio en las características relevantes de las aplicaciones móviles 3D y lograr una ejecución eficiente, se realizó el presente estudio donde los parámetros más representativos de los modelos 3D son puestos a prueba para medir su performance y consumo.

Características tales como el número de polígonos, la aplicación de luces y sombras, la utilización de texturas y/o transparencias, la visualización de sistemas de partículas y el cálculo de la física de objetos que componen la escena, son ejemplos de los principales ítems a evaluar. Se utilizarán entonces la misma serie de pruebas independientes definidas en [5] para evaluar la performance y consumo de energía.

La aplicación de este conjunto de pruebas simplifica considerablemente la tarea de determinar los puntos críticos a optimizar en las aplicaciones desarrolladas, posibilitando así la identificación del punto de equilibrio en la calibración de las características analizadas.

### **3 Propuesta de evaluación**

El principal objetivo es determinar cuáles son las características de una aplicación 3D desarrollada en Unity y Unreal, que generan un mayor impacto en la performance de ejecución y en el consumo de energía en un dispositivo móvil.

Como se ha comentado previamente, las características de las aplicaciones móviles 3D que tienen relación directa con el procesamiento intensivo, también la tienen con el alto consumo de energía.

Con el fin de poder evaluar en forma aislada cada una de estas características en aplicaciones realizadas con el motor Unity, se desarrollaron dos prototipos móviles 3D.

El primer prototipo evalúa cómo las características mencionadas inciden en la performance, tiempo de respuesta y fluidez de ejecución de las aplicaciones desarrolladas con Unity. Esta tarea la realiza mediante un conjunto de pruebas en el que se testean cada una de las características de forma independiente [5]. A lo largo

de cada prueba, se va incrementando el número de objetos en la escena y se analiza el impacto generado.

El segundo prototipo evalúa cómo estas mismas características inciden en el consumo de energía de las aplicaciones desarrolladas con Unity. Esta tarea la realiza mediante una adaptación del conjunto de pruebas definido en [5] en el que también se testean cada una de las características de forma independiente [6].

La metodología de evaluación de consumo de energía implica relevar durante un período de tiempo prefijado el consumo de energía bajo la ejecución de cada una de las distintas pruebas. A diferencia de la metodología utilizada para la evaluación de performance, el número de objetos en la escena es constante durante el tiempo de ejecución de cada prueba. Una vez transcurrido el tiempo prefijado, se obtiene el valor del consumo generado.

Para hacer posible el análisis comparativo entre Unreal y Unity fue necesario implementar también ambos prototipos en el motor Unreal.

## 4 Desarrollo de prototipos

En lo que respecta a los prototipos en Unity, dichos prototipos fueron desarrollados previamente en los trabajos [5] y [6].

Unreal Engine es un framework de desarrollo de aplicaciones 3D que se destaca por su capacidad de manejar hasta un millón de partículas en una sola escena. Como lenguaje de programación utiliza C++ en combinación con un lenguaje visual denominado Blueprints.

La utilización de C++ en Unreal Engine brinda a los desarrolladores un gran control sobre las acciones de todo el sistema, pero a su vez vuelve el proceso de codificación mucho más complejo.

Los Blueprints, son un tipo de programación basada en nodos que permite a los desarrolladores implementar la lógica completa de una aplicación de una forma sencilla. Cada nodo tiene una determinada función con entradas y salidas, y al estar conectarlos entre sí se introduce la lógica necesaria. Los Blueprints pueden contener variables, tipos de datos, matrices, enumeraciones, mapas hash, funciones, macros, eventos personalizados, entre otros.

Utilizar C++ puro es más eficiente para aplicaciones de lógica compleja. Sin embargo, dada la sencillez de las pruebas del prototipo a desarrollar, no existe diferencia apreciable entre ambas alternativas de codificación. Por este motivo, los prototipos fueron desarrollados utilizando Blueprints.

En Unreal Engine existen tres perfiles básicos de calidad (alto, medio y bajo), y es posible configurar manualmente la calidad gráfica que tendrá cada perfil. Para cada familia del procesador y sistema operativo es posible elegir la calidad con la que debe ser ejecutada la aplicación. De todas formas, luego el motor de Unreal puede decidir ejecutar la aplicación en un perfil de calidad inferior si así lo considera necesario. Por esta razón, no es posible modificar la calidad gráfica en tiempo de ejecución, tal como se realiza en el prototipo desarrollado en Unity.

## 5 Pruebas realizadas y resultados obtenidos

Tanto para el estudio de performance como para el de consumo de energía la experimentación consistió en efectuar el conjunto de pruebas sobre una serie de dispositivos móviles con poder de cómputo diferente. Las pruebas se realizaron repetidas veces con cada prototipo y sobre cada dispositivo móvil, a fin de validar la consistencia de la medición.

El conjunto de pruebas es el definido en [5] y son las que se detallan brevemente a continuación:

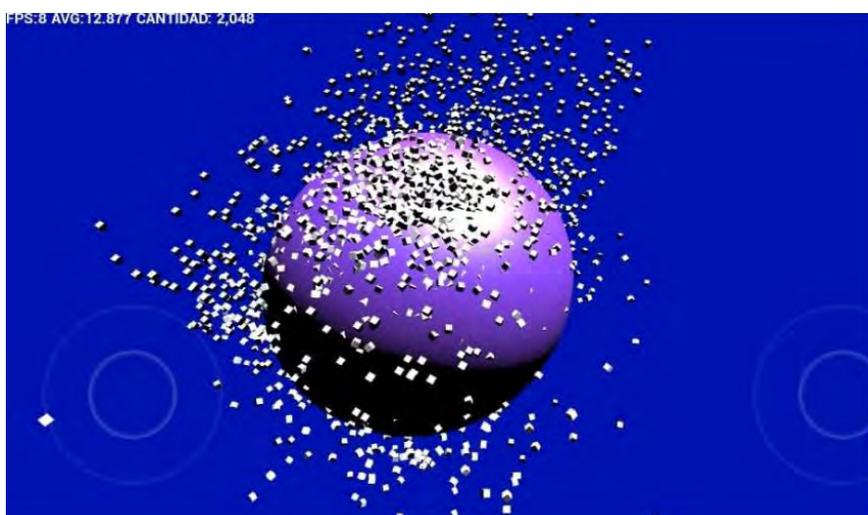
- Prueba 1: renderización de objetos simples. Se presentan progresivamente en pantalla objetos simples sin textura en movimiento en una escena sin iluminación ni sombras. Los objetos deben rotar continuamente a velocidad constante.
- Prueba 2: renderización de objetos complejos. Consiste en visualizar un objeto complejo en movimiento, el cual debe contener un elevado número de polígonos. La distancia de renderizado (*clipping plane*) se va incrementando a medida que avanza la prueba.
- Prueba 3: luces y sombras. Se realiza una simulación similar a la renderización de objetos simples, pero en este caso la escena contiene iluminación y objetos con proyección y recepción de sombras.
- Prueba 4: texturas. Se realiza una simulación similar a la renderización de objetos simples, pero en este caso los objetos poseen texturas complejas, como puede ser transparencias, reflejos, etc.
- Prueba 5: sistemas de partículas. Se crea una escena en donde se presenta progresivamente nuevas instancias de un sistema de partículas (por ejemplo humo, chispas, explosión, etc.).
- Prueba 6: física. Se realiza una simulación similar a la renderización de objetos simples, pero en este caso a los objetos se les aplica reglas de física, como por ejemplo *gravedad*.

En lo que respecta a las pruebas relacionadas con performance, de cada prueba se almacenaron los valores de cuadros por segundo (FPS) logrados a medida que se incrementa la cantidad de objetos en la escena, hasta llegar a un máximo de 8192 objetos. En el caso particular de la prueba 2, en el que sólo se cuenta con un único objeto complejo, se almacenaron los FPS logrados a medida que se visualiza una mayor porción del objeto complejo. En la figura 1 se puede observar la prueba 6 ejecutándose en el nuevo prototipo desarrollado en Unreal Engine.

Las pruebas de consumo de energía se realizaron configurando el número de cuadros por segundo a un número fijo, dado que la cantidad de FPS utilizados tienen una incidencia directa sobre el consumo, tal como se verifica en [6]. A su vez, el número de objetos fijado para cada prueba se estableció con el fin de garantizar que todos los dispositivos usados en un posible futuro, incluso los de menor poder de procesamiento, puedan ejecutar fluidamente el conjunto de pruebas. Las pruebas en cuestión se realizaron en cada ocasión durante varios minutos, para minimizar potenciales errores en la toma de datos debido a eventuales desviaciones temporales. En la figura 2 se puede observar la prueba 1 ejecutándose en el nuevo prototipo desarrollado en Unreal Engine.

Como premisa básica, en todos los casos - y con la finalidad de reducir potenciales errores - se definieron las siguientes condiciones a lo largo de las ejecuciones:

- Al comenzar con el conjunto de pruebas, los dispositivos a utilizar deben tener la carga de la batería mayor al 95%.
- Las ejecuciones deben realizarse con el brillo de la pantalla del dispositivo al mínimo.
- Los dispositivos deben estar configurados en “Modo avión”, para evitar cualquier tipo de recepción de señal que pudiera incidir en las mediciones.
- No deben existir aplicaciones ejecutándose en segundo plano, dado que pueden consumir tiempo de procesamiento y de esa forma afectar las mediciones.



**Fig. 1.** Prototipo de performance en Unreal Engine. Prueba 6.

Un aspecto importante a tener en cuenta bajo Unreal - y que en cierta medida puede afectar tanto a performance como a consumo - es la falta de optimización en los proyectos orientados a dispositivos móviles, básicamente debido a la gestión de Draw Calls [7] que el motor realiza.

La herramienta de profiling seleccionada para realizar las mediciones de consumo de energía en [6] fue Trepro Profiler. Esta herramienta permitía medir en tiempo real el consumo de energía de una aplicación en particular, asegurando una medición precisa de la energía consumida en dispositivos que cuenten con procesadores desarrollados por Qualcomm Technologies Inc., empresa desarrolladora de la herramienta. En la actualidad, dicha empresa reemplazó la herramienta Trepro Profiler por una nueva herramienta denominada Snapdragon Profiler [8]. Sin embargo, utilizando esta herramienta no se obtuvieron resultados estables, por lo que se buscaron alternativas, y finalmente se optó por usar la aplicación AccuBattery. Esta aplicación, brinda información precisa sobre el consumo de batería en miliamperios hora (mAh).

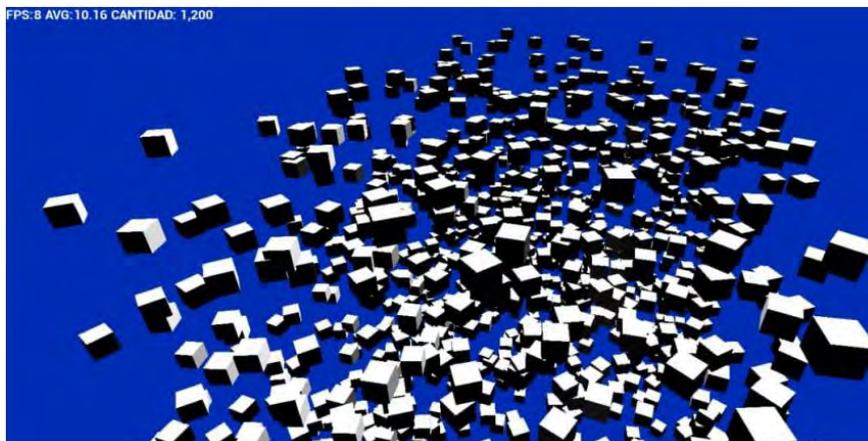


Fig. 2. Prototipo de consumo en Unreal Engine. Prueba 1.

Debido al cambio de la herramienta de medición de consumo, los resultados obtenidos en [6] no pudieron ser utilizados. Se realizaron nuevamente todas las pruebas en el prototipo desarrollado en Unity usando la herramienta de medición AccuBattery. Luego, se realizaron las pruebas en el nuevo prototipo desarrollado en Unreal también utilizando AccuBattery. De esta forma los resultados obtenidos son válidos para el análisis comparativo.

Las tablas 1 y 2 presentan los resultados de la degradación en la performance - medida en cuadros por segundo - en función del aumento de la complejidad computacional de cada una de las pruebas realizadas.

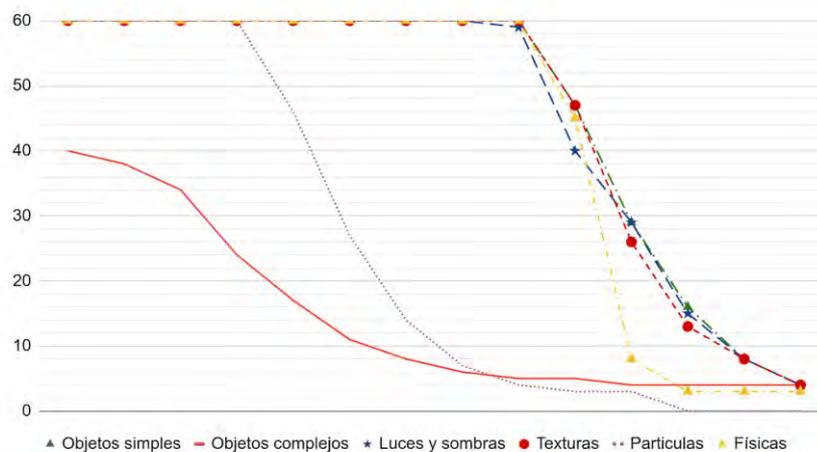
La figura 3 muestra el detalle de la degradación de la performance en Unity para el conjunto de casos de prueba definido, mientras que la figura 4 muestra este mismo detalle bajo Unreal Engine. La tabla 3 muestra los valores obtenidos en el estudio de consumo energético para el conjunto de pruebas realizado sobre ambos frameworks de desarrollo. La figura 5 presenta una comparativa de consumo energético sobre el conjunto de casos de prueba definido, abarcando los valores obtenidos tanto en Unity como en Unreal.

Tabla 1. Degradación de la performance en las pruebas realizadas en Unity.

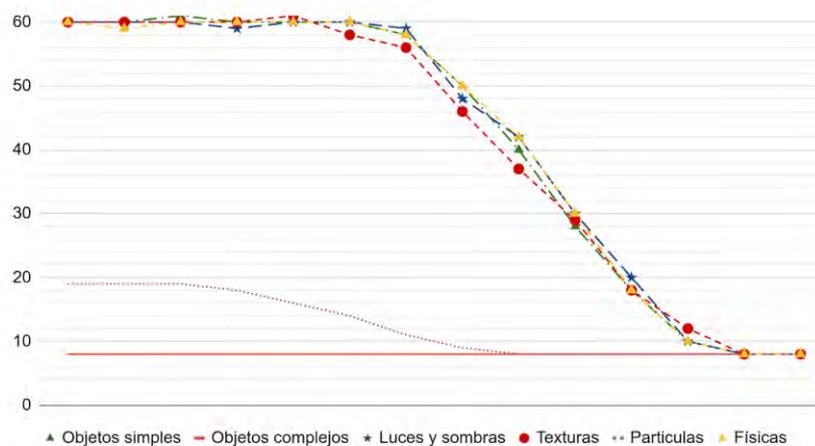
Performance (FPS)	Unity													
Objetos simples	60	60	60	60	60	60	60	60	60	47	29	16	8	4
Objetos complejos	40	38	34	24	17	11	8	6	5	5	4	4	4	4
Luces y sombras	60	60	60	60	60	60	60	60	59	40	29	15	8	4
Texturas	60	60	60	60	60	60	60	60	60	47	26	13	8	4
Partículas	60	60	60	60	46	27	14	7	4	3	3	0	0	0
Físicas	60	60	60	60	60	60	60	60	60	45	8	3	3	3

**Tabla 2.** Degradación de la performance en las pruebas realizadas en Unreal Engine.

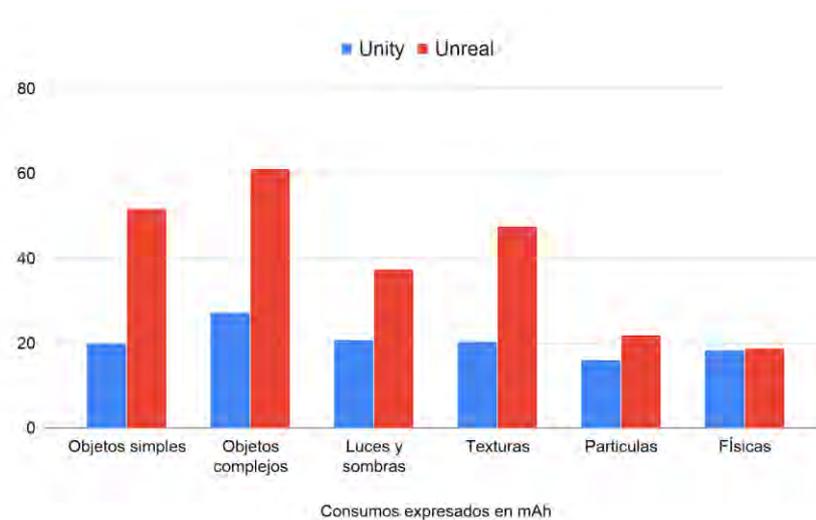
Performance (FPS)	Unreal Engine													
Objetos simples	60	60	61	60	60	60	58	50	40	28	18	10	8	8
Objetos complejos	8	8	8	8	8	8	8	8	8	8	8	8	8	8
Luces y sombras	60	60	60	59	60	60	59	48	42	30	20	10	8	8
Texturas	60	60	60	60	61	58	56	46	37	29	18	12	8	8
Partículas	19	19	19	18	16	14	11	9	8	8	8	8	8	8
Físicas	60	59	60	60	60	60	58	50	42	30	18	10	8	8

**Fig. 3.** Representación gráfica de la degradación de la performance en Unity para el conjunto de pruebas realizado.**Tabla 3.** Consumo en las pruebas realizadas en Unity y Unreal Engine.

Consumo (mAh)	Unity	Unreal
Objetos simples	19.8	51.53
Objetos complejos	27.18	61.06
Luces y sombras	20.73	37.4
Texturas	20.23	47.48
Partículas	15.94	21.78
Físicas	18.36	18.65



**Fig. 4.** Representación gráfica de la degradación de la performance en Unreal para el conjunto de pruebas realizado.



**Fig. 5.** Consumo Unity vs. Unreal Engine.

## 6 Conclusiones y trabajo a futuro

El presente trabajo presenta un conjunto de actividades para evaluar y comparar la performance y el consumo de energía de aplicaciones 3D sobre dispositivos móviles, específicamente para los frameworks Unreal Engine y Unity.

La evaluación propuesta permite determinar la incidencia de cada una de las características principales de las aplicaciones 3D sobre la fluidez de visualización y el consumo energético.

De esta manera, un ingeniero de software puede desarrollar estas aplicaciones teniendo en cuenta los considerandos aquí detallados, tales como cantidad de polígonos, texturas, sombras, entre otras características visuales.

Se desarrollaron dos prototipos en el framework Unreal Engine, que permiten ejecutar una serie de pruebas independientes para cada una de las características típicas de una aplicación 3D. Estos resultados fueron posteriormente contrastados con los obtenidos en las aplicaciones análogas desarrolladas previamente en el framework Unity.

Se determinó que - en términos generales - Unreal Engine tiene la capacidad de desarrollar y procesar gráficos más potentes y realistas que Unity, pero generando un mayor requerimiento de recursos y energía.

A futuro se prevé ampliar esta comparativa a otros frameworks de este tipo, como por ejemplo CryEngine [9], el cual, si bien actualmente no tiene la popularidad de Unity o Unreal, se encuentra en constante crecimiento.

## Referencias

1. Unity. <https://unity3d.com>.
2. Unreal Engine. <https://www.unrealengine.com>.
3. 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.
4. Corbalan L., Fernández J., Cuitiño A., Delia L., Cáseres G., Thomas, P.; Pesado, P. "Development Frameworks for Mobile Devices: A Comparative Study about Energy Consumption". Association for Computing Machinery (ACM). 2018. ISBN 978-1-4503-5712-8/18/05.
5. Cristina, F.; Dapoto, S.; Thomas, P.; Pesado, P. "Performance evaluation of a 3D engine for mobile devices". Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790. A. De Giusti, Springer International Publishing. 2018. ISBN: 978-3-319-75213-6, 978-3-319-75214-3, pages 155-163.
6. Cristina, F.; Dapoto, S.; Thomas, P.; Pesado, P. "Análisis de consumo de energía en aplicaciones 3D sobre dispositivos móviles". XXIV Congreso Argentino de Ciencias de la Computación (CACIC). 2018. ISBN: 978-950-658-472-6, páginas 622-630.
7. Unreal Engine. CPU Profiling. <https://docs.unrealengine.com/en-US/Engine/Performance/CPU/index.html>.
8. Snapdragon Profiler. <https://developer.qualcomm.com/software/snapdragon-profiler>.
9. CryEngine. <https://www.cryengine.com>.

## Modelo de evaluación de datos utilizando el enfoque GQM

Julieta Calabrese <sup>1</sup> , Silvia Esponda , Ariel Pasini , 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 Comisión de Investigaciones Científicas de la Pcia. de Bs. As. (CIC)

<sup>1</sup>Becaria UNLP

{jcalabrese, sesponda, apasini, ppesado}@lidi.info.unlp.edu.ar

**Abstract:** Las organizaciones actuales manejan grandes cantidades de datos. Poder mejorar y mantener la calidad de los mismos es uno de los grandes desafíos que enfrentan dichas organizaciones hoy en día. Para este fin, existen estándares definidos por ISO destinados a medir la calidad de los datos en función de un conjunto de características inherentes y dependientes del sistema. Se presenta MED, un modelo de evaluación de datos destinado a medir las características propuestas por ISO/IEC 25012 mediante el enfoque GQM (Goal, Question, Metric). Dicho enfoque se aplica en el diseño de una evaluación, la cual posteriormente será ejecutada bajo la estructura definida en ISO/IEC 25040.

**Keywords:** Calidad de datos – ISO/IEC 25012 - ISO/IEC 25040 - GQM

### 1. Introducción

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.

Es muy común notar que los datos se ven afectados por factores negativos: ruido, valores perdidos, inconsistencias, un tamaño demasiado grande en cualquier dimensión (número de atributos e instancias), 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 querer destacarse en el mercado actual y de prestar un mejor servicio que cumpla con las exigencias del cliente.

La mayoría de las empresas son conscientes de que el uso de la información genera mayor ventaja competitiva al momento de ofrecer servicios y productos acordes a las necesidades. Por este motivo, las empresas se respaldan en el uso de normas y estándares.

Este contexto centra nuestra atención en los estándares definidos por ISO. La familia ISO/IEC 25000 (mayormente conocida como SQuaRE: Software Product Quality

Requirements and Evaluation) nace en el año 2005 para dar respuesta a las necesidades actuales de las organizaciones. Su objetivo es la creación de un marco de trabajo común para evaluar la calidad de un producto software desde diferentes aspectos. Dentro de la familia ISO/IEC 25000, el estudio se enfoca en las normas ISO/IEC 25012 - “Data Quality Model” [1], la cual define un conjunto de características destinadas a evaluar la calidad de los datos, e ISO/IEC 25040 - “Evaluation process” [2], que define el proceso de evaluación a llevar a cabo.

Lograr medir las características definidas en la ISO/IEC 25012 es de suma importancia, ya que eso conlleva a obtener datos de mejor calidad enfocando la mejora de los mismos en el ámbito que lo requiera. Bajo este contexto, es de interés brindar una solución para lograr medir dichas características mediante la utilización del enfoque GQM (Goal, Question, Metric) [3].

El enfoque GQM es un método orientado a obtener una métrica que mida cierto objetivo a través de la utilización de preguntas. Se desarrolla identificando uno o más objetivos de calidad y elaborando preguntas que detallan a los mismos de la manera más completa posible. Posteriormente, se definen las métricas que tendrán un único resultado en base a las respuestas a dichas preguntas con el fin de identificar el nivel de aceptación con respecto al objetivo definido. Finalmente, se desarrollan mecanismos de validación y análisis de resultados.

El presente trabajo propone brindar un prototipo de herramienta denominado MED (Modelo de Evaluación de Datos) con el fin de diseñar y realizar una evaluación de calidad basada en ISO/IEC 25040 utilizando el modelo definido en ISO/IEC 25012 y el enfoque GQM. El principal objetivo está enfocado en conocer el estado de los datos de una organización mediante el nivel de un conjunto de características de calidad asociadas a dichos datos.

En la siguiente sección se describen las estructuras de las normas ISO/IEC 25012 e ISO/IEC 25040, haciendo hincapié en las características asociadas a los datos. En la tercera sección se define GQM y sus principios. En la cuarta sección se presenta MED, el modelo de evaluación de datos destinado a realizar el diseño de una evaluación utilizando las características definidas en ISO/IEC 25012 y el modelo GQM. Posteriormente, se realiza una evaluación de datos basada en ISO/IEC 25040 aplicando el diseño definido y finalmente se exponen las conclusiones y la bibliografía.

## **2. ISO/IEC 25000**

La familia ISO/IEC 25000 (SQuaRE) propone un conjunto de normas que tiene por objetivo la creación de un marco de trabajo común para evaluar la calidad de un producto de software desde diferentes aspectos. Es de interés destacar en el marco de este artículo las normas ISO/IEC 25012 e ISO/IEC 25040.

### **2.1 ISO/IEC 25012 – Data quality model**

La calidad de los datos es un factor clave en cualquier organización, ya que los mismos representan a la información de una manera formal y adecuada para la comunicación, la interpretación o el procesamiento. ISO/IEC 25012 define un modelo general de calidad

para aquellos datos que están representados en un formato estructurado dentro de un sistema informático, y tiene como objetivo presentar una visión integrada de los mismos para garantizar la interoperabilidad de los sistemas. La norma está compuesta por un conjunto de quince características, las cuales se clasifican en dos grandes grupos:

*Calidad de Datos Inherente:* Grado con el que las características tienen el potencial de satisfacer las necesidades establecidas y necesarias cuando los datos son utilizados bajo condiciones específicas. Desde el punto de vista inherente, la calidad de datos hace referencia a:

- Valores de datos para el dominio y sus posibles restricciones (ej., Reglas de negocio con la calidad requerida por las características en una aplicación)
- Relaciones entre valores de datos (ej., Consistencia)
- Metadatos (ej., Autor, Formato, etc.)

*Calidad de Datos Dependiente del Sistema:* Grado con el que la calidad de datos es alcanzada y preservada a través de un sistema informático cuando los datos son utilizados bajo condiciones específicas. Desde este punto de vista, la calidad de datos depende del dominio tecnológico en el que los datos se utilizan, y se alcanza mediante las capacidades de los componentes del sistema informático tales como: dispositivos hardware (ej., respaldo para alcanzar la Recuperabilidad) y software (ej., herramientas de migración para alcanzar la Portabilidad). Este punto de vista suele ser responsabilidad de los técnicos del sistema.

La *Tabla 1* grafica la clasificación de las características:

*Tabla 1. Clasificación de características definidas en ISO/IEC 25012.*

Característica	Inherente	Dependiente del sistema
Exactitud	X	
Compleitud	X	
Consistencia	X	
Credibilidad	X	
Actualidad	X	
Accesibilidad	X	X
Conformidad	X	X
Confidencialidad	X	X
Eficiencia	X	X
Precisión	X	X
Trazabilidad	X	X
Comprensibilidad	X	X
Disponibilidad		X
Portabilidad		X
Recuperabilidad		X

El grupo de características que conforman la *Calidad de Datos Inherente* está formado por: **Exactitud** (Semántica y Sintáctica), especificando el grado en el que los datos representan correctamente el valor deseado en un contexto específico; **Compleitud**, donde se espera que los datos obligatorios no estén incompletos; **Consistencia**, haciendo referencia a los datos libres de contradicción y coherentes en un contexto específico; **Credibilidad**, incluyendo el concepto de autenticidad, que define el grado en el que los datos se consideran ciertos y creíbles en un contexto específico; y **Actualidad**, la cual define el

grado en el que los datos se encuentran actualizados.

Por otra parte, las características que conforman la *Calidad de Datos Inherente y Dependiente del Sistema* hacen referencia a la **Accesibilidad**, donde se especifica el grado en que los datos pueden ser accedidos en un contexto específico (por ejemplo: personas que necesiten tecnologías de apoyo por algún tipo de discapacidad); **Conformidad**, donde se verifica que los datos correspondientes cumplen con estándares o normativas vigentes; **Confidencialidad** (asociada a la seguridad de la información), donde se asegura que los datos solo son accedidos e interpretados por usuarios específicos autorizados; **Eficiencia**, donde se analiza el grado en el que los datos pueden ser procesados y proporcionados con los niveles de rendimiento esperados; **Precisión**, donde los datos requieren de valores exactos o con discernimiento en un contexto específico; **Trazabilidad**, donde se analiza si los datos proporcionan un registro de los acontecimientos que los modifican; y **Comprensibilidad**, en donde los datos son expresados utilizando lenguajes, símbolos y unidades apropiados y pueden ser leídos e interpretados por cualquier tipo de usuario.

Por último, las características que conforman la *Calidad de Datos Dependiente del Sistema* hacen hincapié en la **Disponibilidad**, que define el grado de los datos para ser obtenidos por usuarios y/o aplicaciones autorizadas; **Portabilidad**, donde se analiza si los datos pueden ser copiados, reemplazados o eliminados al realizar un cambio de un sistema a otro, preservando el nivel de calidad; y **Recuperabilidad**, donde se comprueba que los datos mantienen y preservan un nivel de operaciones en caso de fallos [5].

Existe un conjunto de métricas destinadas a realizar la medición de la calidad de datos en términos de las características definidas. Dichas métricas se encuentran definidas en la ISO/IEC 25024 [4] y no serán de interés para el objetivo del presente artículo.

## 2.2 ISO/IEC 25040 – Evaluation process

ISO/IEC 25040 define el proceso para llevar a cabo la evaluación del producto de software mediante un modelo de referencia, considerando las entradas, restricciones y recursos necesarios para obtener las salidas correspondientes. El proceso para llevar a cabo la evaluación consta de cinco actividades:

- Actividad 1: *Establecer los requisitos de la evaluación*. Esta actividad consiste en establecer el propósito de la evaluación, identificando las partes interesadas en el producto, los riesgos posibles si los hubiese y el modelo de calidad a utilizar.
- Actividad 2: *Especificar la evaluación*. Dentro de esta actividad se especifican los módulos de evaluación (métricas, herramientas y técnicas) junto con los criterios de decisión a aplicar.
- Actividad 3: *Diseñar la evaluación*. En esta actividad se define el plan con las tareas que se deben realizar en la evaluación.
- Actividad 4: *Ejecutar la evaluación*. Actividad destinada a la ejecución de la evaluación, obteniendo las métricas de calidad y aplicando los criterios de decisión.
- Actividad 5: *Concluir la evaluación*. En esta última actividad se culmina la evaluación de la calidad, realizando un informe de resultados finales y conclusiones en base a los valores obtenidos.

### 3. GQM (Goal Question Metric)

GQM es un método orientado a generar una métrica que mida un objetivo de una manera determinada a través de la utilización de preguntas. Proporciona una manera útil para definir mediciones, tanto del proceso como de los resultados de un proyecto.

Está enfocado en que una medición puede ser más satisfactoria si es diseñada teniendo en cuenta las metas, y las preguntas ayudan a medir si se está alcanzando en forma exitosa la meta definida. Se busca mejorar la calidad y confiabilidad reduciendo costos, riesgos y mejorando tiempos [6].

GQM define un objetivo, establece un conjunto de preguntas en base al mismo y genera métricas en base a las respuestas de las preguntas. Asimismo, puede ser utilizado por los miembros individuales de un equipo de proyecto para enfocar su trabajo y determinar su progreso hacia la realización de sus metas específicas. La medición debe ser realizada, en todos los casos, orientada a un objetivo.

El modelo de medición GQM tiene tres niveles:

- Nivel Conceptual (Goal/Objetivo): se identifica a lo que se aspira respecto a los productos, procesos o recursos; con respecto a varios modelos de calidad, desde varios puntos de vista y relativo a un entorno particular.
- Nivel Operativo (Question/Pregunta): se refina un conjunto de preguntas a partir del objetivo con el propósito de verificar su cumplimiento. Las preguntas buscan caracterizar el objeto de medición (producto, proceso o recurso) con respecto a una cuestión de calidad seleccionada y determinar su calidad desde el punto de vista seleccionado.
- Nivel Cuantitativo (Metric/Métrica): se asocia un conjunto de datos para cada pregunta formulando métricas, con el fin de proporcionar una respuesta de manera cuantitativa. Los datos pueden ser objetivos (si dependen únicamente del objeto que se está midiendo y no del punto de vista desde el que se captan) o subjetivos (si dependen tanto del objeto que se está midiendo como del punto de vista desde el que se captan).

### 4. MED: Modelo de Evaluación de Datos

Uno de los elementos claves en cualquier proceso de ingeniería es la medición. Particularmente en SQuaRE, las medidas se emplean para analizar el nivel de diferentes características asociadas a lo que se desea evaluar. Bajo este contexto, se propone un nuevo modelo de medición mediante la utilización de métricas definidas utilizando el enfoque GQM. De esta manera, se deberán responder un conjunto de preguntas con el fin de obtener el valor correspondiente de las métricas asociada al conjunto de características que se esperan medir.

Se define el prototipo de una herramienta con el fin de facilitar el diseño y realización de una evaluación de producto basada en ISO/IEC 25040, utilizando el modelo de calidad definido en ISO/IEC 25012 y realizando la medición bajo el enfoque GQM.

El prototipo se divide en:

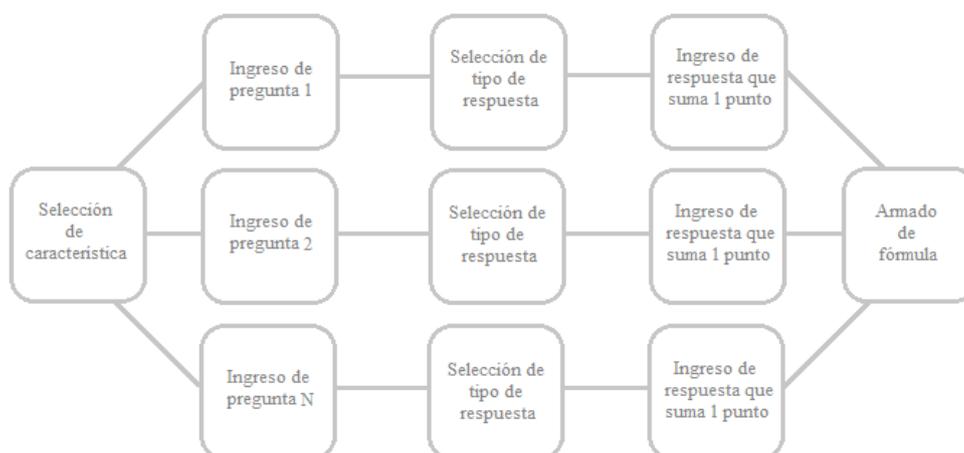
- Diseño de la evaluación
  - Diseño para características generales
  - Diseño para atributos
- Realización de la evaluación

#### 4.1 Diseño de la evaluación

La etapa de diseño de evaluación tiene como objetivo brindarle a un interesado la posibilidad de crear un conjunto de preguntas destinadas a medir los atributos<sup>1</sup> de sus datos en base a las características brindadas por la ISO/IEC 25012 en una futura evaluación. Para ello, la herramienta presentará dos secciones: diseño para características y diseño para atributos.

##### 4.1.1 Diseño para características

Esta sección de diseño está enfocada en medir las características definidas en la ISO/IEC 25012 mediante un conjunto de preguntas genéricas. En base a este objetivo, el interesado tendrá la posibilidad de cargar cada una de las preguntas, sus respuestas permitidas y el armado de la fórmula para la métrica final. La *Fig 1* define el proceso de diseño.



*Fig 1. Proceso de diseño de evaluación para características.*

El proceso está compuesto por:

- **Selección de característica:** Se deberá seleccionar una característica del conjunto de características presentadas por la ISO/IEC 25012. *Por ejemplo: Seleccionaremos las características “Actualidad” y “Compleitud”.*

<sup>1</sup> *Atributo: hace referencia a una característica de una entidad en la base de datos.*

- **Ingreso de pregunta 1..N:** Se ingresarán un conjunto de preguntas de texto libre asociadas a la característica seleccionada.

*Siguiendo con el ejemplo, preguntas asociadas para la característica “Actualidad” podrían ser: (P1) ¿Los datos fueron actualizados en los últimos 6 meses? (P2) ¿Los datos que se utilizan en el entorno de desarrollo se actualizan semanalmente con los datos utilizados en producción?. Para la característica “Compleitud”, preguntas asociadas podrían ser: (P3) ¿Qué porcentaje aproximado de todos los datos se encuentran completos?*

- **Selección de tipo de respuesta:** Se indicará qué tipo de respuesta se espera para la pregunta: SI/NO, numérica, selección de un valor de un rango de valores, entre otros. La respuesta debe poder ser cuantificable.

*Siguiendo con el ejemplo, el tipo de respuesta para ambas preguntas asociadas a Actualidad será “SI/NO”. El tipo de respuesta para la pregunta asociada a Compleitud será “Numérica”.*

- **Ingreso de respuesta que suma un punto:** Se indicará qué respuesta es la esperada para que tenga peso en la fórmula final. Puede ser la coincidencia total con un valor esperado o un valor mayor/menor a lo esperado. En caso de que la respuesta no sea la esperada, no tendrá ningún peso en la fórmula final.

*Siguiendo con el ejemplo, la respuesta que tendrá peso en la fórmula es SI para ambas preguntas asociadas a Actualidad. Para la pregunta asociada a Compleitud, tendrá peso una respuesta mayor o igual a 80% (este valor es establecido por el encargado de diseñar la evaluación).*

- **Armado de fórmula:** Se indicará la fórmula final teniendo en cuenta todas las preguntas realizadas para el armado de la métrica. El resultado de la fórmula deberá generar un valor entre 0 y 1, siendo 1 el valor superior y 0 el valor inferior.

*Siguiendo con el ejemplo, la fórmula para la característica Actualidad será:*

$$0 \leq (P1 + P2) / \text{CANTIDAD DE PREGUNTAS} \leq 1$$

*La fórmula para la característica Compleitud será:*

$$0 \leq P3 \leq 1$$

El valor obtenido en la fórmula determinará el valor de la métrica. Posteriormente, en la evaluación se indicará cuál es el valor esperado para cada una de las características.

#### 4.1.2 Diseño para atributos

Usualmente, existen casos en donde se desea evaluar una característica para un atributo específico de la base de datos. A modo de ejemplo, una persona puede estar interesada en medir la característica “Precisión” del atributo -precio- pero no del atributo -nombre-. Para ello, esta segunda sección de diseño, a diferencia de la primera, permite asociar las preguntas a un atributo específico.

Adaptando el proceso presentado en la primer sección, al seleccionar la característica, se deberán cargar cada uno de los atributos que se deseen medir mediante preguntas para dicha característica. Posteriormente, el proceso funciona de forma similar a la primer

sección.

*Siguiendo con el ejemplo, se crearán preguntas específicas para medir la característica “Completitud”, puntualmente de los atributos -DNI- y -CUIL-.*

*(P4) ¿Qué porcentaje aproximado de los números de DNI se encuentran completos? | Tipo de respuesta: Numérica | Respuesta que suma un punto: >= 90%*

*(P5) ¿Qué porcentaje aproximado de los números de CUIL se encuentran completos? | Tipo de respuesta: Numérica | Respuesta que suma un punto: >= 60%*

*Fórmula:  $0 \leq (P4+P5) / CANTIDAD\ DE\ PREGUNTAS \leq 1$*

#### 4.2 Realización de la evaluación

Al momento de finalizar el diseño de la evaluación, la misma queda disponible para su realización y ejecución. Un evaluador será el encargado de llevar a cabo el proceso y completar cada sección de la estructura definida basada en la ISO/IEC 25040. La herramienta permitirá incluir lo definido en las etapas de diseño al momento de realizar y ejecutar la medición.

Para ello, se seguirá con el ejemplo presentado en el diseño y se tomará un conjunto de datos de prueba definidos en la *Tabla 2*.

*Tabla 2. Datos de afiliados.*

DNI	CUIL	Apellido	Nombre	Domicilio	Teléfono
36111555	27361115554	DIAZ	ALEJANDRA	16-1616	
12546548	20125465488	GARCIA	LUIS	18-1201	13-6422
	23456584541	FERNANDEZ	PEDRO	14-5675	
37394444		RODRIGUEZ	SOFIA	13-5456	45-7865
	20342581659	PEREZ	CAROLINA		67-5698

#### Establecer los requisitos de la evaluación

En primer lugar, se define el propósito de la evaluación.

*El propósito de la evaluación es determinar qué tan actualizados y qué tan completos se encuentran los datos, haciendo hincapié en el DNI y en el CUIL de los afiliados.*

Basándose en el propósito, se seleccionan las características de interés para la evaluación, definidas en la norma ISO/IEC 25012.

*Las características de la ISO/IEC 25012 seleccionadas para la evaluación serán: Completitud y Actualidad.*

#### Especificar la evaluación

En esta etapa de la evaluación, es necesario definir criterios de decisión<sup>2</sup> para cada una de las características y para la evaluación final. Para ello, se deberá definir para cada característica cuáles son los valores considerados (siempre deberán ser entre 0 y 1) para los rangos: inaceptable, mínimamente aceptable, rango objetivo y excede los requerimientos. Posteriormente, deberán indicarse los valores considerados para la evaluación final.

*La Tabla 3 define los criterios de decisión de cada una de las características evaluadas.*

*Tabla 3. Criterios de decisión de las características.*

	Característica: Actualidad	Característica: Completitud
Excede los requerimientos	$0.8 \leq \text{valor\_métrica} < 1$	$0.9 \leq \text{valor\_métrica} < 1$
Rango objetivo	$0.4 \leq \text{valor\_métrica} < 0.8$	$0.6 \leq \text{valor\_métrica} < 0.9$
Mínimamente aceptable	$0.2 \leq \text{valor\_métrica} < 0.4$	$0.4 \leq \text{valor\_métrica} < 0.6$
Inaceptable	$0 \leq \text{valor\_métrica} < 0.2$	$0 \leq \text{valor\_métrica} < 0.2$

### **Diseñar la evaluación**

Al momento de diseñar la evaluación, debe especificarse si la evaluación será realizada con una cantidad específica de usuarios respondiendo las preguntas o si las preguntas serán respondidas por un único usuario.

*Las preguntas serán respondidas por una única persona, la cual posee acceso a los datos de prueba.*

### **Ejecutar la evaluación**

Esta etapa presentará todas las preguntas definidas previamente. No se hará distinción sobre la característica que se está evaluando. Deberán ser respondidas en su totalidad para poder llevar a cabo la medición.

*(P1) ¿Los datos fueron actualizados en los últimos 6 meses?*

*RTA: SI (equivale a 1 en la fórmula por lo establecido en el diseño)*

*(P2) ¿Los datos que se utilizan en el entorno de desarrollo se actualizan semanalmente con los datos utilizados en producción?*

*RTA: NO (equivale a 0 en la fórmula por lo establecido en el diseño)*

*(P3) ¿Qué porcentaje aproximado de todos los datos se encuentran completos?*

*RTA: 80% (equivale a 1 en la fórmula por lo establecido en el diseño)*

*(P4) ¿Qué porcentaje aproximado de los números de DNI se encuentran completos?*

*RTA: 60% (equivale a 0 en la fórmula por lo establecido en el diseño)*

*(P5) ¿Qué porcentaje aproximado de los números de CUIL se encuentran completos?*

*RTA: 80% (equivale a 1 en la fórmula por lo establecido en el diseño)*

---

<sup>2</sup> Rangos de valores clasificados en: Inaceptable, Mínimamente Aceptable, Rango Objetivo y Excede los Requerimientos.

### Finalización de la evaluación

Una vez que finalizó la ejecución, la herramienta ejecutará las fórmulas correspondientes y mostrará los resultados para cada una de las características evaluadas. Para ello, se tiene en cuenta la respuesta del usuario, la fórmula especificada en el diseño de la evaluación y los criterios de decisión para definir el rango de la métrica final. Esta información se espera que sea de utilidad para los desarrolladores, los cuales podrán mejorar sus futuros desarrollos para evitar generar y almacenar datos de baja calidad en sus sistemas.

#### **ACTUALIDAD – Resultados**

*Fórmula:*

$$(P1 + P2) / \text{CANTIDAD DE PREGUNTAS} = (1 + 0) / 2 = 0,5 \implies \text{Rango objetivo}$$

#### **COMPLETITUD – Resultados**

*Subfórmula general (F1):*

$$P3 = 1$$

*Subfórmula por atributos (F2):*

$$(P4 + P5) / \text{CANTIDAD DE PREGUNTAS} = (0 + 1) / 2 = 0,5$$

*Fórmula:*

$$(F1+F2) / \text{CANTIDAD DE SUBFÓRMULAS} = (1 + 0,5) / 2 = 0,75 \implies \text{Rango objetivo}$$

## 5. Conclusiones

Se presentó MED, un modelo de evaluación de datos basado en el enfoque GQM tomando como base las características definidas en la norma ISO/IEC 25012. GQM parte de un objetivo concreto para luego crear preguntas asociadas a dicho objetivo y, mediante la combinación de las respuestas de las mismas, se obtiene una fórmula destinada a crear la métrica asociada.

MED aborda la evaluación de calidad de datos de una manera mucho más sencilla y ágil, permitiendo a una organización conocer el estado de sus datos logrando así información más confiable.

Se realizó una evaluación de calidad de datos basada en ISO/IEC 25040 utilizando el diseño de evaluación definido en el modelo y tomando como ejemplo dos características: “Actualidad” y “Compleitud”. Para las respuestas a las preguntas definidas, se tomaron como referencia un conjunto de datos de ejemplo que sirvieran de apoyo para llevar a cabo la evaluación.

Se proyecta ampliar el modelo, generando una herramienta usable por cualquier tipo de persona facilitándole la tarea de diseño y realización de evaluaciones de calidad de datos.

## 6. Bibliografía

1. ISO/IEC 25012:2008. *Software engineering -- Software product Quality Requirements and Evaluation (SQuaRE) -- Data quality model.*
2. ISO/IEC 25040:2011. *Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Evaluation process.*

3. *“The goal question metric approach”*. Basili, Victor R; Caldiera, Gianluigi; Rombach, H. Dieter.
4. *ISO/IEC 25024:2015. Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Measurement of data quality*.
5. *“Guía para evaluar calidad de datos basada en ISO/IEC 25012”*. Calabrese, Julieta; Esponda, Silvia; Pasini, Ariel, Boracchia, Marcos; Pesado, Patricia. Congreso Argentino de Ciencias de la Computación - CACIC 2019.
6. *“Asistente para la evaluación de calidad de producto de software según la familia de normas ISO/IEC 25000 utilizando el enfoque GQM.”*. Tesina de Licenciatura en Sistemas. Autoras: Muñoz, Rocío; Calabrese, Julieta. 2018.

# Workshop Arquitectura, Redes y Sistemas Operativos

Coordinadores

Jorge Ardenghi (UNS)  
Carlos Buckle (UNPSJB)  
Marcelo Arroyo (UNRC)

## PARÁMETRO H: ANÁLISIS DE LA AUTOSIMILARIDAD DE UN TRÁFICO VoIP

Sergio A. Villiani<sup>1</sup>, Diego R. Rodríguez Herlein<sup>2</sup>

<sup>1</sup> Facultad Ciencias Económicas y Sociales – UCALP  
[sergio.villiani@ucalpvirtual.edu.ar](mailto:sergio.villiani@ucalpvirtual.edu.ar)

<sup>2</sup> Unidad Académica Río Gallegos - UNPA  
Campus universitario - Oficina B 18  
[dherlein@uarg.unpa.edu.ar](mailto:dherlein@uarg.unpa.edu.ar)

**Resumen.** Los modelos tradicionales de tráfico permiten controlar la variabilidad de la demanda y, se puede ejercer control de manera de garantizar algunos niveles de Calidad de Servicio (QoS). La naturaleza autosimilar del tráfico multimedia impone nuevos desafíos en las estrategias para configurar y garantizar una calidad de servicio extremo a extremo. Dada su creciente utilización, resulta interesante conocer las características del tráfico que genera. Estudios previos ponen en evidencia las características autosimilares del tráfico. Esta propiedad se puede cuantificar a través del parámetro de Hurst. El presente trabajo es un estudio del comportamiento del tráfico de VoIP en una red LAN real y las recomendaciones de configuración de los elementos de la red, mediante la evaluación del parámetro H. La plataforma de VoIP está basada en Microsoft® Lync Server 2010 y SIP gateways para la interconexión con la red de telefonía pública. En cuanto a los clientes, existen teléfonos IP con clientes Lync incorporados y teléfonos USB (handset USB), con lo cual el tráfico de la PC y el tráfico de voz compiten sin ningún tipo de priorización. El escenario en el cual se tomaron las muestras fue solamente las llamadas salientes a la PSTN y las conferencias ya que las llamadas entre internos son punto a punto.

**Palabras clave:** VoIP, Tráfico, Autosimilaridad, Hurst, QoS

### 1 Introducción

La Autosimilaridad se define como la apreciación de una estructura estadísticamente similar a lo largo de un amplio rango de escalas, es decir, que la señal en términos estadísticos posee un aspecto parecido, aunque se varíe la escala. Es un concepto relacionado a la invariabilidad o persistencia de las características estadísticas de un proceso estocástico, frente a los cambios de escala temporales. Se puede clasificar como Autosimilaridad Determinística y Estadística.

La noción de fractal [1], introducida por Mandelbrot, se utiliza principalmente para el estudio de estructuras irregulares y fragmentadas que se presentan invariantes en diferentes escalas. Esta característica se denomina autosimilitud, ya que cada una de las partes, cualquiera sea su resolución, se presenta semejante al todo. Esto significa que observando tales objetos con diferentes niveles de aproximaciones las

características geométricas se preservan [2]. Ejemplos de fractales determinísticos son Curva de Koch, Conjunto de Cantor [3], Triángulo de Sierpinski y Alfombra de Sierpinski.

Se dice que un proceso posee autosimilaridad estadística [4] si las estadísticas no cambian cuando cambia la escala temporal. Los procesos que presentan este tipo de autosimilitud permiten modelar el tráfico autosimilar. Para medirla se pueden utilizar estadísticas de segundo orden, que permiten capturar la variabilidad de los procesos. Dicha invarianza a la escala se puede definir en términos de la autocorrelación.

La *autosimilitud estadística en el sentido estricto* [1], se define de la siguiente manera: dado un proceso estocástico  $\{X(t), t \in \mathbb{R}\}$ , un factor escalar  $a > 0 \in \mathbb{R}$  y un parámetro  $H$  tal que  $0,5 < H < 1$  donde se cumpla la condición (1),

$$\{X_a(t_1), X_a(t_2), \dots, X_a(t_n)\} = \{a^H X(t_1), a^H X(t_2), \dots, a^H X(t_n)\} \quad (1)$$

Se dice que dicho proceso es autosimilar con parámetro  $H$ , lo que significa que ambas secuencias poseen la misma distribución de probabilidad.  $H$  se denomina **Parámetro de Hurst** y es una medida de la persistencia del comportamiento del proceso y de la *Dependencia de Largo Rango*.

Si en lugar de establecer las equivalencias a nivel de Distribución de Probabilidad [5] se lo realiza a nivel de estadísticas de 2º orden, como lo son la Función de Autocorrelación  $R$  y la Varianza  $Var[X]$ , de la secuencia temporal que representa el proceso estocástico, se está frente a una *Autosimilaridad en Sentido Amplio*.

Cuando se trata de estudiar el tráfico de una red de datos, se pueden considerar procesos estocásticos definidos en puntos discretos del tiempo. Lo que significa el promedio de los valores de cada bloque no solapados de tamaño  $m$ . El promedio del  $k$ -ésimo bloque es el valor de la  $k$ -ésima variable aleatoria del proceso agregado  $X_k^{(m)}$ .

Siendo  $\gamma_k^{(m)}$  la Autocovarianza de  $X_k^{(m)}$  en términos de estacionariedad de segundo orden se puede afirmar lo siguiente:

- *Autosimilaridad exacta de segundo orden*:  $X(t)$  es exactamente autosimilar de segundo orden con parámetro de Hurst  $H$ , donde  $0,5 < H < 1$

$$\text{si } \rightarrow \gamma(k) = \frac{\sigma^2}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \quad (2)$$

- *Autosimilaridad asintótica de segundo orden*:  $X(t)$  es asintóticamente autosimilar de segundo orden con parámetro de Hurst  $H$ , donde  $0,5 < H < 1$

$$\text{si } \rightarrow \lim_{m \rightarrow \infty} \gamma^{(m)}(k) = \frac{\sigma^2}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \quad (3)$$

De acuerdo a las definiciones anteriores, las distribuciones de  $X$  y  $X^{(m)}$  están vinculadas a través de una relación de escala que involucra al parámetro de Hurst (4).

$$X =_d m^{1-H} X^{(m)} \quad (4)$$

$=_d$ : significa que ambos procesos poseen la misma distribución de probabilidades y la ecuación se satisface tanto para  $m \geq 0$  como para  $m \rightarrow \infty$ .

La relación del parámetro de Hurst con la varianza (5) es la siguiente:

$$\begin{aligned} & Var(X^{(m)}) \\ & = m^{2H-2} Var(X) \end{aligned} \quad \begin{cases} Var(X^{(m)}) = \frac{1}{m} Var(X) & \text{para } H = \frac{1}{2} \\ Var(X^{(m)}) = \frac{1}{m^\beta} Var(X) & \text{para } \frac{1}{2} < H < 1 \quad 0 < \beta < 1 \rightarrow H = 1 - \frac{\beta}{2} \end{cases} \quad (5)$$

$$\text{función de autocorrelación } \rightarrow r(k) = \frac{\gamma(k)}{Var(k)} \quad (6)$$

Cuya aproximación asintóticamente exacta es la siguiente expresión (7),

$$r(k) = \frac{1}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \approx H(2H-1)k^{2H-2} \quad k \rightarrow \infty \quad (7)$$

En particular

$$r(k) = ck^{-\beta} \quad \text{para } \frac{1}{2} < H < 1 \quad \wedge \quad 0 < \beta < 1 \quad \wedge \quad c > 0 \quad (8)$$

La función decae muy lentamente (en forma hiperbólica), lo que conduce a la propiedad de que la función de autocorrelación es no sumable. Cuando esto se cumple el proceso estacionario  $X(t)$  tiene una *Dependencia de Rango Largo (LRD)* [7]. La LRD, es una característica relacionada con el comportamiento de la correlación de un proceso estocástico, en particular con su invariabilidad frente a las traslaciones temporales.

Si el proceso es estacionario y ergódico (la media de su conjunto es igual a la media temporal de una sola realización), el valor de  $\beta = 1$  o lo que es lo mismo  $H = \frac{1}{2}$ , con lo cual la varianza de la media temporal decrece hasta cero con una pendiente  $\frac{1}{m}$  y no presenta LRD. Al ser la variable independiente, el coeficiente de correlación  $\rho = 0$ , la función de autocorrelación  $R_X = 0 \Rightarrow R_{X^{(m)}} \rightarrow 0$  en forma rápida o exponencial cuando  $m \rightarrow \infty$ , con lo cual no existe el efecto memoria, es decir se está en presencia una *Dependencia de Rango Corto o SRD*.

### 1.1 Métodos de estimación del Parámetro H

Un problema fundamental en el análisis series temporales que exhiben LRD, es la estimación del grado de autosimilaridad, o parámetro de Hurst. Algunos de los métodos propuestos para el cálculo de dicho parámetro [9] se observan en la siguiente figura (Fig.1)

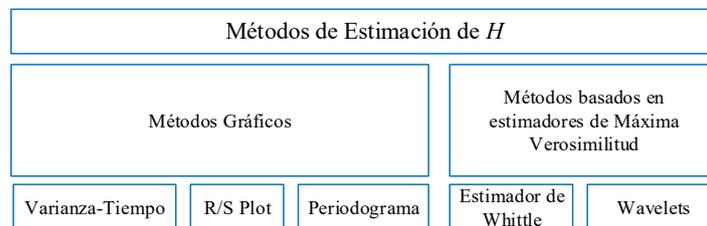


Fig. 1. Métodos de estimación del parámetro H

- *Método de Varianza-Tiempo*: La varianza de un proceso autosimilar de una serie agregada posee esta expresión (9) para valores grandes de m

$$Var[X^{(m)}] \approx \frac{Var(X)}{m^\beta} \quad (9)$$

Como  $H = 1 - \frac{\beta}{2}$  y tomando el logaritmo

$$\log\{Var[X^{(m)}]\} \approx [\log Var(X)] - \beta \cdot \log(m) \quad (10)$$

Como  $\log Var(X)$  es una constante independiente  $m$  de si se grafica  $Var[X^{(m)}]$  vs  $m$  en un gráfico doble logarítmico, se obtiene una línea recta de pendiente  $-\beta$ . Si dicha recta está en el rango (-1,0) se está en presencia de autosimilaridad.

- *Método de Re-escalado:* El método del rango de re-escalado calcula el estadígrafo R/S para diferentes intervalos temporales y permite establecer una relación entre la amplitud del intervalo de desviaciones acumuladas (rango estadístico R) y la desviación estándar de la misma (S), medido para un intervalo de tiempo (t). Para un proceso estocástico  $X(t)$  el rango re-escalado de dicho proceso sobre un intervalo de tiempo N se define la relación  $\frac{R}{S}$  de la siguiente manera (11):

$$\frac{R}{S} = \frac{\max_{1 \leq j \leq N} \{\sum_{k=1}^j [X_k - \mu_N]\} - \min_{1 \leq j \leq N} \{\sum_{k=1}^j [X_k - \mu_N]\}}{\sqrt{\frac{1}{N} \sum_{k=1}^j [X_k - \mu_N]^2}} \quad (11)$$

Siendo  $\mu_N$  la media del proceso, medida sobre el período N

Para un proceso autosimilar si N es grande

$$\frac{R}{S} \cong \left(\frac{N}{2}\right)^H \text{ con } H > 0.5 \quad (12)$$

Aplicando logaritmos

$$\log\left(\frac{R}{S}\right) \cong H \cdot \log(N) - H \cdot \log(2) \quad (13)$$

Graficando  $\left[\frac{R}{S}\right]$  vs N en un gráfico doble logarítmico se obtiene una recta cuya pendiente es el parámetro de Hurst. Los resultados posibles son:

- ✓ Si el proceso es un movimiento aleatorio,  $H = \frac{1}{2}$  el gráfico será una línea horizontal.
  - ✓ Si es un proceso persistente es decir  $H > \frac{1}{2}$  se obtendrá una línea creciente.
  - ✓ Si el proceso es antipersistente es decir  $H < \frac{1}{2}$  se obtendrá una línea decreciente.
- *Método de Periodograma:* Para un proceso estocástico estacionario discreto, la autocorrelación (14) y la densidad espectral (15) se definen:

$$R(k) = E[X(t)X(t+k)] \quad (14)$$

$$S(\omega) = \sum_k R(k)e^{-jk\omega} \quad (15)$$

Si el proceso es ergódico en correlación, es decir, los promedios temporales son iguales a los promedios en conjunto, la función de autocorrelación se la puede estimar mediante (16),

$$R_N(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} X(n+k)X(n) \quad (16)$$

La densidad espectral de un proceso estocástico  $X(t)$  definido en instantes discretos, se puede estimar a lo largo de un período N de la siguiente manera (17),

$$I_N(\omega) = \frac{1}{2\pi N} \left\{ \sum_{k=1}^N X_k e^{jk\omega} \right\}^2 \quad (17)$$

Dicho estimador se conoce como **Periodograma o función de intensidad**.

- *Estimador de Whittle:* [11] Si las series observadas provienen de un proceso autosimilar con parámetro H y que se supone que posee una forma particular como el FBM (Movimiento Browniano Fraccionario), la densidad de potencia del proceso se puede expresar como  $S(\omega, H)$ , entonces H se puede estimar encontrando aquel valor de H que minimiza la siguiente expresión (18),

$$\int_{-\pi}^{\pi} \frac{I_N(\omega)}{S(\omega, H)} d\omega \quad (18)$$

Este es el Estimador de Whittle, si la serie  $\{X_k\}$  tiene una longitud N, la integral se convierte en una suma discreta sobre las frecuencias  $\omega = \frac{2\pi}{N}, \frac{4\pi}{N}, \dots, 2\pi$ .

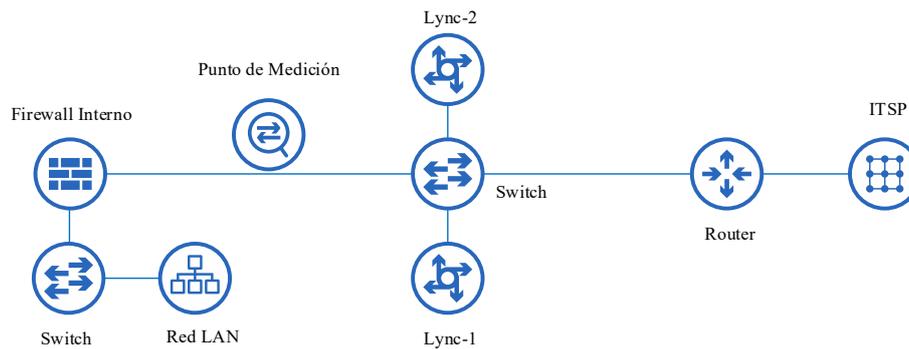
Además, se obtiene la varianza de la muestra, lo que permite calcular intervalos de confianza. La varianza de la muestra se expresa de la siguiente manera (19),

$$\text{Var}(\hat{H}) = 4\pi \left[ \int_{-\pi}^{\pi} \left[ \frac{\partial \log S(\omega)}{\partial H} \right]^2 d\omega \right]^{-1} \quad (19)$$

El estimador de Whittle supone que la serie temporal es un proceso autosimilar de una determinada forma y proporciona una estimación del parámetro de Hurst dentro de un intervalo de confianza [8].

## 2 Captura del tráfico

En la siguiente figura (Fig.2) se observa el escenario que representanta la topología y el punto donde se hicieron las capturas de tráfico.



**Fig. 2.** Topología y punto de captura de tráfico

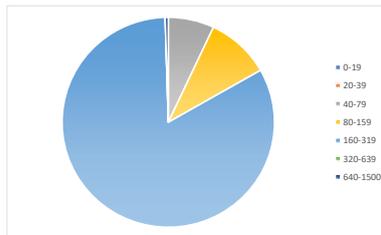
El Punto de Medición es donde concurren todas las llamadas entrantes y salientes (desde y hacia el proveedor ITSP) desde los clientes de la red LAN, todo el tráfico de VoIP pasa por los Servidores Lync además de la señalización SIP. En una llamada entre internos de la red, los servidores Lync intervienen en la señalización y la comunicación se realiza punto a punto entre estaciones.

La metodología empleada para capturar el tráfico se basó en la herramienta tcpdump filtrando el tráfico UDP desde y hacia los hosts Lync-1 y Lync-2. Posteriormente dicha captura se analizó con el software Wireshark, las estadísticas básicas de la captura se observan en la siguiente tabla (Tabla 1).

**Tabla 1.** Estadísticas básicas de la captura de tráfico.

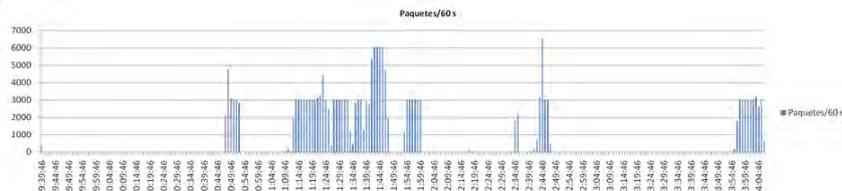
<i>Descripción</i>	<i>Valor</i>
Primer Paquete	<b>09:39:46 a.m.</b>
Último Paquete	<b>02:06:58 p.m.</b>
Tiempo Transcurrido	<b>04:27:12</b>
Paquetes Capturados	<b>205276</b>
Tiempo entre el primer y último paquete en segundos	<b>16032,285</b>
Promedio paquetes por segundo	<b>12,804</b>
Promedio tamaño del paquete en bytes	<b>205,974</b>
Tamaño de la captura en bytes	<b>42281469</b>
Promedio bytes por segundo	<b>2637,27</b>
Promedio Mbit por segundo	<b>0,021</b>

En la siguiente figura (Fig.3) se puede observar la distribución del tamaño de los paquetes capturados

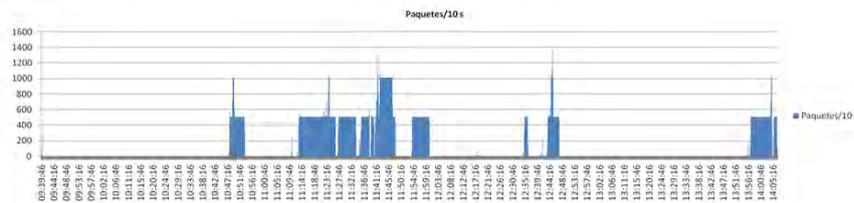


**Fig. 3.** Distribución del tamaño de los paquetes capturados (bytes)

En las siguientes figuras (Fig.4, 5 y 6) se presentan las gráficas de paquetes recibidos con diferentes longitudes de intervalo de tiempo, lo que permite apreciar la existencia de autosimilaridad [4] [7] en el tráfico capturado, ya que se observa un comportamiento similar para las distintas escalas de tiempo [12].



**Fig. 4.** Distribución de la recepción de los paquetes – Intervalo 60s.



**Fig. 5.** Distribución de la recepción de los paquetes – Intervalo 10s

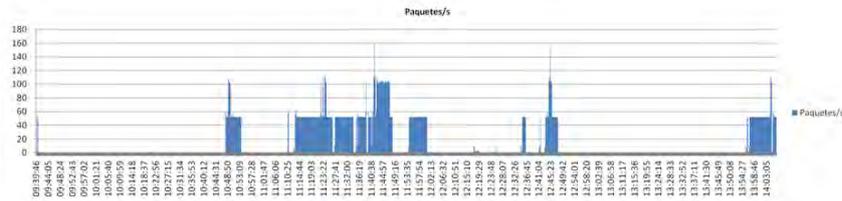
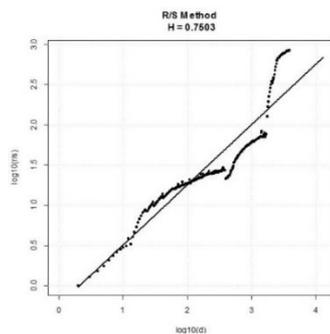


Fig. 6. Distribución de la recepción de los paquetes – Intervalo 1s

### 3 Análisis Estadístico

De la captura original descrita en Captura de Tráfico, se procesó para filtrar el tamaño de los paquetes (bytes) en el intervalo de tiempo correspondiente a 11:41:00 y 11:48:00, dicho intervalo de tiempo es el que posee más cantidades de llamadas concurrentes. Para obtener las estadísticas se utilizó la herramienta informática R Project (<http://www.r-project.org/>), en su versión de 64 bits (2.15.3) [10].

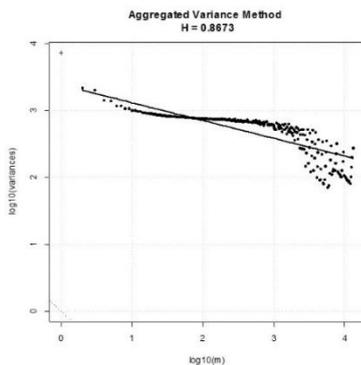
En las siguientes figuras (Fig.7, 8, 9, 10 y 11) se observa del lado derecho el cálculo del parámetro H en RStudio y en la izquierda la gráfica correspondiente, para cada uno de los métodos utilizados.



#### Cálculo de H

```
> rsFit(data[,1], levels = 400, cut.off = 10^(0.1,
4.58), doplot=TRUE)
Title:
Hurst Exponent from R/S Method
Call:
rsFit(x = data[, 1], levels = 400, cut.off =
10^(0.1, 4.58), doplot = TRUE)
Method:
R/S Method
Hurst Exponent:      H      beta
                   0.7502812 0.7502812
Hurst Exponent Diagnostic:
  Estimate Std.Err t-value Pr(>|t|)
X 0.7502812 0.01176258 63.78541 2.84788e-183
Parameter Settings:
  n levels minnpts cut.off1 cut.off2
39006 400 3 1 38018
```

Fig. 7. Cálculo de H – Método R/S



#### Cálculo de H

```
> aggvarFit(data[,1], minnpts = 3, levels = 400,
cut.off = 10^(0.1, 4.58), doplot = TRUE, trace =
FALSE, title = NULL, description = NULL)
Title:
Hurst Exponent from Aggregated Variances
Call:
aggvarFit(x = data[, 1], levels = 400, minnpts =
3, cut.off = 10^(0.1, 4.58), doplot = TRUE, trace =
FALSE, title = NULL, description = NULL)
Method:
Aggregated Variance Method
Hurst Exponent:      H      beta
                   0.8673216 -0.2653568
Hurst Exponent Diagnostic:
  Estimate Std.Err t-value Pr(>|t|)
X 0.8673216 0.004678798 185.3727 0
Parameter Settings:
  n levels minnpts cut.off1 cut.off2
39006 400 3 1 38018
```

Fig. 8. Cálculo de H – Varianza

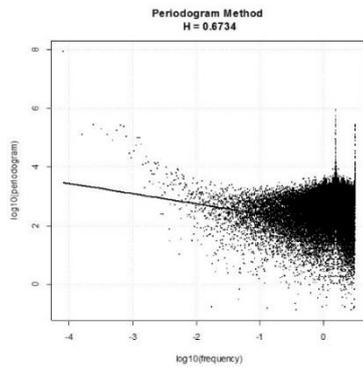


Fig. 9. Cálculo de H – Periodograma

Cálculo de H

```
> perFit(data[,1], cut.off = 0.1, method =
c("per", "cumper"),doplot = TRUE, title = NULL,
description = NULL)
Title:
Hurst Exponent from Periodgram Method
Call:
perFit(x = data[, 1], cut.off = 0.1, method =
c("per", "cumper"), doplot = TRUE, title = NULL,
description = NULL)
Method:
Periodogram Method
Hurst Exponent:      H      beta
                   0.6734411 -0.3468821
Hurst Exponent Diagnostic:
      Estimate Std.Err t-value Pr(>|t|)
X 0.6734411 0.01071573 62.84603      0
Parameter Settings:  n cut.off
                   39006      10
```

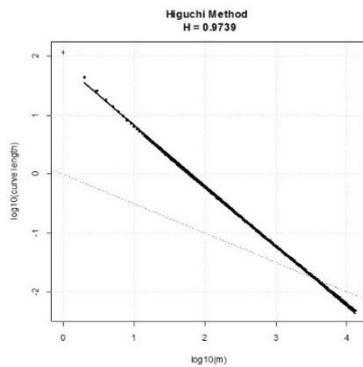


Fig. 10. Cálculo de H – Higuchi

Cálculo de H

```
> higuchiFit(data[,1], levels = 400, minnpts = 3,
cut.off = 10^c(0.1, 4.58), doplot = TRUE, trace =
FALSE, title = NULL, description = NULL)
Title:
Hurst Exponent from Higuchi Method
Call:
higuchiFit(x = data[, 1], levels = 400, minnpts =
3, cut.off = 10^c(0.1, 4.58), doplot = TRUE, trace =
FALSE, title = NULL, description = NULL)
Method:
Higuchi Method
Hurst Exponent:      H      beta
                   0.9739041 -1.0260959
Hurst Exponent Diagnostic:
      Estimate Std.Err t-value Pr(>|t|)
X 0.9739041 0.003075101 316.7064      0
Parameter Settings:  n levels minnpts cut.off1 cut.off2
                   39006      400      3      1      38018
```

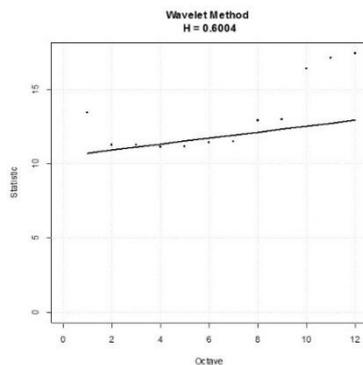


Fig. 11. Cálculo de H – Wavelet

Cálculo de H

```
> waveletFit(data[,1], length = NULL, order = 2,
octave = c(2, 8),doplot = TRUE, title = NULL,
description = NULL)
Title:
Hurst Exponent from Wavelet Estimator
Call:
waveletFit(x = data[, 1], length = NULL, order =
2, octave = c(2, 8), doplot = TRUE, title = NULL,
description = NULL)
Method:
Wavelet Method
Hurst Exponent:      H      beta
                   0.6003966 0.2007933
Hurst Exponent Diagnostic:
      Estimate Std.Err t-value Pr(>|t|)
X 0.6003966 0.04656874 12.8927 5.000241e-05
Parameter Settings:  length order octave1
octave2
                   32768      2      2
                   8
```

En la siguiente tabla (Tabla 2) se observan los valores obtenidos en los distintos métodos de cálculo del parámetro de Hurst.

**Tabla 2.** Cálculos del parámetro H por los distintos métodos.

<i>Método</i>	<i>Valor</i>
R/S	<b>0.7502812</b>
Varianza	<b>0.8673216</b>
Periodograma	<b>0.6734411</b>
Higuchi	<b>0.9739041</b>
Wavelet	<b>0.6003966</b>

Los métodos gráficos son de carácter práctico [9], ya que permiten obtener una idea aproximada si un conjunto de datos se ajusta a características autosimilares ( $H > 0,5$ ) o si posee una dependencia de corto alcance ( $H \approx 0,5$ ). El método de Wavelet presenta un mayor rigor estadístico y en este caso tiene un valor  $H = 0.6003966$ , con lo cual se puede afirmar que los datos obtenidos de la captura presentan un carácter autosimilar.

#### 4 Conclusiones

El tráfico autosimilar en una red de datos puede tener impacto negativo en la performance de la misma incluyendo retrasos mayores en las colas de los elementos de conectividad y pérdida de paquetes.

Una consecuencia práctica del efecto de la autosimilaridad es la necesidad de tener buffers más grandes en routers y switches que los previstos mediante el análisis y simulación de colas tradicional.

Pero estos buffers de mayor tamaño crean mayores retrasos en los flujos de datos. El producto retardo por ancho de banda causa inconvenientes en el tráfico multimedia.

Debido a la rafagosidad de tráfico inherente al tráfico de VoIP, la pérdida de paquetes y los retardos de la red son los problemas comunes que experimentan las aplicaciones multimedia tales como la VoIP.

En particular en este trabajo la plataforma analizada es basada en Microsoft® Lync Server 2010, la única manera de implementar Calidad de Servicio es habilitando DiffServ tanto en los servidores como en las PC clientes y todos los elementos de conectividad que existan entre los clientes y los servidores para garantizar el entorno de QoS.

Las directivas QoS se aplican a una sesión de inicio de usuario o a un equipo como parte de un objeto de directiva de grupo (GPO) que se vincula a un contenedor de Active Directory como un dominio, un sitio o una unidad organizativa (OU).

QoS basada en directiva permite:

- Definir la prioridad del tráfico: Para marcar el tráfico de red saliente mediante un valor específico de Punto de código de servicios diferenciados (DSCP), tal como se describe en RFC 2474.
- Administrar el uso del ancho de banda: Puede limitar de velocidad para el tráfico saliente. Con este límite, los componentes de QoS limitan el tráfico de red

saliente agregado que coincide con la velocidad especificada en la configuración de la directiva QoS.

## References

1. Porras Castro, Stalin Efrain, Rivera Moreno, Paulo Cesar, *Análisis del algoritmo algebraico de asignación dinámica del buffer para sistemas de encolamiento bajo condiciones de tráfico autosimilar*, 2008.
- 2- MA. Alzate Monroy, 2001, *Introducción al tráfico autosimilar en redes de comunicaciones*, Ingeniería, ISSN-e 0121-750X, Vol. 6 Núm. 2 (2001), págs. 6-17.
- 3- S. Plaza, 2013, “*Introducción a los procesos infinitos y Fractales*”
- 4- W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, *On the self-similar nature of Ethernet traffic (extended version)*, in IEEE/ACM Transactions on Networking, vol. 2, no. 1, pp. 1-15, Feb. 1994, doi: 10.1109/90.282603.
- 5- O.I. Sheluhin, S.M. Smolskiy and A.V. Osin, 2007, *Self-Similar Processes in Telecommunications*, Print ISBN:9780470014868 |Online ISBN:9780470062098 |DOI:10.1002/9780470062098, John Wiley & Sons, Ltd.
- 6- P. Bravo García, 2004, *Modelado fractal de tráfico de la red de Ingeniería Telemática*, Universidad Carlos III de Madrid. Departamento de Ingeniería Telemática.
- 7- K.Park and W.Willinger, 1999, *Self-Similar Network Traffic: An Overview*, Wiley Online Library
- 8- P.ULANOVVS ET AL., *Modeling Methods of self-similar traffic for network performance evaluation*, Scientific Proceedings of RTU. Series 7. Telecommunications and Electronics, 2002.
- 9- RG. Clegg, *A Practical Guide to Measuring the Hurst Parameter*, arXiv, Nov. 2006
- 10-Package ‘fArma’, “Rmetrics - Modelling ARMA Time Series Processes”, RStudio, <http://cran.nexr.com/web/packages/fArma/fArma.pdf> (Dic2019)
- 11-Christian Grimm., Georg Schluchtermann; *IP Traffic Theory and Performance*; SpringerVerlag; ISBN 978-3-540-70603-8; 2008.
- 12-Gagandeep Kaur, Dr. Vikas Saxena, Prof. J. P. Gupta; *Characteristics Analysis of Web Traffic with Hurst Index*; Proceedings of the World Congress on Engineering and Computer Science 2010 Vol I WCECS 2010, October 20-22, 2010, San Francisco, USA

## Interfaz Humano-Máquina web amigable para dispositivo IoT.

Rocío Madou<sup>1</sup>, Federico N. Guerrero<sup>1</sup>, Enrique M. Spinelli<sup>1</sup>

<sup>1</sup> Grupo de Instrumentación Biomédica Industrial y Científica (GIBIC),  
LEICI (UNLP-CONICET), Facultad de Ingeniería UNLP,  
116 y 48 S/N, La Plata, Argentina  
{rocio.madou, federico.guerrero, spinelli}@ing.unlp.edu.ar

**Resumen.** Este trabajo presenta una solución para vincularse y configurar un sistema embebido que tiene conectividad WIFI pero que no dispone de ningún dispositivo de entrada ni visualización local. El equipo, diseñado bajo el paradigma IoT y denominado WIMUMO, permite adquirir señales biomédicas y transmitir las utilizando el protocolo OSC a través de una red inalámbrica. Por razones de seguridad eléctrica, se alimenta mediante baterías y está dispuesto en un gabinete plástico cerrado, sin ningún tipo de conector accesible. Sin embargo, esta característica dificulta interactuar con el equipo hasta que logra realizarse su configuración de red. Para facilitararlo, al encenderse WIMUMO genera una red propia y cuenta con un servidor web. Esto permite, mediante un browser de propósitos generales, configurar el equipo para que se conecte a una red existente, definir las direcciones a las cuales debe enviar las señales, entre otros parámetros. Toda la configuración se realiza mediante una interfaz web amigable que implementa criterios de usabilidad establecidos para el buen diseño de interfaces, pudiendo ser realizada sin problemas por usuarios sin más conocimiento que conectar su celular a una red WIFI. La solución propuesta se basa en un sistema multicapas. La capa más baja es capaz de interactuar, utilizando el lenguaje Bash, con los servicios a nivel sistema operativo, modificando los archivos de configuración siguiendo la lógica de la máquina de estados de conexión desarrollada. En las capas de alto nivel se encuentra el servidor desarrollado bajo NodeJS cuyo diseño visual está dado por la librería Bootstrap, encargado de la interacción, recolección y comunicación de datos del usuario, a través de la herramienta Socket.io.

**Palabras clave:** Internet de las cosas (IoT), interfaces humano-máquina (HMI), usabilidad.

### 1 Introducción

Un dispositivo dentro del paradigma de la internet de las cosas (IoT, en inglés) tiene la capacidad de transferir datos a través de una red sin requerir de interacciones humano a humano o humano a computadora. Esta tecnología, que comenzó en el ámbito académico-científico para un reducido conjunto de elementos, y llevada masivamente a los objetos del hogar u oficina por las empresas, se está estableciendo

en todos los ámbitos de la vida cotidiana, desde sistemas de monitoreo de la salud [1][2], hasta performances artísticas [3].

WIMUMO es un dispositivo de medida no invasivo, para procesamiento y visualización de señales electrofisiológicas o biopotenciales. Dentro de su amplio espectro de aplicaciones, puede utilizarse para performances artísticas. En este ámbito, permite tomar medidas de señales de electromiograma (EMG) generadas por un performer, que son captadas mediante electrodos y transmitidas dentro de una misma red LAN, utilizando el protocolo OSC, a un dispositivo generador de multimedia [4]. Se trata de un equipo diseñado con tecnología disponible comercialmente, el cual no posee una interfaz para la visualización o el ingreso de datos (como por ejemplo una pantalla), pero cuenta con conexión inalámbrica WiFi. En una primera versión, la instalación y configuración del equipo se realizaba a través de una conexión por consola remota y segura (SSH) por una persona con alto conocimiento técnico. Esta solución no es apropiada para este equipo, que será usado por personas sin conocimiento de la tecnología. Esto exige que su puesta en funcionamiento sea simple, amigable al usuario y cumpla con la heurística que define los criterios de usabilidad del dispositivo [5].

Teniendo en cuenta el ámbito al que está orientado WIMUMO y que debe operar en distintas locaciones, en algunos casos sin conexión a internet, debe ser utilizado y calibrado sin requerir una tecnología puntual por parte del usuario, i.e. requiere un sistema multiplataforma que cuente con un modo de configuración mínimo.

Además, debe explotar el conocimiento informal adquirido del usuario. Es decir, aprovechar el aprendizaje previo sobre la manipulación de un elemento que posea características similares, como lo puede ser la conexión de un celular a una red inalámbrica.

En la literatura se encuentran estrategias para dispositivos similares, que permiten la configuración usando una conexión USB [6] u otros medios que no son los tradicionales, como la configuración a través del sonido [7], que involucran cierta dependencia con la conectividad web, aplicaciones propietarias [8] o hardware específico. El desarrollo de aplicaciones locales multiplataforma puede ser muy costoso, ya que los elementos de software utilizados varían según el dispositivo en donde se va a trabajar. En cambio, las aplicaciones web corren sobre los navegadores ya incorporados por los fabricantes del equipo, las dependencias se limitan sólo a las características del navegador, sin requerir la instalación de herramientas adicionales obteniendo estabilidad y soporte independientemente de la plataforma.

Se propone como solución un sistema de capas que involucra la interacción del algoritmo como parte del sistema operativo, con un desarrollo de alto nivel encargado del manejo de la interfaz con el usuario a través de un servidor web, capaz de cumplir con los criterios de usabilidad considerados necesarios.

## 2 Materiales y métodos

En las siguientes subsecciones se describe la implementación de un sistema de configuración de redes a través de un servidor web, comenzando por la especificación de las herramientas necesarias para su desarrollo, y la descripción de una situación de uso promedio de un usuario, seguido de una descripción detallada de los pasos

necesarios para llevar a cabo dicha configuración. Luego, desde un punto de vista técnico, en primer lugar, se explica la lógica de la estructura de funcionamiento y, posteriormente, la implementación del mecanismo. Finalmente, se identifican los criterios a cumplir por la heurística de usabilidad.

## 2.1 Plataforma

WIMUMO se desarrolló a partir de una minicomputadora Raspberry Pi Zero W, controlada por el sistema operativo Raspbian Buster lite, versión septiembre 2019, siendo su comunicación con el usuario posible sólo a través de una interfaz de comandos. Raspbian cuenta con un conjunto de procesos, librerías y herramientas llamado systemd, encargado de la administración del sistema y los servicios. A éste, se le integró el servicio de avahi-daemon, para el reconocimiento del dispositivo a través de su nombre de host.

La Raspberry Pi Zero W posee conectividad inalámbrica WiFi, pero su hardware se limita a la comunicación con un único perfil de red, y no permite su configuración como access point (AP) y cliente al mismo tiempo.

Para implementar un servidor descripto más adelante, se utilizó Node.js v.12.16.2, el gestor de paquetes de node, npm v.6.14.4, y paquetes correspondientes al módulo de comunicación paralela entre servidor y cliente socket.io. Éste fue diseñado para coexistir y trabajar en conjunto con la muestra y el envío de datos desarrollados previamente.

## 2.2 Descripción del funcionamiento

WIMUMO es un equipo portátil que puede ser llevado a un laboratorio, consultorio, sala de ensayo, o cualquier locación. Quien lo utiliza podría ser cualquier persona interesada en el uso de las señales de biopotencial, hasta el momento han sido usuarios ligados a la performance artística. Al llegar a la locación, el usuario debe poder conectar WIMUMO a la infraestructura WiFi existente con relativa facilidad, o como mínimo poder acceder a la interfaz del equipo a través de una red provista por el equipo mismo.

Al presentarse por primera vez al usuario, WIMUMO se encuentra en modo access point. Por este motivo, para comenzar su configuración, el usuario se conecta a la red perteneciente al WIMUMO desde un dispositivo con conexión WiFi, ingresando las credenciales correspondientes. Una vez dentro de la misma, abre una ventana del navegador disponible e ingresa la dirección IP '192.168.99.1', elegida como la receptora de pedidos del servidor. En los dispositivos que cuenten con el servicio de red Bonjour, es posible utilizar el nombre de host wimumo00x.local.

Una vez dentro del servidor, el usuario se dirige a la sección "Redes" la cual cuenta con una interfaz similar al modo de conexión conocido y utilizado por los usuarios en gran parte de los dispositivos comerciales, en donde se solicita los datos de conexión de la red a la que se desea conectar. Una vez completados los campos, hace click al botón "conectar" que genera un mensaje indicador de que la configuración de red se está realizando. Si la configuración ha sido correcta, el usuario puede ver un mensaje

en fondo verde con instrucciones para comenzar su uso. Caso contrario, un indicador rojo le solicita que vuelva a la pantalla principal y reinicie la configuración.

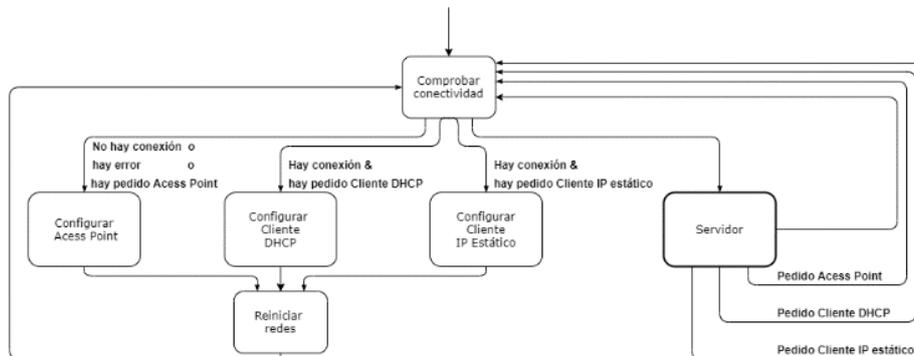
A partir de este momento se encuentra una situación crítica para usuarios poco experimentados ya que enfrentan la desafiante tarea de averiguar la dirección IP que el equipo tomó en la red. Siguiendo las indicaciones presentadas en el “aviso verde de configuración exitosa”, el usuario debe lograr conectarse a la red que indicó previamente y luego de esperar a que WIMUMO esté también conectado, selecciona una de las tres opciones para encontrar la IP recibida por el router, en caso de que no se haya indicado una dirección estática:

- Usar un servicio disponible en la página de WIMUMO en internet.
- Escanear la red con una aplicación disponible en el *store* del sistema operativo (en el propio cartel se recomienda la utilización de *Bonjour Browser* por su facilidad de uso).
- Acceder utilizando el nombre de host *wimumo00x.local* desde un dispositivo que cuente con los servicios de red Bonjour.

A partir de este momento, el usuario tiene acceso a todas las prestaciones que el WIMUMO tiene para ofrecer desde su servidor.

### 2.3 Descripción del algoritmo

Para lograr el modo de funcionamiento descrito previamente, se implementó una máquina de estados que se integró al inicio del sistema operativo de la Raspberry.



**Fig. 1** Máquina de estados que representa el funcionamiento del algoritmo de configuración de modo de conexión del WIMUMO.

Como se puede ver en la Figura 1, la configuración comienza con la comprobación de conectividad del dispositivo. La primera vez que se inicia la Raspberry, no cuenta con ninguna configuración preexistente, por lo que no puede poseer un modo de conexión. Al detectar este estado, se llama a la solicitud de configuración como access point. Una vez finalizado, se reinicia el sistema de red y se vuelve a comprobar la existencia de una conexión. Esta devuelve una confirmación, lo que permite la puesta en funcionamiento del servidor.

El modo descrito se definió como *modo default*, ya que en caso de no posibilidad de conexión o falla en los archivos involucrados, siempre vuelve al modo AP.

Con el servidor en funcionamiento el usuario puede solicitar el modo de cliente de red, utilizando los protocolos DHCP (IP asignada de forma dinámica por el router) o IP estática (asignada por el usuario). Ingresados los datos correspondientes a cada protocolo, se inicia la configuración y, una vez finalizado, el reinicio de red.

Por última vez, se vuelve a comprobar de forma exitosa la conexión a la red y se reinicia el servidor disponible para comenzar con la adquisición.

## 2.4 Descripción de la implementación

El desarrollo de la solución de la problemática se dividió en dos capas. La primera, de bajo nivel, en donde la combinación de algoritmos se transforma en servicios del sistema operativo para una mejor sincronización y estabilidad ante la posibilidad de errores.

Simulando el funcionamiento de comandos desde una consola, Bash es el lenguaje elegido para los algoritmos correspondientes a la máquina de estados, programa principal y funciones independientes que contienen el accionar de cada estado, y la administración servidor. Se utilizaron *flags* o banderas de estado, a través de la presencia o ausencia de archivos vacíos sin extensión. Los scripts principales de la máquina y el servidor tienen asociado un servicio particular, a través del cual se desarrolla su puesta en funcionamiento o detención. Por ello, la máquina de estados puede llamar a un script de configuración, o iniciar o detener cualquier servicio del sistema operativo, según corresponda, al igual que lo hacen el resto de los bloques administrados por Systemd.

En una capa superior, donde se encuentra la aplicación web, los elementos de comunicación cliente-servidor y sus aspectos funcionales y estéticos, los datos de conexión son obtenidos, procesados y almacenados mediante una interfaz que combina la utilización de formularios, listas y botones, junto con avisos de estado e instructivos.

En el momento en que el usuario ingresa la IP asignada al WIMUMO en la barra de búsqueda del navegador, se inicia la ejecución de una serie de etapas de interacción cliente-servidor [2]. El servidor recibe un pedido de conexión, emitiendo una respuesta que incluye la renderización de una página HTML, cuyo funcionamiento está determinado por código JavaScript embebido en la misma, y su diseño por los elementos de la librería Bootstrap personalizados en archivos CSS. Además, se abre un canal de comunicación asincrónica paralela a través del uso de socket.io.

Al cargarse la pestaña de “configuración de redes”, comienza el llamado sincrónico de funciones que derivan en la solicitud al sistema operativo, transferencia por sockets y muestra de una lista de redes visibles por la antena de la placa, como primer elemento del formulario que compone la interfaz. Una vez seleccionada la SSID, o nombre de red, por el usuario, es necesario que se complete el campo de texto oculto correspondiente a la contraseña de la misma. Bajo el mismo, se muestra un checkbox, que habilita la muestra de los campos y el ingreso de los datos necesarios para la configuración del protocolo IP estático (dirección IP deseada, máscara de red, dirección del router y DNS). También es posible activar la solicitud de modo access

point, activando el checkbox correspondiente, el cual oculta todos los campos de ingreso de datos y selecciones, dejando sólo visible el botón de confirmación del formulario.

Al hacer clic a este último, se recolectan los datos correspondientes, dependiendo de la configuración deseada por el usuario, y se envían a través del websocket a la primera función de configuración de parámetros dentro de servidor, mostrándose una nueva pestaña, la cual contiene una barra de carga y un mensaje que indica que el proceso está en curso. A partir de este momento, se comienza a trabajar con un módulo que permite la ejecución de comandos en una consola, y la modificación de archivos de cualquier directorio del sistema operativo. Dichos datos son copiados en los archivos de redes correspondientes, los flags con creados o eliminados, y el servicio de la máquina de estados es iniciado.

Si la secuencia de elementos fue realizada con éxito, el usuario visualiza en pantalla un instructivo de acciones para poder volver a conectarse al servidor y hacer uso del mismo. En caso contrario, se le indica el fallo de la operación, solicitando que vuelva a iniciarla.

## 2.5 Descripción de la métrica para determinación de usabilidad

El estándar ISO 9241-11 [9] define el término usabilidad como “el alcance en el que un producto puede ser utilizado por usuarios específicos para alcanzar metas determinadas con efectividad, eficiencia y satisfacción en un contexto particular de uso”, por lo que los criterios de calidad en la experiencia de usuario son variables dependiendo la aplicación.

En el caso de los dispositivos electrónicos o aplicaciones móviles, existen una serie de factores que pueden afectar la usabilidad del instrumento, desde la interfaz gráfica, o fisiología del usuario, hasta los aspectos sociales [10]. Es por ello que se estableció cierta heurística o principios del “buen diseño de interfaces” [5].

De acuerdo a la implementación que se buscaba para la tecnología, según la heurística establecida en el artículo citado previamente, las características que se consideró necesario su cumplimiento fueron: Diseño visual intuitivo, Visibilidad del estado del sistema, Correspondencia entre el sistema y el mundo real, Flexibilidad y eficiencia de uso, Prevención de errores, Ayuda y documentación.

## 3 Resultados

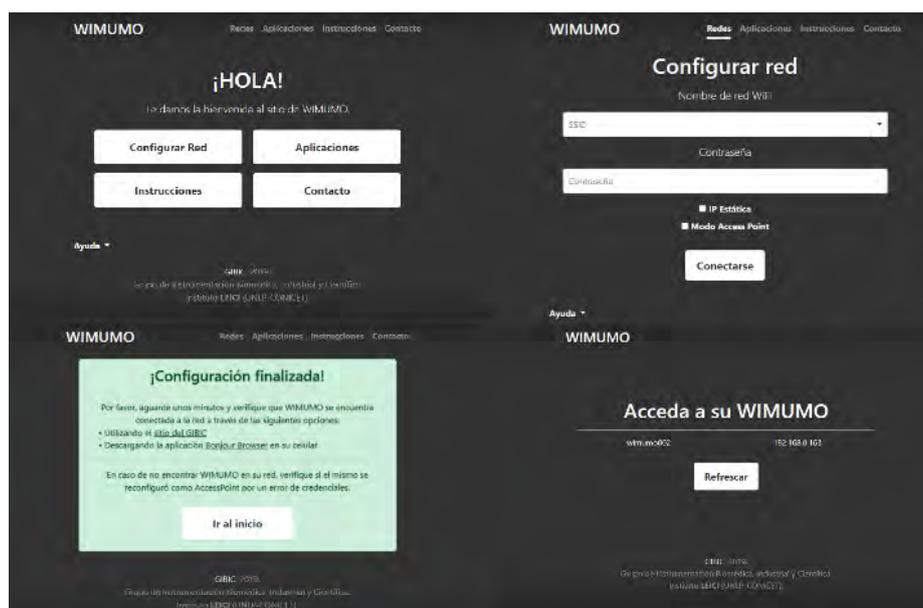
La estrategia adoptada permitió obtener un dispositivo capaz de modificar sus configuraciones a través de un sistema multicapa con comunicación dentro y entre las mismas. En donde, la capa de más alto nivel, encargada de la interpretación directa del usuario, obtiene los datos del nuevo modo de operación, y la de más bajo nivel, integrada dentro del sistema operativo, administra los archivos y servicios logrando la nueva configuración.

Previamente, si el equipo contaba con una versión de configuración previa, era imposible la comunicación sin la intervención de periféricos. Al momento de escribir este artículo, se dispone de un equipo que, ante la imposibilidad de conectarse, auto-

configura su modo de funcionamiento en access point, logrando una comunicación remota no sólo a través de una consola de comandos, sino con un servidor que contiene una interfaz web amigable al usuario, e intuitiva.

Buscando cumplir con la heurística de usabilidad de interfaces de software, se obtuvo un conjunto de pantallas aptas para la explotación del conocimiento previamente adquirido por el usuario en sistemas con similares características de conexión, las cuales, con un diseño minimalista evitando la sobrecarga de información, se pueden definir como auto-explicativas y a partir de su interacción, orientan al usuario retornando un feedback dinámico sobre el estado del equipo en cada momento, Figura 2. El soporte back-end del servidor, con una comunicación asincrónica a través de sockets, asegura que la necesidad de participación del usuario se limite al ingreso de credenciales y a la espera correspondiente a la demora de la configuración.

Al desarrollo de la tecnología se le agregó la redacción de un instructivo mínimo en formato papel, indicando los pasos a seguir por el usuario, junto con los tiempos de retardo que corresponden a cada etapa, evitando de esta forma que la impaciencia del mismo pueda llevar a un apagado abrupto causado por la persona en medio de una configuración.



**Fig. 2:** Diseño de las pantallas que componen el camino de configuración exitoso del WIMUMO.

Gracias a estas características, personas con distinto nivel de conocimiento técnico pudieron cambiar la red fácilmente. Se puso a prueba el rendimiento del equipo y la comunicación con otros dispositivos siendo cliente de una red o administrador de la propia. Además, se comprobó la robustez del sistema ante el ingreso de credenciales incorrectas o la mala señal de red por distanciamiento entre dispositivos.

Además, fue posible la identificación aproximada de los tiempos de funcionamiento del equipo. Al encenderse, si la conexión es exitosa, el servidor tarda 90 segundos en estar disponible y accesible. El peor caso, donde no se puede obtener una IP ya que la red no está disponible y se realiza un cambio de configuración a AP, se aproxima a los 3 minutos. La demora de la configuración de una nueva red a través del servidor toma 2 minutos en promedio.

### 3.1 Implementación de usabilidad

Los criterios de usabilidad fueron satisfechos con éxito a partir de las siguientes implementaciones:

Un *diseño visual intuitivo* se logró a partir de la utilización de elementos simples, con nombres auto-explicativos y textos de ayuda cortos. Los mensajes fueron ideados para ser explicativos y junto con los colores de las alertas y la barra dinámica de carga, permiten *visibilizar el estado del sistema*. Cada pieza dispuesta en un diseño de interfaz conocido previamente por el usuario lleva a que *el sistema pueda corresponderse con el mundo real*. La utilización de las herramientas dinámicas de Bootstrap permitieron la *flexibilidad* necesaria para cada pantalla y *eficiencia de uso* en las características de los distintos dispositivos.

En cuanto a la *prevención de errores*, se implementó con la verificación de que todos los datos necesarios sean ingresados en los campos de texto y la adopción del modo AP ante el ingreso de datos de conexión erróneos por parte del usuario.

Finalmente, se incorporaron dos instructivos al equipo como *ayuda y documentación*. El primero virtual disponible como una pestaña del servidor ante la consulta necesaria al momento de ingreso de las credenciales nuevas. El segundo, anexo al equipo en formato papel, indica los pasos necesarios para lograr una conexión exitosa.

## 4 Discusión

A partir de las pruebas realizadas fue posible distinguir una clara relación entre el conocimiento técnico de una persona y la necesidad de documentación para la configuración del equipo. Para los usuarios con mayor grado de tecnicidad, la documentación sólo fue necesaria al comenzar con la ejecución. En el caso contrario, el no conocimiento sobre el funcionamiento de las redes, limitó su progreso en todos los pasos que involucraban el accionar con cambio de configuración y dirección IP; momentos como la conexión inicial en modo access point, el ingreso al servidor con una dirección particular, la búsqueda de la nueva dirección IP por la solicitud de conexión como cliente con el protocolo DHCP, y el ingreso de dicha dirección para poder volver acceder al servidor. Sin una documentación clara y detallada, los usuarios con este nivel de conocimiento no podrían utilizar el dispositivo.

Existen otras plataformas, como el ESP32, que poseen doble perfil de red. De esta forma es posible la comunicación con el equipo a través de una red externa o la propia del WIMUMO simultáneamente. Esto permite que no sea necesario el cambio de configuración, por lo que un usuario con pocos conocimientos técnicos puede

interaccionar con el servidor desde su dirección de administrador de red, mientras se realiza la comunicación con otros instrumentos a través de la red externa, como cliente.

La retroalimentación del usuario, al igual que la documentación, es un factor que ganó importancia a medida que desarrollaron los componentes. Al presentarse como una caja cerrada, el equipo no cuenta con una respuesta que le permita al usuario conocer el estado del mismo. Esto generó necesidad de implementar notificaciones dentro de la interfaz, indicando, por ejemplo, si se encuentra dentro de los pasos de procesamiento de la información para la configuración de un nuevo perfil de red o, en caso de finalizado, el resultado del proceso. Sin embargo, existe una porción de tiempo en donde el servidor no se encuentra disponible, por lo que se considera importante, con posibilidad de adoptar a futuro, la incorporación de un led RGB indicador de estado.

Lo dicho anteriormente toma mayor relevancia al considerar los tiempos de funcionamiento del mismo. Actualmente, los tiempos de operación son tolerables gracias a los elementos indicadores de estado propios de la interfaz. Sin embargo, se busca reducir dicha tardanza optimizando el software buscando hacer mejor uso de la capacidad computacional de la placa, e incluso se propone pasar a una plataforma como el ESP32.

Es importante destacar que los algoritmos como servicios permitieron una buena integración al sistema operativo estableciendo explícitamente el momento en que se ejecutan los mismos dentro de la secuencia de inicio.

Finalmente, se encontró con una limitación en la flexibilidad del funcionamiento del equipo. Si la red externa que fue configurada por el usuario no se encuentra disponible al momento de encendido, este se auto-configura a su red propia olvidando la conexión anterior, teniéndose que volver a realizar el proceso de cambio de red. En el caso donde se trabaja siempre en la misma red, es esencial que se respeten los pasos de encendido de cada uno de los elementos involucrados, volviéndolo impráctico. Por lo que es necesario la creación de perfiles de red, comprobando por la última configuración como cliente al encender el equipo.

## 5 Conclusiones

El trabajo desarrollado consiste en un sistema de configuración remota de un dispositivo IoT, a través de un servidor web. La utilización de dicha tecnología, permite el acceso a una interfaz amigable e intuitiva con cualquier plataforma con conexión a inalámbrica, que contenga un navegador, pudiendo explotar el conocimiento informal adquirido previamente por el usuario, en dispositivos similares.

El desarrollo de los algoritmos de modificación de configuración en lenguaje Bash manipulados como servicios, posibilita una mejor integración con el sistema operativo Raspbian, añadiendo robustez ante fallos. En cuanto a la interfaz de usuario brindada por el servidor, la utilización de la herramienta Node.js permite el desarrollo rápido y de bajo costo, utilizando de paquetes estables con buen soporte.

Además, se cumple con los principios para un correcto diseño de interfaces, destacando la importancia de la incorporación de documentación clara y ordenada para la interacción con el dispositivo.

Dada la relevancia dentro del tema, se establece como trabajo futuro la realización y documentación de pruebas de campo con usuarios reales.

## 6 Agradecimientos

Este trabajo fue financiado en parte mediante los subsidios I-219 de la Universidad Nacional de La Plata, PIP0558 del CONICET y PICT-2015/2257 de la Agencia Nacional de Promoción Científica y Tecnológica. El trabajo forma parte del proyecto de extensión acreditado FI UNLP “Fabricación de equipos vestibles de medición de señales bioeléctricas para uso en performances artísticas inclusivas”.

## 7 Referencias

1. M. Giannetto, V. Bianchi, S. Gentili, S. Fortunati, I. De Munari, and M. Careri: An integrated IoT-Wi-Fi board for remote data acquisition and sharing from innovative immunosensors. Case of study: Diagnosis of celiac disease. In: *Sensors Actuators, B Chem.*, vol. 273, no. June, pp. 1395–1403 (2018), doi: 10.1016/j.snb.2018.07.056.
2. V. I. Kubov, Y. Y. Dymyrov, and R. M. Kubova: Wireless Devices HTML-interface for Medical Applications. In: 2019 8th Mediterr. Conf. Embed. Comput. MECO 2019 - Proc., no. June, pp. 1–4 (2019) doi: 10.1109/MECO.2019.8760057.
3. A. Ceriani, R. Madou, M. Haberman, and E. Spinelli: Bioelectric signals of the body: from electronic engineering to artistic performance. In: *EAI Endorsed Trans. Creat. Technol.*, vol. 7, no. 22, p. 163481 (2020), doi: 10.4108/eai.13-7-2018.163481.
4. F. N. Guerrero et al.: WIMUMO Project: A Wearable Open Device for Physiological Signals Acquisition. In: *Revista Argentina de Bioingeniería*, vol. 24, no. 3, pp. 13–18 (2020).
5. A. Alexandru, M. Ianculescu, and D. Coardos: Improved patient engagement in self-management of health, a key to sustainable preventative healthcare systems. In: vol. 247, Springer International Publishing (2018).
6. M. O. Jewell, E. Costanza, and J. Kittley-Davies Agents: Connecting the things to the internet: An evaluation of four configuration strategies for wi-fi devices with minimal user interfaces. In: *UbiComp 2015 - Proc. 2015 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput.*, pp. 767–778 (2015), doi: 10.1145/2750858.2807535.
7. L. Liu, Z. Han, L. Fang, and Z. Ma: Tell the device password: Smart device wi-fi connection based on audio waves. In: *Sensors (Switzerland)*, vol. 19, no. 3 (2019), doi: 10.3390/s19030618.
8. D. T. Nguyen and T. Kim: An SDN-Based Connectivity Control System for Wi-Fi Devices. In: *Wirel. Commun. Mob. Comput.*, vol. 2018 (2018), doi: 10.1155/2018/9359878.
9. ISO/IEC. ISO 9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs). Part 11. Guidance on Usability, ISO/IEC 9241-11:1998(E) (1998).
10. D. Svanæs, O. A. Alsos, and Y. Dahl.: Usability testing of mobile ICT for clinical settings: Methodological and practical challenges. In: *Int. J. Med. Inform.*, vol. 79, no. 4 (2010), doi: 10.1016/j.ijmedinf.2008.06.014.

# Extended Petri Net Processor for Embedded Systems

Luis Orlando Ventre, Orlando Micolini, and Emiliano Daniele

Depto. Computación Lab. de Arquitectura de Computadoras  
Facultad de Ciencias Exactas Físicas y Naturales - U.N.C  
Córdoba, Argentina

{luis.ventre,orlando.micolini}@unc.edu.ar,emiliano.daniele@mi.unc.edu.ar  
<https://fcefyn.unc.edu.ar/>

**Abstract.** The evolution of technology and electronic devices, the widespread use of IoT, and the compliance with specific regulatory requirements of the industry have made the process of designing embedded systems more complex and challenging. These systems are generally parallel, concurrent, reactive, and/or event-driven. In these systems, the data and events are heterogeneous and non-deterministic as they interact with the external environment. Extended Petri nets constitute an elective platform and system-independent modeling language, which makes it appropriate for modeling embedded systems. To take full advantage of the modeling efforts, it is desirable to use the built models to obtain part of the system implementation. This paper presents the design and implementation of an Extended Petri Processor and its modular architecture. This processor makes use of the extended state equation of Petri Nets, executing the model of the mentioned systems, intending to mitigate the time needed for development, and to reduce programming errors.

**Keywords:** Petri Processor, Petri Nets, Synchronization, IoT, Code Generation, FPGA.

## 1 Introduction

Nowadays, critical, reactive (RS) and event-driven embedded systems (EDA) [1] are in high demand, especially by Industry 4.0 [2]. The design of these systems must meet strict non-functional requirements since they are parallel, concurrent systems and interact with variables and events from the system itself and from the outside world, where the data and events are heterogeneous and non-deterministic [3]. This article presents the design and implementation of an extended Petri processor and its modular architecture, for the development of these types of systems. This processor aims to reduce development time and mitigate programming errors.

The design phases of an embedded system include the development of the model based on a set of requirements [4]. This model is the basis for other stages, including the application building stage [5].

## 2 Extended Petri Net Processor for Embedded Systems

Non-autonomous and extended PN is a general-purpose modeling language that supports the modeling of reactive, concurrent, and parallel systems regardless of the platform. In the design of RS and EDA, the transformation of the model into software implies a translation work that leads to interpretation and implementation errors. In order to mitigate these errors, this processor capable of running the model regardless of the platform has been developed.

According to our research documents, there is no history of automatic code generation or model execution performed by a processor. The innovation of this proposal includes the design and implementation of a modular heterogeneous architecture processor, which executes the extended state equation of PN [12], and which has the expression capacity of a Turing machine. This processor keeps all the properties verified in the model, since the model is essentially not interpreted and/or transcribed in code, but rather executed.

As a precedent to this work, and as part of this same research project, the development of a modular Petri processor (PP) that executes ordinary PN can be found in [6]. The main difference is the expression capacity, which is not that of a Turing machine since it cannot execute models with different types of arms or guards. Also, in [7], a study was carried out on more than 70 references that make use of PN for the solution of SR and EDA.

The following section sets out the objectives of this work; while in section 3 a brief theoretical framework is presented. Then, in section 4, the architecture of the Extended Petri Processor (PPX) is described as the proposed solution. In section 5, the results, the tests carried out and finally the conclusions are exposed.

## 2 Objectives

The system's logic model contains the necessary information for implementing the software logic. The coding stage involves successive iterative efforts to refine, interpret, and transcript the model into code, detecting and correcting errors in the process. This entails an overload of effort and time in the development stages.

The main objective of this work is the design and implementation of a processor that executes the logic model as code, thus unifying the modeling and coding stages. As a secondary objective, a modular design and interconnection with a traditional processor is carried out.

## 3 Methodology and Tools

There are several modeling tools, among which are: UML diagrams [8] and Petri Nets (PN) [9, 10]. UML diagrams provide the necessary characteristics, partially, but they essentially lack mechanisms for formal verification that strictly guarantee compliance with critical requirements, which are fundamental aspects to be implemented in the SR and EDA.

Since the extended, non-autonomous, PN [11] allows to model concurrency, local and global state, and parallelism, it is possible to verify them formally.

They are executable [7] and scalable when expressed with the extended state equation [12]; they have been considered to be the most convenient formalism and have been selected as a modeling tool and a processor execution language.

The development tools and implementation of the PPX are explained and described in subsection 4.5.

### 3.1 Petri Nets

A marked PN, denoted as PN, is a quadruple [5] defined by:

$$PN = (P, T, I, M_0) \quad (1)$$

Where:

- $P = p_1, p_2, \dots, p_n$  is a finite, non-empty set of places.
- $T = t_1, t_2, \dots, t_m$  is a finite, non-empty set of transitions.
- $I$  is the incidence matrix that relates places with transitions and vice versa.
- $M_0$  is the initial markup of the PN.

### 3.2 Synchronized or Non-Autonomous PN

This type of PN introduces events into the model and it is an extension of autonomous PN [11] [7]. Non-autonomous PNs model systems in which firings are synchronized with external discrete events. Events are associated with transitions, and the firing occurs when two conditions are met: the transition is enabled and the event associated with the transition had taken place.

External events correspond to changes in the state of the system's environment (including time) while internal events are changes in the state of the system itself. Synchronized PNs can then be defined as a triplet:

$$PN_{sync} = (PN, E, sync) \quad (2)$$

Where:

- $PN$  is a marked PN,
- $E$  is a set of external events and
- $sync$  is the function that relates the transitions  $T$  with  $E \cup \{e\}$ , where  $\{e\}$  is the null event, that is, those transitions that are automatically fired.

#### Perennial, Non-Perennial and Null Events

There are different types of events. A detailed description can be found in [7].

#### Extended Equation of State

In order to mathematically represent the existence of the new inhibitor, reader and reset arms, a matrix is required for each type of arc. These matrices are similar to the matrix  $I$ . When the arcs have a weight equal to one, the terms of the matrix are binary. A transition can be enabled if it meets the following conditions: if it has an inhibitor arm that does not have a token in the associated place; if it has a reading arm, and the place associated has one or more tokens; if

#### 4 Extended Petri Net Processor for Embedded Systems

it has a guard and the guard value is equal to true; if it has an event associated, and one or more events were queued and if it has a label with a time interval, and the counter is in the valid time range.

So the extended equation of state is now defined as:

$$M_{(j+1)} = M_j + I * (\sigma \text{ and } Ex) \# A \quad (3)$$

In this expression  $Ex$  is the extended enabled vector, represented by:

$$Ex = E \text{ and } B \text{ and } L \text{ and } G \text{ and } Z \quad (4)$$

Where  $E, B, L, G$  y  $Z$  are the enabled vectors of the different arcs. The details of the calculation of (3) and (4) are found in [12].

## 4 Architecture of the Solution

The hypothesis of this work is based on the fact that a model made with a non-autonomous PN is a set of instructions, equations and restrictions or rules to generate the I/O behavior of a system. That is, the model is described as state transitions and mechanisms to accept input trajectories and generate output trajectories depending on their state.

The defining, in terms of system specifications, has the advantage of a solid mathematical basis and unequivocally defined semantics. To specify a behavior, the model needs an agent. This is basically a computer system capable of executing the model. The same model, expressed in a formalism, can be executed by different agents, thus enabling portability and interoperability at a high level of abstraction.

In this project, the PPX is the agent in charge of executing the model by making use of (3), so that it can generate the desired behavior. Extended PNs allow to model systems of events or stimuli, states, logic, policy and actions, which means it can be decoupled and they manage the control and execution of the whole system.

### 4.1 Architecture of the PPX

The main blocks of the PPX implement (3) which are: matrix-program, calculation-state, queues, and policies.

### 4.2 PPX modules

The Fig.1 represents a synthesized version of the processor.

**Matrix-Program Modules:** responsible for defining the processor program with the matrices and vectors of the state equation. They are represented in Fig.1, identified with \*, and they are: the matrices I, H, R, Rst, A and Time comparison window, and the vector of automatic firings.

**Calculation-Status Modules:** responsible for calculating and maintaining the status of the PN. Its components are marked in Fig.1 with #, which are: the state vector, the new state vector, L, B, V, E, G, S, the timers array, the vector of possible firings, the Adder and the Calculation-reset module.

**Queues module:** responsible for storing the input and output events and communicating the results of the PN execution with the traditional processor, Microblaze (MCS)[13]. These components are marked in Fig.1 with @ (the firing and exit request queues).

**Policies Module:** responsible for selecting the transition with the highest priority from the vector of possible firings. Its components have been identified on the Fig.1 with &, they are the Firing Policy Matrix and the highest priority Firing Vector.

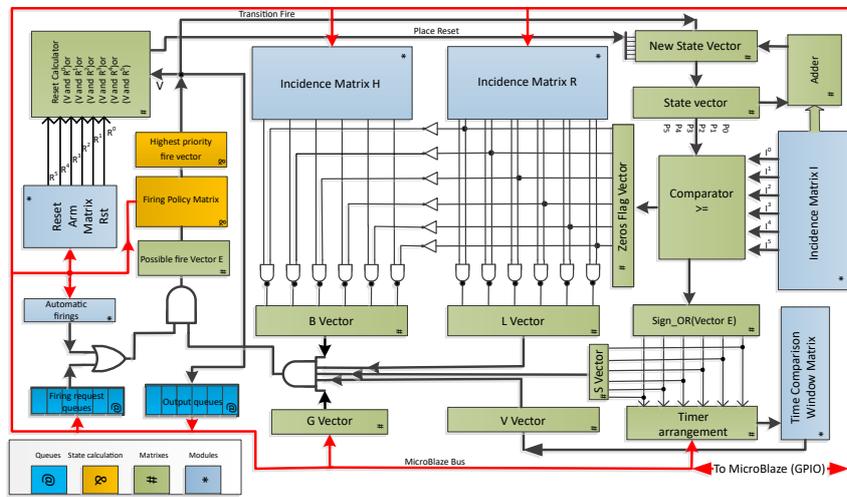


Fig. 1. PPX architecture.

### 4.3 Modules Description

The parts and functions of each component correspond to (3). They are consistent with the proposed modular structure and the main ones are:

**Incidence Matrix  $I$**  - array of integers. Its dimension is  $|T| \times |P|$ .

**Inhibitor Arms Matrix  $H$**  - binary matrix. Its dimension is  $|T| \times |P|$ .

**Reader Arms Matrix  $R$**  - binary matrix. Its dimension is  $|T| \times |P|$ .

**Arm Matrix Reset  $Rst$**  - binary matrix. Its dimension is  $|T| \times |P|$ .

**Time Comparison Window Matrix** - stores the alpha and beta values which correspond to the lower and upper time limits [5].

**State Vector** - vector of positive integers. Its dimension is  $|P|$ .

## 6 Extended Petri Net Processor for Embedded Systems

**Vector  $L$**  - binary vector. Its dimension is  $|T|$ . Inhibits the transition if the place is not marked.

**Vector  $B$**  - binary vector. Its dimension is  $|T|$ . Inhibit the transition if the place is marked.

**Vector  $V$**  - binary vector. Its dimension is  $|T|$ . Enable sensitive transition if your timer is in the window range (between alpha-beta).

**Guardian Vector  $G$**  - binary vector with the guard values. Its dimension is  $|T|$ .

**Sensitized Vector  $S$**  - binary vector, where each position corresponds to each column of the matrix of possible next states. Its dimension is  $|T|$ .

**Timers arrangement** - vector of integers. Its dimension is  $|T|$ . They are counters that are activated when the associated transition is enabled and are reset when it is disabled or fired.

**Firing Request Queues** - its interface exposes an input vector to the PPX, where each position of the vector corresponds to a transition.

**Exit Queues** - its interface exposes an output vector from the PPX, where each position of the vector corresponds to a transition.

**Firing Policy Matrix** - binary matrix of dimension  $|T| \times |T|$ . Its values indicate the relative priority between transitions.

**Highest Priority Firing Vector** - binary vector that represents the transition to be fired.

#### 4.4 Processor Algorithm

Single-server semantics have been adopted in this work, so only one transition is fired at a time. Two cycles are required for each firing to determine and report the new status.

**Cicle 1 - Calculations** - In this cycle, the necessary calculations are performed to determine which transition to fire.

The tagging vector is compared with each column of the incidence matrix, denoted in Fig.1 as  $I^i$ , to get the sign bit of each element of the array of possible next states. The results are binary columns where the  $i$ -th column contains the signs of the values of the next marking vector, in case the transition  $i$  is executed. This matrix contains only the signs of the possible next states; given a column, if any of the values is negative, it means that the next state will not be reachable by the PN (negative values in a tagging vector indicate that the transition is not enabled). To obtain this value, a logical disjunction is performed between all the elements of each column. The vector  $S$  is built using these values.

For each element of the state vector, a zero bit is determined. With these bits, the zeros flag vector is constructed. This indicates whether or not the place is marked and it is indicated in Fig.1 as Zeros Flag. The vectors  $B$  y  $L$  are calculated with the product of the Zeros Flag vector and the matrices  $H$  and  $R$ , respectively.

The Guards Vector  $G$  is updated from the MCS processor, since it represents all the conditions that are external to the PPX processor.

The vector  $V$  indicates whether a transition is enabled (based on the values of the vector  $S$ ), and if it is in the time window programmed in the Array of Timers.

The transitions that are possible to be fired are obtained from the logical conjunction between the vectors  $B, L, G, V$  and  $S$ . The logical disjunction between the Firing Request Queue vector and the Automatic Firings indicates the transitions that are requested to fire. Next, the logical conjunction between these last two vectors is carried out, which indicates the possible firings, noted as  $E$  in Fig.1. This vector contains a value of 1 in the positions of the transitions that are possible to fire. Since single-server semantics have been adopted, it is necessary to determine the highest priority transition to fire. To achieve this, the Firing Policies Matrix is used. This returns the highest priority Firing Vector, which contains a value equal to 1 in the transition to fire. In this cycle it is also calculated if any place should be zeroed, so the internal product is performed between the columns  $R^i$  of the matrix of reset arms ( $Rst$ ) and the highest priority Firing Vector. This calculation is carried out by the module named `Reset_calculator`. If the value is one, the place reset bus sets the place value to zero.

**Cycle 2 - Update** - This is the cycle that computes the firing of the transition that was selected in the previous cycle. Here, the firing takes effect and the value of the marking vector is updated. To achieve this, the selected transition works as a column selector.

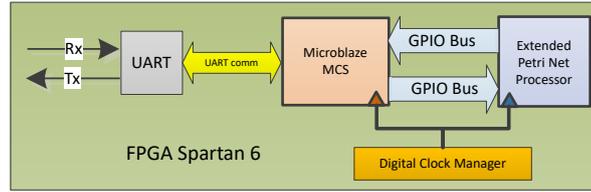
The adder, which is at the top of the Fig.1, performs the sum of the marking vector with the matrix column  $I$  selected by the firing vector. The result of that sum is stored as the new marking vector. In this cycle, if applicable, the space indicated on the reset bus is reset. Additionally, the queues are updated; increasing the output queue counter and decrementing the input queue counter.

#### 4.5 PPX, FPGA, and Microblaze MCS

The PPX solves the logic of the system so it operates with an associated processor [7][6] which executes the actions required. To achieve this, the PPX was interconnected with the MCS processor as shown in Fig.2. It was implemented in a Spartan 6 FPGA from Xilinx (Atlys) [14], in which an IP-core MCS was installed since it is included in the ISE tool [15] and it has a low impact on the resources required for its implementation. A communications module (UART) was also installed in order to carry out the testing and a clock management module (DCM). This configuration has been selected to establish comparisons with the work carried out in [6].

#### 4.6 Queues

The input and output queues of the PPX are configurable, each transition has an input queue and an associated output queue. A detailed description is found in [6].



**Fig. 2.** Interconnection between processors.

#### 4.7 Priorities and Conflicts between Transitions

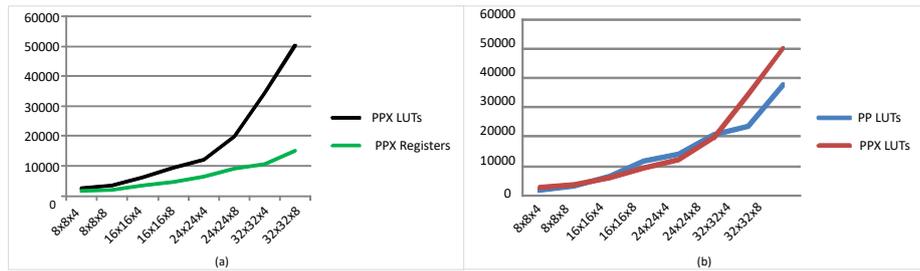
The PPX does not detect conflict states [11] between transitions, this is why it treats all enabled transitions as if they were in conflict (single server). This semantics, in conjunction with the priority policy module, also solves the problem of conflicts and makes it deterministic. The Firing Policies Matrix module determines the transition to fire, this module is configurable at runtime.

## 5 Results

With the PPX-MCS heterogeneous architecture, different application cases were executed to evaluate its performance. The successful executions of the cases were raised in [16] [17]. The comparison of resources has been carried out taking into account the results obtained in [6]. For the purposes of this comparison, the configuration, FPGA and development tools selected were the same.

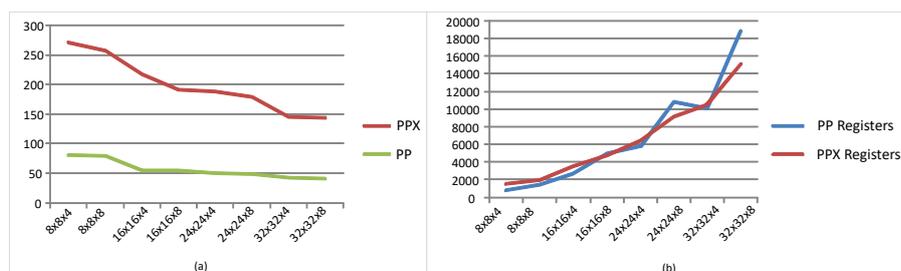
**FPGA Resource Consumption** - The processor was installed with different configurations of vector and matrix elements. Each configuration is expressed with a triplet of integers, which are:  $P \times T \times Pa$ , where  $P$  is the number of places,  $T$  the number of transitions, and  $Pa$  the length of the word that represents the weight of the arcs and the amount of tokens that a place supports. Multiple kernel syntheses were performed using different word-lengths (4-bit and 8-bit data). In Fig.3 a) the amount of resources that were used from the FPGA is displayed for synthesized configurations. In Fig.3 a) the exponential increase in the consumption of resources is observed, as the number of elements of the matrices and vectors increase. In Fig.3 b) and Fig.4 b) the consumption of resources of the PPX is compared to the PP, where the average increase of LUTs is 10% while the increase in register consumption is 18%. Overall, the impact of the MCS processor in resources is 12.72% for LUTs, and 21.09% for registers. These resources are the same for all the synthesized configurations of the PPX.

**Frequency Analysis** - The maximum theoretical frequencies for the different processor instances are shown in Fig.4 a). Since optimizations have been made to the modules and interconnects, the PPX has achieved a substantial improvement in frequency over the PP. It is observed that for a configuration of  $8 \times 8 \times 8$ , the maximum frequency is 257MHz while for  $8 \times 8 \times 4$  it is 271MHz; for a  $16 \times 16 \times 8$  configuration the frequency is 190 MHz and for  $16 \times 16 \times 4$  it is 216 MHz. It should be noted that the decrease in frequency, with respect to the length



**Fig. 3.** a) Consumption of LUTs and PPX Registers. b) Comparison of the use of LUTs between PPX and PP.

of the word, is 18% in average. On the other hand, if the length of the word is maintained and the number of places and transitions is increased, that is, the size of the PN, the difference in frequency is significantly larger, on average 30%.



**Fig. 4.** a) Comparison of theoretical maximum frequencies between PPX and PP. b) Comparison of Records between the PPX and the PP.

## 6 Conclusion

In this project, a processor (PPX) was designed and implemented for executing the extended state equation, with a modular architecture. This processor extends the semantic capacity of the PP developed in [6]. The inclusions of different types of arcs, temporal semantics, and guards in the PPX give the PPX the expression capacity of a Turing machine without significantly increasing the necessary resources. Queue scheduling and support for different types of events have been maintained, as well as the communications module. The results show that the PPX is suitable, for the FPGA selected, for embedded systems that require up to 32 composite logical conditions, 32 logical variables, and 32 events that must be evaluated simultaneously. From the data obtained in the frequency

analysis, the substantial improvement in the maximum theoretical frequency with respect to the PP stands out. The serial communication module has facilitated the configuration, debugging, and programming tests from a console. The modular implementation of the PPX implies a breakthrough for maintenance, scalability and future autoconfiguration.

## References

1. N. Halbwachs, *Synchronous programming of reactive systems*. Springer Science & Business Media, 2013, vol. 215.
2. K. Schwab, *The fourth industrial revolution*. Currency, 2017.
3. A. Munir, A. Gordon-Ross, and S. Ranka, *Modeling and optimization of parallel and distributed embedded systems*. John Wiley & Sons, 2015.
4. B. P. Zeigler, A. Muzy, and E. Kofman, *Theory of Modeling and Simulation: Discrete Event & Iterative System Computational Foundations*. Academic press, 2018.
5. M. Diaz, *Petri nets: fundamental models, verification and applications*. John Wiley & Sons, 2013.
6. O. Micolini, E. N. Daniele, and L. O. Ventre, "Modular petri net processor for embedded systems," in *Argentine Congress of Computer Science*. Springer, 2017, pp. 199–208.
7. O. Micolini, "Arquitectura asimétrica multicore con procesador de petri," Ph.D. dissertation, Facultad de Informática, UNLP 2015.
8. B. Selic and S. Gérard, *Modeling and analysis of real-time and embedded systems with UML and MARTE: Developing cyber-physical systems*. Elsevier, 2013.
9. M. Zhou and N. Wu, *System modeling and control with resource-oriented Petri nets*. Crc Press, 2018.
10. S. Siewert, *Real-time embedded components and systems*. Cengage Learning, 2016.
11. R. David and H. Alla, *Discrete, continuous, and hybrid Petri nets*. Springer, 2005, vol. 1.
12. O. Micolini, L. O. Ventre, and M. I. Schild, "Generalized state equation for non-autonomous petri nets with different types of arcs," 2016.
13. P. D. Group and ATLAS, "Xilinx, microblaze processor reference guide."
14. C. Digilent. Atlys spartan-6 fpga trainer board. [Online] Available: <https://store.digilentinc.com/>
15. C. XILINX. Ise webpack design software. [Online]. Available: <https://www.xilinx.com/products/design-tools/ise-design-suite/ise-webkit.html>
16. O. Micolini, L. O. Ventre, and M. Ludemann, "Methodology for design and development of embedded and reactive systems based on petri nets," in *2018 IEEE Biennial Congress of Argentina (ARGENCON)*. IEEE, 2018, pp. 1–7.
17. O. Micolini, L. O. Ventre, M. Ludemann, J. I. R. Viano, and C. C. Bien, "Case study of reactive and embedded system design modeled with petri nets," in *2018 IEEE International Conference on Automation(ICA-ACCA)*. IEEE, 2018, pp. 1–7.

## **Análisis de rendimiento y consumo para sistema embebido con requisitos de tiempo explícitos**

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

Universidad Nacional de La Matanza,  
Departamento de Ingeniería e Investigaciones Tecnológicas,  
Florencio Varela 1903 - San Justo, Argentina

{ecarnuccio, gdeluca, wvaliente, mvolker}@unlam.edu.ar  
{raul.villcasd, mati.adagio}@gmail.com  
www.unlam.edu.ar

**Resumen.** El propósito de este trabajo consiste en el estudio y análisis del rendimiento de un sistema embebido que emplea el microcontrolador STM32F103C8T6, con el objetivo de reducir el consumo de energía originado desde una batería. Para ello se basó en patrones de diseño de máquinas de estados, que permiten mejorar los tiempos de ejecución y respuesta del sistema. Esto también se logra a través de la realización de distintas ejecuciones variando la velocidad del microcontrolador, al modificar la frecuencia del reloj. El desarrollo ejecuta módulos que poseen características de tiempo real, debido a que fueron diseñados para que un sistema de monitoreo de personas que detecta caídas, realiza geolocalización y ofrece control de pánico, entre otras funcionalidades más. Por ese motivo debe cumplir con un tiempo límite de ejecución estricto. Para garantizar estos requisitos, se aplicaron distintas técnicas de medición del rendimiento utilizando pulsos de reloj y señales de temporizador.

**Palabras Clave.** Análisis de Rendimiento, Consumo de Energía, STM32, Tiempo Real

### **1 Introducción**

Los sistemas embebidos (SE) actuales requieren cada vez más conexiones, tanto para sistemas externos, como también sensores y actuadores. Además ejecutan algoritmos cada vez más complejos, cuya completitud se ve aumentada si se le suma como requerimiento no funcional, la premisa del bajo consumo energético. Esto se debe a que por lo general su fuente de energía puede ser una batería. Por ese motivo se planteó aplicar patrones de diseño de máquina de estado, que proporciona un enfoque práctico para atender a distintos estímulos externos con un único hilo de ejecución [11]. De esta forma, además de tener impacto sobre el rendimiento y los tiempos de respuesta [5], los patrones de diseño ayudan en la mantenibilidad del código y seguridad del dispositivo. En este sentido la presente investigación se basa en algoritmos de tiempo

real, donde se tiene requisitos estrictos sobre los tiempos de respuesta. Por ejemplo, para que funcione correctamente un algoritmo que se encarga de detectar caídas, mínimamente debe leer mediciones del acelerómetro cada 24 milisegundos. De lo contrario, no se conseguirá detectar los picos de aceleración característicos que generan las caídas, el algoritmo de detección empleado fue analizado en profundidad en [7]. Además se necesitaba que la batería tenga una duración de su carga activa mayor a 24hs, para poder así asegurar su disponibilidad y comodidad de uso del dispositivo por parte del usuario monitoreado. Los requisitos no funcionales antes comentados, se completaron como parte de un sistema de tiempo real. En el cuál, la tarea crítica debe culminar correctamente antes de cumplirse su tiempo límite de respuesta. Esto es como consecuencia de que el ciclo completo de funcionamiento incluye el tiempo desde que solicita los datos, su normalización y finalmente su procesamiento, junto con las actividades necesarias a realizar sobre los otros sensores y actuadores.

En el análisis se optó por una arquitectura que no utiliza Sistema Operativo, de modo que el programa se ejecuta siempre sobre un único hilo de ejecución. Cabe aclarar que se podría haber utilizado bibliotecas que dan funcionalidades de múltiples tareas concurrentes, como *FreeRTOS* [12]. Sin embargo se determinó conveniente emplear patrones de diseño de máquina de estado, dado que la opción de utilizar bibliotecas externas aumenta el tamaño del binario ejecutable, debido a que agregan muchas líneas de código adicionales. De esta forma terminan consumiendo los recursos acotados del espacio disponible de memoria, dado que el embebido está limitado a solo 64 Kbyte para poder ejecutar el programa.

Por consiguiente en este documento, a los fines de garantizar el cumplimiento crítico de los requisitos de tiempo de ejecución mencionados, se detalla el análisis de rendimiento del sistema. De forma tal, que permita ajustar las configuraciones de la arquitectura para lograr las reducciones del consumo energético y validar los tiempos de ejecución en los algoritmos desarrollados.

## 2 Desarrollo

### 2.1 Máquina de estados

Los algoritmos utilizados basan su funcionamiento en un conjunto de máquinas de estados, que se encuentran interconectadas entre sí. Se planifico usar esta metodología, ya que al ejecutar sobre un único procesador, es indispensable que el sistema no caiga en esperas activas sobre los distintos componentes de software, controladores y algoritmos. Por eso se planteó como objetivo, que el sistema pueda leer correctamente los valores obtenidos por el acelerómetro, luego ejecute satisfactoriamente el algoritmo de caídas, lea las coordenadas de ubicación del GPS<sup>1</sup> y envíe notificaciones vía SMS<sup>2</sup> y por mensajes REST<sup>3</sup> al servidor sin retrasos. Para asegurar que se pueda satisfacer dicho comportamiento, se ha realizado un estudio exhaustivo del rendimien-

<sup>1</sup> GPS: *Global Positioning System*, Sistema de posicionamiento global.

<sup>2</sup> SMS: *Short Message Service*, Servicio de mensaje corto de la telefonía móvil.

<sup>3</sup> REST: *REpresentational State Transfer*, Interfaz de servicios de Internet.

to global del sistema. A continuación se expone gráficamente el diagrama del sistema (Fig. 1), en donde se detalla la máquina de estados central, llamado *Controlador\_Principal*.

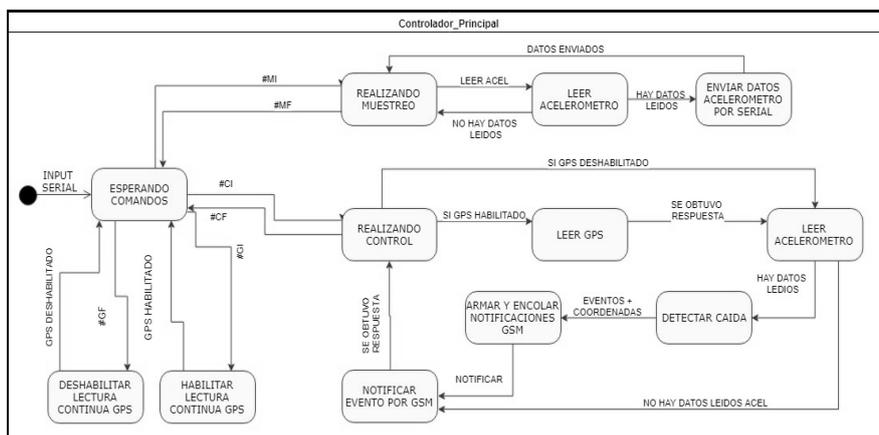


Fig. 1. Máquinas de estados del Controlador Principal.

El funcionamiento del SE es controlado por comandos a través de *bluetooth*<sup>4</sup>, y dependiendo de la instrucción recibida ejecutará determinadas operaciones. La función principal, comienza cuando se recibe el comando “#CI”, la cual iniciará la secuencia de ejecución de estados encargados de realizar el control principal del dispositivo embebido. En este punto es importante mencionar, que al ejecutar determinados estados de la (Fig. 1), se produce la invocación a las funcionalidades que internamente contienen otras máquinas de estados diferentes a la figura anterior. En consecuencia, al iniciar el ciclo *Realizando Control*, que se encarga de monitorear a la persona que lleva el dispositivo, primeramente se comprueba si se encuentra habilitada la lectura continua del GPS. Si esto es verdadero, se ejecuta el estado *LeerGPS*, que es la responsable de obtener las coordenadas de ubicación del SE. Luego pasa al estado *LeerAcelerómetro*, responsable de leer los valores del acelerómetro a través de su propia máquina de estados, que se encarga de estabilizar y normalizar las mediciones. Posteriormente se continúa con el estado *DetectarCaídas*. En él, se ejecuta el algoritmo responsable de detectar las caídas, que se encuentra analizado en detalle en [4]. Finalmente, en caso de detectarse una caída, se cambia al estado *NotificarEventoPorGSM*, con la finalidad de enviar mensajes al servidor. El cuál distribuye la notificación a los usuarios responsables, informándoles de los eventos que hayan ocurrido en el sistema. Estas notificaciones se realizan a través de mensajes SMS y protocolo REST.

<sup>4</sup> Bluetooth: Protocolo inalámbrico de comunicación.

Es importante mencionar que en cada una de las etapas mencionadas, poseen su propia máquina de estados para lograr realizar tareas cortas y asincrónicas. De este modo no se generan esperas activas que bloqueen el flujo de ejecución del programa.

## 2.2 Configurar la frecuencia del reloj del microcontrolador

El SE se construyó utilizando el microcontrolador STM32F103C8T6. Una de las ventajas que tiene, es poder configurar la frecuencia de oscilación de su reloj principal (HCLK), ver (Fig. 2). De esta manera se puede configurar el sistema para que este componente funcione a distintas a frecuencias, desde 72 MHz como máximo a 8 MHz de mínimo [1]. En este contexto se realizaron distintas pruebas modificando HCLK, para luego medir el desempeño, donde se analizaron las variaciones de tiempo en los algoritmos detallados anteriormente. De esta forma se llevaron a cabo mediciones de tiempo en distintos puntos específicos de esos algoritmos, junto con las evaluaciones de la variación del consumo energético. Por ende se realizaron configuraciones para que HCLK funcione con frecuencias de 8MHz, 16MHz y 72 MHz. Como caso particular para la frecuencia de 8 MHz, se seleccionó al oscilador interno HSI como origen del pulso. Mientras que en los casos de 16 MHz y 72 MHz, se utilizó el oscilador externo HSE.

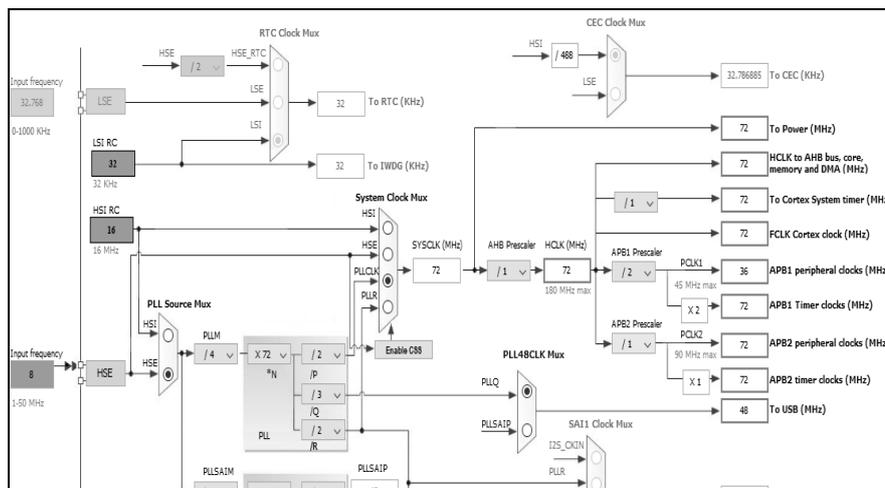


Fig. 2. Ventana de configuración frecuencia de microcontrolador SMT32.

En las pruebas se observó que con distintos valores del HCLK, se afectan indirectamente a la comunicación con los sensores y actuadores externos que se encuentren conectados. Esto se debe a que los periféricos trabajan a distinta velocidad que el microcontrolador, por eso existen dos buses independiente de periféricos (APB1 y APB2)<sup>5</sup>, que realizan la comunicación con dispositivos que posean frecuencias diferentes a las del HCLK. No obstante estos son dependientes del mismo, debido a que

<sup>5</sup> ABP1 y ABP2: *Advanced Peripherals Bus 1 and 2: Bus dedicado a periféricos 1 y 2.*

se configuran utilizando divisores de frecuencia con respecto al valor del HCLK. En este sentido, al realizar las pruebas a distintas frecuencias del HCLK, el funcionamiento de los periféricos conectados al sistema se vio afectado. Por ese motivo al establecer el HCLK a 72 MHz, para un correcto funcionamiento de bus I2C<sup>6</sup> (que utiliza el acelerómetro), se configuró a una frecuencia de 400 KHZ. Mientras que al utilizar el HCLK a 8 MHz y 16 MHz, la frecuencia del bus I2C debió ser disminuida a los 100 KHZ.

### 2.3 Metodologías de medición de tiempo en un sistema embebido

En el análisis de requisitos no funcionales, el tiempo de respuesta para un SE es una tarea compleja. Esto es como consecuencia de que para obtener información sobre el rendimiento de la ejecución, se deben realizar las lecturas de sensores dentro del propio SE, almacenarlas y posteriormente analizarlas en una computadora fuera del SE. Por ende, para realizar dichas tareas, se pueden utilizar herramientas que recolectan la información en forma externa o también pueden realizarse manualmente mediante lecturas internas.

#### 2.3.1 Análisis comparativo de herramientas que miden desempeño

En esta investigación se analizaron las herramientas que están disponibles para la familia de microcontroladores STM32 [2]. Por ende, para lograr realizar mediciones de desempeño del sistema, básicamente se pueden llevar a cabo empleando dos aplicaciones. Ambas poseen licenciamiento comercial, debido a que otorgan la posibilidad de uso evaluativo por un breve periodo de tiempo, pero con funcionalidades limitadas. Estas herramientas se llaman: *IAR™ EWARM* y *Keil® MDK-ARM μVISION*. En las especificaciones de la herramienta *μVISION* de *Keil®*, se observó que al ejecutar el código con distintas velocidades de reloj, se pueden obtener lecturas estadísticas incorrectas sobre el rendimiento de la ejecución [3]. También existe otro tipo de herramienta, con similar tipo de licenciamiento. Este es el caso de *Tracealyzer* de *Percepio* [4], que se integra al programa que ejecuta en el SE como una biblioteca. Sin embargo, esta se encuentra especialmente preparada para Sistemas Operativos de Tiempo Real (RTOS). En este tipo de arquitectura, se puede ejecutar el S.O *FreeRTOS*, ya que es totalmente compatible con dicha herramienta. No obstante se descartó su utilización, dado que consume recursos y acota el limitado espacio disponible de memoria. Por estos motivos se decidió realizar las mediciones manualmente en forma interna.

#### 2.3.2 Lectura de desempeño interna

Existen dos técnicas para tomar mediciones de tiempo, desde el código fuente del propio programa a ejecutar. La primera de ellas utiliza el temporizador de cuenta descendiente *Systick*. Este funciona mediante interrupciones, que se activan cuando el

<sup>6</sup> I2C: *Inter Integrated Circuits*, Circuito Inter-Integrado que funciona como bus serie topología maestro-esclavo.

valor de su registro contador interno llega al valor cero. De esta manera se puede utilizar *Systick* para que genere una interrupción cada un determinado tiempo, por ejemplo cada un milisegundo. Esto permite que se pueda utilizar el temporizador para medir el tiempo de ejecución, debido a que se registra la interrupción desde que se inicia el sistema, de forma tal que sea utilizado como un temporizador dedicado. Así mismo una de las características del modelo empleado, es que posee 24 bits de representación para reinicio automático. Además puede enmascarar la interrupción cuando el contador descendiente llega a cero y la fuente de reloj es programable [13]. Por otra parte, la segunda forma de medición de tiempo, es a través de componentes internos. Para ello se debe utilizar el registro CYCCNT<sup>7</sup>, que es un contador de ciclos del reloj del CPU, que forma parte de la unidad de depuración DWT<sup>8</sup>, proporcionada por la arquitectura STM32. Esta es una técnica que es utilizada en *profiling*, desarrollada por *Serj Bashlayev* [8]. La cual permite generar marcas de tiempo a través del valor inicial del CYCCNT, que contenga en un momento dado. De forma tal, que después se contrarresta el valor que posee ese registro en distintos instantes de tiempo. En consecuencia para realizar ese cálculo se debe tener en cuenta la frecuencia del reloj principal, en base a la cantidad de pulsos por microsegundos que este genera.

#### 2.4 Mediciones de tiempos realizadas durante la ejecución del sistema

Debido a las limitaciones mencionadas sobre las herramientas de desempeño externas, en esta investigación se han utilizado los componentes internos del microcontrolador para medir el desempeño del sistema. Para ello se han utilizado las dos técnicas antes mencionadas, la desarrollada por *Serj Bashlayev* y en determinadas ocasiones se ha empleado el método de *Systick*. En ese sentido se han realizado mediciones en diferentes puntos críticos del SE, con el objetivo de analizar su rendimiento y comprobar que se puedan cumplir los límites de tiempo. A continuación se enumeran los puntos claves en donde se han realizado las mediciones de tiempo pertinentes, para las distintas velocidades de reloj de sistema.

- Lectura del tiempo demandante desde que se solicitan los datos al acelerómetro, hasta el momento que son leídos, dentro de *LeerAcelerómetro*.
- Interiormente en la máquina de estado de *LeerGPS*, se midió el tiempo que le demanda al sistema para empezar a leer una línea de datos del GPS, a través de un puerto serial, hasta que se determina si esa línea contiene un formato NMEA<sup>9</sup> válido. Como así también, cuanto se tarda para la detección del NMEA completo.
- Dentro del *NotificarEventoPorGSM*, se obtuvo el tiempo que necesita el algoritmo para ejecutar la secuencia de comandos AT. La cual es necesaria para poder enviar un SMS o ya sea un mensaje del tipo REST.

<sup>7</sup> CYCCNT: *CYcle Clock CouNter*, Contador de ciclos de reloj de CPU.

<sup>8</sup> DWT: *Debug Watchpoints data Tracing*, Unidad de depuración de datos.

<sup>9</sup> NMEA: *National Marine Electronics Association*, es una especificación instrumentos marítimos y receptores GPS.

- La cuantificación de tiempo que requiere la máquina de estados del Controlador Principal en ejecutar un ciclo completo, en donde se obtenga la lectura de acelerómetro, las coordenadas de GPS y se envíe un mensaje por GSM.

Estos puntos fueron seleccionados debido a que en ellos se pueden producir demoras en su ejecución.

### 3 Análisis de resultados

En la tabla 1 se detallan los resultados de los tiempos obtenidos, luego de la realización de las mediciones en los puntos críticos previamente mencionados. En ella se puede observar la cantidad de mediciones que fueron efectuadas por cada uno de los eventos descritos, como así también los valores máximos y mínimos de tiempo registrados en esas mediciones. Los valores mostrados dependen de la configuración del reloj principal seleccionada en ese momento.

Eventos	CLOCK 72 MHZ			CLOCK 16 MHZ			CLOCK 8 MHZ		
	Número Mediciones	Valor max [us]	Valor min[us]	Número Mediciones	Valor max [us]	Valor min[us]	Número Mediciones	Valor max [us]	Valor min[us]
Lectura Acelerómetro	3881	1802	1108	2420	4402	2857	1004	6840	3995
Lectura GPS Línea NMEA	65	63996	54995	69	80009	75975	51	4675005	730004
Lectura GPS Línea Serial	1908	254	74	2735	984	203	192	4489	406
Lectura GPS carácter NMEA	482	103	1	650	312	7	1060	871	14
Envío SMS	5	5404263	4495155	3	4810253	4404262	4	4549081	4448007
Envío REST	8	15504805	8434725	4	8962651	8052403	6	16448783	8424796
Bucle principal	15791	782	1	13023	649	3	15717	272	6

**Tabla 1.** Tiempos de ejecución según frecuencia 72/16/8MHz del reloj principal.

En la tabla 1 se puede observar que al aumentar la frecuencia del reloj, para los distintos puntos críticos de los diferentes algoritmos, los tiempos de ejecución van disminuyendo. Esto ocurre en la mayoría de los casos a excepción de las pruebas en donde el GSM<sup>10</sup> realiza el envío de un SMS y mensajes tipo REST. Dado que los tiempos requeridos para ejecutar las secuencias de comandos AT<sup>11</sup> para dichas operaciones se mantienen constantes, por más que se modifique la velocidad del reloj principal. La razón de esta linealidad en los tiempos de ejecución de las operaciones GSM, se debe a que siempre se realizó la transferencia de datos por los puertos seriales con la misma velocidad en baudios, por lo que estos tiempos no se ven afectados

<sup>10</sup> GSM: *Groupe Spécial Mobile*, Sistema estándar digital de telefonía móvil.

<sup>11</sup> AT: *for getting the modem's ATtention*, obtener atención del modem.

significativamente. Otra excepción que se observó luego de las mediciones, es que el valor máximo de tiempo obtenido al ejecutar el bucle principal, aumenta a mayor frecuencia del reloj. Sin embargo, en los valores mínimos de tiempo registrados se produce el efecto contrario, por lo que en ese caso si se cumple la regla mencionada inicialmente. Luego del análisis se descubrió que la causa de este suceso se debe a que durante el bucle principal se ejecutan instrucciones que demoran instantes con tiempos máximos muy breves, generando así picos máximos de duración en microsegundos. No obstante el tiempo promedio de la duración del bucle principal se mantiene cercano a los valores mínimos registrados, en cada una de las configuraciones mostradas.

En la Tabla 2 se muestran las especificaciones del fabricante del STM32F103C8T6 [1], con respecto a la variación del consumo del microcontrolador para distintas frecuencia del reloj externo ( $f_{HCLK}$ ). Como aclaración estas mediciones corresponden a la temperatura 25 °C y un voltaje de entrada 3,3 V. Al contrarrestar esto valores con los consumos de las placas de desarrollo Arduino Uno [9] y Arduino NANO [10], el consumo resultante del STM32 es menor. Esto se debe a que las dos placas Arduino mencionadas trabajan a una frecuencia definida de 16 MHz, en donde su consumo energético es de 46 mA y 15 mA respectivamente. Ambos superiores a los 9,3 mA que consume esta arquitectura configurada para la misma frecuencia.

Parámetro	$f_{HCLK}$ [MHz]	Consumo [mA]
Modo ejecución con los dispositivos conectados.	72	36
	48	24,2
	36	19
	24	12,9
	16	9,3
	8	5,5

**Tabla 2.** Consumo máximos según frecuencia del STM32F103C8

Utilizando la herramienta *STM32CubeMx*, se puede predecir cuanto será la duración de la batería. Esto se describe en la Tabla 3, en donde la primera fila demuestra cómo se puede prolongar la vida útil de la batería para bajas frecuencias del microcontrolador. Mientras que en la fila dos, se comparan cuando se agregan al cálculo del consumo los sensores y los mecanismos de comunicación utilizados. Para ello se configuraron los puertos *USART1*<sup>12</sup> para utilizar bluetooth, *USART2* para GPS y *USART3* para GSM, Además se utilizó el puerto *I2C* para la conexión con el sensor acelerómetro *MPU6050* [6]. Esta configuración produce un consumo adicional de 118 mA, el cual se ve reflejado en los consumos promedios que se pueden observar en la tabla. En la

<sup>12</sup> USART: *Universal Asynchronous Receiver-Transmitter*, Transmisor-Receptor Asíncrono Universal.

predicción se tomó en cuenta el uso constante de los dispositivos y se descartaron los picos de consumo, como así también los modos en ahorro de energía.

Eventos	CLOCK 72 MHZ		CLOCK 16 MHZ		CLOCK 8 MHZ	
	Duración Batería (hrs)	Consumo Promedio (mA)	Duración Batería (hrs)	Consumo Promedio (mA)	Duración Batería (hrs)	Consumo Promedio (mA)
Solo micro-controlador	68	31,94	258	8,5	427	5,15
Micro-controlador+ dispositivos I/O	14	149,94	17	127,5	18	124,15

**Tabla 3.** Duración estimada de la batería en horas.

## 4 Conclusiones

Se han logrado establecer varios puntos de comparación con el uso de técnicas de medición de tiempos y rendimiento del sistema embebido. Con respecto a las mediciones que se han llevado a cabo, se evaluaron herramientas y técnicas, a partir de ello se han elegido emplear dos métodos. El primero de ellos por contadores de ciclos, que permite medir intervalos de tiempo en determinadas secciones específicas del sistema. Mientras que el segundo método empleado, *Systick*, se ha utilizado para ejecutar eventos de mediciones luego de un determinado tiempo. Por un lado se constató que al hacer uso de máquinas de estados en toda la solución, permite atender simultáneamente a varios dispositivos, dejando ciclos de reloj del microcontrolador para atender con prioridad a los sensores más críticos. También se pudo constatar que ajustando la velocidad de la arquitectura utilizada se permite minimizar el consumo energético. Además se verifica que si bien varía el tiempo de respuesta de las funciones críticas, se logra conservar el límite de tiempo requerido.

En base a las mediciones realizadas, se considera que la solución adecuada para implementar en el sistema consiste en configurar la velocidad del reloj principal para que trabaje a 16 MHZ. Esta elección de configuración permite un adecuado ahorro energético, inclusive menor que la arquitectura Arduino, con tiempos de respuestas acorde a los que son requeridos en la solución. En contraste, para frecuencias mayores a la elegida, se producirá mayor consumo energético. Mientras que por debajo de esta velocidad podría ocurrir que los algoritmos necesarios, en conjunto, no puedan cumplir con sus requerimientos de funcionamiento.

En futuros trabajos se plantea la posibilidad de armar un entorno que permita estos tipos de pruebas generales con diferentes sistemas embebidos, ya sean reales o virtuales. De esta formase podrá elegir el hardware más conveniente a utilizar de acuerdo a las necesidades de un proyecto que se desee llevar a cabo.

## 5 Referencias

1. STMicroelectronics. STM32F103x8 DataSheets.l. : STMicroelectronics, 2015. pág. 48, STM32F103x8, STM32F103xB - Medium-density performance line ARM®-based 32-bit MCU with 64.
2. STMicroelectronics. STM32 microcontroller debug toolbox. 2017.
3. Arm Limited.  $\mu$ Vision User's Guide - Execution Profiler. [En línea] 2019. [Citado el: 24 de 07 de 2019.] [http://www.keil.com/support/man/docs/uv4/uv4\\_db\\_dbg\\_execprofiler.htm](http://www.keil.com/support/man/docs/uv4/uv4_db_dbg_execprofiler.htm).
4. Percepio. Percepio Tracealyzer. [En línea] <https://percepio.com>.
5. Douglass, Bruce. Software Engineering for Embedded Systems. Software Design Architecture and Patterns for Embedded Systems. s.l. : Elsevier Inc, 2013, 4.
6. InvenSense, Inc. MPU-6000 and MPU 6050 Product Specification Revision 3.4. . Sunnyvale : Inc.InvenSense, 2013.
7. De Luca, Graciela, y otros. Desarrollo de un prototipo detector de caídas utilizando la placa Intel Galileo Generación I y el sensor MPU6050. Anales del XXIII Congreso Argentino de Ciencias de la Computación. La Plata : s.n., Octubre 2017, págs. 954-963.
8. Bashlayev, Serj. <https://github.com/Serj-Bashlayev>. [En línea] 4 de Octubre de 2018.
9. Arduino. Arduino UNO - DataSheet. [En línea] 1998. <https://www.arduino.cc/en/uploads/Tutorial/595datasheet.pdf>.
10. Arduino. DataSheet Arduino NANO. [En línea] 2008. <https://www.arduino.cc/en/uploads/Main/ArduinoNanoManual23.pdf>.
11. Valiente, Waldo, y otros. Uso colaborativo del procesador en sistemas embebidos para múltiples interfaces. San Juan : WICC2019, 2019. 978-987-3619-27-4.
12. Amazon Web Services, Inc. FreeRTOS. [En línea] 2019. <https://www.freertos.org/>.

# Arquitectura Cliente-Servidor de Alto Rendimiento para servicio RTK

José H. Moyano<sup>1,2</sup>, Karina M. Cenci<sup>1,2</sup>, and Jorge R. Ardenghi<sup>1,2</sup>

<sup>1</sup> Laboratorio de Investigación en Sistemas Distribuidos

<sup>2</sup> Laboratorio de I+D en Ing. de Software y Sistemas de Información (UNS-CIC

Provincia de Buenos Aires)

Departamento de Ciencias e Ingeniería de la Computación

Universidad Nacional del Sur

Bahía Blanca, Argentina

{jose.moyano,kmc,jra}@cs.uns.edu.ar

**Resumen** La exactitud de los *Global Navigation Satellite Systems* (GNSS) varía entre 2 y 10 metros. Es posible mejorar la exactitud a nivel de centímetros en tiempo real utilizando *Real Time Kinematic* (RTK). Con esta técnica, un receptor con posición conocida (*base station*), calcula los errores y transmite información de corrección al *rover*, para que se localice con precisión de centímetros. RTKLIB es una biblioteca de software *open source* que permite implementar aplicaciones BS o *rover*. Este trabajo propone la implementación de una arquitectura cliente-servidor de bajo costo, que provea información de corrección en tiempo real a dispositivos *rovers*, utilizando RTKLIB y hardware comercial estándar.

## 1. Introducción

*Global Navigation Satellite Systems* (GNSS) es el conjunto de sistemas de georreferenciamiento satelital, donde *space vehicles* (SV) en órbita transmiten información a través de radiofrecuencia a receptores en tierra. Con esta información, el receptor calcula la distancia geométrica a cada SV seguido, y utilizando trilateración, puede localizarse en la superficie terrestre [9].

El error en el posicionamiento es provocado por incertidumbre en las órbitas y *clocks* de los SV; retrasos en la señal provocados por la atmósfera [7]; y ruido y *multipath* en el receptor [5]. Por este motivo, se logra una exactitud horizontal de entre 6 y 10 metros [17] para receptores de simple constelación y simple frecuencia, y de entre 2 y 6 metros para receptores con soporte para múltiples constelaciones o frecuencias. Estas precisiones resultan insuficientes para tareas topográficas y geodésicas, por lo que existen distintos aumentos [9] que mejoran el desempeño.

*Real Time Kinematic* (RTK) es una técnica que utiliza diferencias simples y dobles sobre las medidas de fase de las señales satelitales capturadas por dos receptores, resuelve la ambigüedad de ciclos de la portadora utilizando algoritmos como LAMBDA, y logra posiciones con centímetros [15, 12, 10] de exactitud en tiempo real.

RTK requiere la instalación de un segundo dispositivo receptor GNSS, de posición conocida, que a partir de la información satelital obtiene parámetros de error en el posicionamiento, y genera información de corrección [6]. A este dispositivo se lo llama *estación base* (BS). Esta información de corrección es transmitida a un *rover*, que es el nombre con el que se designa al objeto de posición desconocida. El *rover* también cuenta con un receptor de GNSS, pero necesita la información de corrección para ajustar su localización y lograr la exactitud buscada.

La información provista por una estación base tiene un rango de utilidad acotado por las condiciones atmosféricas. En la práctica, bajo condiciones ionosféricas promedio, estos errores pueden desprejarse en distancias entre receptores de hasta 10km [14]. Cuanto más lejos se encuentre el *rover* de la estación base, menor precisión tendrán las correcciones. Es por ello que es deseable que la estación base no se encuentre muy alejada del *rover*, y también lo que motiva instalar varias BS si se desea proveer el servicio en un área amplia de terreno.

Las estaciones base fijas de operación ininterrumpida, que se instalan para proveer servicio de corrección RTK de forma permanente, se denominan *Continuous Operation Reference Stations* (CORS). *Network RTK* (NRTK) es la técnica de utilizar una red de CORS, para proveer a *rovers* dentro del área de cobertura de las CORS con información de corrección RTK. Los *rovers* utilizan una conexión a internet, generalmente GSM, para solicitar a un servidor la información de corrección de la CORS más cercana [2].

El precio de adquirir BS particulares, los montos elevados de servicios de CORS, o el costo y disponibilidad acotada de las comunicaciones GSM de alta velocidad para conectar con NRTK, resultan un impedimento para el uso masivo de las tecnologías de posicionamiento de alta precisión, y motivan buscar alternativas de bajo costo.

RTKLIB es una biblioteca de software escrita en lenguaje C con licencia BSD2-clause [8], que ofrece la generación de información de corrección, o el uso de información de corrección para obtener soluciones de posición con centímetros de precisión. Esto permite implementar una BS en el primer caso, o un *rover* en el segundo. Takasu y Yasuda [20] demostraron un desempeño razonable utilizando un módulo u-blox y RTKLIB.

El objetivo de este trabajo es analizar las debilidades en el diseño cliente-servidor de RTKLIB, y proponer mejoras o adaptaciones que pueden aplicarse siguiendo los criterios de aplicaciones cliente-servidor de alto rendimiento.

En **Estructura del software RTKLIB** se revisa el diseño de la biblioteca y los puntos importantes para este trabajo. **Análisis de RTK y RTKLIB como servicio de alto desempeño** evalúa RTK desde la perspectiva cliente-servidor, y si RTKLIB respeta criterios de alto desempeño para una implementación. **Propuesta de arquitectura cliente servidor de alto desempeño para posicionamiento RTK** presenta una propuesta de implementación de servicio. En **Discusión** se presentan las observaciones y resultados de una primera implementación.

## 2. Estructura del software RTKLIB

RTKLIB ofrece funciones para calcular posicionamiento de precisión con distintas técnicas (PPP, DGPS, RTK), en simple y doble frecuencia, en tiempo real y para pos-procesamiento, para las constelaciones GPS, GLONASS, Galileo, QZSS, BeiDou, SBAS [19]; a través de código fuente C, y adaptado para sistemas POSIX y WIN32.

Entre sus utilidades, se incluyen programas para compilar en Borland C++ para sistemas Windows, que permiten hacer pos-procesamiento con datos de *rover* y BS, y con servicios de NTRIP a través de internet. Otras herramientas permiten hacer *plots* de un recorrido con un mapa como fondo, y convertir distintos formatos de intercambio de mensajes.

**RTK:** Del conjunto de prestaciones que incluye RTKLIB, es de relevancia para este trabajo la función de RTK, que se encuentra implementada en el archivo fuente `rtksvr.c`.

`rtksvr.c`, entre sus funciones, cuenta con dos rutinas que realizan los cálculos de corrección: `rtksvrstart` y `rtksvrthread`, donde la primera inicia mediante llamadas al sistema (*syscalls*) un hilo de ejecución que corre la segunda.

`rtksvrthread` itera sobre las observaciones satelitales, y para cada observación genera la información de corrección, en secuencia. Cada una de estas observaciones tiene que ser válida a partir de la información de *fix* de GNSS.

`rtksvr.c` requiere otros archivos fuente para implementar las comunicaciones con los dispositivos, operaciones matemáticas, *logging* y exclusión mutua. Para generar los datos corregidos, requiere de `rtkpos.c`, el cual invoca a las funciones que implementan el filtro de Kalman extendido y el algoritmo MLAMBDA.

## 3. Análisis de RTK y RTKLIB como servicio de alto desempeño

Como arquitectura cliente-servidor de alto rendimiento, RTKLIB puede analizarse desde dos puntos de vista: Las virtudes y falencias que presenta como servicio a clientes, y las propiedades que tiene como proceso que se ejecuta en un servidor.

**Análisis como servicio:** Como servicio provisto por un servidor de alto desempeño [11], RTKLIB carece de una interfaz adecuada. La biblioteca sólo provee un hilo de ejecución que, utilizando las observaciones satelitales, genera información de corrección o soluciones de posición y la deja disponible en estructuras internas.

Esta implementación, aunque cuenta con la capacidad de ofrecer una solución *open source* a la técnica de RTK, no avanza sobre la creación de un servicio de posicionamiento de alta precisión que pueda abastecer a más de un cliente, sea local o remoto, sin que estos clientes incorporen como parte de su código fuente a la biblioteca, enlacen con ella, y cuenten con acceso a hardware adecuado para realizar y transmitir las correcciones.

Con el advenimiento de los dispositivos móviles y las tecnologías IoT, es de esperar se implemente un sistema específico de BS, formado por un sistema embebido, que responda a solicitudes de información de corrección de más de un *rover*, siendo los *rovers* dispositivos móviles diversos, como pueden ser teléfonos móviles, tabletas, dispositivos de rastreo, *smartwatches*, sensores, sistemas IoT, etc.

Resulta por ello importante proponer una estructura superior de gestión de solicitudes y comunicaciones, que pueda aprovechar los cálculos y el hardware especializado asociado a una BS, para suministrar de forma eficiente, respetando las cotas de tiempo real, información de corrección RTK.

*Comunicaciones:* La implementación clásica de estación base simple para posicionamiento de precisión, consiste en la BS con su antena GNSS, y una conexión a un *transceiver* de radio. Esta implementación está limitada a *rovers* que cuenten con receptores de radio compatibles con el *transceiver*.

Actualmente, las comunicaciones de radio son digitales, soportando internamente protocolos de red avanzados como TCP/IP. A nivel de capa de enlace, se pueden citar como ejemplos los estándares inalámbricos IEEE 802.11 y 802.16 en todas sus variantes, y el estándar GSM.

En este escenario, una BS moderna que pretenda proveer servicio de RTK a un conjunto de dispositivos de diversas prestaciones, contará con un enlace de red TCP/IP, que posee la característica de ser independiente del medio, y puede recibir solicitudes a través de redes cableadas o inalámbricas.

Es de esperar entonces que un servidor RTK implementado con RTKLIB, tenga una interfaz TCP/IP, y la forma en la que esa interfaz se comunica con sus clientes dependa de la implementación, esto es, el hardware de comunicaciones con el que cuenta el dispositivo que funciona como servidor.

**Análisis como proceso:** El hilo `rtksvrthread`, procesa cada una de las observaciones del *rover*, generando la posición RTK con la función `rtkpos`. Este procesamiento se realiza en secuencia, procesando una observación luego de la otra de forma sucesiva (figura 1).

```

1  ....
2  for (i=0; i<fobs[0]; i++) { /* for each rover observation data */
3      obs.n=0;
4      for (j=0; j<svr->obs[0][i].n&&obs.n<MAXOBS*2; j++) {
5          obs.data[obs.n++]=svr->obs[0][i].data[j];
6      }
7      for (j=0; j<svr->obs[1][0].n&&obs.n<MAXOBS*2; j++) {
8          obs.data[obs.n++]=svr->obs[1][0].data[j];
9      }
10     rtksvrlock(svr);
11     rtkpos(&svr->rtk, obs.data, obs.n, &svr->nav);
12     rtksvrunlock(svr);
13     ....

```

**Figura 1.** Procesamiento de posiciones satelitales en Segmento `rtksvrthread`

Dado que las observaciones pueden ajustarse independiente una de la otra, el diseño de este algoritmo, no considera aprovechar las posibilidades que ofrece un sistema multiprogramado de tiempo compartido, como es un sistema Windows o Linux para los cuales está preparado, o las posibilidades que ofrecen los procesadores de múltiple núcleo, que se encuentran disponibles incluso en dispositivos simples.

Esta implementación se construye pensando en un hilo de ejecución, que corre como parte de un programa principal, el cual incluye a RTKLIB como parte de su código fuente, y está pensado también para procesamientos *offline* (en el caso de programas de escritorio), o con hardware limitado (para implementaciones en sistemas embebidos).

#### **Análisis de RTK:**

*Cliente:* El cliente de una BS es el responsable de solicitar la información de corrección, procesarla y generar la solución con precisión de centímetros. En un escenario donde un área es cubierta con estaciones base para ofrecer servicio de RTK, también debe escoger la BS adecuada. Involucra determinar BS disponibles en el radio de alcance del *rover*, y escoger la que combine cercanía, calidad de información y menor carga.

Desde la perspectiva de una arquitectura cliente-servidor, podemos decir que el *rover* es un cliente pesado (*fat*) [21], que contiene la lógica de negocio del servicio RTK.

*Servidor:* Analizando implementaciones comerciales actuales de RTK, el servidor (BS), nunca es *fat*. Las responsabilidades de la estación base están limitadas a generar información de corrección a partir de observaciones y posición conocida. Luego, esta información es transmitida.

*Otras características del servidor:* En el caso de transmisiones públicas desde el servidor, es conveniente mantener un enlace sin estado. No necesita el servidor conocer al *rover*, sólo responder sus solicitudes.

Cuando la transmisión desea restringirse, como es el caso de información de corrección que se destina a una suscripción paga, o a usos particulares de una organización, es necesario cifrar la comunicación y exigir credenciales de ingreso. En este caso, puede mantenerse la comunicación sin estado con relativa simplicidad, con un *rover* enviando en su solicitud sus credenciales.

## **4. Propuesta de arquitectura cliente servidor de alto desempeño para posicionamiento RTK**

Actualmente, pueden encontrarse en el ambiente científico, comercial y en comunidades tecnológicas, intentos RTK a bajo costo [20, 16]. Estos intentos se concentran en la implementación de una BS que soporte uno o dos clientes, pero no se consideran las necesidades de un sistema que pueda proveer posicionamiento de precisión a un gran número de dispositivos, como son los dispositivos IoT, ni toman en cuenta los requisitos de rendimiento y comunicaciones de los sistemas cliente-servidor que implementan estas funciones.

El objetivo de este trabajo es abordar el próximo paso en la implementación de una solución de RTK, teniendo en consideración una arquitectura cliente-servidor de alto rendimiento que pueda entregar a *rovers* información de corrección confiable, realizando los ajustes que requiera RTKLIB para mejorar su desempeño y soportar múltiples enlaces de comunicación.

#### 4.1. Topología

El uso de RTK con BS simple, corresponde a la topología de cliente único. Es la metodología tradicional para agrimensura y geodesia, con un usuario propietario de la estación base y el dispositivo receptor. El usuario es encargado de la instalación y configuración de ambos elementos.

Esta topología desperdicia la capacidad de una BS de proveer a varios *rovers*; y no se considera la posibilidad de ofrecer servicio RTK a terceros mediante estaciones de referencia permanentes.

En contrapartida, una arquitectura con múltiples clientes [21], se cuenta con una BS como servidor, y cualquier número de clientes pueden conectar con ella para obtener correcciones.

Finalmente, la estructura más adecuada para implementar la arquitectura cliente-servidor RTK, es contar con múltiples clientes y múltiples servidores pensando el servicio como cobertura en un área de estaciones base.

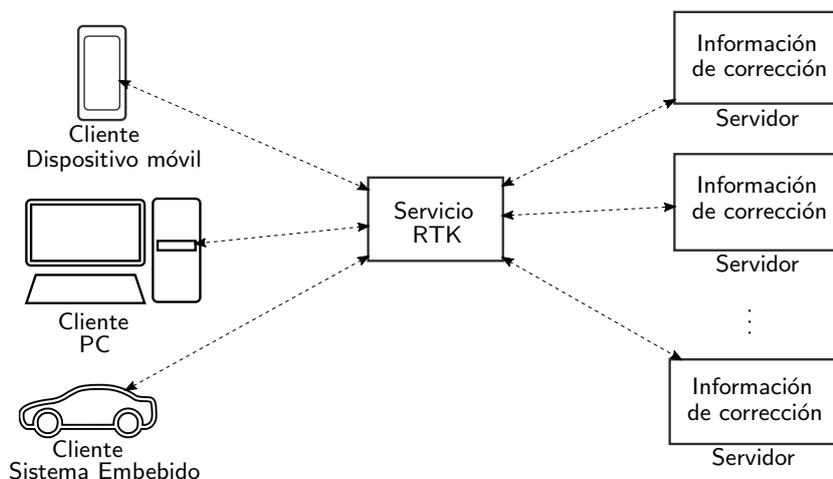
Lo que permite es la escalabilidad del servicio. La calidad de los datos de corrección está limitada por la distancia entre el *rover* y la estación base. Es por este motivo que para ampliar el área de cobertura del servicio, es necesario instalar servidores que puedan ofrecer la información de corrección a clientes cercanos.

#### 4.2. Tiers y arquitectura

Desde el punto de vista de una aplicación cliente-servidor, definimos para RTK:

- **Lógica de aplicación:** Determinar la posición en tiempo real de un *rover* a partir de la información de corrección.
- **Lógica de negocio:** Existiendo más de una BS, consiste en encontrar y escoger la más cercana geográficamente, con mejor calidad de correcciones y disponible; y establecer la conexión.
- **Lógica de datos:** El servicio de BS generando la información de corrección válida para su zona de influencia.

Por lo expuesto anteriormente, una arquitectura de 2T, delega en el cliente seleccionar el servidor más cercano. Esto es, la lógica de negocio debe estar necesariamente en el cliente. A esta responsabilidad se añade también mantener una lista de servidores RTK con sus localizaciones, actualizar esta lista, y determinar cuál es el servidor más adecuado a utilizar.



**Figura 2.** Arquitectura 3T

Estas tareas son demandantes en procesamiento y comunicaciones. Si consideramos que el *rover* es un dispositivo móvil, funcionando a baterías, posiblemente conectado a través de una red de datos móviles paga, la arquitectura 2T no es la más adecuada para proveer un servicio RTK.

Si se implementa una arquitectura 3T (figura 2), el cliente se comunica directamente con un servidor que maneja la lógica de negocio. Éste conoce las BS, su localización y disponibilidad, y sería posible incorporar balance de carga entre ellas. Además, aumenta la flexibilidad pudiendo agregar nuevas estaciones de forma transparente al cliente, y facilita la implementación de conexiones con estado para transacciones más eficientes.

### 4.3. Implementación

A nivel de lógica de datos, se implementa un servicio BS en Linux, que compila con RTKLIB, abriendo un puerto de comunicaciones TCP/IP. Para la lógica de negocio, se crea un servicio para Linux que se conecta al servicio BS. La función de la lógica cliente es obtener correcciones de la lógica de negocio y enviar a través de una interfaz USB/RS232 a su módulo GNSS. Este módulo tiene soporte para correcciones.

**Plataformas:** Raspberry Pi 3 Model B+, con procesador de arquitectura ARM Cortex A53, y sistema operativo Linux, distribución Raspbian Buster.

Los receptores son placas de evaluación Smart GPS revisión 1.11. El *chip* de GNSS fue desoldado y reemplazado por u-blox LEA-6T-0 de igual *pinout* con soporte para mensajes *raw* con formato propietario ublox y RTCM. El funcionamiento se validó utilizando el software u-center provisto por el fabricante del módulo, con información de corrección del servicio NTRIP provisto por RAMSAC [13].

Se utilizaron dos modelos de antenas, ambas activas, Garmin GA25MCX y Taoglas AGGP .25F.

**Código fuente:** Se escribe en lenguaje C versión C89 [1], compilado para la combinación de SO y procesador.

**Condiciones de operación:** Las primeras pruebas fueron realizadas con el modo *moving baseline* en condiciones acotadas. Está planificada una segunda etapa de pruebas utilizando puntos geodésicos conocidos para las antenas.

## 5. Discusión

Las pruebas iniciales se realizaron con una PC como *rover* y Raspberry Pi como BS, funcionando con *moving baseline*. `htop` para verificar desempeño de los procesos. La exactitud de las mediciones se determinó midiendo la distancia geométrica entre antenas.

**Exactitud:** La exactitud entre 2 y 6 metros con baja precisión, no resultó la esperada u obtenida en otras pruebas realizadas sobre RTKLIB [20]. Sin embargo, esto puede estar vinculado al hardware utilizado. Se necesita un análisis más exhaustivo, reducir el *multipath* utilizando planos a tierra para las antenas, utilizar versiones recientes de los módulos GNSS (M8T, M8N), hacer verificaciones con base estacionaria, comparar con medidas de pos-procesamiento.

**Procesamiento y comunicaciones:** Bajos requisitos de procesamiento y comunicaciones. La carga generada por los hilos de ejecución no resultó significativa a la operación del sistema, así como las comunicaciones. En pruebas posteriores, es necesario simular un volumen alto de solicitudes, a través de un enlace inalámbrico, para verificar los efectos de la interferencia.

**Problemas observados:** RTKLIB no implementa una arquitectura de software definida. Sufre de *code smells*: Archivos de código fuente extensos, funciones con decenas de parámetros, líneas con múltiples sentencias, código condicional, compilación con *warnings*. Muchas de sus variables tienen nombres de una o dos letras. El autor de este trabajo supone que se pretende relacionar los algoritmos implementados con sus ecuaciones, como son el filtro de Kalman extendido que suaviza la posición, y la implementación de MLAMBA para resolver la ambigüedad. Estas variables de una letra designan matrices de covarianza, actualización de estado, predicción, observaciones, entre otros. El problema de este diseño yace en que el código fuente no cuenta con recursos tipográficos que hagan legibles a estas declaraciones. Deberían reemplazarse con nombres que declaren intencionalidad.

Es recomendable realizar un *refactoring* del código, estableciendo primero *test harnesses* [3], y reemplazar las declaraciones de tipo por `stdint.h`. Las verificaciones realizadas en [20] utilizaron con un procesador de arquitectura de 32 bits, mientras que las implementaciones evaluadas en este trabajo utilizan un procesador con arquitectura de 64 bits, generando incertidumbres en el funcionamiento al no utilizar tipos de datos estandarizados.

Estas modificaciones son difíciles de realizar. RTKLIB cuenta con una batería de tests con un *coverage* limitado, que resultado difícil de evaluar debido a la

falta de un criterio estándar de validación. Estos tests están implementados como programas C independientes que ejecutan fragmentos de código.

## 6. Conclusiones

Es indiscutible la ubicuidad de sistemas electrónicos y de software en actividades industriales, comerciales y particulares. Desde dispositivos móviles (*smartphones*, tabletas, *smartwatches*, etc.); hasta sistemas de *Internet of Things* (IoT) que actúan y miden condiciones ambientales con algún propósito, como detectar situaciones de riesgo (contaminación, incendios), recopilar información (estaciones meteorológicas), o automatizar vehículos y máquinas. Una de las necesidades frecuentes en estos sistemas es determinar la localización. La tecnología de RTK ofrece posicionamiento y localización con precisión de centímetros en tiempo real, valiéndose sólo de un receptor y antena de GNSS comercial, si se cuenta con un servicio que provea la información de corrección.

Como se mostró en este trabajo, es posible incorporar RTK a sistemas con GNSS, sin un aumento significativo de costo, con versiones alternativas compatibles de los módulos GNSS ya utilizados. Sin embargo, esta tecnología requiere un proveedor de información de corrección. Aunque existen servicios gubernamentales de NRTK, suelen estar restringidos, y tener como objetivo su uso en agrimensura [13]. Además sufren limitaciones de cobertura, por la distancia a la CORS más cercana, o por la ausencia de conexión a internet de alta velocidad.

RTKLIB permite implementar con hardware genérico, dispositivos de bajo costo que generan la información de corrección. En los últimos años, estos elementos han provocado un creciente interés en la comunidad tecnológica y científica con necesidades de RTK de bajo costo [4, 22, 18]. Este trabajo inicia los primeros pasos para la implementación de un servicio RTK escalable, de alta disponibilidad, construido en software libre, que haga realidad el objetivo de posicionamiento satelital de alta precisión para aplicaciones de IoT.

## Referencias

- [1] Computer y Business Equipment Manufacturers Association. *ISO/IEC 9899:1990*. Inf. téc. International Organization for Standardization, 1990.
- [2] Paolo Dabove y col. «Network Real Time Kinematic (NRTK) Positioning – Description, Architectures and Performances». En: *Satellite Positioning - Methods, Models and Applications*. Mar. de 2015, págs. 23-46.
- [3] M.C. Feathers. *Working Effectively with Legacy Code*. Martin, Robert C. Prentice Hall PTR, 2004.
- [4] María S. Garrido-Carretero y col. «Low-cost GNSS receiver in RTK positioning under the standard ISO-17123-8: A feasible option in geomatics». En: *Measurement* 137 (2019), págs. 168-178.
- [5] André Hauschild. «Basic Observation Equations». En: *Springer Handbook of Global Navigation Satellite Systems*. Ed. por Peter J.G. Teunissen y Oliver Montenbruck. 2017. Cap. 19.

- [6] William Henning. *User Guidelines for Single Base Real Time GNSS Positioning*. Abr. de 2014.
- [7] Thomas Hoblger y Norbert Jakowski. «Atmospheric Signal Propagation». En: *Springer Handbook of Global Navigation Satellite Systems*. Ed. por Peter J.G. Teunissen y Oliver Montenbruck. 2017. Cap. 19.
- [8] Open Source Initiative. *The 2-Clause BSD License*. 2020. URL: <https://opensource.org/licenses/BSD-2-Clause>.
- [9] Elliott D. Kaplan. «Introduction». En: *Understanding GPS/GNSS Principles and Applications*. Ed. por Elliot D. Kaplan y Christopher J. Hegarty. third. Artech House, 2017. Cap. 1.
- [10] J. Liu y col. «Review of GNSS ambiguity validation theory». En: *Geo-mat. Inform. Sci.* (2014).
- [11] C. Loosley y F. Douglas. *High-Performance Client/Server*. Wiley, 1997.
- [12] Y Lou y col. «An algorithm and results analysis for GPS+ BDS inter-system mix double-difference RTK». En: *Geodesy Geodyn* (2016).
- [13] Instituto Geográfico Nacional. *Red Argentina de Monitoreo Satelital Continuo*. 2020. URL: <http://www.ign.gob.ar/NuestrasActividades/Geodesia/Ramsac>.
- [14] Dennis Odijk. «Positioning Model». En: *Springer Handbook of Global Navigation Satellite Systems*. Ed. por Peter J.G. Teunissen y Oliver Montenbruck. 2017. Cap. 19.
- [15] Dennis Odijk y Lambert Wanninger. «Differential Positioning». En: *Springer Handbook of Global Navigation Satellite Systems*. Ed. por Peter J.G. Teunissen y Oliver Montenbruck. 2017. Cap. 26.
- [16] Fan Ouyang y col. «Automatic delivery and recovery system of Wireless Sensor Networks (WSN) nodes based on UAV for agricultural applications». En: *Computers and Electronics in Agriculture* 162 (2019), págs. 31-43.
- [17] Brent A. Renfro y col. *An Analysis of Global Positioning System (GPS) Standard Positioning Service Performance for 2019*. Inf. téc. The University of Texas at Austin, 14 de mayo de 2020.
- [18] Rosendo Romero-Andrade y col. «Comparative analysis of precise point positioning processing technique with GPS low-cost in different technologies with academic software». En: *Measurement* 136 (2019), págs. 337-344.
- [19] Tomoji Takasu. *RTKLIB: An Open Source Program Package for GNSS Positioning*. 2020. URL: <http://www.rtklib.com/>.
- [20] Tomoji Takasu y Akio Yasuda. «Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB». En: *International Symposium on GPS/GNSS* (ene. de 2009).
- [21] S. Chandra Yadav y S. Kumar Singh. *An Introduction to Client Server Computing*. New Age International Pvt. Ltd., Publishers, feb. de 2009.
- [22] Yun Zhang y col. «Static and kinematic positioning performance of a low-cost real-time kinematic navigation system module». En: *Advances in Space Research* 63.9 (2019). Multi-GNSS: Methods, Benefits, Challenges, and Geosciences Applications, págs. 3029-3042.

# Workshop Innovación en Sistemas de Software

Coordinadores

Pablo Fillottrani (UNS)  
Marcelo Estayno (UNLZ)  
Alicia Mon (ITBA)  
Dante Zanarini (UNR)

# Motor de Reglas desacoplado orientado a formato JavaScript Object Notation.

Marcos Maciel

CAETI, Buenos Aires, Argentina

Mmaciel03@hotmail.com

**Resumen:** Las empresas afrontan el desafío de competir en escenarios que varían de acuerdo con nuevas tendencias, clientes que buscan calidad a menor precio, o a un repentino cambio de interés desde un producto y/o servicio a otros. El acceso a dispositivos móviles con internet como medio de comunicación generan tanta curiosidad que los usuarios continuamente navegan buscando alternativas. Para satisfacer con las demandas de estos clientes tecnológicos las compañías deben cambiar sus reglas de negocios periódicamente y adaptarse inmediatamente al medio que los rodea. Con el propósito de cumplir a la demanda de cambios dinámicos y exageradamente rápido, se presenta en este trabajo un modelo orientado a crear reglas de negocios en lenguaje natural del dominio para evaluar las condiciones de verdad y modificar la información resultante. Esta propuesta está orientada a sistemas con arquitecturas orientadas a servicios que conectan aplicaciones front-end con back-end a través del formato Json (JavaScript Object Notation) como medio de intercambio de información. Las estadísticas realizadas demuestran una performance aceptable en la validación de atributos Json y una ganancia considerable entre el desarrollo de la regla en código contra la parametrización de la regla.

**Palabras clave:** Motor de reglas, BRMS, reglas de negocios, Json, microservicios, tecnología, lenguaje del dominio del negocio.

## I. Estado del arte.

Los motores de reglas tienen una larga trayectoria, la cual comienza con la inteligencia artificial, donde fueron aplicados como método para representar conocimiento [01]. Un motor que evalúa reglas es un sistema experto específico porque la regla representa un entendimiento del dominio del negocio. Estos sistemas inteligentes son ágiles cuando gestionan la lógica de negocio de forma independiente de la aplicación [01]. Un motor de reglas apoya los cambios volátiles porque separa la regla de los procesos [02]

En el mercado existe motores de reglas comerciales y open source tales como

FICO Blaze Advisor, las reglas en Blaze Advisor son escritas en su propio lenguaje llamado SRL (Lenguaje de reglas estructurado). Previamente a crear una regla un modelo de objeto debe ser creado. Blaze es orientado al desarrollador y como producto comercial tiene un costo asociado.

Jena es un framework open source [04] fuertemente orientado dentro del marco de la Web Semántica, incluye un motor de reglas genérico, aunque incluye un ambiente

de programación para interactuar con RDF, RDFS, OWL y SPARQL [05].

Drools [06] es un sistema de administración de reglas de negocios, incluye el motor de inferencia basado en reglas de tipo inferencia hacia adelante para Java. Su lenguaje declarativo es lo suficientemente flexible para soportar cualquier dominio y está orientado a objetos, por otro lado, también es necesario contar con conocimientos de Eclipse Java®

Clips [07] es una herramienta para construir sistemas expertos de dominio público. Utiliza un lenguaje orientado a objetos llamado. Está orientado a desarrolladores, es necesario contar con conocimientos técnicos.

Oracle Policy Automation [08] son un grupo de productos para modelar e implementar reglas. Las reglas pueden ser creadas sin un modelo, en un lenguaje natural y en un editor tipo Word, esto puede causar algún problema derivado de la flexibilidad del lenguaje. La implementación de las reglas lleva una serie de pasos que requieren cierto conocimiento de herramientas externas al ecosistema del motor de regla.

## II. Introducción.

En un mundo globalizado por internet las personas tienen muchas alternativas al alcance de la mano. Todos pueden acceder a teléfonos móviles en busca de opciones y de todas estas, elegir la que mas se ajusta a sus necesidades. La competencia entre organizaciones es mas variada y por ello la flexibilidad para cambiar ondemand debe ser proactiva y dinámica. Para las organizaciones que cuentan con sistemas orientados a servicios se propone un gestor de reglas desacoplado del sistema Core de negocios, orientado a Json [14], sin compilación, ni despliegue. El propósito es brindar una solución controlada por el negocio e independiente del equipo técnico. Otros aspectos para destacar de este motor de reglas son que al ser orientado a servicios no existe dependencia del lenguaje de programación con el sistema Core, No es orientado a objetos y por lo tanto no es necesario trabajar con clases o entidades del dominio, es también factible de implementar sobre un sistema ya productivo.

La ventaja de los servicios api rest es que usan un formato estándar para el intercambio de información, llamado Json cuyos atributos contienen valores no tipados. Otros servicios pueden recibir una estructura Json no tipado y procesarlo sin importar el lenguaje con el que haya sido construido. Esta propuesta implementa el uso de Json para parametrizar los distintos tipos de atributos que modelan las entidades de negocios y en conjunto forman el dominio del conocimiento de la organización. Por ejemplo, una tabla representa un nivel y un atributo de la tabla representa una entidad. Esta forma permite agrupar las entidades por niveles, aspecto importante a la hora de armar una regla. Cuando un sistema genera una representación en Json explícitamente lo modela orientado a objetos, es decir el conjunto de atributos se encuentra ajustado al modelo de dominio que los usuarios del negocio entienden, esta característica es aprovechada por el motor propuesto para evitar la dependencia de las clases, objetos y lenguaje de programación seleccionado para desarrollar el sistema Core. El modelo de datos que sustenta al motor de reglas puede ser generado desde tablas de base de datos o directamente desde los distintos tipos de response Json que el sistema genere. Si bien el estándar Json no es tipado, la propuesta incluye relacionar una entidad del negocio con un tipo de datos para mejorar la validación entre <término><hechos><término>. Desde el punto de vista técnico, la ejecución de la regla se inyecta entre el modulo front-end y back-end. En los sistemas orientados a servicios el front-end realiza una serie de pasos que llevan a los usuarios finales a componer un conjunto de datos, estos son enviados con una estructura formulada por el equipo de sistemas. Esta estructura Json es recibida por el módulo back-end quien previo a ejecutar las rutinas Core, reenvía la estructura al servicio de rule engine. Este ejecuta la/s regla/s previamente configuradas, si corresponde modifica el/los valor/es de los atributos y devuelve la estructura al back-end. Este último paso califica al motor de reglas como desacoplado, porque no es necesario incluir código intruso en el sistema Core. Existen distintas variantes para la implementación, dando lugar por ejemplo a la construcción de un servicio Facade [20] que capture las llamadas al Core, derive al motor de reglas, obtenga una respuesta y envíe al sistema Core.

Este paper está organizado de la siguiente forma: descripción del proceso de implementación de cambios tradicional en la sección III, en la sección IV se desarrolla la propuesta de implementación de cambios con motor de reglas, Testing y estadísticas en la sección V, y por último se hallan conclusiones y trabajos futuros en sección VI.

### **III. Proceso de implementación de cambios tradicional.**

Las aplicaciones Core del negocio tomadas como base para esta investigación tienen un ciclo de vida para incorporar cambios que se ajusta a una metodología de grandes compañías.

A continuación, se lista las tareas y horas involucradas en el ciclo de vida de un cambio:

- Definición del requisito por parte de usuario final del negocio.
- Análisis de la necesidad por parte del analista del negocio.
- Reunión de análisis entre el analista del negocio y del analista de sistemas.
- Reunión de transferencia del conocimiento entre el analista de sistemas y el líder de sistemas.
- Reunión de transferencia del conocimiento entre el líder de sistemas y desarrolladores front-end y back-end.
- Reunión de transferencia del conocimiento entre el líder de sistemas y tester de sistemas.

- Crear de una rama del repositorio.
- Armar ambiente local para desarrollo.
- Desarrollo de los cambios.
- Pruebas del desarrollador de sistemas en ambiente local.
- Publicación de los cambios en ambiente QA (pruebas).
- Testing del tester de sistemas
- Validación del analista de sistemas.
- Pruebas del analista de negocios y aceptación.
- Crear de una rama del repositorio.
- Planificación y armado de implementación.
- Revisión y aprobación de pasaje a producción.
- Implementación en servidores productivos.
- Merge branch-trunk en repositorio.
- Comunicación de los cambios implementados al usuario final del negocio.
- Almacenamiento y cierre del ticket.

El tiempo medio que toma implementar un cambio se ajusta al siguiente cuadro.

	Usuario Negocio	Analista del Negocio	Analista de Sistemas	Líder de Sistemas	Desarrolladores	Tester de Sistemas
Definición del requisito	10					
Análisis de requisito		5				
Reunión funcional		4	4			
Reunión técnica			3	3		
Reunión Diseño				4	4	
Reunión testing				2	2	
Gestion repositorio codigo QA				0.5		
Armado ambiente desarrollo					2	
Desarrollo de los cambios					4	
Pruebas de sistemas					1	
Publicación de los cambios en QA					0.5	
Validación tester					2	
Validación funcional					1	
Gestion repositorio codigo a producción					0.5	
Planificación y armado de implementación					0.5	
Revisión y aprobación de pasaje a producción					0.5	
Implementación producción					1	
Merge branch-trunk en repositorio					1	
Comunicación					0.5	
Almacenamiento y cierre del ticket					0.5	
Sub-Total	10	9	7	9.5	21	0
Total	56.5					

**Tabla. 1.** Tabla de tareas involucradas en un cambio menor.

La Tabla. 1 describe los pasos involucrados en el proceso de cambio, desde la definición de un usuario del negocio pasando por sistemas hasta la implementación en producción.

Esta medición es para un cambio del tipo básico:

Cuando Cliente es Cliente Afinidad

Cuando Importe  $\geq$  \$ 1000 entonces Importe = Importe \* 0.9

Es decir, se quiere “premiar al cliente que compra asiduamente y por valor igual o superior a un valor  $x$  con el 10% de descuento”.

En código de sistemas se puede representar de la siguiente forma

```
Function bool IsVIP (object item)
{
    var client = repository.getClient(item);
    return client.vip;
}
Function decimal GetDiscount (object item)
{
    return repository.getDiscount(item);
}
Function decimal GetConstant ()
{
    return repository.getConstant ();
}

Function main(objeto item)
{
    If(IsVip(item))
    {
        if(item.importe >= GetConstant)
        {
            Var discount = GetDiscount (item);
            Item.importe = Item.import * discount;
        }
    }
    return item;
}
```

La carencia de flexibilidad para incorporar cambios de forma rápida obliga a involucrar tiempo de distintos roles, quienes aportan valor al proceso, pero también agregan horas de trabajo.

En el código de ejemplo se puede cuantificar 4 métodos, al menos 2 parámetros para % de descuento y un valor para comparar cuando se supera un límite, además de una rutina principal. Es probablemente necesaria la creación de tablas o atributos sobre tablas existente para almacenar estos valores que el desarrollador necesita codificar.

En resumen, el ciclo tradicional de implementación de cambios necesita de varios actores y pasos para cumplir los requerimientos del negocio. El tiempo dedicado puede variar en función de una a otra metodología, pero el resultado global es un valor horario que esta propuesta propone acotar. Para lograrlo es necesario separar el diseño y gestión de las reglas de los procesos de sistemas [15]. Esta división de responsabilidades evita depender exclusivamente del profesional informático y la carga se reparte con los analistas del negocio.

#### **IV. Proceso de implementación de cambios con Motor de Reglas.**

El modelo propuesto en este documento utiliza como tecnologías de construcción Node.js® [13], ReactJs© [09], Redux© [10], npm© [11], PostgreSQL© [12] y una arquitectura orientada a microservicios. El prototipo se encuentra codificado en Node.js® para el backend y ReactJs©, Redux© para el front-end. Existen otros frameworks de desarrollo que permiten construir el módulo, pero se decide continuar una línea de desarrollo adoptando las nuevas tendencias del mercado.

El motor se encuentra completamente apoyado sobre Json como estructura de intercambio de datos, entre sus características se pueden listar: no usa tag de fin, es más legible y liviano (ver tabla 2).

<pre> {"employees":[   { "firstName":"John",     "lastName":"Doe" },   { "firstName":"Anna",     "lastName":"Smith" },   { "firstName":"Peter",     "lastName":"Jones" } ]}                 </pre>	<pre> &lt;employees&gt;   &lt;employee&gt;     &lt;firstName&gt;John&lt;/firstName&gt;     &lt;lastName&gt;Doe&lt;/lastName&gt;   &lt;/employee&gt;   &lt;employee&gt;     &lt;firstName&gt;Anna&lt;/firstName&gt;     &lt;lastName&gt;Smith&lt;/lastName&gt;   &lt;/employee&gt;   &lt;employee&gt;     &lt;firstName&gt;Peter&lt;/firstName&gt;     &lt;lastName&gt;Jones&lt;/lastName&gt;   &lt;/employee&gt; &lt;/employees&gt;                 </pre>
<p>Tamaño: 157 bytes (157 bytes)</p>	<p>Tamaño: 291 bytes (291 bytes)</p>

Tabla. 2. Tabla de comparación entre formatos Json y xml.

#### IV.A. Definición del modelo de datos del motor de reglas.

El modelo de datos propuesto esta implementado en 7 tablas de base de datos. En la Tabla. 3 se explican 2 tablas de sistemas: function y type las cuales contiene:

Tabla type	numeric	date	string	float	boolean	calculo	Json
Tabla function	numeric	= > <= >= <>	asignar				
	date	= > <= >= <>	asignar				
	string	contiene - igual - comienza con - termina con - asignar					
	boolean	= asignar					
	Json	asignar					
	calculo	+= -= *= asignar					

Tabla. 3. Tabla de tipos de datos y funciones.

Las tablas de sistemas tienen como finalidad definir de acuerdo al tipo de datos que función se puede ejecutar sobre una entidad de negocio. Por ejemplo, de una entidad Código de Cliente se le asigna el tipo numeric y la funciones a comparar están determinada por comparaciones numéricas.

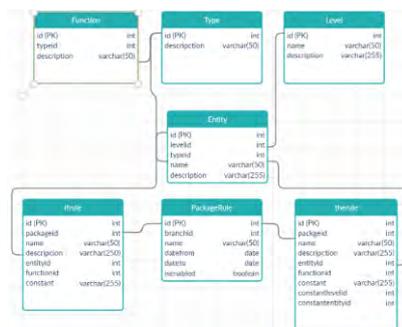


Fig. 1. Diagrama de entidad relación usado por el motor de reglas.

El resto de las tablas son PackageRule, Level, Entity, IfRule, y ThenRule. La tabla PackageRule tiene por objetivo agrupar reglas ifrule y thenrule. Level agrupa entidades y tabla Entity agrupa columnas de las entidades de negocios. Una regla esta conformada por al menos un registro en la tabla IfRule y un registro en la

tabla ThenRule. Este modelo de datos denominado de negocios esta orientado a usuarios que conocen el dominio.

### IV.B. Definición de datos.

Los datos necesarios para parametrizar reglas se encuentran relacionado con las tablas de base de datos del negocio, es así como las tablas (ej. Clientes, productos, etc.) son registradas como niveles en la tabla Level y los atributos (ej. Nombre cliente, nombre producto) de cada tabla son registrados como entidades en la tabla Entity. Un registro de una tabla contiene un objeto del negocio representado en los sistemas Core como clases. Este motor se abstrae de las clases y el lenguaje de programación para usar las tablas y columnas como método para representar objetos. Estos pueden ser tablas y/o vistas de base de datos.



Fig. 2. Proceso de transformación de modelo de negocio en tabla del motor de reglas.

id	level	level_description	id	levelid	typeid	name	description
1	1 cotizacion	header	1	2	1	1 ramo	Ramo
2	2 item	Item o Riesgo	2	3	1	1 periodo_refacturacion	Periodo de Refacturacion
3	3 asegurado	Tomador	3	4	1	2 fec_desde	Fecha Vigente Desde
4	4 cobertura	Cobertura	4	5	1	2 fec_hasta	Fecha Vigente Hasta
5	5 amparo	Amparo	5	6	1	1 fronting	Número de Poliza
6	6 importe	Importes	6	7	1	1 tipo_poliza	Tipo de Poliza
7	7 anexo	Anexos					

Fig. 3. Datos de tabla Level.

Fig. 4. Datos de tabla Entity.

En la Fig. 2 se describe el proceso general para entender como el motor de reglas relaciona los datos del dominio. En principio las tablas candidatas de la base de datos (las tablas de parámetros en general son excluidas) son enumeradas y su nombre es usado para llenar la tabla Level (ver Fig. 3). En este caso se toma de ejemplo un modelo de cotización de autos de la industria de seguros, los datos necesarios son un encabezado o cotización, datos del asegurado, vehículo, coberturas y amparos contratados, clausulado etc. En segundo lugar, por cada tabla candidata se listan las columnas y su tipo de datos que en conjunto llenan la tabla Entity (ver Fig. 4). Este par Nivel y Entidad da sentido a la programación de la regla que esta orientada al lenguaje natural y permite ordenar las entidades según el nivel que corresponda al dominio. Para un usuario del negocio es simple de entender una sentencia del tipo Ramo de la Cotización o Cotizacion.Ramo ya que es parte diario de su lenguaje de comunicación.

Cada registro de la tabla Entity tiene asociado a un tipo de datos (Fig. 4). Cuando el usuario del negocio está creando una regla y elige una entidad, su tipo de datos filtra las funciones que se pueden aplicar. Por ejemplo, una entidad numérica solo puede ser comparada por es mayor, es menor, es igual, es distinta pero no puede ser comparada por una función alfabética, como se muestra en Tabla. 4.

Cotizacion.Ramo es Autos	Cotizacion.Ramo es Autos
Cotizacion.Ramo = 4	Cotizacion.Ramo comienza con 4
<b>correcto</b>	<b>incorrecto</b>

Tabla. 4. Tabla de comparación de funciones.

### IV.C. Servicio de Ejecución de Regla.

La regla se ejecuta en un servicio api rest [16], debido a la elección de los equipos de sistemas por sobre SOAP, RPC etc, [17][18] que tiene las siguientes tareas:

- Recibe una estructura en formato Json desde un cliente.
- Recupera los paquetes de reglas almacenados en la base de datos
- Evalúa regla a regla
- Si corresponde modifica la misma estructura recibida en punto a.

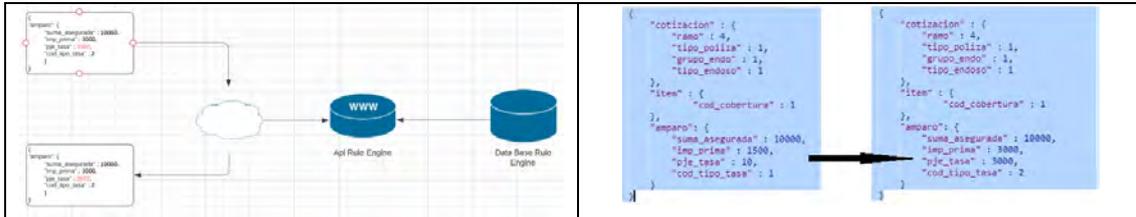


Fig. 5. Servicio Api rest Motor de regla

La estructura recibida es almacenada en memoria y se ejecuta la validación la sintaxis de Json [19], si esto es correcto continua con la recuperación de los paquetes de reglas. Cada paquete contiene un conjunto de reglas ifrule y thenrule o condición y acción. El paso siguiente es validar regla a regla, este proceso se ejecuta encadenado hacia adelante [20], el grupo de condiciones -como un todo- debe ser verdadero para desencadena una o varias acciones, estos pasos están graficados en Fig. 5. La descripción de proceso de validación de regla en pseudocódigo se encuentra resumida en Fig. 6.

```

Procedimiento main(json)
si formato de json es invalido se termina la ejecucion
j <- json
k <- buscar paquetes de regla en la base de datos
mientras(k)
  m <- asigna las reglas ifrule(condicion)
  mientras(m)
    se obtiene el valor del atributo de m
    se busca el atributo con j[m.level] y j[m.entity] (json de entrada)
    si existe el atributo en j entonces
      se obtiene el tipo de datos de m
      se obtiene la funcion de comparacion de m
      si valida si regla es verdadero (j[level].j[entity].value funcion j[m.level].j[m.entity].value)
        continuar
  n <- asigna las reglas thenrule(accion)
  if(ifrule es verdadero)
    mientras(n)
      asignar al atributo json j[m.level].j[m.entity] = funcion n.constante

```

Fig. 6. Servicio Api rest Motor de regla

El método de comparación usado es primero buscar la regla parametrizada en base de datos, entonces se busca en el cuerpo de la estructura Json recibida por coincidencia, es decir por pares atributo padre y atributo hijo, si se encuentra este par se toma el valor contenido en el atributo. El tipo de datos y su función asociada es recuperada del par atributo hijo almacenado en base de datos (en la regla) entonces se aplica la función al valor recibido contra el valor almacenado en la regla. Si el valor recibido en Json no se corresponde con el tipo de datos asociado en el atributo hijo de la regla, entonces la validación se toma como no verdadera, dejando al paquete de regla como falso y no se aplican cambios.

#### IV.D. Front-end para gestionar las reglas.

La interfaz para componer o crear regla se encuentra implementada en forma simple y clara, el concepto principal es llamado paquete de reglas, este contiene un conjunto de n reglas agrupadas en dos grupos independientes en cuanto a su gestión, pero dependientes en relación con los aspectos funcionales del dominio de la regla.

El módulo tiene por objetivo crear una o varias reglas del tipo condición que se evaluara en un futuro proceso, y que puede ser verdadera o falsa. Cuando se desarrolla un paquete de reglas, este tiene una o varias condiciones del tipo *Si esto es verdadero Entonces cambiar el valor* de una variable del dominio. Este proceso encadena una lista de condiciones, si todas ellas se cumplen entonces se ejecuta su proceso dependiente cuya finalidad es modificar valores existentes incluso en la misma condición previamente evaluada. Una regla de negocio se define de la siguiente forma:

“Si <condición> entonces <acción>” o “if <condition> then <action>”

Regla A		Regla B	
Condición	Acción	Condición	Acción
1=1	A asignar B	1=2	
2=2		2=3	
3=3			
4=4			

Tabla. 5. Proceso de validación y modificación de datos.

Como se describe en Tabla 5, todas las condiciones deben ser verdaderas para desencadenar una acción, ej. Regla A. Si una de las condiciones es falsa no se desencadena acción, ej. Regla B

A continuación, se presenta la interfaz de usuario propuesta para gestionar las reglas.

+	Name	Description	Enabled	Date From	Date To
	Nuevo Negocio	Nuevo Negocio Plan Normal	✓	01/01/2020	31/12/2020

« < Page 1 of 1 > » 1/1 Display 1-10 of 10

Fig. 7. Interfaz de usuario para administrar paquetes de reglas.

La Fig. 7 contiene la lista de paquetes de reglas gestionadas por la interfaz con nombre y descripción, estos dos campos combinados deben describir en forma concreta la validación de negocio completa. Un caso concreto puede ser Paquete: Nueva póliza de Autos, donde las reglas contenidas validan todos los aspectos para vender una póliza del ramo automotores.

+	Name	Description	Enabled
	Grupo de Cobertura	Si Grupo de Cobertura es RC	✓
	Prima Mínima	Si Prima Mínima es menor a \$ 1500	✓
	Ramo Autos	Si es ramo es Autos	✓
	Plan Normal	Si el plan es Normal	✓
	Grupo de Endoso	Si Grupo de Endoso es Nuevo Negocio	✓
	Tipo de Endoso	Si Tipo de Endoso es Nuevo Negocio	✓

Name: Grupo de Cobertura, Descripción: Si Grupo de Cobertura es RC,  Enabled

Level: Cobertura, Entity: Código de Cobertura, Function: =, Constant: 1

Validator:

```

cobertura :
{
  cod_cobert : 1
}
        
```

Fig. 8. Interfaz de usuario para administrar paquetes de reglas if o condición.

+	Name	Description	Enabled
	Tipo de Tasa	Entonces asignar Tipo de Tasa FIJO	✓
	Tasa de la Cobertura	Entonces asignar Prima Mínima	✓
	Prima Mínima	Entonces asignar \$ 1500 a Prima Mínima	✓

Name: Tipo de Tasa, Descripción: Entonces asignar Tipo de Tasa FIJO,  Enabled

Level: Amparo, Entity: Tipo de Tasa, Function: asignar, Constant: 2

Select a Level:  Select an entity:

```

amparo :
{
  tipo_tasa : 2
}
        
```

Fig. 9. Interfaz de usuario para administrar paquetes de reglas then o acción.

Las Fig. 8 y 9 muestran la interfaz del usuario del negocio en el lenguaje que este comprende, lado izquierdo de las imágenes y superior derecho. En este caso la regla valida que una venta para un automotor cuando el usuario selecciona la cobertura obligatoria por ley: Responsabilidad civil, tenga un valor mínimo. En caso de validación verdadera (un valor x es menor a \$1500) se ejecuta la acción de asignar el \$1500, estas validaciones son estándar en la industria y el valor continuamente es modificado de acuerdo con la inflación país. El almacenamiento de esta regla en base de datos respeta el lenguaje del área informática y que

internamente el código entiende como estructura Json, este formato está expuesto en las imágenes inferiores derecha de las mismas figuras.

### IV.D. Implementación con Json

El modelado de datos que soporta el motor de reglas descrito en primer lugar es expuesto en la interfaz de usuario que permite administrar los paquetes de reglas. El servicio como modulo independiente ejecuta en orden los paquetes habilitados cuando es invocado por un cliente externo. Este paso es llamado desacoplado: el servicio de ejecución de reglas puede existir sin clientes que lo consuman, sin clases previamente programadas, sin importar el lenguaje de programación y/o modelo de base de datos de los sistemas Core clientes que consuman el servicio.

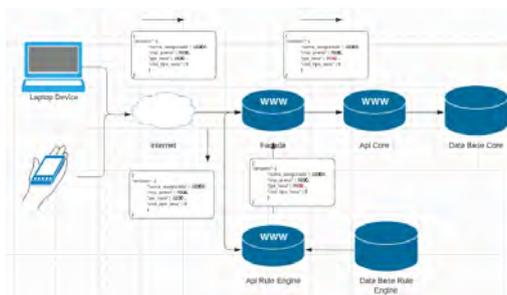


Fig. 10. Proceso de transformación de modelo de negocio en tabla del motor de reglas.

Un cliente en su dispositivo ya sea desktop, laptop o móvil hace un request o pedido a una página html o aplicación móvil. Este request puede ser un formulario que deba completar para solicitar una cotización online. Una vez completa el formulario el cliente hace un postback o envío de datos a su servicio de dominio con la estructura definida por el desarrollador informático. En este punto de la arquitectura, se propone implementar una capa de servicios por delante de Api Core que actúe como un Facade [21] que intercepte el mensaje, lo derive al api de motor de reglas. El servicio de motor de reglas ejecutara la validación sobre el/los paquete/s de regla/s parametrizado/s por usuarios del negocio. El paso posterior a la validación puede resultar en la modificación o no de la estructura json y como ultima tarea se retorna la respuesta, como se ejemplifica en la Fig. 10.

### V. Testing

Se realizan pruebas de performance para determinar si el peso en el tamaño de un Json estándar puede ser un limitante para la ejecución de reglas. Las estructuras dependen de cada sistema cliente y también los valores contenidos por los atributos, para el caso de estudio el módulo del sistema tomado de ejemplo tiene un peso total máximo de 11256kb. El peso mínimo del mismo modulo es de 254kb, este valor es tomado como limite inferior para las pruebas. El test se ejecutó sobre un ambiente controlado en un pc desktop de desarrollo con las siguientes especificaciones:

Procesador Intel® Core™ i5-8350U CPU @ 1.70GHz 1.90 GHZ 8.00 GB procesador 64bits.

Peso kb	Iteraciones 50	Iteraciones 100	Iteraciones 150	Iteraciones 200	Iteraciones 300	Iteraciones 400
245kb	11ms	10ms	10ms	10ms	9ms	10ms
398kb	11ms	10ms	9ms	10ms	9ms	9ms
736kb	11ms	10ms	10ms	10ms	10ms	9ms
984kb	10ms	13ms	10ms	10ms	10ms	9ms
1180kb	11ms	11ms	12ms	10ms	10ms	10ms
1716kb	11ms	11ms	10ms	11ms	11ms	10ms
11265kb	14ms	13ms	14ms	11ms	11ms	11ms

Tabla. 6. Test de performance sobre ejecución de reglas.

En la tabla 6 se detalla el resultado de las pruebas, se realizaron 6 ciclos de pruebas, en cada ciclo se ejecutaron distintas iteraciones secuenciales y con un aumento en la cantidad de ejecuciones que van desde las

50 iteraciones hasta 400 iteraciones. Las aplicaciones de negocios son complejas y la estructura Json enviada por un front-end a un servicio back-end puede variar dependiendo de esta complejidad, por este motivo se toma el peso de la estructura Json (atributos y valores) como un factor importante para entender cómo puede degradar la performance. El rango de los distintos pesos de las estructuras Json que pueden llegar al motor de reglas va desde 245kb hasta 11256kb para este modelo de aplicación. Luego de las pruebas se puede concluir que el peso no es un factor de degradación para el motor de reglas propuesto, el tiempo de ejecución medio de un paquete de reglas se encuentra entre los 9ms y los 14ms.

## VI. Conclusión y trabajos futuros

En este primer trabajo sobre motor de reglas desacoplado orientado a servicios con formato de intercambio de información Json, se presenta un modelo independiente del lenguaje de programación de clientes que lo implementa, no es necesario embeber el motor en un programa o modulo existente, no orientado a objetos, de fácil parametrización e instalación. El principal aporte de este modelo es un motor de reglas simple, liviano e independiente que tiene como objetivos maximizar los cambios dinámicos del negocio, acercar a usuarios del negocio una idea que los ayude a autogestionarse y disminuir el esfuerzo-tiempo de programación y mantenimiento de los sistemas Core. La evolución de este trabajo está orientado a usuarios del negocio, que puedan gestionar reglas de forma sencilla y práctica. Desde el punto de vista de la complejidad, mejorar el desarrollo de reglas agregando rangos numéricos para contemplar por ejemplo en una condición valor entre x e y, misma característica para las fechas. Para los casos de valores constantes, la alternativa de invocar a un servicio que retorne este valor. Agregar el versionado de reglas para registrar la evolución en el tiempo de los cambios en el negocio.

## Referencias

- [01] Bajec, Marko & Krisper, Marjan. (2005). Issues and Challenges in Business Rule-Based Information Systems Development.. 887-898.
- [02] Business Rule Group. <http://www.businessrulesgroup.org/theBRG.htm> [Acceso 01/02/2020].
- [03] FICO® Blaze Advisor® <https://www.fico.com/es/products/fico-blaze-advisor-decision-rules-management-system> [Acceso 06/04/2020].
- [04] Jena Apache para Java. <http://jena.sourceforge.net/> [Acceso 05/02/2020].
- [05] SPARQL Query Language for RDF <https://www.w3.org/TR/rdf-sparql-query/> [Acceso 05/02/2020].
- [06] Drools. <https://www.drools.org/> [Acceso 24/02/2020].
- [07] Clips C Language Integrated Production System <http://www.clipsrules.net/> [Acceso 20/02/2020].
- [08] Oracle Policy Automation. <https://www.oracle.com/applications/customer-experience/service/intelligent-advisor/policy-automation.html> [Acceso 03/01/2020].
- [09] React <https://reactjs.org/> [Acceso 10/10/2019].
- [10] Redux <https://redux.js.org/> [Acceso 15/10/2019].
- [11] NPM <https://www.npmjs.com/> [Acceso 11/10/2019].
- [12] Postgresql <https://www.postgresql.org/> [Acceso 04/11/2019].
- [13] Nodejs <https://nodejs.org/es/> [Acceso 10/10/2019].
- [14] Json Javascript Object Notation [https://www.w3schools.com/Js/js\\_json\\_intro.asp](https://www.w3schools.com/Js/js_json_intro.asp) [Acceso 28/06/2019].
- [15] Ross, R. G. The Business Rule Approach. Computer 36 (2003), 85–87.
- [16] Servicio Api Rest <https://docs.microsoft.com/en-us/azure/architecture/best-practices/api-design> [Acceso 08/07/2019].
- [17] A. Neumann, N. Laranjeiro and J. Bernardino, "An Analysis of Public REST Web Service APIs," in IEEE Transactions on Services Computing, doi: 10.1109/TSC.2018.2847344.
- [18] Rodríguez Carlos et al., "REST APIs: A Large-Scale Analysis of Compliance with Principles and Best Practices", International Conference on Web Engineering, 2016.
- [19] Json Syntax [https://www.w3schools.com/js/js\\_json\\_syntax.asp](https://www.w3schools.com/js/js_json_syntax.asp) [Acceso 28/06/2019].
- [20] Giarratano, J., & Riley, G. (1998). Expert systems: Principles and programming (3rd ed.). Brooks/Cole Publishing Co. Pacific Grove, CA, USA.
- [21] Gamma, E., Helm, R., Johnson, R., Vlissides, J. M. (1994). Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Professional. ISBN: 0201633612

## Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística.

Carlos Barrionuevo, Jorge Ierache , Iris Sattolo

Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica. SECYT  
Escuela Superior de Ingeniería, Informática y Ciencias Agroalimentarias  
Universidad de Morón Cabildo 134, Buenos Aires, Argentina  
{cbarionuevo,jierache,isattolo}@unimoron.edu.ar

**Abstract.** A través de las expresiones faciales se transmite más de la mitad del significado de un mensaje. Diferentes estudios, han demostrado además que algunas de estas tienen el carácter de universales. Sin embargo, una de las características más interesantes de las expresiones faciales es que revelan emociones. Esta cualidad, ha hecho que diversas disciplinas se hayan volcado a su estudio con diferentes objetivos, la computación no fue la excepción. La detección de rostros y sus partes principales, ha sido uno de los grandes avances en el área de la visión por computadora. Esto, sumado al auge de la última década del aprendizaje automático, posibilitó el desarrollo de sistemas capaces de detectar emociones a través del análisis de expresiones faciales. En este trabajo describiremos las diferentes etapas del desarrollo, entrenamiento y prueba de un algoritmo de regresión logística para la detección de emociones.

**Keywords:** Expresiones faciales, emociones, visión por computadora, aprendizaje automático, regresión logística.

### 1 Introducción

A partir del nuevo siglo, mejorar la interacción entre el ser humano y las máquinas se ha convertido en uno de los principales objetivos de la computación. Con este fin, diferentes líneas de investigación se centran en crear nuevas formas de interactuar. La computación afectiva [1] consiste en brindarles a las computadoras la capacidad de interpretar el estado emocional del usuario. El análisis de las expresiones faciales constituye una de las formas más eficaces para revelar el estado emocional de un individuo. 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]. El presente trabajo tuvo como objetivo describir el proceso de construcción de un sistema de detección de emociones capaz de reconocer las 7 expresiones emocionales catalogadas como universales por la teoría de Paul Ekman. En la sección 2 se describe la base de datos de imágenes de rostros utilizada, y se detalla el proceso de extracción de las características necesarias para entrenar un algoritmo de aprendizaje automático.

En la sección 3 se presenta el algoritmo utilizado mientras que en la sección 4 se muestran los experimentos realizados con el algoritmo implementado en los que se compara su rendimiento con otro servicio de similares características, y por último en la sección 5 se brindan las conclusiones y futuras líneas de trabajo.

## 2 Enfoque categórico- Base de datos de imágenes de rostros

En el marco del enfoque categórico, Paul Ekman plantea en su teoría la existencia de seis expresiones faciales universales que trascienden el idioma y las diferencias regionales, culturales y étnicas; a las que relaciona con seis emociones basales: enojo, asco, felicidad, miedo, tristeza y sorpresa (en inglés “anger”, “disgust”, “fear”, “happiness”, “sadness” y “surprise”). Posteriormente en su trabajo [3] Ekman adiciona una séptima expresión facial que representa la emoción “desprecio” (en inglés “contempt”). En la figura 1a se muestran 7 imágenes representativas de las 7 expresiones faciales universales tomadas de la página web de Paul Ekman [4] con el titulado de la foto entre paréntesis Para la presente investigación se utilizó la base de datos de rostros “RaFD”[5] desarrollada por la Universidad de Radbound de la ciudad de Nijmegen (Holanda). Esta se compone de imágenes de rostros de 67 modelos en su mayoría de raza caucásica, de ambos sexos, adultos y niños. De acuerdo con el Sistema de codificación facial [6] (Facial Action Coding System, “FACS”) desarrollado por Paul Ekman, cada modelo aparece en el set fotografiado desde cinco ángulos distintos (0°,45°,90°,135° y 180°) y las imágenes frontales (90°) tomadas con el sujeto dirigiendo la mirada hacia 3 direcciones diferentes (izquierda, al frente y hacia la derecha). Este set considera las 7 expresiones faciales universales definidas anteriormente y adiciona una octava expresión catalogada como “neutral”. Todas estas características lo hacen apto para múltiples campos de investigación tales como señales faciales de atención y procesamiento de expresiones faciales. Si bien existen otras bases de datos de rostros tales como “Jaffe” [7] también conocida como “la base de imágenes de mujeres japonesas” y el “Cohn-Kanade dataset” [8], se eligió “RaFD” por las siguientes razones: a) Alta Calidad y resolución de las imágenes (1024 x 681 píxeles). Las sesiones de fotos se desarrollaron en un ambiente altamente controlado, con condiciones óptimas de luz. b) Variedad de características de los modelos tales como raza, género y edades, c) Cada una de las imágenes se encuentra debidamente etiquetada con la expresión facial emocional mostrada en ella a través del nombre del archivo, d) Sesión de fotos dirigida y asistida por especialistas certificados en el Sistema de codificación facial. Si bien esta característica es compartida entre diversos sets de imágenes, no es menor y aporta confiabilidad para la investigación. Dentro del nombre del archivo de cada una de las imágenes se puede encontrar el nombre de la base de datos, el ángulo de la toma, el número de modelo, la raza, el género, la emoción mostrada y la dirección de la mirada. Por ejemplo, si el nombre de la imagen es el siguiente “Rafd090\_07\_Caucasian\_male\_sad\_left”, indica que la imagen fue tomada a 90 grados (de frente), el número de modelo es el “07”, la raza del mismo es caucásica, el género es masculino, la emoción mostrada es “tristeza” (en idioma inglés “sad”) y la mirada del sujeto apunta hacia la izquierda (en idioma inglés “left”).

Como el set incluye imágenes tomadas desde cinco ángulos diferentes solo se utilizaron las tomadas a noventa grados (de frente) obteniendo un total de 460 imágenes. De este subconjunto se reservó un 90% de las mismas (414 imágenes) para el entrenamiento [9] del algoritmo de regresión logística que se aplicó en el reconocimiento de emociones. El restante 10% (46 imágenes) se guardó para la realización de pruebas y experimentos de predicción [9]. Esta última parte se eligió de forma tal que se eliminó por completo del conjunto de imágenes que se utilizaron para entrenamiento a los modelos número 16,18,23,24,25,30,32,36,38 y 54.

Presentada la base de datos de rostros con la que se trabajó, el siguiente paso fue desarrollar una rutina que posibilite la extracción de las características necesarias de cada imagen del set. El lenguaje de programación elegido fue “Python”, ya que cuenta con la biblioteca “dlib” [10] que proporciona las herramientas necesarias para el desarrollo. La primera tarea, consistió en obtener del nombre del archivo la etiqueta de la emoción asociada al mismo. En base a esta se asoció un código numérico que consistió en un valor entre uno y ocho. De esta manera, se representó de forma unívoca cada emoción. En la tabla 1 se detalla la codificación de emociones utilizada:

**Tabla 1.** Codificación de emociones.

Etiqueta	Traducción	Código
Angry	Enojo	1
Contemptuous	Desprecio	2
Disgusted	Asco	3
Fearful	Miedo	4
Happy	Feliz	5
Neutral	Neutral	6
Sad	Tristeza	7
Surprise	Sorpresa	8

Como resultado de esta tarea se obtuvo un vector identificado con la letra “Y” de “K” dimensiones, donde “K” es el número de imágenes del set que se utilizaron para entrenar el algoritmo de regresión logística. El vector final, fue almacenado en un archivo con extensión “.txt”. El siguiente paso fue para cada una de las imágenes, detectar la región del rostro (también conocida como “región de interés”). Esto permitió poder identificar sobre esta una serie de puntos característicos (también conocidos como “landmarks”). Existen una gran variedad de detectores de puntos característicos del rostro que varían (entre otras cosas) en la cantidad de “landmarks” que identifican; pero coinciden generalmente en la localización de las regiones de la boca, cejas, nariz, ojos y contorno de la cara. Para este trabajo, se utilizó el detector incluido en la biblioteca “dlib” [10]. El mismo es una implementación del algoritmo desarrollado por Vahid Kazemi y Josephine Sullivan [11]. Este permite obtener las coordenadas en los ejes “x” e “y” de 68 puntos característicos de la cara. La localización de estos se puede observar en la figura 1b [12]. Una vez localizados los landmarks, considerando los pares (x,y) de cada uno se totalizan 136 elementos con los cuales se generó una fila por imagen. De esta forma las características faciales de cada una de las imágenes se vieron representadas por su correspondiente tupla. De forma general, el resultado de este proceso es una matriz de “m” filas y “n” columnas. Siendo “m” el número de ejemplos de entrenamiento y “n” el número de

características. Para el presente caso de estudio, “m” fue igual a 414 (número de imágenes que se utilizó para entrenamiento del algoritmo) y “n” como se acaba de mencionar a 136. Esta matriz se identificó con la letra “X” y se conoce como “matriz de características”. Al igual que el vector “Y” se almacenó en un archivo con extensión “txt”. Una vez descritas ambas matrices, resulta importante mencionar que ambas fueron generadas en simultaneo, por lo que la fila “i” de la matriz “X” se correspondió con el elemento “i” del vector “Y”. Finalizado este proceso, ya se cuenta con la entrada necesaria para el entrenamiento del algoritmo de aprendizaje supervisado.



Fig. 1a Siete Expresiones faciales universales según la teoría de Paul Ekman.

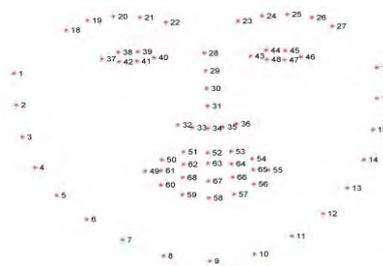


Fig. 1b Distribución de landmarks en el rostro (fuente pyimagesearch).

### 3 Predicción de emociones con el empleo de Regresión Logística.

Dada una nueva imagen del rostro de una persona, para predecir cual es la emoción que está expresando, una de las posibilidades es utilizar un algoritmo de aprendizaje supervisado. Se le dice “supervisado” ya que para su entrenamiento necesita de un conjunto de datos previamente etiquetado y clasificado. Sobre este grupo de datos conocido como “conjunto de datos de entrenamiento” el algoritmo realizará predicciones y las comparará con las etiquetas, con el error obtenido y a través de sucesivas iteraciones irá ajustando el modelo logrando así un aprendizaje progresivo. Existen una gran variedad de algoritmos de estas características tales como la regresión lineal, regresión logística, redes neuronales, máquina de soporte de vectores, K vecinos más próximos, etc [13]. La utilización de uno u otro generalmente depende de las dimensiones del problema [14]. Para este trabajo se optó por un algoritmo de regresión logística [15]. La implementación del mismo se llevó a cabo en el lenguaje de programación Octave [16]. Este algoritmo de clasificación nos permite a través de un clasificador estimar la probabilidad de que un nuevo ejemplo pertenezca a una clase. Como para el presente problema se tuvo un total de ocho clases, se aplicó además una técnica denominada “one vs all” [14] (uno contra todos). De esta forma, con la ayuda de este método fue posible entrenar ocho clasificadores. Uno por cada una de las emociones consideradas. Cada clasificador se ve representado por su correspondiente función hipótesis. Primero, para simplificar la explicación definiremos la función polinómica “z” (que forma parte de la hipótesis) según la propuesta de Andrew Ng [14]:

$$z(x) = \theta^T \cdot X \tag{1}$$

Donde:

- “ $\theta$ ” es una matriz de “ $i$ ” filas y “ $j$ ” columnas. Siendo “ $i$ ” el número de clases y “ $j$ ” el número de parámetros del polinomio. Para el caso en estudio, “ $i$ ” será igual a ocho (número de emociones) y “ $j$ ” equivalente al número de características (136) que conforman los “landmarks”. Inicialmente, todos los elementos de esta matriz fueron iguales a 0 (cero). Estos son los parámetros que se ajustaron a través del entrenamiento para luego poder realizar predicciones.
- “ $X$ ” es la matriz de características definida en el apartado anterior.

Definida la función “ $z$ ”, la expresión de la función hipótesis para regresión logística [14] será la siguiente:

$$h_{\theta}^i(x) = \frac{1}{(1 + e^{-z})} \quad (2)$$

Donde:

- “ $i$ ” se refiere a la  $i$ -ésima clase.

La función anterior consiste en aplicar la función sigmoide a “ $z(x)$ ”. Esto produce que el resultado de esta sea un valor entre 0 y 1. Este representa la probabilidad de que un nuevo ejemplo de entrenamiento (para este caso las características de un nuevo rostro) pertenezca a la clase (emoción) “ $i$ ”. La etiqueta de la emoción del clasificador que arroje la mayor probabilidad será la salida final del algoritmo.

Por otra parte, Se denomina “costo” a la penalización que pagará el algoritmo dado un valor de probabilidad calculado por la función hipótesis en caso de que la etiqueta sea “ $y$ ”. El costo para regresión logística se define a través de la siguiente expresión [14]:

$$\text{costo}(h_{\theta}(x), y) = -y \cdot \log(h_{\theta}(x)) - (1 - y) \cdot \log(1 - h_{\theta}(x)) \quad (3)$$

Si  $y = 1$  (estimamos la probabilidad de que pertenezca a esa clase), la representación del costo se observa en la figura número 2. Por ejemplo, si dada una nueva imagen que muestra la emoción “enojo” ( $y=1$ ) pueden darse las siguientes situaciones:

- Si el valor calculado por la función hipótesis  $h(x)$  asociada al clasificador de la emoción “enojo” calcula un valor de 1 (100% de probabilidad de que la emoción mostrada sea “enojo”), el costo que pagará el algoritmo será 0.
- Si el valor calculado por la función hipótesis  $h(x)$  asociada al clasificador de la emoción “enojo” tiende a 0 (0% de probabilidad de que la emoción mostrada en la imagen sea “enojo”) el costo que pagará el algoritmo tenderá a infinito.

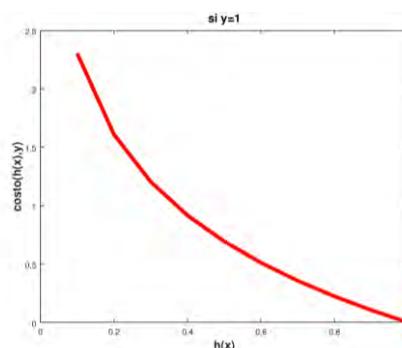
Definido el costo, la expresión de la función de costos es la siguiente [14]:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))] \quad (4)$$

Donde:

- “ $m$ ” es igual al número de ejemplos de entrenamiento. (en nuestro caso 414 instancias de rostros)

Como es posible observar, en la expresión (4), la función de costos calcula el error promedio entre todos los ejemplos del set de entrenamiento.



**Fig. 2** Representación del costo.

Respecto de la matriz " $\theta$ ", tal como se mencionó antes, inicialmente todos sus elementos serán iguales a 0(cero). El objetivo es entonces encontrar aquellos parámetros " $\theta$ " que minimicen el error de la función de costos. Para esta tarea existen diferentes técnicas, tales como el descenso de gradiente cuya implementación fue considerada para este trabajo, pero finalmente se optó por una función de optimización avanzada provista por "Octave" denominada "fminunc" ("function minimization unconstrained") [17]. "fminunc" recibe como entradas un puntero a una función " $f$ ", una matriz de parámetros " $x$ " y un vector de opciones clave-valor adicionales entre las que es importante mencionar el número máximo de iteraciones (clave "MaxIter"). Esta función retorna la matriz de parámetros " $x$ " de forma tal que " $f(x)$ " es un mínimo local [18]. Para el presente trabajo se parametrizó dicha función con: a) un puntero a la función costos, b) la matriz de parámetros " $\theta$ " y c) la clave "MaxIter" definida con el valor 1200 que indica la cantidad de veces que se iteró sobre el set de datos buscando converger en un mínimo local de la función de costos. La salida de esta función es la matriz de parámetros ajustada en base a nuestros datos de entrenamiento.

#### 4 Experimentos y resultados obtenidos

El set de pruebas que se utilizó, como se mencionó al momento de describir la base de datos de rostros, se compone de 46 imágenes, las cuales corresponden a los sujetos (modelos) 16,18,23,24,25,30,32,36,38 y 54. En la tabla 2 se muestra la descripción del set de pruebas:

**Tabla 2.** Descripción del set de pruebas.

Código	Descripción	Cantidad de imágenes	Modelos utilizados
1	Enojo (angry)	5	16,18,30,32 y 36.
2	Desprecio (contemptous)	7	16(2),18,23,25,30 y 32.
3	Asco (disgusted)	7	18(2),25,30(2),32 y 38.
4	Miedo (fearful)	8	16(2),18,25(3),30 y 32.
5	Felicidad (happy)	4	24,25,30 y 54.
6	Neutral (neutral)	4	24,25,30 y 32.
7	Tristeza (sad)	5	16,18,25 y 30(2).
8	Sorpresa (surprised)	6	16,25,30,32,36 y 38.

En la tabla 2 en la primera columna se indica el código de emoción utilizado en el presente trabajo (de acuerdo con lo definido en la tabla 1), en la segunda columna la descripción de dichos códigos (nombre de la emoción), en la tercer columna se encuentra la cantidad de imágenes presentes en el set de pruebas relacionadas a cada emoción y en la última columna “Modelos utilizados” se indica el número de modelo y entre paréntesis la cantidad de imágenes del mismo (si fuera más de una). Al igual que en la etapa de entrenamiento, fue necesario generar la matriz de características para el set de datos de pruebas. La misma fue de 46 filas (número de imágenes) y 136 columnas (cantidad de características). La matriz resultado, fue almacenada también en un archivo con extensión ".txt". Esta tarea es realizada por una rutina desarrollada en lenguaje Python, como se indicó anteriormente. Una vez obtenida la matriz "X", el siguiente paso fue evaluar las 136 características de cada una de las imágenes en las funciones hipótesis correspondientes a los 8 clasificadores emocionales. El resultado de esta operación fue una matriz de 8 filas y 46 columnas. La cantidad de filas se corresponde con la cantidad de clasificadores (de las distintas emociones), el número de fila indica el código de emoción de acuerdo con lo definido en la tabla 1. De esta forma, en la fila 1 se encontrarán las salidas del clasificador asociado a la emoción con código 1 (Enojo). La cantidad de columnas por su parte es igual al número de imágenes del set de pruebas. Así, por ejemplo, en la columna 1 se obtuvieron las salidas de los 8 clasificadores para la primera imagen. La salida final del algoritmo es para cada imagen el número de fila (que como acabamos de mencionar se corresponde con el número de clasificador emocional) para el que se calculó el máximo valor de probabilidad. Los resultados obtenidos posibilitaron la construcción de la matriz de confusión [19] de la figura 3:

		Esperado								Precisión
		1	2	3	4	5	6	7	8	
Predicido	1	5	0	0	0	0	1	1	0	0,83
	2	0	4	0	0	0	1	0	0	0,8
	3	0	0	7	0	0	0	0	0	1
	4	0	0	0	7	0	0	0	0	1
	5	0	0	0	0	4	0	0	0	1
	6	0	1	0	0	0	2	0	0	0,67
	7	0	2	0	1	0	0	4	0	0,57
	8	0	0	0	0	0	0	0	6	1
Recall		1	0,57	1	0,88	1	0,5	0,8	1	85%

Fig. 3 Matriz de confusión construida con los resultados obtenidos para la prueba con 46 imágenes.

De la matriz presentada en la figura 3 y en función de la cantidad de imágenes por emoción indicadas en la tabla 2, es posible deducir que el sistema desarrollado en el presente trabajo falló en: a) 3 imágenes donde el resultado esperado era “desprecio” (código 2) y el valor predicho fue en 2 ocasiones la emoción “tristeza” (código 7) y en la restante la emoción “neutral” (código 6), acertando en las cuatro restantes imágenes correspondiente a desprecio, para un total de 7 imágenes (desprecio); b) en una imagen donde el resultado esperado era “miedo” (código 4) el valor predicho fue “tristeza” (código 7) para un total de 8 imágenes (miedo); c) en 2 imágenes en las cuales se esperaba como resultado “neutral” (código 6) el algoritmo las confundió con las emociones “enojo” (código 6) y “desprecio” (código 2), para un total de 4 imágenes (Neutral); d) en una imagen en donde el resultado esperado era “tristeza”

(código 7) el valor predicho fue “enojo” (código 1) para un total de 5 imágenes de tristeza . Totalizando 7 predicciones erróneas y 39 predicciones correctas. Además, con las predicciones realizadas se calcularon 3 métricas diferentes que nos permitieron evaluar el rendimiento del modelo predictivo propuesto (algoritmo de regresión logística). La primera de ellas se denomina “Recall” [20] y mide para cada una de las clases de emociones que fracción del total de imágenes disponibles para esa clase fueron correctamente predichas. Tomando como ejemplo la emoción “Desprecio” y observando la columna relacionada a la misma en la figura 3, notamos que de 7 imágenes disponibles para dicha emoción el algoritmo acertó en la predicción de 4, obteniendo un “recall” de 0,57 (57%) para dicha emoción. La segunda métrica que se calculó fue la “Precisión”, esta indica para cada uno de los 8 clasificadores emocionales, de la cantidad de predicciones realizadas por estos cuantas fueron correctas. Observando por ejemplo la fila relacionada a la emoción “Tristeza” (código 7) en la figura 4, se nota que de 7 ocasiones en las que el valor predicho fue esta emoción, se acertó en 4, obteniendo una precisión de 0,57 (57%). Finalmente, la última métrica que se calculó fue la llamada “Accuracy” [20] (exactitud). Esta a diferencia de las dos métricas anteriores mide el desempeño global del algoritmo y no el particular de cada emoción. Se calcula como el total de predicciones realizadas correctamente sobre el total de predicciones realizadas. Como se comentó anteriormente el total de predicciones correctas en esta prueba fue de 39 sobre 46 predicciones totales, obteniendo un “accuracy” del 85%. Realizados los testeos con el set de datos, se llevó a cabo además una prueba independiente con imágenes externas al mismo, en este caso fueron seleccionadas las imágenes presentadas en la figura 1a que como se mencionó antes, fueron tomadas de la página web de Paul Ekman [4] y consideran distintos sujetos para la representación de las distintas emociones. Para esta prueba, las imágenes fueron clasificadas con el sistema propuesto y se realizó una comparación con el servicio de inteligencia artificial “Face” de Microsoft [21], con independencia de la cantidad de landmarks de cada modelo, el cual también considera las mismas ocho emociones planteadas para este trabajo. Los resultados obtenidos se muestran en la figura 4, donde es posible observar los valores de probabilidad calculados por los 8 clasificadores emocionales de ambos sistemas para cada una de las imágenes. se puede observar que el sistema de Microsoft Face falló en la predicción de la emoción “desprecio” mientras que el modelo predictivo desarrollado lo hizo en las emociones “enojo” y “miedo”. En la figura 4, en la primera fila se muestran los títulos de las imágenes y entre paréntesis se aclara el código de emoción asociado al resultado esperado. Para cada imagen de la primera fila, se detallan los resultados obtenidos por el servicio “Face” y por el sistema desarrollado propuesto en este artículo (S.P.). La primera columna por su parte muestra el código de emoción seguido por el nombre de esta presentado en la tabla 1 que fue con el que se trabajó en este artículo. La salida final del algoritmo de regresión logística que se construyó en este trabajo fue la que se muestra en la figura 5. Como segunda prueba independiente, se tomó uno de los sujetos del set de pruebas inicial compuestos por 46 fotografías de la base de datos de rostros “RaFD”[5]. El modelo elegido fue el número 30 [9], ya que se contaba con al menos una imagen de cada emoción (8 imágenes en total), y se realizó nuevamente una comparación con el servicio de Microsoft. Los resultados obtenidos se muestran en la figura 6.

	(1) anger		(2) contempt		(3) disgust		(4) fear		(5) happiness		(7) sadness		(8) surprise	
	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.
1 - Enojo	0,472	0	0	0	0,006	0	0	0	0	0	0	0	0	0
2 - Desprecio	0,006	0	0,213	0,01	0	0	0	0	0	0,012	0	0,018	0	0
3 - Asco	0,001	0	0	0	0,993	0,043	0	0	0	0	0	0	0	0,02
4 - Miedo	0,038	0	0	0	0	0	0,991	0,153	0	0	0	0,001	0,006	0
5 - Feliz	0	0	0,004	0	0	0	0	0,001	1	0,465	0	0	0	0
6 - Neutral	0,089	0,056	0,772	0,007	0	0	0	0,071	0	0	0,124	0,005	0,017	0
7 - Tristeza	0,001	0,002	0,011	0	0	0	0,001	0,916	0	0	0,876	0,762	0	0
8 - Sorpresa	0,392	0	0	0	0	0,002	0,008	0	0	0	0	0	0,977	0,124

Fig. 4 Resultados obtenidos entre el sistema propuesto y el servicio “Face” de Microsoft.

Resultado:

Ordinal (1) Resultado: 6 - Neutral  
 Ordinal (2) Resultado: 2 - Desprecio  
 Ordinal (3) Resultado: 3 - Asco  
 Ordinal (4) Resultado: 7 - Tristeza  
 Ordinal (5) Resultado: 5 - Felicidad  
 Ordinal (6) Resultado: 7 - Tristeza  
 Ordinal (7) Resultado: 8 - Sorpresa

Fig. 5 Salida final del sistema propuesto para la prueba con el set de imágenes de la página web de Paul Ekman.

	(1) angry		(2) contemptous		(3) disgusted		(4) fearful		(5) happy		(6) neutral		(7) sad		(8) surprised	
	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.	Face	S.P.
1 - Enojo	0,291	1	0	0	0,66	0,005	0	0	0	0	0,074	0,006	0	0	0	0
2 - Desprecio	0,004	0,931	0,002	0,046	0,001	0	0,014	0,001	0	0	0,05	0,094	0	0	0	0
3 - Asco	0	0	0	0	0,339	0,999	0,225	0	0	0,085	0	0,013	0	0	0	0
4 - Miedo	0	0	0	0	0	0	0,679	1	0	0	0	0,016	0,354	0	0	0
5 - Feliz	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
6 - Neutral	0,704	0	0,998	0,957	0	0,004	0	0	0	0	1	0,993	0	0,1	0	0
7 - Tristeza	0	0	0	0	0	0	0,007	0,713	0	0	0	0	0,644	0,998	0	0,137
8 - Sorpresa	0,001	0	0	0	0	0	0,074	0,002	0	0	0	0,026	0	1	1	0

Fig 6 Resultados obtenidos de la comparación entre el sistema propuesto y el servicio “Face” de Microsoft para la prueba con el sujeto número 30.

Es posible observar en la figura 6 que tanto el sistema desarrollado para el presente trabajo como el servicio de Microsoft fallaron en la imagen que expresaba “desprecio” calculando ambos sistemas el mayor valor de probabilidad en la emoción “neutral” (99,8% calculado por el servicio “Face” y 95,7% calculado por nuestro sistema). Además, el servicio de Microsoft erró en la predicción de la fotografía que expresaba “enojo” confundiéndola también con la expresión “neutral”.

## 5 Conclusión y futuras líneas de trabajo

De acuerdo con el análisis realizado sobre los resultados obtenidos, podemos concluir que el sistema construido presentó un desempeño correcto tanto en las pruebas con imágenes del set de entrenamiento como con imágenes externas al mismo. Existe un margen de mejora que podría darse entrenando al algoritmo con un número mayor de imágenes, con modelos con mayor diversidad de características, o modificando el proceso de extracción de características, obteniendo por ejemplo el grado de apertura de los ojos, boca, etc. En el marco de futuras líneas de trabajo se evaluará el rendimiento de otras técnicas de aprendizaje supervisado (tales como redes neuronales o máquina de soporte de vectores) en la tarea de detección de emociones y elaborar una comparación con la aplicación construida en la presente investigación. Además, se contemplará la integración en sistemas multimodales [22] [23] combinado

diferentes sensores (variación de ritmo cardíaco, conductancia de piel, eeg a través de interfaces cerebro-máquina), reforzando de esta forma la determinación categórica de emociones realizada a través de la detección de rostros. Esto permitirá la integración en diversos dominios de aplicación en el marco de la computación afectiva.

## 6 Referencias

1. Picard, R.W., et al.: Affective Computing (1995)
2. A. Mehrabian, "Communication without Words", *Psychology Today*, vol. 2, no 4, (68)
3. Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), 712–717.
4. Paul Ekman Group, <https://www.paulekman.com/>. Accedido en junio 2020.
5. Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, 24(8), 1377—1388. DOI: 10.1080/02699930903485076
6. Ekman, P., Friesen, W. V., & Hager, J. C. (2002a). *Facial Action Coding System: The manual*. Salt Lake City, UT: Research Nexus
7. Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998).
8. Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression.
9. Conjuntos de imágenes de rostros usados para entrenamiento y pruebas, <https://drive.google.com/drive/folders/1AhYfPoBoBp0oU4WBWgSmXeD8xTPKqXLJ?usp=sharing>.
10. Dlib, <https://pypi.org/project/dlib/>
11. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees, 2014
12. Facial landmarks with dlib, OpenCV, and Python. *Pyimagesearch* <https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/>
13. Tipos ML, <https://medium.com/soldai/tipos-de-aprendizaje-autom%C3%A1tico-6413e3c615e2>.
14. Andrew Ng. Machine Learning. Coursera. <https://www.coursera.org/learn/machine-learning>
15. La Regresión logística, <https://www.analyticslane.com/2018/07/23/la-regresion-logistica/>.
16. Octave, <https://www.gnu.org/software/octave/>.
17. fminunc, <https://octave.sourceforge.io/octave/function/fminunc>. Accedido en junio 2020.
18. Minimizers, <http://octave.org/doc/v4.4.1/Minimizers.html>. Accedido en septiembre 2020.
19. Multi-Class Metrics Made Simple, Part I: Precision and Recall. Towards data science. <https://towardsdatascience.com/multi-class-metrics-made-simple-part-i-precision-and-recall-9250280bdc2>. Accedido en junio 2020.
20. Confusion Matrix For Your Multi-Class Machine Learning Model. Towards data science. <https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-f9aa3bf7826>. Accedido en junio 2020.
21. Face, Microsoft. <https://azure.microsoft.com/es-mx/services/cognitive-services/face/>. Junio 2020.
22. Ierache, J., Nicolosi, R., Ponce, G., Cervino, C., & Eszter, E. (2018). Registro emocional de personas interactuando en contextos de entornos virtuales. XXIV CACIC 2018, (págs. 877-886).
23. 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, Libro de actas pp 417-426

## **Proposal of a Multimodal Model for emotional assessment within affective computing in gastronomic settings**

Jorge Ierache, Facundo Nervo, Iris Sattolo, Rocio Ierache, Gabriela Chapperón  
Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica  
(ISIER) Secretaria de Ciencia y Tecnología -Universidad de Morón  
Escuela Superior de Ingeniería, Informática y Ciencias Agroalimentarias  
Cabildo 134, Buenos Aires, Argentina  
jierache@unimoron.edu.ar

**Abstract.** This paper deals with the problem of emotions recognition in a multimodal way, presents a model of fusion of emotional records coming from EEG signals provided by the use of brain-machine interfaces, face image capture and the record of the response from the user before stimuli induced by images (IAPS, OLAF). It particularly approaches a proposal of multimodal fusion which validates the user's potential subjective responses facing those stimuli.

**Keywords:** Affective Computing, Food, Biometric Data, Physiological Data, Multimodal Framework, Neuromarketing.

### **1 Introduction**

This paper follows the research of I+D ISIER-UM, within the targeted research project submitted in the PICTO-UM-2019-00005. It is part of the institute of intelligent systems and experimental robotics teaching ISIER UM. It capitalizes the ground basis of the PING/17-03-JI-002 UM project, since, in the context of cognitive and affective computing, research using Brain-computer interface applied to Domotics and Robotics [1][2][3]; Influence on the biometric emotional state of people, EEG Data Exploitation and physiological parameters of users interacting in virtual settings [4][5]; Evaluation of the degree of users' attention in classroom settings [6] were conducted in this order.

R Picard [7] defines affective computing as "the IT related to emotions, not only to those considered as most important, such as joy or sadness, but also to those related to computers, such as interest, boredom or frustration". S. Baldassarri [8] suggests that "affective" systems must be able to capture and recognize the user's emotional states through readings of signals from the face, voice, body or any other reflex of the emotional process taking place. It would be logical to think that different biometric information has to be simultaneously recorded, creating a Multimodal (MM) context in order to deduce an individual's emotional state.

Some papers connected to this approach are presented in the Review Calvo R et al. They link up face and voice, speech and text by using SVM [9]. This paper makes a difference by adding up a model of fusion MM between the subjective assessment of the users' emotional experience records (SAM) and the objective records (EEG, face) by validating the users' responses to the emotional stimuli.

In the lines of I+D ISIER-UM, the conducted research focuses on passive EEG-BCI, which gather information about the user's state (from biomarkers) in order to assess/improve the interaction between men and virtual, real or mixed settings and the integration of physiological sensors (ECG, Heart rate, Conductance). In the research for emotional stimulation of the test subjects, a set of real pictures from the IAPS (International Affective Picture System) database was used [10],[11].

As for nutrition, the relationship between emotions and food intake is called "emotional eating". Our emotions have a powerful effect over our food choice and eating habits. It has been proposed that food assessments are not valid among individuals and groups unless feelings towards food signals are compared to feelings towards strong experiences unrelated to food, which are used as reference points [12]. In this research, the set of standardized food pictures OLAF (Open Library of Affective Foods) with their subjective assessments is used [13][14]. To record the degree of pleasure and its emotional intensity, SAM (Self-Assessment Manikin) surveys are used [15] based on the Russell test [16]. The emotion models are introduced in the second section; the problem and solution (system architecture and solution model) are introduced in the third section; the tests performed are introduced and the results obtained are discussed in the fourth section. Finally, conclusions and future lines of work are introduced in the fifth section.

## 2 Model Emotions

Over the last fifty years, different ways of modelling emotions represented from several perspectives were proposed. Two of them are widely used – the dimensional approach [17] and the categorical approach. *The categorical approach*, initially proposed by psychologist Paul Ekman [18], states that there is a set of six basic and universal emotions which are not determined by culture:

*"We experience emotions as we feel them not as we have chosen them"* – P. Ekman.

He then updated this set to seven emotions [19], to include contempt. So, these emotions are a) anger, b) disgust, c) happiness, d) fear, e) sadness, f) surprise, g) contempt

On the other hand, *the dimensional approach* indicates that affective states are distributed in a continuous space, whose dimensional axes indicate the quantification of a characteristic. Even though there is no current consensus about the definition of emotion, there is an agreement about emotions being classified in three continuous dimensions: 1) **Valence** (Assessment): a bipolar construct from pleasant to unpleasant, 2) **Arousal**: from calmed to activated (or aroused), and 3) **Dominance** (Control): value appearing with less internal consistency. One of the most accepted models is the Circumplex Model of Affectivity by J. Russell [18] [20], [6], known as

Arousal-Valence Circumplex or Model (bidimensional model), whose axes are Arousal (relaxed vs. aroused) and Valence (pleasure vs. disgust). It allows understanding how people control their emotions, and this activity is represented as the result of two strength vectors, which take the individual from an emotional situation to a more pleasant and better valued one. Russell represents emotions through a spatial model in which affective concepts are graphed in the following order, in which the number represents the vector or radial: pleasure (0), arousal (45), activation (90), anguish (135), disgust (180), depression (225), somnolence (270) and relaxation (315).

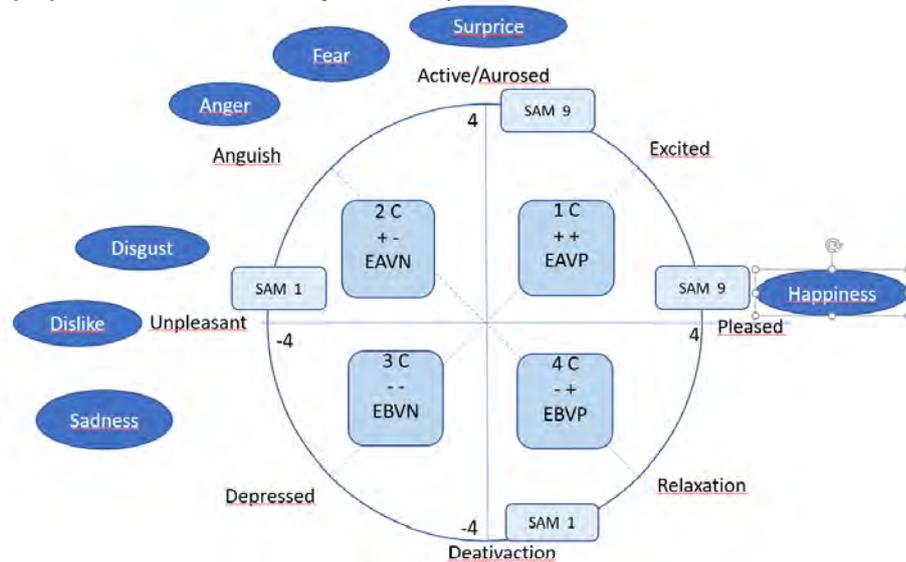
To create the emotional context, the OLAF picture set is used [13], [14]. This set specifically aims at studying emotions towards food and is the most used set of standardized stimuli in IAPS experimental research [11]. A set of over 1000 photos in color grouped in 20 sets of about 60 pictures which represent objects, people, landscapes and situations from daily life is used for emotional stimulation of the test subjects. In our case, it was used to get a reference value of the individual's emotional state. The decision to standardize IAPS pictures based on subject assessments in the affective dimensions of valence, activation and dominance is based on both theoretical and experimental data [11]. A tool used to capture emotions through surveys is SAM proposed by Lang in 1985. As described by its authors [15], SAM is a non-verbal, easy-to-apply method for swift assessment of pleasure (Valence), arousal and dominance associated with the emotional reaction of a person in a certain situation. SAM spans from a happy, smiling to a frowned, unhappy image when representing the pleasure dimension, and from an aroused, open-eyed to relaxed, sleepy image representing the arousal dimension. The dominance dimension represents control changes with SAM size variations, in which a big image indicates maximum situation control. Dimensional values provided by SAM do not need to be processed. There are two scales. One is in the 1–5 scale and the other scale allows adding four intermediate values to obtain a possible scale of 1–9; in this case, the image is not drawn nine times but in the possible middle point.

In the field of brain-computer interfaces, devices which translate neuronal information from specific brain areas to software or external hardware-controllable data have been developed [21]. BCI are often used as devices to assist people with motor or sensory disabilities, videogames, and robot control among others. Specifically, for this research work, BCI Emotiv EPOC [22] is used [25], which includes an Affective Suite, allowing visualization of emotion metadata: Engagement/Boredom, Frustration, Meditation, Instantaneous Excitement, Long-Term Excitement. Regarding raw EEG signal values, (F3, F4, AF3, AF4) are considered. To calculate the arousal and valence values, the algorithm proposed by R. Ramirez and Z. Vamvakousis [23] is used, exploiting the application of Rossi F. et al [24]. Equation 1 represents the arousal algorithm and equation 2 represents the valence algorithm:

$$Arousal = \frac{\beta AF_4 + \beta AF_3 + \beta F_4 + \beta F_3}{\alpha AF_4 + \alpha AF_3 + \alpha F_4 + \alpha F_3} \quad (1)$$

$$Valence = \frac{\beta AF_4 + \beta F_4}{\alpha AF_4 + \alpha F_4} - \frac{\beta AF_3 + \beta F_3}{\alpha AF_3 + \alpha F_3} \quad (2)$$

A necessary transformation is added to Russell's proposal for the integration of the different approaches (see Figure 1). Four areas deriving from the combination of high arousal (HA) or low arousal (LA) and positive valence (PV) or negative valence (NV), of SAM scales and MDET (Transformed BCI Emotiv Metadata) are identified. The proposed table 1 will be analyzed in the performed tests section.



**Fig.1** Russell's model with Ekman's 7 basic emotions, SAM, transformed BCI (Emotiv) Metadata proposal.

### 3 Problem

As mentioned in section 2, there are two basic approaches in terms of emotional models; however, there are no detailed proposals allowing their association in a multimodal context in order to achieve an emotional assessment within affective computing. Specially, gastronomic settings are chosen to assess the emotional response of individuals so that an emotional profile and its corresponding comparison for different dishes can be developed. A multimodal approach of the consumer's biometric-emotional information is proposed by using a Brain-Computer interface (BCI) together with a facial capture of the consumer during the observation of a dish picture.

#### 3.1 Solution description

A system architecture [25] whose main entries are both brainwave captures (by using BCI) and dish pictures and user's face captures is designed. This information is kept in a specific app database, which identifies each moment with the synchronized "time" and works as a unique identification value. The app allows labelling wave

levels for later analysis and placing the moments in which each of the waves or levels is higher or lower than specific configurable signal values or thresholds. Hence, the biosignal capturing app has an interface which manages necessary data for EEG data acquisition from BCI EMOTIV [19] and includes local variables where each of them is saved. It also saves time and date of the specific execution for integrating them with dish pictures and individual's face capturing app data. The picture capturing system was used in C# language, and records sections, identifiers, date, picture location and description in the database. The data structure is the following: Section, Identifier, Date, Picture, Description. The database is used in the SQL Server and includes the BCI EMOTIV metadata, face pictures, and the IAPS and OLAF picture sequence used during the test. Power BI [26] and OpenViBE [27] are used for BCI's EEG signal exploitation. In particular, the developed MM system allows us to apply filters in BD. In the annex "Concept Model of the Multimodal architecture of emotional assessment in affective computing used applied to gastronomic settings" [25], a concept diagram of the system architecture is included.

**Experimental model of Multimodal Fusion:** To perform the test and analyze the obtained results, IAPS and OLAF pictures (stimulus) have been previously selected and will be shown to the test subject during the session. During the session, the face and EEG information from the brain-computer EMOTIV interface are constantly captured. So, the average Emotiv metadata values (arousal and meditation or relaxation) are obtained during the display of each food picture from the OLAF set. In post-processing, the **dimensional approach** is first used to determine the arousal and valence values from the raw EEG data (F3, F4, AF3, AF4); then, 128 values, which are filtered with OpenViBE [27], are obtained. These arousal and valence values are consistent with Russell's dimensional model [20]. Then, the **categorical approach** proposed by P.Ekman [18] is used to determine the emotional state from the test subject's face capture. To determine emotions, MS Face [28] app was used. This app detects one or several human faces with attributes such as age, emotion, gender, position, smile and facial hair, including 27 reference points for each face of the picture. In particular, the emotions consistent with the categorical approach are "Anger", "Contempt", "Disgust", "Fear", "Happiness", "Neutral (calm)", "Sadness", and "Surprise". To perform the multimodal fusion, face capture, BCI Emotiv's EEG data, maximum value determined by affective suite, which we will call MDEmv, and the SAM survey results are combined with the dimensional model of emotions. Emotions are linked to the affiliation quadrant within Russell's circumplex model. Table 1. The assessment of SAM survey with or without pictures is converted into the dimensional model for which SAM values corresponding to High Arousal / High Valence (HA/HV) in the 6-9 range are categorized with 1-4 values. SAM survey values in the 1-4 range corresponding to Low Arousal / Negative Valence (LA/NV) are categorized with -1 to -4 values. The 5 value in the SAM survey for Arousal/Valence is considered zero, represented by the origin point of Russell's circumplex quadrants. For the association of BCI Emotiv Epoc metadata, the emotions from the affective suite -arousal, engagement, boredom, frustration, meditation, which we particularly name Emotional State Metadata (ESM) within the 0-1 range, are considered. For this, the highest value obtained which we will call MDEmv is considered.

**Table 1** Multimodal Fusion (Face Capture, EEG) and Approach Integration to the dimensional model of emotions

Emotion Ekman	Face Capture	Fusion Proposal to Russell's Dimensional Circumplex Approach. Emotion and Associated Value (Fig 2)								
		Emotion	MDEmv	Quadrant	Arousal			Valence		
					Value	SAM	EEG	Value	SAM	EEG
Anger	Anger	Rage	Frustration	+ - 2C	EA	+1 +4	+	VN	-1 -4	-
Contempt	Contempt	Anger/ Disgust	Frustration	+ - 2C	EA	+1 +4	+	VN	-1 -4	-
Disgust	Disgust	Disgust	Frustration	+ - 2C	EA	+1 +4	+	VN	-1 -4	-
Fear	Fear	Fear	Frustration	+ - 2C	EA	+1 +4	+	VN	-1 -4	-
Joy	Happiness	Joy	Arousal	++ 1C	EA	+1 +4	+	VP	1 4	+
Sec. emotion	Neutral	Relaxed /calm	Meditation/pleasant	-+ 4C	EB	0 -4	-	VP	0 4	+
Sadness	Neutral Sadness	Sadness	Boredom	-- 3C	EB	-1 -4	-	VN	-1 -4	-
Surprise	Surprise	Surprise	Arousal	++ 1C	EA	+1 +4	+	VP	+1 +4	+

#### 4 General case of Tests and Obtained Results

The experimental protocol sequence is made up by the following phases, which allow the proposed model validation.:

**First Phase:** Test subject is shown a black screen picture (stimulus) for 10 seconds to assess the initial emotional state of the test subject. The resulting recorded values are shown in table 3.

**Second Phase:** The IAPS picture (stimulus: cemetery code 9220) related to quadrant 3 of Russell's model is shown for 10 seconds, and the SAM survey is then conducted. Its relation to the mean value published for this picture is considered. Table 3.

**Third Phase:** The OLAF picture (stimulus: tortilla code 5515) is shown for 10 seconds with its mean values. The SAM survey is then conducted to assess the subject's emotional state. The resulting values are shown in table 3.

**Table 3** Results obtained in phases 1/2/3

<b>Phase 1 Picture</b>	Stimulus :Black screen
SAM	[5(0);5(0)]
Face Attributes	"Anger":0.0, "Contempt":0.005, "Disgust":0.0, "Fear":0.0, "Happiness":0.0, "Neutral":0.984, "Sadness": 0.011, "Surprise": 0.0
Affective Suite	Arousal (0.35); Meditation (0.33); 0.78 (Frustration); 0.57 (Attention); resulting MDEmv: Frustration and MDE (MDEF1): [Excitemen (Arousal): 0.35(-1.5); Meditation (Valence): 0.33(-1.7)]
EEG	[E (-); V (-)] threshold 0.4 Russell's Third Quadrant
<b>Phase 2 Picture</b>	Stimulus :IAPS 9220: Arousal 2.27 (SD 1.61), Valence 3.83 (SD: 2.33)
SAM	[5 (0); 1 (-4)]
Face Attributes	"Anger":0.0, "Contempt": 0.001, "Disgust": 0.0, "Fear": 0.0, "Happiness": 0.0, "Neutral": 0.997, "Sadness": 0.002, "Surprise": 0.0
Affective Suite	Instantaneous Excitement (0.49); Meditation (0.33); 0.83(Frustration); 0.56(Attention); resulting MDEmv: "Neutral"; MDE (MDEF2): [Excitemen (Arousal): 0.49(-0.1); Meditation (Valence): 0.33(-1.7)]
EEG	[E(-); V(-)] Russell's third quadrant
<b>Phase 3 Picture</b>	Stimulus :OLAF Cod. 5515 E: 4.11, SD (2.71), V: 7.06, SD (1.67)
SAM	[9(4); 9(4)]
Face Attributes	"Anger": 0.0, "Contempt": 0.001, "Disgust": 0.0, "Fear": 0.0, "Happiness": 0.983, "Neutral": 0.015, "Sadness": 0.0, "Surprise": 0.0
Affective Suite	Arousal (0.97); Meditation (0.33); 0.86 (commitment/Frustration); 0.55 (Attention); resulting MDEmv: commitment/Frustration; MDE (MDEF3): [Arousal: 0.97 (4.7); Meditation: 0.33 (-1.7)]
EEG	[E(+); V(+)] threshold 0.2 Russell's first quadrant

Consistent with phase 3, figure 2 shows the captured face picture of the test subject using the brain computer interface, and arousal and meditation values during the observation of food picture fat5515 (tortilla) from the OLAF set. The phase results are shown in table 4 according to the fusion proposal adopted in table 1.



**Fig. 2** Visualization of Stimulus (OLAF Picture 5515), Face Capture, Arousal/Meditation values indication of the test subject during a moment in phase 3

**Table 4.** Result of MM fusion of Arousal and Valence

Phas	Stimulus	EK MA N	Face Capt	Fusion proposal to Russell's Dimensional Circumplex Approach. Emotion and Associated Value (Fig 2)	
				Emotion	Value
				Arousal	Valence

							Value	SAM	EEG	Value	SAM	EEG
1	NO	Sadness	Neutral	Sadness	Frustration	--3C	EB	+5	-	VP	+5	-
2	IAPS 9220	Sadness	Neutral	Sadness	Frustration	--3C	EB	+5	-	VN	1	-
3	OLAF 5515	Joy	Happiness	Arousal	Arousal	++1C	EA	+9	+	VP	+9	+

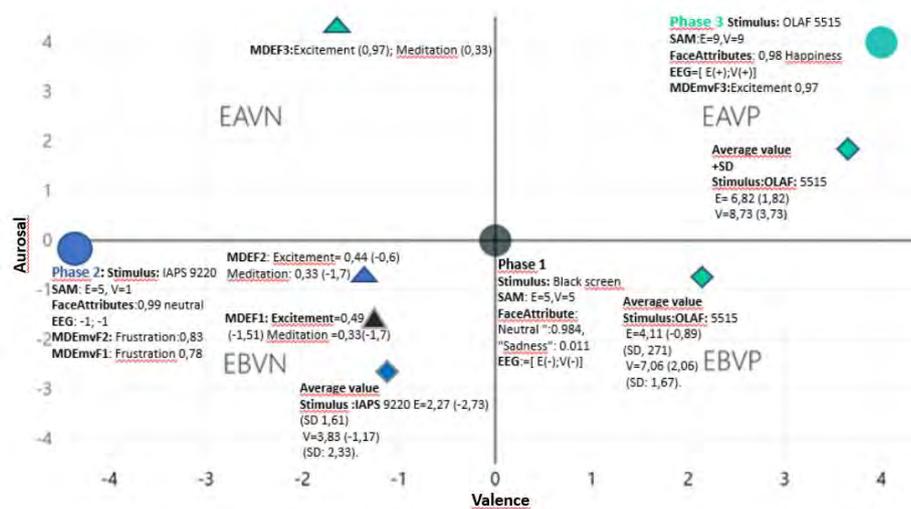


Fig.3. Result representation within the multimodal fusion of Arousal/Valence

It is consistent with the column “EEG Arousal/Valence Determination” in table 1. Figure 3

Figure 3 shows the results of MM validation for each of the emotional states (neutral, sadness, happiness) the test subject suffers, they match with stimuli of images and are later validated with the representative parameters of the populations which are part of the sets of data applied (IAPS, OLAF).

Figure 3 presents the resulting emotional record related to the subject’s obtained arousal and valence values according to the results from face capture, EEG, Emotiv Metadata (EMD) and SAM survey within the dimensional model of emotions in each phase shown in table 4. In particular, two types of representation are identified for EMD; one by selecting the highest value obtained (MDEmvfx) and EMD by considering arousal and meditation both converted to the dimensional model for each phase of the test, in this order, to adjust the different origin values found between zero and one as metadata value of transformed emotional states (MDEF) = (MDE\*10)-5.

In this case, valence is identified with meditation and arousal is identified with the emotion named short-term arousal.

## 5 Conclusions and future lines of work

Even though the results related to the base tests were successful, it is worth mentioning what H. Gunes [16] et al. have proposed. “Even though there have been great breakthroughs in the research field of affective computing, modelling, analyzing, interpreting and responding to naturalist affective human behavior is still challenging for automated systems since emotions are complex constructions with unclear limits and significant particular variations in expressions and experiences”. Promising results which contribute to the development of affective computing settings related to the emotional profile of the test subject representing the consumer have been achieved through the fusion model. The fusion of multimodal and multidimensional data (dimensional and categorical emotional approach) is achieved in the representation scheme of Arousal/Valence and thus allowing for a referencing framework of an individual’s emotional state before, during and after the test supporting the aforementioned since food assessments are not valid among individuals and groups unless feelings towards food signals are compared to feelings towards strong experiences unrelated to food, which are used as reference points. Future lines of work will involve the first attempts to validate the Geneva Emotion Wheel, however, there have been some difficulties to reproduce the intensity scale theoretically predicted for some terms of Geneva Wheel’s rays. [29]. Sets of faces images which allow the MM integration of Geneva Wheel [30] [31], including the Emotional Journey [32] concept will be explored in order to attain an increase of the user’s emotional experience. As future lines of research, the integration of the test subject’s physiological parameters is being developed by using sensors of heart rate variation, galvanic skin responses and emotion inference from vocal expression

## References

- [1] Ierache, J., Pereira, G., Iribarren, J., & Sattolo, I. (2012). Robot Control on the Basis of Bioelectrical Signals. En *Robot Intelligence Technology and Applications 2012* (pp. 337-346). Korea: Springer
- [2] Ierache, J., Pereira, G., Sattolo, I., Guerrero, A., D’Álto, J., & Iribarren, J. (2011). Control vía Internet de un Robot ubicado en un sitio remoto aplicando una Interfase Cerebro-Máquina. XVII CACIC 2011 (págs. 1373-1382). La Plata: REDUnci
- [3] Ierache, J., Nervo, F., Pereira, G., & Iribarren, J. (2014). Estado Emocional Centrado en Estímulos, Aplicando Interfase Cerebro-Maquina. XX CACIC, 2014. Bs.As.: REDUnci
- [4] Ierache, J., Cervino, C., Eszter, E., Fortin, D., & Castro Menna, A. (2018). Influencias del estado biométrico emocional de personas interactuando en contextos de entornos virtuales. WICC, pp 785-789. Corrientes
- [5] Ierache, J., Nicolosi, R., Ponce, G., Cervino, C., & Eszter, E. (2018). Registro emocional de personas interactuando en contextos de entornos virtuales. XXIV CACIC 2018, (págs. 877-886).

- [6] 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, Libro de actas pp 417-426
- [7] Picard, R. (2000). *Affective Computing*. En T. M. press, *Affective Computing* (págs. 4-8). Cambridge Massachusetts: The Mit Press
- [8] Baldasarri, 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>
- [9] Rafael A. Calvo, Senior Member, IEEE, and Sidney D'Mello, *Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications* IEEE transactions on affective computing, vol. 1, no. 1, january-june 2015.
- [10] Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (1997) *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. NIMH
- [11] Lang, P. J., Bradley, M. M., & Cuthbert, B.N. (2008). *Affective ratings of pictures an instruction manual*. International affective picture system (IAPS) Technical Report A-8. Gainesville: University of Florida.
- [12] García Campayo J. et al (2017), *Mindful Eating sabor de la atención* Ed Siglantana S.L.
- [13] Miccoli, L., Delgado, R., Rodríguez-Ruiz, S., Guerra, P., García, E., & Fernández-Santaella, M. C. OLAF, the Open Library of Affective Foods in ADULTS: <http://digibug.ugr.es/handle/10481/41499>
- [14] Miccoli, L., Delgado, R., Rodríguez-Ruiz, S., Guerra, P., García, E., & Fernández-Santaella, M. C. (2014). OLAF, Open Library of Affective Foods: A tool to investigate the emotional impact of food in adolescents. *PLOS ONE*. Zenodo. <http://doi.org/10.5281/zenodo.10202>
- [15] Lang, P. J. (1985) *The cognitive psychophysiology of emotion: Fear and anxiety*. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp.131-170). Hillsdale, NJ, US.
- [16] J. Posner, J.A. Russell y B.S. Peterson, (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, ISSN: 0954-5794.
- [17] H. Gunes, B. Schuller, M. Pantic et al., (2011) "Emotion representation, analysis and synthesis in continuous space: A survey", en *Face and Gesture 2011, IEEE*, 2011, págs. 827-834, ISBN 978-1-4244-9140-7.
- [18] P. Ekman, "Basic Emotions", *Handbook of Cognition and Emotion*, T. Dalgleish y M.J. Power, eds., Chichester, UK: John Wiley & Sons, Ltd, 2005, cap.3, págs.45-60, ISBN: 978-0-47197-836-7
- [19] Paul Ekman conjunto de siete emociones básicas y universales. <https://www.paulekman.com/universal-emotions/> vigente julio 2020
- [20] Russell, J. (1980) A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178.
- [21] Hamadicharef, "Brain Computer Interface Literature- A bibliometric study", in 10th International Conference on Information Science, Signal Processing and their Applications, Kuala Lumpur, 2010, pp. 626-629.
- [22] Emotiv Epoc <https://www.emotiv.com/product/emotiv-epoc-14-channel-mobile-eeeg/>
- [23] Ramirez, Rafael & Vamvakousis, Zacharias. (2012). Detecting Emotion from EEG Signals Using the Emotive Epoc Device. 7670. 175-184. 10.1007/978-3-642-35139-6\_17.
- [24] Rossi, F. Tesis Educación de estados emocionales basada en captura multimodal de bioseñales dirección J. Ierache. 2019, LSIA- FIUBA.
- [25] Anexo "Modelo Conceptual de la arquitectura Multimodal de valoración emocional en el marco de la computación afectiva aplicado en ambientes gastronómicos". Ierache et al 2020 [https://drive.google.com/file/d/1-pTV317PY\\_Z5R6IuVZE9QyifpthQ0LuN/view?usp=sharing](https://drive.google.com/file/d/1-pTV317PY_Z5R6IuVZE9QyifpthQ0LuN/view?usp=sharing)
- [26] <https://powerbi.microsoft.com/es-es/> vigente julio 2019.

- [27] <http://openvibe.inria.fr/> vigente julio 2019.
- [28] <https://azure.microsoft.com/es-mx/services/cognitive-services/face/#features> junio 2020.
- [29] Scherer, Klaus R. (2005): “¿Qué son las emociones? ¿Y cómo pueden ser medidas?” *Social Science Information*, vol 44, n°4: 695–729. SAGE Publications
- [30] Baenziger, T., Tran, V. and Scherer, K.R. (2005) “The Emotion Wheel. A Tool for the Verbal Report of Emotional Reactions”, conference of the International Society of Research on Emotion, Bari, Italy
- [31] Yenny Alexandra Méndez Alegría, César Alberto Collazos Ordóñez, Rosa Gil y Antoni Granollers I Saltiveri Rueda de emociones de Ginebra+: Instrumento para la valoración emocional de los usuarios mientras participan en una evaluación de sistemas interactivos. *Dyna (Spain)* ISSN:0012-7361, 2016 vol:91 fasc:2 págs: 151-155.DOI: /10.6036/7572
- [32] Du, Shen & Shu, Edouard & Tong, Feifei & Ge, Yinghao & Li, Lu & Qiu, Jingbo & Guillotel, Philippe & Fleureau, Julien & Danieau, Fabien & Muller, Daniel. (2016). Visualizing the Emotional Journey of a Museum. 7-14. DOI:10.3384/ecp10302.

# Estrategias a nivel de macro y micro-planificación en un prototipo de verbalización de ontologías

Martín Bermudez and Sandra Roger

Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA),  
Facultad de Informática. Universidad Nacional del Comahue  
Buenos Aires 1400, (8300) Neuquén  
`martin.bermudez@est.fi.uncoma.edu.ar`  
`roger@fi.uncoma.edu.ar`

**Resumen** En este trabajo se presenta un primer prototipo de un trabajo de generación de lenguaje natural capaz de procesar el contenido de una ontología y generar un texto en lenguaje natural. El texto resultante debe estar organizado y expresado con una sintaxis aceptable, para que sea accesible a cualquier usuario cuyo interés sea el dominio modelado, pero no tenga las herramientas necesarias para comprender las lógicas subyacentes.

**Keywords:** Verbalización, Ontología, OWL, Generación del Lenguaje Natural

## 1. Introducción

La Web Semántica puede ser vista como una colección de estándares y tecnologías que permiten a las máquinas entender el significado (semántica) de la información disponible en la Web [10]. Para describir el dominio de interés en la Web Semántica, se adoptó el uso de ontologías. Algunos lenguajes conocidos para representar ontologías son RDF y OWL2.

Como las ontologías estructuran los datos de manera formal, son poco comprensibles por usuarios no expertos, y no brindan una cómoda visualización de la información para aquellos que quieran beneficiarse del uso de las tecnologías semánticas. Por este motivo, expresar el contenido formal en Lenguaje Natural (LN) resulta atractivo, brindando la capacidad de documentar y expresar ontologías en un lenguaje accesible por usuarios no entrenados. Un Sistema de Generación de Lenguaje Natural (SGLN) puede tener diferentes objetivos en cuanto a la salida que producen, ya sea enfocarse en la facilidad de lectura, la capacidad de poder generar diferentes oraciones para expresar la misma información, enfocarse en la organización del contenido, entre otras. En el ámbito de las ontologías existen algunos trabajos que involucran el procesamiento de lenguaje natural [4][5][8]. Como veremos más adelante, cada enfoque posee ventajas y desventajas, por ejemplo en cuanto a la integración con las ontologías, debido

a la necesidad de modificar la ontología subyacente o de requerir convenciones de nombrado; como en su potencial para ser de utilidad al usuario final, debido a la complejidad de las oraciones, la falta de flexibilidad léxica, y la organización del contenido del texto, lo que resulta en una pobre comprensión del texto.

Teniendo en cuenta estas ventajas y desventajas, en este trabajo se propone el primer prototipo de un sistema de generación de texto que mantenga un equilibrio entre la dificultad de integrar el sistema con las ontologías, y la capacidad del sistema de generar texto que resulte útil al usuario final. El Sistema de Generación de Texto debe ser capaz de recibir una ontología de la Web Semántica y generar un documento de texto organizado y en lenguaje natural. No será necesario agregar información a la base de conocimiento, mientras que se espera alcanzar una salida que sea sintácticamente aceptable.

Los dos puntos claves en este trabajo son: la ORGANIZACIÓN DE LA INFORMACIÓN, para estructurar el texto con base en las relaciones semánticas, con el fin de reducir la carga cognitiva que exige reconocer las relaciones del dominio; y la GENERACIÓN DEL TEXTO EN LENGUAJE NATURAL, maximizando la cohesión de las oraciones.

## 2. Organización de la información

En esta sección describimos el diseño de un módulo para organizar y estructurar la información de una ontología de forma más significativa, de manera que favorezca la estructura inicial de un texto, teniendo en cuenta el preprocesamiento para el análisis e interpretación de datos de Reiter E. [7] dando como resultado resultado una jerarquía de tópicos que será usado para poder abordar la coherencia global del texto.

Considerando la suposición de la Teoría de Veins [2], las referencias entre unidades textuales son posibles si se encuentran en una relación estructural. Por lo cual, creamos una estructura jerárquica que facilite la relación entre unidades textuales evitando que una entidad en un nivel  $N$  haga referencia a entidades que se encuentren en niveles  $N+1$  (niveles inferiores).

Las entidades que tienen mayor impacto en la ontología aparecen en los niveles superiores de la jerarquía, con el objetivo de usarlas como hilo conductores para estructurar el texto. Por ello, trataremos dos problemas principales: la obtención de los tópicos más relevantes para tratar en el texto y el recorrido del grafo para obtener las unidades de información adecuada que maximicen la relación semántica respecto a los tópicos del texto (para abordar la coherencia global).

### 2.1. Diseño

Para organizar la información, se crearon cuatro módulos que transforman la ontología de entrada en un árbol con toda la información contenida en la ontología. **Traductor**: recibe la entrada y la traduce a una representación interna de la ontología. **Clasificador de Entidades**: clasifica y extrae las entidades más

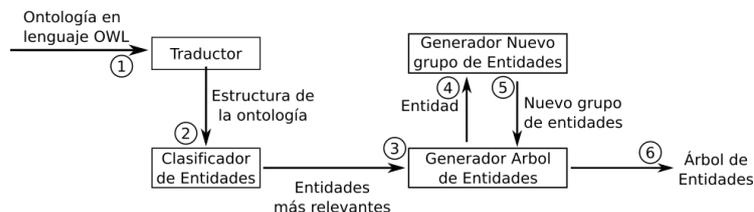


Figura 1: Módulos que componen el Organizador de Información

relevantes, utilizando como criterio las medidas de centralidad. **Generador de Nuevo Grupo de Entidades:** recibe una Entidad y recorre sus relaciones para crear un nuevo grupo de entidades relacionado semánticamente. **Generador del Árbol de Entidades:** se encarga de crear los niveles del Árbol de Entidades.

En la Figura 1 se muestra la transformación de la ontología OWL a través de los módulos. Los pasos 4 y 5 son iterativos, el proceso finaliza cuando no queden más Entidades que explorar, lo cual es definido por el Generador del Árbol de Entidades.

*Criterio de clasificación* Consideramos que una clase tiene más información que otra (y por lo tanto que es más relevante) si está presente en más cantidad de dominios de *ObjectProperties*. Esto permite reconocer las clases que tengan mayor cantidad de conexiones con otras clases, y que a su vez sean el núcleo de la relación. El cálculo de información de cada clase, utilizamos el grafo subyacente a la ontología y calculamos el *outdegree*<sup>1</sup> de cada nodo que represente una clase. Para el cálculo se tiene en cuenta únicamente la relación *rdf:domain*.

Algunas ventajas de este enfoque son: sencillez de implementación<sup>2</sup>; no se requiere agregar información extra al dominio; no es necesario utilizar sobre la ontología un razonador que requiera una complejidad computacional que sea intratable. La única observación es que hay que inferir el dominio y rango de las *ObjectProperties*. Sin embargo, la inferencia de dominio y rango se realiza teniendo en cuenta la jerarquía de *ObjectProperties*, sin necesidad de inferir clases equivalentes o disjuntas.

La eficacia de este enfoque depende fuertemente de que las *ObjectProperties* más importantes tengan el dominio explícito.

*Seleccionando las principales clases* Para obtener el tema principal (o temas principales), se seleccionan las clases que superen el valor promedio entre cantidad de propiedades y cantidad de clases. Si ninguna supera el valor promedio, se selecciona la o las clases con el valor más alto. Cuando se selecciona más de una clase que pertenece a una misma jerarquía, se eliminan las subclasses, dejando como representante a la superclase en común.

<sup>1</sup> El *outdegree* es la medida de la cantidad de conexiones salientes de un nodo.

<sup>2</sup> Únicamente se debe recorrer el grafo calculando el valor de cada nodo. El recorrido del grafo tiene a lo sumo una complejidad polinomial.

*Agrupando la información de las Entidades* El tercer módulo crea nuevos grupos de entidades para agregar como un nuevo nivel al Árbol. Los elementos de cada grupo serán insertados como hijos de la Entidad correspondiente.

Dada una Entidad  $E$ , debe recorrer las relaciones de  $E$  en la ontología, para seleccionar Entidades que serán agregadas al nuevo grupo. El contenido de cada grupo depende del tipo de  $E$ . Si  $E$  es una Clase, el nuevo grupo estará constituido por Propiedades, Individuos y Subclases. Si  $E$  es una Propiedad, el nuevo grupo estará constituido por Subpropiedades cuyo dominio contenga a  $E$ , o por las clases del Rango de  $E$  si no posee Subpropiedades. Si por el contrario  $E$  es un Individuo, no se creará ningún grupo nuevo.

Cuando existen más de una Propiedad en un grupo, se busca reemplazarlas por propiedades en común de más alto nivel en la jerarquía de propiedades.

*Generador Árbol de Entidades* El cuarto módulo crea la estructura del Árbol, sincronizando la interacción entre los demás módulos. Su objetivo es formar la jerarquía, comenzando desde las Entidades más Relevantes, y agregando nuevos niveles con ayuda del Generador de Nuevo Grupo de Entidades. Si hay clases que no fueron utilizadas, invoca un algoritmo para agregar estas clases en una rama especial.

### 3. Generación del documento de texto

Para el proceso de generación del documento de texto, se usó el Árbol de Entidades del organizador de información detallado anteriormente para dar una estructura inicial al texto, se agrupó el contenido de cada entidad de la ontología en párrafos y secciones, y luego se verbalizaron los axiomas, teniendo en cuenta la coherencia local dentro de los párrafos generados. Para ello, se crearon tres módulos principales: **Macroplanificador**, **Microplanificador** y **Realizador**, los cuales implementan los procesos asociados a las etapas de Macroplanificación, Microplanificación y Realización gramatical [9], respectivamente.

Para mantener la coherencia local según la Teoría de Centrado [6], y evitar cambiar continuamente los focos de atención, para cada tópico tratado en el texto, agrupamos toda su información en párrafos consecutivos. Dado una entidad  $E$ , para recuperar la información que la describe, se recorre la estructura de grafo a partir del nodo que representa a  $E$ , navegando por las relaciones del nodo, a solo un nivel de profundidad.

La relación entre el Organizador de Información y el Generador del Documento de Texto se puede ver en la arquitectura general de nuestro sistema (Figura 2).

#### 3.1. Diseño Macroplanificador

Para crear el documento en la etapa de Macroplanificación, se reemplazó el módulo Generador del Árbol de Entidades del Organizador de Información, por un nuevo módulo específico del Macroplanificador.

La representación interna del documento de texto tiene en cuenta los siguientes criterios: Cada *Entidad* presente en la estructura que resulta del Organizador de Información es considerada un tópico. A cada *Tópico* del primer nivel le corresponde una sección. Los tópicos anidados (Subclases, Individuos y Relaciones) se corresponden a subsecciones (secciones anidadas). Cada *Sección* está compuesta de al menos un párrafo y un título. Cada *Párrafo* trata un único tópico, el cual puede ser una clase o un individuo. También, un párrafo puede o no tener alguna sentencia. Esto se debe a que no todas las entidades tienen asociados axiomas que la describan.

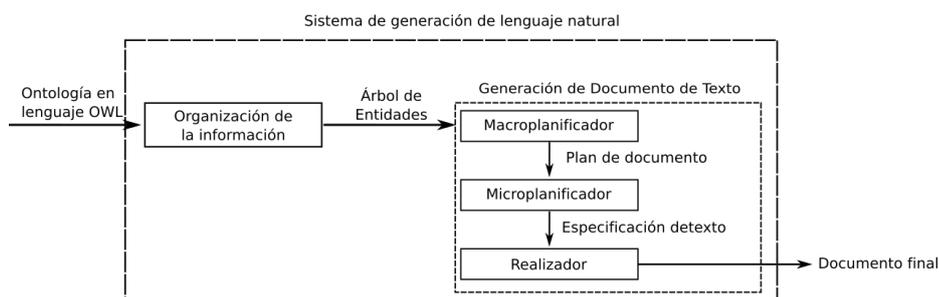


Figura 2: Arquitectura de los módulos que componen nuestro Sistema de Generación de Lenguaje Natural

**Secciones y párrafos** Dentro de una sección para una Clase, se verbalizarán sus axiomas y se enumerarán sus Individuos, Subclases y Relaciones, todo en un párrafo. En una sección asociada a una Relación, solo habrán subsecciones acerca de las Entidades que pertenecen a la Relación. En una sección para Individuos, se creará un párrafo para describir sus propiedades.

**Subsecciones** Las subsecciones son secciones anidadas dentro de otras secciones. Cuando surge a partir de un Individuo, la subsección trata como tópico al Individuo; cuando surge a partir de una Subclase, el tópico principal es la Subclase; y cuando surge a partir de una Relación, el tópico principal resulta ser el rango de la Relación, es decir, otra Clase.

**Oraciones** Las oraciones pertenecen a los párrafos por lo que tratan un solo tópico. Cada oración que tenga como tópico a una Clase, aborda un solo tipo de información, por lo que existe una oración para las clases disjuntas, una oración para las clases equivalentes, y así con los demás tipos de información. Cada oración correspondiente a un Individuo contiene las propiedades con sus valores declaradas sobre el Individuo.

Al tratar un solo tipo de información por oración, se maximiza la cohesión dentro de cada oración.

### 3.2. Diseño Microplanificador

En esta etapa se llevarán a cabo las tareas para describir una Entidad en lenguaje natural a partir de la información que contiene cada grupo creado en la Macroplanificación. Para simplificar, se asoció a cada constructor OWL solo algunas formas de verbalización. Se utilizó *Stanford POS Tagger*<sup>3</sup> para etiquetar las palabras y poder trabajar con patrones de la gramática del lenguaje humano.

El enfoque adoptado para crear la sintaxis de las oraciones se basa en un recorrido bottom-up de la jerarquía de un *Axioma*. Se comienza desde las hojas, construyendo oraciones parciales a partir de las *Entidades*, luego se procede a subir por los niveles a través de los constructores de *Expresiones de Clases*, componiendo nuevas oraciones parciales, hasta alcanzar la raíz de la jerarquía, donde se termina de construir la oración final. Adicionalmente, se realizan algunos tratamientos morfológicos para agregar fluidez y coherencia al texto, tal como inserción de artículos, y manejo de concordancia de género y número.

Con el objetivo de comenzar a formar las oraciones parciales, el primer paso a realizar es la verbalización de las *Entidades*, ya que están directamente relacionadas con las IRIs (o tienen acceso al *label*, en caso de extraer los nombres desde los *labels*). En el nivel de constructores de *Expresiones de Clase*, es posible generar oraciones complejas subordinando las expresiones que se encuentran anidadas, o componer nuevas oraciones parciales teniendo en cuenta los tipos de componentes involucrados. En el nivel más abstracto se encuentran los constructores de *Axiomas*. Estos constructores tienen la característica de no ser recursivos entre ellos, por lo que es posible generar oraciones independientes, yuxtapuestas o coordinadas.

**Componentes y oraciones parciales** Durante la creación de una oración, se crean oraciones parciales y se componen entre ellas. Se asoció a cada constructor un tipo de componente oracional para evitar una exhaustiva programación de composiciones. Los tipos de componentes son los siguientes: *Término (T)*: caracterizado por no poseer verbo. Pueden contener adverbios, sustantivos y adjetivos. *Sintagma Verbal (SV)*: debe poseer un verbo. *Oración Negativa (ON)*: representa la negación de un componente. *Unión*: representa una disyunción de oraciones parciales. *Intersección*: representa una adición o subordinación de oraciones parciales.

Los componentes que retorna cada constructor se definen a continuación: Componente Término (T): (Clase, Individuo). Componente Sintagma Verbal (SV): (Propiedad, Constructores de cuantificación, Constructor *hasValue*, Constructores con cardinalidad). Componente Unión: (*UnionOf*, *OneOf*). Componente Intersección: (*IntersectionOf*). Componente Oración Negativa (ON): (*ComplementOf*).

**Verbalización de constructores OWL** Esta tarea se encarga de componer las oraciones según la información de cada constructor OWL. En algunos constructores se diseñó más de una forma de composición, mejorando la variabilidad, fluidez e interpretación de las oraciones.

<sup>3</sup> Un software para asignar la categoría gramatical a las palabras de un texto <https://nlp.stanford.edu/software/tagger.html>

Para reconocer cómo componer las oraciones en cada constructor, se revisó empíricamente los axiomas de algunas ontologías, y se buscó utilizar oraciones que sean lo más genéricas posibles, que permitan la comprensión de los axiomas.

**Expresiones de referencia** Las expresiones de referencia solo fueron implementadas para las *Clases*. Para llevar a cabo la referenciación, se tuvo en cuenta el uso de pronombres demostrativos (“este/esta”), a veces, acompañados por el primer sustantivo (si posee) del nombre de la *Clase*.

**Agregación de sentencias** Basaremos la agregación en la conjunción por componentes compartidos [1]. El objetivo es que los elementos en común aparezcan una sola vez, mediante elipsis del componente repetido. La agregación de sentencias ocurre en dos lugares. Uno es durante el proceso de producción de una oración en el constructor *N-Ary*, es decir dentro de una oración; el otro es durante el proceso de creación de párrafos, es decir, entre oraciones.

### 3.3. Diseño del realizador

Esta última etapa se encarga de dar formato al Documento de Texto para presentar todas las secciones y párrafos de manera más agradable. Reduiremos las secciones en párrafos, para que el Documento Final sea más compacto, tratando de mejorar la estética y el proceso de lectura, sin perjudicar la coherencia y la segmentación de la información.

Para llevar a cabo la reducción, se midió la información verbalizada, usando como métrica la cantidad de oraciones presentes en la descripción de un tópico.

**Criterio de reducción de Secciones** Para decidir si un tópico es planificado como una Sección, se utilizan los siguientes criterios: debe poseer cinco o más oraciones; debe poseer alguna Subsección con cinco o más oraciones; en cualquier otro caso será planificado como un Párrafo.

Cuanto menor sea la cantidad de oraciones requeridas para ser considerada una sección, la planificación del documento será más similar al Documento Inicial; mientras que si la cantidad de oraciones requeridas es muy grande, la planificación perderá niveles de jerarquía, y únicamente contemplará como secciones a los tópicos del primer nivel. Ninguno de los dos extremos es conveniente: o se pierde coherencia, o se obtiene un texto con demasiados niveles de secciones, que resulta antinatural. Sin embargo, un documento en el que predominan las subsecciones (aunque tengan poca información), ayuda a los humanos a reconocer y establecer relaciones entre los tópicos, siendo el objetivo principal no desaprovechar la semántica del dominio modelado, y teniendo como intención principal que el lector comprenda el dominio modelado en la ontología.

## 4. Una aplicación sobre la ontología *wine*

Analizamos el comportamiento del sistema utilizando como entrada la ontología Wine. Esta ontología tiene como objetivo describir un dominio de vinos y comidas<sup>4</sup>, por lo que esperamos que las entidades que se consideren más rele-

<sup>4</sup> [https://protege.stanford.edu/publications/ontology\\_development/ontology101-noy-mcguinness.html](https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html)

vantes sean aquellas afines a los vinos y comidas. A continuación se presentan algunos fragmentos del documento generado.

1 Wine A wine is a potable liquid. A wine has wine sugar as sugar, has wine flavor as flavor, has wine body as body, has wine color as color, has wine descriptor and made from grape a wine grape. This located in some region. Also, wine has only winery maker. Also, wine has exactly one maker, made at least one wine grape, wine body, wine color, wine flavor and wine sugar. There are the following kinds of wines: italian wine, . . . , late harvest and alsatian wine.  
 Full bodied wine: a full bodied wine is a wine that has full body.  
 Burgundy: a burgundy is a wine that located in bourgogne region. This has dry sugar. There are the following kinds of burgundies: white burgundy and red burgundy.  
 . . . Otros vinos realizados como párrafos.  
 1.1 Italian wine  
 A italian wine is a wine that located in italian region. chianti is the only kind of wine.  
 Chianti: a chianti is a italian wine. This has only light or medium body. Also, chianti has red color, moderate flavor and dry sugar. located in chianti region. made from grape sangiovese grape. chianti classico is a type of chianti.  
 Chianti classico: a chianti classico has mc guinnesso maker and medium body.  
 1.2 ... 1.7  
 1.8 Red wine  
 A red wine is a wine that has red color. There are the following kinds of wines: dry red wine, red burgundy, port and red bordeaux.  
 Port: a port is a red wine. This has full body, strong flavor and sweet sugar. located in portugal region. taylor port is a type of port. Taylor port: a taylor port has taylor maker.  
 1.8.1 Red burgundy  
 A red burgundy is a burgundy and red wine. This made from grape pinot noir grape. Also, red burgundy made at most one wine grape. cotes d or is the only kind of burgundy. Cotes d or: a cotes d is a red burgundy that located in cotes d or region. This has moderate flavor. clos de vougeot cotes d or is a type of cotes d or. Clos de vougeot cotes d or: a has clos de vougeot maker.  
 . . . Otros vinos realizados como secciones  
 1.14 Wines descriptor  
 Wine sugar: a wine sugar is a wine taste. a wine sugar is a sweet, off dry or dry. there are the wines sugar: dry, off dry and sweet.  
 Wine color: a wine color is a wine descriptor. a wine color is a rose, red or white. there are the wines color: white, red and rose.  
 Wine flavor: a wine flavor is a wine taste. a wine flavor is a moderate, delicate or strong. there are the wines flavor: moderate, strong and delicate.  
 Wine body: a wine body is a wine taste. a wine body is a light, medium or full. there are the wines body: medium, full and light.

Para identificar el formato del texto, a las secciones se le incluye el número de sección a la izquierda del título de la sección, y cada párrafo comienza con sangría.

La ontología Wine cuenta con 138 clases y 16 propiedades. De las 16 propiedades, 12 fueron usadas para verificar qué clases superan el promedio para ser seleccionadas como las principales. El promedio fue de 2.5 y la única clase que lo superó fue Wine, lo que cumple parcialmente el resultado esperado, ya que parte del objetivo de la ontología es describir el dominio de los vinos. Respecto a la sección que describe las comidas, quedó desplazada a *otras secciones*, siendo un resultado no esperado según el objetivo de la ontología. Sin embargo, analizando manualmente la ontología, se puede apreciar que el porcentaje de información que describe a las comidas es significativamente menor en relación a la información referida a los vinos, factor por el cual resulta aceptable que no aparezcan como sección principal.

Analizando la coherencia global y el sentido del texto, la estructuración resultó razonable, enfocada en los vinos y las propiedades de los vinos.

La implementación de la sintaxis presenta algo de redundancia y repetitividad, aunque las oraciones son claras. El resultado de la realización final mantiene la coherencia global y establece las relaciones pertinentes entre los tópicos.

A continuación se presenta una comparación de la verbalización de la clase *Beaujolais* de la ontología *Wine*, entre el trabajo [3] y el nuestro. En [3] proponen listar sus características de la siguiente manera: “*A Beaujolais is a Wine that:*

- *is made from at most 1 grape, which is Gamay Grape,*
- *has Delicate flavor,*
- *has Dry sugar,*
- *has Red color,*
- *has Light body ”*

En nuestro caso, la oración resultante es la siguiente: “**Beaujolais has light body, red color, delicate flavor and dry sugar. Made from grape gamay grape.**”

Existe clara diferencia, no solo en el criterio de presentación lista frente a prosa, sino también en el hecho de elegir cuánta elipsis emplear (en nuestro caso, además de omitir *Beaujolais*, también omitimos el verbo “has” para evitar redundancia), y de separar en diferentes oraciones las características asociadas al “has” y las asociadas al “made from”. Decidir qué representación usar depende de muchos factores, por ejemplo, la versión en prosa no resulta muy compleja, pues la omisión del “has” solo ocurre en tres características, por lo que no recae tanta carga sobre el lector. Quizá en descripciones más largas sea conveniente un formato en lista, pero también hay que tener en cuenta cómo impactaría visualmente representar todas las clases de una gran ontología de esta forma. De acuerdo con [3], pueden existir métodos de generación más complejos que mejoren la calidad del texto de salida.

## 5. Conclusiones

Se diseñó un Sistema de Generación de Lenguaje Natural, orientado a generar un texto organizado a partir de una ontología (OWL) de la Web Semántica.

Para maximizar la relación semántica entre los tópicos, se propuso una solución basada en el análisis del grafo subyacente a la ontología, midiendo las relaciones entre sus nodos. De esta manera se obtuvieron las Entidades de la ontología que resultaban más sobresalientes, con el fin de usarlas como hilo conductor de los temas del texto generado. Esta solución se vio motivada por la característica que poseen las ontologías de la Web Semántica, de ser grafos libres de escala [11].

El resultado obtenido del caso de prueba retornó una estructura organizada, que se correspondía con los resultados esperados. Sin embargo, la solución propuesta depende de la cantidad de *ObjectProperties* y del contenido de sus dominios, por lo que queda pendiente poder analizar la solución con mayor diversidad de ontologías, variando la densidad de sus relaciones.

Además de maximizar la relación entre los tópicos, se buscó maximizar la relación entre los elementos de los párrafos. Para esto se usó como base la coherencia local, a través de los postulados de la Teoría de Centrado.

Como se mostró en la sección 4, la solución basada en el desarrollo de estos módulos resultó en textos legibles y organizados. Si bien validar textos puede resultar subjetivo, podemos apreciar que existe suficiente relación semántica en el orden de los temas que se presentan, tanto entre las Secciones como entre las Oraciones, por lo que cumplen con cierta Coherencia Global y Local.

Aún resulta conveniente explorar otras técnicas para desarrollar las gramáticas y la composición de las oraciones. Desarrollar las gramáticas manualmente es un trabajo costoso, y debido a la naturaleza recursiva de las construcciones de los axiomas y de la semántica de sus constructores, es difícil mantener el control de la oración compuesta, tanto a nivel sintáctico como semántico. Queda pendiente probar el sistema con otras ontologías que presenten mayor complejidad en sus axiomas, con el fin de desarrollar nuevas gramáticas que abarquen más casos de prueba.

## Referencias

1. S. Bernardos. Marco metodológico para la construcción de sistemas de generación de lenguaje natural. *Unpublished doctoral dissertation*). *Informatica*, 2003.
2. D. Cristea, N. Ide, and L. Romary. Veins theory: A model of global discourse cohesion and coherence. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1*, pages 281–285. Association for Computational Linguistics, 1998.
3. D. Hewlett, A. Kalyanpur, V. Kolovski, and C. Halaschek-Wiener. Effective nl paraphrasing of ontologies on the semantic web. In *Workshop on end-user semantic web interaction, 4th int. semantic web conference, galway, ireland*, 2005.
4. D. C. Moreno and M. V. Lombardo. Ontología y procesamiento de lenguaje natural. *KnE Engineering*, pages 492–501, 2018.
5. M. C. Pérez Hernández. Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento. *Estudios de Lingüística del español*, 18:000–0, 2002.
6. M. Poesio, R. Stevenson, B. D. Eugenio, and J. Hitzeman. Centering: A parametric theory and its instantiations. *Computational linguistics*, 30(3):309–363, 2004.
7. E. Reiter. An architecture for data-to-text systems. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 97–104. Association for Computational Linguistics, 2007.
8. M. Váñez. La web semántica y las tecnologías del lenguaje humano. In *Web semántica y sistemas de información documental*, pages 155–180. Trea, 2009.
9. M. Vicente, C. Barros, F. S. Peregrino, F. Agulló, and E. Lloret. La generación de lenguaje natural: análisis del estado actual. *Computación y Sistemas*, 19(4):721–756, 2015.
10. L. Yu. *A developer's guide to the semantic Web*. Springer Science & Business Media, 2011.
11. H. Zhang. The scale-free nature of semantic web ontology. In *Proceedings of the 17th international conference on World Wide Web*, pages 1047–1048. ACM, 2008.

# GeoPerfil Profesional: una herramienta automática de información sobre profesionales

Emanuel Balcazar, Lucas Bobadilla, Waldo Fusiman, and Leo Ordinez

Laboratorio de Investigación en Informática (LINVI), FI - UNPSJB  
Bvd. Brown 3051, Puerto Madryn, Argentina  
{emanuelbalcazar13, lucasboba, wfusiman, leo.ordinez}@gmail.com

**Resumen** La Provincia del Chubut cuenta con características poblacionales de muy baja densidad general y alta concentración en pocas ciudades. A la vez, la actividad industrial y comercial requiere un agregado de valor diferencial para poder competir con los grandes centros urbanos o incluso lograr exportar sus productos. En este sentido, no se dispone de un registro de profesionales calificados que puedan aportar su conocimiento a dichos sectores. La presente propuesta busca contribuir al fortalecimiento del sector PyME de la provincia del Chubut mediante la construcción de una herramienta de extracción y procesamiento automático de información, para la consolidación de un mapa de perfiles profesionales de la provincia. Se pretende extraer información pública de medios digitales, analizarla en términos de campos disciplinares, identificar personas y proveer una forma de visualización georreferenciada.

**Keywords:** perfiles profesionales, extracción automática de información, medios digitales, NLP

## 1. Introducción

La Provincia del Chubut presenta una especialización productiva estrechamente ligada a la explotación y aprovechamiento de sus riquezas naturales. En términos generales, la economía de Chubut tiene una presencia notable de producción de *commodities* que, a su vez, posibilitan la existencia de un conjunto de industrias relacionadas[3][7].

**Contexto productivo provincial.** De acuerdo a datos del Observatorio PyME Regional Provincia del Chubut, el 43 % de las PyME industriales del Chubut se localiza en Comodoro Rivadavia, en tanto el 57 % restante reside en las localidades de Trelew, Puerto Madryn, Rawson y Gaiman [3]. Respecto a la contratación de personal, las principales dificultades de estas empresas parecen concentrarse en la contratación de operarios calificados y/o técnicos no universitarios, con el 80 % de las firmas con un grado de dificultad medio o alto. En segundo lugar, aparecen las dificultades en la contratación de universitarios, con el 73 % de

las firmas con dificultades medias o altas. Transversalmente debido a la característica de explotación de recursos naturales a gran escala se debe mencionar el desarrollo de base tecnológica para la industria metalmeccánica.

La necesidad de formación de recursos humanos especializados, el mejoramiento de procesos productivos, la incorporación de nuevas tecnologías y el apoyo al desarrollo de nuevos productos, son algunos de los principales desafíos identificados. Se trata de problemáticas que comparten los proveedores de la industria petrolera, del aluminio y naval pesquera, y que deben a su vez ser atendidas en estrecha relación con las áreas vinculadas al desarrollo productivo de la Provincia. El principal motivo señalado por el cual los empresarios tienen dificultades para contratar personal es la escasa preparación o formación inadecuada para cubrir los puestos [3]. En este sentido, se ve la necesidad de aprovechar el limitado capital humano disponible en la Provincia, poniendo a disposición de los empresarios un mapa de perfiles profesionales calificados.

**Trabajos relacionados.** La extracción y el procesamiento automático de información mediante técnicas de *Procesamiento de Lenguaje Natural* (NLP) [4] tiene una gran variedad de campos de aplicación. En [9], los autores analizan artículos periodísticos a fin de detectar relatos dentro de narrativas más complejas y de ellos obtener información de los actores involucrados en esos relatos. En [5], los autores analizan la cobertura periodística respecto al bullying y cyberbullying en un período de seis años en diarios de Estados Unidos. En [6], se estudia y compara el desempeño de diferentes consultores económicos polacos en base a sus informes. Un caso cercano al propuesto en este trabajo, aunque para un estadio superior, es presentado en [2]. Allí se desarrolla una medida para cuantificar la preparación de los empleados que pertenecen a una gran empresa con respecto al paradigma de la Industria 4.0. El proceso permite la identificación de tecnologías, técnicas y habilidades relacionadas contenidas en las descripciones de trabajo. A partir de estos, se mide el impacto de la Industria 4.0 en cada perfil de trabajo. Por otro lado, en [1] se plantea un sistema de recomendación académica en base a los conocimientos previos manifestados por los candidatos, una clasificación de conocimiento generada a partir de Wikipedia.

**Objetivo del trabajo.** En virtud de la ausencia de un registro sistematizado, en particular, de graduados universitarios en la provincia<sup>1</sup>, se planteó la posibilidad de construir un prototipo de software que realice la tarea de consolidar la información dispersa respecto a los mismos, mediante la búsqueda, extracción, clasificación y presentación de información de manera automática. En este proceso, el insumo principal serán notas periodísticas que den cuenta de nuevos graduados o colaciones de grado.

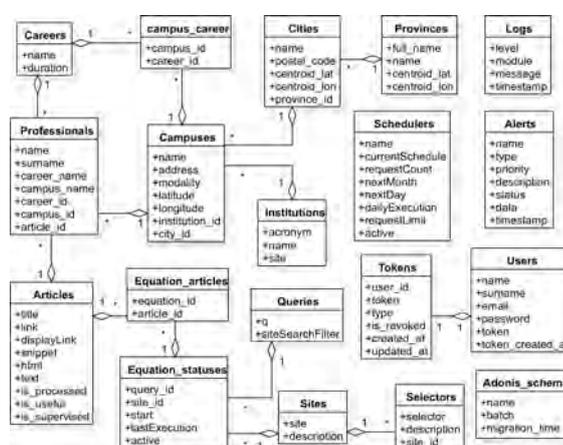
---

<sup>1</sup> En Chubut se encuentran dictando carreras presenciales la Universidad Nacional de la Patagonia San Juan Bosco, la Universidad Tecnológica Nacional y la Universidad del Chubut (provincial).

**Organización.** El resto del trabajo se organiza de la siguiente manera: la Sección 2 presenta el modelo de datos general de la propuesta; en la Sección 3 se expone el mecanismo de recuperación de la información de medios periodísticos; el mecanismo de procesamiento de la información obtenida se muestra en la Sección 4; una síntesis de los resultados obtenidos a partir del prototipo desarrollado se expone en la Sección 5; finalmente, la Sección 6 presenta las conclusiones y trabajos futuros.

## 2. Modelo de Datos

El esquema de datos del sistema se detalla en la Figura 1:



**Figura 1.** Esquema de datos

Las principales entidades del modelo son:

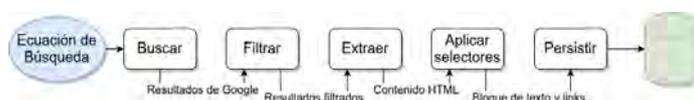
- campuses:** representa las sedes institucionales. Sus atributos permiten guardar su nombre, dirección, modalidad de cursado (presencial o virtual), su posición georeferenciada y posee además una *foreign key* a la institución a la cual pertenece.
- institutions:** representa las instituciones. Por lo general una institución posee una o varias sedes por lo cual se decidió separar la sede de la institución para diferenciarlos bien.
- careers:** representa las carreras, solo se guardan el nombre y su duración como información básica.
- campus\_careers:** hace de intermediario entre las tablas campus y careers dado a que su relación se definió como un *muchos a muchos* debido a que muchas carreras pueden dictarse en múltiples sedes a su vez que una sede dicta múltiples carreras.

- equations:** representa una búsqueda de Google, en donde el atributo “q” almacena las palabras a utilizar en el motor.
- equation\_statuses:** representa el estado de una ecuación, en la cual se guarda una referencia a la consulta de búsqueda y el sitio web utilizado, el último mes en el que fue ejecutada la ecuación y el último índice obtenido de la búsqueda.
- equation\_articles:** hace de intermediario entre las entidades *equation\_statuses* y *articles* para registrar qué ecuaciones se utilizaron para los artículos que fueron encontrados.
- articles:** se guardan los artículos obtenidos por el extractor en cada ejecución, es utilizada por el módulo NLP para obtener y procesar el contenido de los mismos.
- sites:** aquí se guardan los sitios webs de interés para ser utilizados por las ecuaciones de búsqueda, de esta forma es más sencillo agregar o quitar determinados sitios.
- selectors:** guarda los distintos selectores utilizados para cada sitio web, el cual la ecuación de búsqueda también hace uso para aplicarlos sobre los HTML y obtener así los textos de cada artículo.
- users:** se guardan los usuarios del sistema, con su correo electrónico y contraseñas encriptadas.
- tokens:** es creada por el framework AdonisJS por defecto.
- schedulers:** se guarda la configuración del planificador, sus atributos permiten guardar la periodicidad actual del mismo, cuál es la periodicidad a ejecutar por día y por mes, cuál es el límite de solicitudes por día que se pueden realizar y si el planificador debe estar o no activo al iniciar la aplicación.
- adonis\_schema:** es usada por el framework AdonisJS.

### 3. Recuperación de la Información

#### 3.1. Extracción de Artículos

Para la recuperación de los artículos que contengan información sobre profesionales graduados, se construyó un extractor de artículos que, en conjunto con el planificador, realizan todo el proceso de extracción y persistencia de los mismos. El esquema general del extractor se muestra en la Figura 2 y a continuación se describen sus componentes.



**Figura 2.** Esquema del proceso de extracción de artículos.

**A. Ecuación de búsqueda:** El extractor recibe como entrada una ecuación de búsqueda, se entiende por ecuación a un conjunto de parámetros que contienen los valores necesarios para que Google pueda realizar su búsqueda. El esquema general de una ecuación de búsqueda se muestra a continuación:

```
{"extractor": <nombre del extractor>
  "eq": {"q": <palabras a buscar>,
        "siteSearch": <sitio donde buscar>,
        "cx": <id del buscador personalizado>,
        "key": <key del usuario de Google>,
        "siteSearchFilter": "i"},
  "selectors": <array de selectores a aplicar>}}
```

**extractor:** permite manejar múltiples extractores.

**q:** consulta de búsqueda, Google lo utiliza como las palabras “clave” en su buscador.

**siteSearch:** sitio en donde Google realizará su búsqueda, esto permite filtrar los sitios webs y orientar las búsquedas en determinados lugares de interés (en este caso educativos y de noticias).

**siteSearchFilter:** un indicador de que el parámetro “siteSearch” debe tenerse en cuenta, por defecto su valor es siempre “i” (*include*).

**start:** un número que indica a partir de qué índice se debe comenzar la búsqueda, este índice se agrupa de 10 en 10 por cada página de los resultados de Google por lo cual haciendo cálculos se puede obtener resultados a partir de determinadas páginas.

**selectores:** un conjunto de valores que son utilizados para extraer efectivamente el texto incluido en el artículo. Cada página posee un conjunto de selectores que permiten obtener partes de la misma.

**cx:** ID del buscador personalizado. Éste se obtiene al crear un buscador en Google CSE en donde además se nos otorga una *key* para hacer uso de la misma.

**B. Búsqueda:** Luego de recibir la ecuación, el extractor hace la llamada a la API de Google CSE para obtener los resultados, además se le provee de otros parámetros (como el id de cliente y la key) para poder hacer uso de la misma. La API retorna un JSON con los resultados obtenidos a causa de la búsqueda.

**C. Filtro:** Tras recibir los resultados de búsqueda, se filtran algunos enlaces de sitios webs que contengan determinadas palabras para evitar manipular información innecesaria. En nuestro caso, se filtran los enlaces que contengan la palabra “tags” ya que dichas páginas no brindan ningún tipo de información. Además se filtran los artículos que ya hayan sido obtenidos anteriormente para evitar procesar repetidas veces una misma información.

**D. Extracción:** En este paso se realiza la extracción iterando sobre cada resultado de búsqueda de Google y obteniendo todo el contenido de los enlaces *en crudo* para su posterior procesamiento.

La extracción se realiza accediendo secuencialmente a los “links” devueltos por Google sobre cada artículo encontrado en el resultado de búsqueda, mediante una librería se hace una llamada HTTP al enlace obteniendo así el sitio completo en formato HTML.

**E. Aplicación de selectores:** Se aplican sobre los HTML obtenidos en el paso anterior, quienes dividen y obtienen las porciones de texto de interés de cada artículo. Los selectores aplicados son diferentes para cada sitio por lo que cada resultado de búsqueda de Google devuelve todo lo recuperado en un determinado sitio para facilitar su procesamiento.



**Figura 3.** Ejemplo de uso de selectores Diario El Chubut

Los selectores se obtienen analizando la estructura HTML del sitio web que se desea observando en qué secciones se suelen ubicar los textos de los mismos. Para ilustrar el proceso se usan las noticias de la página <https://www.elchubut.com.ar> (ver Figura 3). Utilizando la herramienta de “inspeccionar elemento” del navegador, se puede identificar el bloque correspondiente al texto del artículo, en este caso con el atributo `class=notapage_cuerpo`, así que éste es el selector de dicha página.

Al aplicar el selector y guardarlo en una variable, podemos ver que posee una lista de un elemento con el bloque completo de texto, accediendo a ella se puede obtener el texto del mismo así como otra información (metadatos) que se almacenan en el sitio. Como último paso este selector se guarda en el sistema para su uso durante las extracciones automáticas. Este proceso debe ser realizado con cada página distinta ya que se suelen utilizar distintos selectores y la estructura HTML varía dependiendo del sitio.

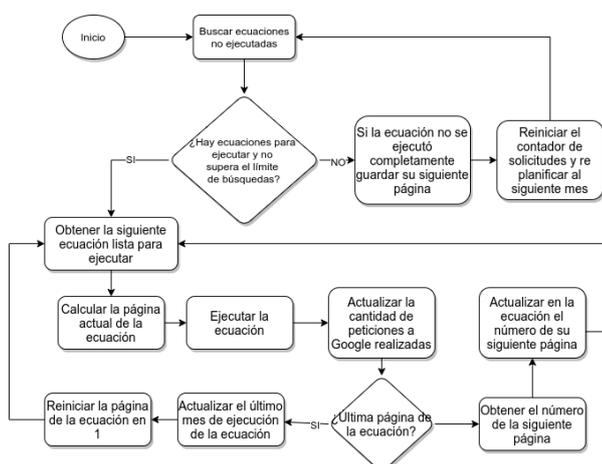
**F. Persistencia:** Como último paso, se persisten los artículos extraídos, en particular los textos obtenidos así como el enlace de donde vino, la ecuación

utilizada, etc. Estos artículos se marcan con una bandera que indica que están listos para ser procesados por el NLP quien es el siguiente en actuar.

### 3.2. Planificación

Para poder realizar la extracción de artículos de forma automatizada, se tomó la decisión de implementar un planificador para la ejecución periódica del extractor y la obtención constante de información. Esta ejecución se ajusta automáticamente a los parámetros de ejecución preconfigurados al iniciar la aplicación y los mismos pueden ir variando dependiendo del estado de la ejecución como por ejemplo: se ejecutaron todas las ecuaciones o se llegó al límite de búsquedas por día.

Esta decisión da como resultado la implementación de un planificador automatizado que se ejecuta en el servidor en intervalos configurables de tiempo. Esto permite que no se necesite de una persona involucrada en la búsqueda y extracción sino que se deja la responsabilidad a dicho proceso. En la Figura 4 se detalla como es el funcionamiento del planificador.



**Figura 4.** Esquema de funcionamiento del planificador

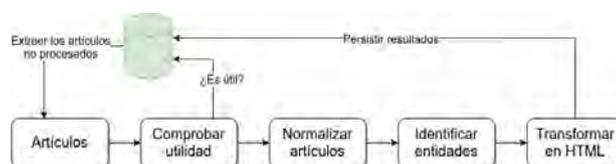
El planificador se implementó para acceder a tantos resultados de búsqueda como sea posible, esto significa que puede ocurrir una situación en la cual una búsqueda quede “por la mitad” y debe retomarse al día siguiente debido al límite de 100 búsquedas por día, impuesto por Google. Un inconveniente que posee la API de Google en su uso gratuito es que no se puede acceder más allá de la página 10 de los resultados, por lo que es un factor a tener en cuenta al momento de hacer la búsqueda.

Como periodicidad, se definió que el planificador deberá ejecutar una vez al día hasta llegar al límite de las 100 peticiones a la API de Google por mes

mientras aún haya ecuaciones sin ejecutar. En caso de que se hayan ejecutado todas las ecuaciones disponibles, el planificador se auto-replanaifica para ser ejecutado al mes siguiente. Estos parámetros pueden ser modificados para ser utilizados por ejemplo en testeos del planificador, haciendo que su ejecución sea en intervalos de pocos minutos para asegurar su correcto funcionamiento.

#### 4. Procesamiento de Artículos

Para el procesamiento del texto se optó por utilizar una librería llamada SpaCy [8], al ser una librería de Python el módulo completo de procesamiento de texto se realizó en este lenguaje. En la Figura 5 se muestra el esquema inicial que se planteó para este módulo.



**Figura 5.** Esquema del procesamiento de artículos

Una vez que los artículos son extraídos y almacenados en la base de datos por el planificador, se dispone a ejecutar el módulo de NLP. Los pasos que se ejecutan en el módulo son:

1. Los artículos se guardan con una bandera `is_processed` que indica si el artículo ya fue o no procesado por el módulo NLP. Se obtiene de la base de datos todos los artículos hasta el momento sin procesar para comenzar con el análisis y procesamiento.
2. Se comprueba la utilidad del artículo, esto se hace simplemente observando el contenido del título del mismo, si presenta las palabras “colación”, “graduados” u otras similares, se lo tiene en cuenta como un artículo de utilidad y se lo marca como un artículo útil en la base de datos. Esto se debe a que muchos artículos obtenidos por el extractor si bien pueden contener las palabras de búsqueda deseadas, quizás su contenido no es del todo preciso o es diferente a lo esperado. Por lo que realizar un análisis y filtro de los artículos basado en su contenido es importante para poder diferenciarlos.
3. Como tercer paso, se eliminan los caracteres que puedan perjudicar al NLP al momento de identificar las entidades. En este paso se detectó que en las distintas fuentes de los artículos, el carácter coma (,) generaban ruido al NLP, dado que si están mal ubicadas o hay mucha cantidad de ellas, puede generar confusión al momento de obtener un nombre y apellido de una persona.
4. Una vez quitado en el artículo los caracteres ruidosos, en este paso se ejecuta el NLP, para que realice la identificación de nombres y apellidos de profesionales.

Para la identificación de nombres y apellidos, se usa como base es el dataset de SpaCy `es_core_news_md`, donde en la configuración se le indica que identifique solamente las entidades personas `PER` y se le pasa un `template` para que cuando identifique a la persona genere el HTML correspondiente.

5. Cuando se detectan los apellidos y nombre, se transforma el artículo a un formato HTML, que resalta los nombres encontrados y permitirá desde la interfaz de usuario confirmar, y asignar una carrera e institución a cada uno de los profesionales.

## 5. Resultados



Figura 6. Interfaz de usuario para supervisión.

**Supervisión de resultados.** Como parte del procesamiento final de los artículos, se incluye una interfaz de usuario donde se muestran los resultados obtenidos por el procesamiento, y a partir de allí, mediante un asistente, se solicita la confirmación o corrección de la información de los profesionales, como así también la asignación de la carrera/título e institución de cada uno de ellos. Una captura se muestra en la Figura 6.

**Desempeño.** El funcionamiento de los extractores y el planificador obtuvieron buenos resultados así como la información que se lograba obtener al utilizar el módulo NLP con la librería SpaCy. La ejecución de una ecuación de búsqueda demoró aproximadamente entre 1 y 3 minutos, esto depende de varios factores como: la conexión a internet, la cantidad de páginas que devuelva la búsqueda (no siempre se puede alcanzar el límite de 10 páginas por ecuación) y la cantidad de veces en la que la ecuación sea ejecutada por encima de las demás debido a que no se ejecuta una ecuación completa de una sola vez sino que una vez extraída una página se cambia a la siguiente ecuación lista para ejecutar.

La cantidad de artículos obtenido por ecuación rondan entre los 30 y 100 artículos generalmente, y el número de profesionales puede ser muy disperso ya

que se han encontrado artículos con más de 200 profesionales y otros casos con 20 o 30.

## 6. Conclusiones

La implementación del proyecto, logró mostrar la factibilidad de construir una herramienta basada en técnicas de recuperación de información y procesamiento de lenguaje natural, para asistir a las PYMEs de Chubut en su mejora de competitividad.

Como trabajos futuros se propone la ampliación del espectro de búsquedas para identificar profesionales no a partir de la noticia de su graduación sino mediante sus menciones en noticias y la vinculación de dichos nombres con potenciales usuarios de redes sociales, así como la implementación de un mecanismo de *verificación (externa) y validación (propia) de identidad*.

## Referencias

1. Amini, B., Ibrahim, R., Othman, M.S., Selamat, A.: Capturing scholar's knowledge from heterogeneous resources for profiling in recommender systems. *Expert Systems with Applications* 41(17), 7945 – 7957 (2014), <http://www.sciencedirect.com/science/article/pii/S0957417414003807>
2. Fareri, S., Fantoni, G., Chiarello, F., Coli, E., Binda, A.: Estimating industry 4.0 impact on job profiles and skills using text mining. *Computers in Industry* 118, 103222 (2020), <http://www.sciencedirect.com/science/article/pii/S0166361519309327>
3. Ibañez, J., Ball, F.: Industria manufacturera año 2010 : Observatorio pyme regional provincia del chubut. Tech. Rep. CDD 338.47, Fund. Observatorio Pyme; Bononiae Libris; GECSEA Patagonia - Universidad Nacional de la Patagonia San Juan Bosco, Buenos Aires (Sept 2011)
4. Jackson, P., Schilder, F.: Natural language processing: Overview. In: Brown, K. (ed.) *Encyclopedia of Language Linguistics (Second Edition)*, pp. 503 – 518. Elsevier, Oxford, second edition edn. (2006), <http://www.sciencedirect.com/science/article/pii/B0080448542009275>
5. Moreno, M., Gower, A., Brittain, H., Vaillancourt, T.: Applying natural language processing to evaluate news media coverage of bullying and cyberbullying. *Prevention Science* 20 (08 2019)
6. Rybinski, K.: Ranking professional forecasters by the predictive power of their narratives. *International Journal of Forecasting* (2020), <http://www.sciencedirect.com/science/article/pii/S0169207020300601>
7. Selva, R.A.: Ventajas competitivas de la provincia del chubut, herramientas para la toma de decisiones. Tech. Rep. 49246, CONSEJO FEDERAL DE INVERSIONES, Buenos Aires (Nov 2011)
8. SpaCy: Industrial-Strength Natural Language Processing. <https://spacy.io/> (2020), online; accessed 20 julio 2020
9. Zhang, H., Boons, F., Batista-Navarro, R.: Whose story is it anyway? automatic extraction of accounts from news articles. *Information Processing Management* 56(5), 1837 – 1848 (2019), <http://www.sciencedirect.com/science/article/pii/S0306457318306101>

## COVID-19: un análisis comparativo de Apps

Juan Fernández Sosa <sup>1</sup> , Verónica Aguirre <sup>1</sup> , Lisandro Delia <sup>1</sup> , Pablo Thomas <sup>1</sup> , Leonardo Corbalán <sup>1</sup> , Patricia Pesado <sup>1</sup> 

<sup>1</sup> Instituto de Investigación en Informática LIDI (III-LIDI). Facultad de Informática – Universidad Nacional de La Plata, La Plata, Argentina.  
Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)  
{jfernandez, vaguirre, ldelia, pthomas, corbalan, ppesado}@lidi.info.unlp.edu.ar

**Resumen.** El virus que produce la enfermedad de COVID-19 aún continúa extendiéndose por el planeta. Hasta el momento, no se cuenta con una vacuna o cura efectiva, por lo que contener al virus en las primeras fases es un trabajo crucial que deben realizar aquellos países afectados. Se han propuesto múltiples soluciones tecnológicas para detener y contener el avance del virus y no saturar los sistemas de salud. Las aplicaciones móviles resultan oportunas porque son ejecutadas en dispositivos que están extendidos en la sociedad. En este trabajo se realiza un análisis comparativo de diferentes aplicaciones de software para dispositivos móviles promovidas por los gobiernos como herramientas tecnológicas para combatir la pandemia de COVID-19.

**Keywords:** Dispositivos móviles, aplicaciones móviles, apps, COVID-19, coronavirus.

### 1. Introducción

El 31 de diciembre de 2019 la Oficina de la OMS (Organización Mundial de la Salud) en China fue informada de varios casos de neumonía de causas desconocidas que posteriormente se identificaría como una nueva enfermedad, denominada COVID-19. El 11 de marzo de 2020 el director general de la OMS informó que por los niveles de propagación y gravedad, el COVID-19 puede caracterizarse como una pandemia [1]. Esto significa que el virus que provoca dicha enfermedad logró extenderse hacia una gran cantidad de países y continentes, afectando a un estremecedor número de personas.

El primer caso registrado fuera de China fue en Tailandia el 13 de enero de 2020 y 7 días más tarde se confirmaba el primero del continente americano, en Estados Unidos. Brasil notificó el primer caso en América Latina y del Caribe el 26 de febrero del 2020. Al 18 de Julio del 2020 se han registrado más de 14 millones de casos alrededor del mundo, sumando más de 600 mil muertes.

Las diferentes formas de contagio del virus hacen que su velocidad de transmisión sea muy alta [2]. Por el momento no se ha encontrado una vacuna, por lo que los esfuerzos por parte de los países afectados están puestos en contener y desacelerar el número de contagios para no saturar sus sistemas de salud. Esto incluye el cierre de fronteras, cuarentenas, distanciamiento social e identificación de los casos positivos. Estudios basados en modelos matemáticos han demostrado que las personas

asintomáticas han causado alrededor del 80% de las infecciones [3]. La detección y respuesta temprana son la clave para contener el brote inicial.

Las nuevas tecnologías han permitido cambiar la forma en que las enfermedades son identificadas, rastreadas, atendidas, tratadas y percibidas, logrando mejores resultados que en experiencias pasadas [4]. En este contexto y con el fin de frenar los contagios, se recurrió a los dispositivos móviles. Dichos dispositivos que están extensamente difundidos en la sociedad poseen una gran capacidad de procesamiento, además de otras características de hardware que permiten pensar múltiples soluciones en pos de contener el avance de la pandemia.

En este sentido, se han realizado diferentes desarrollos atacando múltiples aristas. Por ejemplo, aplicaciones de rastreo de contactos que permiten aprovechar las capacidades de los dispositivos móviles para detectar si existió cercanía con alguna persona portadora del virus [3][5]; aplicaciones destinadas a informar y asistir a la sociedad, proveyendo de información fidedigna de la situación en su región; servicios; información de la pandemia y métodos de prevención.

El presente trabajo, tiene como objetivo realizar un análisis comparativo de diferentes aplicaciones de software para dispositivos móviles promovidas por los gobiernos como herramientas tecnológicas para combatir la pandemia de COVID-19.

En la sección 2 se detallan los criterios de selección de las aplicaciones y países de estudio. La sección 3 presenta información general y técnica de cada una de las aplicaciones seleccionadas. En la sección 4 se realiza una comparación en función de los permisos requeridos por dichas aplicaciones y en la sección 5 se analizan aspectos funcionales de cada aplicativo. Finalmente se presenta un análisis de resultados, conclusiones y trabajos futuros.

## 2. Selección de Apps

En la actualidad, el mundo se encuentra batallando contra un virus para el cual no se ha desarrollado aún una vacuna efectiva. El objetivo por parte de los gobiernos de diferentes países, regiones y territorios es el de detener la velocidad de contagios para no saturar los sistemas de salud.

Desde el sector público se han desarrollado múltiples soluciones tecnológicas, con el fin de detener el avance del virus en sus territorios, focalizándose en cuatro puntos [6][7]:

- Generar conciencia, educar y comunicar activamente sobre la pandemia.
- Encontrar casos activos, y determinar cadenas de contagio para de este modo aislar a personas sospechadas de haber contraído el virus.
- Monitorear el comportamiento y crecimiento de la pandemia.
- Testear a todas las personas que presenten síntomas compatibles con la enfermedad.

Un ejemplo del uso de la tecnología en este contexto son las aplicaciones de móviles, que permiten a los gobiernos ofrecer multiplicidad de servicios a las personas a través de sus dispositivos móviles.

Este trabajo analiza 22 aplicaciones móviles promovidas por los gobiernos de 22 países pertenecientes a todos los continentes alcanzados por la pandemia. Estas se encuentran mencionadas en la tabla 1.

12 de las 22 aplicaciones analizadas en este artículo provienen de países americanos, justificando su elección en que, al momento de su escritura, América es la región del planeta más afectada por la pandemia. Se han analizado 10 aplicativos de la región latinoamericana incluyendo los siguientes países: Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Perú y Uruguay. Por otra parte, se ha estudiado el caso de Estados Unidos y Canadá correspondiente a América del Norte

En lo que respecta a la región asiática se han seleccionado diferentes países utilizando diversos criterios. Se incluyó a la India, por ser el país de la región con más casos en la actualidad. Por otra parte, se eligió Corea del Sur por ser uno de los primeros países donde el virus ingresó inmediatamente después de China y por su estrategia basada en el rastreo de casos, testeo y tratamiento [8]. Finalmente se consideró a Singapur por ser pionero en la adopción del uso de dispositivos móviles para el rastreo de contactos mediante tecnología bluetooth [9].

En el caso de China, país donde se originó la pandemia, no se ha podido encontrar una aplicación móvil destinada exclusivamente al contexto de la pandemia a nivel nacional. En dicho país se han empleado aplicaciones preexistentes con mucha adopción en la sociedad como lo son Alipay y WeChat [10]. En ambas aplicaciones se agregó una nueva funcionalidad para que cada usuario pueda escanear un código QR y conocer si estuvo en contacto con algún caso positivo de COVID-19. En este país, el rastreo de contactos se realiza utilizando Big Data sobre datos proporcionados por las autoridades de transporte y salud. Al escanear los códigos QR, se controla y gestiona el acceso a diferentes zonas o lugares del país. Dicho código cambia de color entre verde, amarillo y rojo dependiendo del riesgo que representen los usuarios por haber estado en contacto con algún caso positivo de COVID-19.

El continente europeo fue el siguiente en enfrentarse a la pandemia, y uno de los más afectados por ésta. Se seleccionaron aplicaciones promovidas por los gobiernos de España, Francia, Italia y Alemania, países gravemente afectados. Si bien el Reino Unido está posicionado como uno de los países con más contagios a nivel mundial, no está contemplado en este trabajo debido a que no existe, hasta el momento, una aplicación oficial. Sin embargo, ha habido intentos de implementación de una aplicación móvil para el rastreo de contactos utilizando la misma tecnología que Singapur. Dicha aplicación fue probada en la Isla de Wight, pero a partir de los resultados obtenidos se optó por revisar y estudiar una solución aplicable a nivel nacional, implementando la tecnología anunciada por Apple y Google [11][12].

Se han seleccionado de Oceanía los dos países con más casos registrados: Australia y Nueva Zelanda.

Por último, y siguiendo la misma regla, los dos países del continente africano con más casos de COVID-19 son Sudáfrica y Egipto. Este trabajo analiza la aplicación móvil adoptada en Egipto, impulsada por el sector público. Sudáfrica tuvo un intento de aplicativo para el rastreo de contactos, llamado Covid-ID [13] desarrollado por un equipo multidisciplinario incluyendo científicos voluntarios, banqueros, empresarios

y estudiantes. Sin embargo, no resultó de interés en este país por lo que no pudo convertirse en un proyecto nacional [14].

**Tabla 1.** Listado de aplicaciones móviles analizadas en este trabajo.

<b>Aspectos Generales</b>			
<b>Región</b>	<b>País</b>	<b>Nombre aplicativo</b>	<b>Desarrollador</b>
<b>Latinoamérica</b>	<b>Argentina</b>	<i>CuidAr COVID-19 Argentina</i>	Presidencia de la Nación Argentina
	<b>Bolivia</b>	<i>Bolivia Segura</i>	Agetic Bolivia
	<b>Brasil</b>	<i>Coronavirus - SUS</i>	Gobierno de brasil
	<b>Chile</b>	<i>CoronaApp</i>	Gobierno de Chile - Gobierno Digital
	<b>Colombia</b>	<i>CoronaApp-Colombia</i>	Instituto nacional de salud y gobierno de colombia
	<b>Costa Rica</b>	<i>EDUS</i>	Caja Costarricense de Seguro Social
	<b>Ecuador</b>	<i>SaludEC</i>	Prichsouth Tecnologías de Sur
	<b>México</b>	<i>COVID-19MX</i>	Secretaría de Salud MX
	<b>Perú</b>	<i>Peru en tus manos</i>	Presidencia del Consejo de Ministros
	<b>Uruguay</b>	<i>Coronavirus UY</i>	AGESIC
<b>América del Norte</b>	<b>Estados Unidos</b>	<i>Apple COVID-19</i>	Centros para el Control y la Prevención de Enfermedades (CDC), la Casa Blanca y la Agencia Federal para el Manejo de Emergencias (FEMA).
	<b>Canadá</b>	<i>Canada COVID-19</i>	Salud de Canadá
<b>Asia</b>	<b>India</b>	<i>Aarogya Setu</i>	Gobierno de India
	<b>Corea de Sur</b>	<i>self-quarantine safety protection</i>	Ministerio de Administración Pública y Seguridad
	<b>Singapur</b>	<i>TraceTogether</i>	Agencia de tecnología del

			gobierno
<b>Europa</b>	<b>España (gral)</b>	<i>Asistencia COVID-19</i>	Gobierno de España. Ministerio de Asuntos Económicos y Transf. Digital
	<b>Alemania</b>	<i>Corona-Warn-App</i>	Instituto Robert Koch
	<b>Francia</b>	<i>StopCovid France</i>	Gobierno de Francia
	<b>Italia</b>	<i>Immuni</i>	Presidencia del Consejo de Ministros
<b>Oceanía</b>	<b>Australia</b>	<i>COVIDSafe app</i>	Departamento de Salud de Australia
	<b>Nueva Zelanda</b>	<i>NZ COVID Tracer</i>	Ministerio de Salud de NZ
<b>África</b>	<b>Egipto</b>	<i>Salud de Egipto</i>	Ministerio de Salud de Egipto

### 3. Análisis de información técnica

El análisis aquí presentado comienza con el relevamiento de información técnica de cada una de las 22 aplicaciones móviles seleccionadas. La mayoría de esta información se encuentra accesible desde las diferentes tiendas de distribución de dichos aplicativos. A continuación se desarrollará cada una de las características analizadas, cuyos resultados pueden apreciarse en la tabla 2.

- **Plataformas:** Se busca conocer si las aplicaciones están disponibles para Android y para iOS, plataformas líderes actualmente en el mercado.
- **Enfoque y Tecnología de desarrollo:** en el desarrollo de aplicaciones de software para dispositivos móviles, pueden emplearse diferentes estrategias [15] [16]. La estrategia denominada “nativa” consiste en el desarrollo de una aplicación específica para cada una de las plataformas en las que se desea distribuir el aplicativo. Por otro lado, la estrategia “multiplataforma” consiste en desarrollar un solo proyecto, reutilizando el mismo código fuente para ser ejecutado en las diferentes plataformas. Las aplicaciones multiplataforma pueden ser desarrolladas empleando una gran diversidad de tecnologías.
- **Fecha de lanzamiento:** indica la fecha en que se publicó la aplicación en las tiendas de distribución para cada una de las plataformas en las que el aplicativo está disponible.
- **Repositorio público:** en el caso de aquellas aplicaciones que esté disponible, se muestran los repositorios públicos en donde reside el código fuente de cada una de ellas.
- **Tamaño:** especifica el tamaño en MB que requiere la aplicación para ser instalada en los dispositivos, este dato está discriminado por plataforma.

### 3.1 Recolección de datos

Para recolectar la información de las 22 aplicaciones analizadas se procedió de diferentes formas: con el objetivo de comprobar la disponibilidad de las aplicaciones en las plataformas Android e iOS, se buscó en las tiendas respectivas por el nombre del aplicativo.

Para determinar el enfoque y la tecnología de cada una de las aplicaciones, se analizaron los diferentes archivos apk de la versión de la aplicación para Android. Estos son archivos comprimidos de los cuales se puede obtener información sobre las librerías empleadas en el desarrollo. Para lograr esto último se utilizó la herramienta “dex2jar” [17] la cual permite convertir un archivo .apk en uno con extensión .jar de Java que contiene dichas librerías. Para finalizar el proceso, se utilizó la herramienta Java Decompiler [18] con el objetivo de examinar ese archivo resultante y determinar la tecnología de desarrollo empleada.

La fecha de lanzamiento y el tamaño de cada aplicativo se obtuvo de las fichas de descripción de las aplicaciones en las respectivas tiendas de distribución (App Store para iOS y Google Play Store para Android). Algunas aplicaciones presentan versiones con fecha anterior a la pandemia, es decir que algunos países se han inclinado por adaptar aplicaciones que ya se encontraban funcionando. En ese caso, se ha tomado como fecha de lanzamiento la fecha en que dicha aplicación ha sido modificada con funcionalidades para afrontar la pandemia.

Por último, los repositorios públicos fueron localizados a partir de la mención explícita en los sitios webs oficiales que promueven las aplicaciones, o como resultado de la búsqueda del código de la aplicación en GitHub, una de las plataformas de gestión de proyectos más conocida.

**Tabla 2.** Aspectos técnicos analizados.

\* Países que utilizaron aplicaciones preexistentes

Aspectos Técnicos								
País	Plataformas		Enfoque y Tecnología de desarrollo	Fecha de lanzamiento		Repositorio público	Tamaño del aplicativo [MB]	
	Android	iOS		Android	iOS		Android	iOS
<b>Argentina</b>	Si	Si	Nativo	23/3/2020	28/3/2020	Si	11	17,3
<b>Bolivia</b>	Si	Si	Multiplataforma (NativeScript)	12/3/2020	13/3/2020	No	15	67
<b>Brasil</b>	Si	Si	Multiplataforma (Ionic)	14/3/2020 *	28/2/2020 *	Si	22	60
<b>Chile</b>	Si	Si	Multiplataforma (Flutter)	16/4/2020	15/4/2020	No	26	104,3
<b>Colombia</b>	Si	Si	Nativo	8/3/2020 *	15/3/2020	No	9,5	70
<b>Costa Rica</b>	Si	Si	Nativo	26/3/2020 *	28/3/2020 *	No	7,4	37,6

<b>Ecuador</b>	Si	Si	Multiplatafor ma (Xamarin)	19/3/2020	4/4/2020	No	42	100
<b>México</b>	Si	Si	Multiplatafor ma (Appcelerator)	28/3/2020	6/4/2020	No	3,36	18,5
<b>Perú</b>	Si	Si	Nativo	3/4/2020	6/4/2020	No	6,6	39
<b>Uruguay</b>	Si	Si	Nativo	20/3/2020	20/3/2020	No	23	94
<b>Estados Unidos</b>	No	Si	Nativo	-	27/3/2020	No	-	17,1
<b>Canadá</b>	Si	Si	Multiplatafor ma (Capacitor)	30/3/2020	30/3/2020	No	8,2	26
<b>India</b>	Si	Si	Nativo	2/4/2020	2/4/2020	Si	4	35,9
<b>Corea de Sur</b>	Si	Si	Nativo	7/3/2020	13/3/2020	No	14	26,1
<b>Singapur</b>	Si	Si	Nativo	19/3/2020	18/3/2020	Si	23	62,1
<b>España (gral)</b>	Si	Si	Multiplatafor ma (ReactNative)	5/4/2020	6/4/2020	No	22	26,5
<b>Alemania</b>	Si	Si	Nativo	12/06/2020	15/06/2020	Si	31	18,9
<b>Francia</b>	Si	Si	Nativo	30/05/2020	02/06/2020	No	22	26,5
<b>Italia</b>	Si	Si	Nativo	29/05/2020	01/06/2020	Si	27	40,5
<b>Australia</b>	Si	Si	Nativo	25/4/2020	26/4/2020	Si	11	39,9
<b>Nueva Zelanda</b>	Si	Si	Nativo	15/5/2020	19/5/2020	No	16	54,9
<b>Egipto</b>	Si	No	Nativo	7/4/2020	-	No	12	-

#### 4. Análisis de permisos

El propósito de un permiso es proteger la privacidad del usuario [26]. Las aplicaciones solicitan de manera explícita diferentes permisos a los usuarios para acceder a datos privados (como por ejemplo contactos) así como también a funciones del sistema (como por ejemplo la cámara y GPS) con el fin de que una aplicación no impacte negativamente a otras aplicaciones, el sistema operativo o el usuario.

A continuación, se listan y explican los permisos analizados en este trabajo que conllevan a exponer datos de la privacidad de los usuarios. Para determinar los

permisos requeridos por las aplicaciones estudiadas se utilizó la ficha de la aplicación en su versión para Android. Los resultados se presentan en la tabla 3.

- **Ubicación:** este permiso detalla información confidencial del usuario, ya que permite identificar lugar de trabajo, vivienda, entre otros. La ubicación puede ser determinada empleando diferentes técnicas como por ejemplo las redes de telefonía móvil, Wi-Fi, redes de sistema de posicionamiento global (GPS) y Bluetooth para determinar la ubicación aproximada.
- **Cámara:** específica el acceso a la cámara o a la captura de imágenes/videos desde el dispositivo.
- **Micrófono:** otorga permiso para el acceso al audio del micrófono desde el dispositivo.
- **Bluetooth:** permite a los dispositivos intercambiar datos de manera inalámbrica con otros dispositivos Bluetooth. Para esto, un dispositivo se habilita para recibir solicitudes de conexión entrantes. Después otro dispositivo (luego de un proceso de detección) acepta la solicitud de sincronización, y se completa un proceso de conexión en el que se intercambian claves de seguridad.
- **Almacenamiento de datos:** permite leer y escribir información desde el almacenamiento USB y almacenamiento externo del dispositivo.
- **Teléfono:** se utiliza para funciones de telefonía como por ejemplo la realización y desvío de llamadas, detección de una llamada entrante, etc.
- **ID del dispositivo:** el ID es un identificador único de un dispositivo. Se crea en el primer inicio del dispositivo y se mantiene a menos que se haga un restablecimiento de fábrica. Permite compartir esta identificación con aplicaciones o dispositivos.

**Tabla 3.** Permisos solicitados por las aplicaciones.

\* Para el caso de Estados Unidos, no se pudo determinar los permisos solicitados, al no poder tener acceso a la aplicación por limitaciones geográficas.

País	Ubicación	Cámara	Micrófono	Bluetooth	Almacenamiento de datos	Teléfono	ID del dispositivo
Argentina	X	X	-	-	-	-	-
Bolivia	X	-	-	-	X	X	-
Brasil	X	-	-	-	-	X	-
Chile	X	-	-	-	X	-	-
Colombia	X	-	-	X	-	X	X
Costa Rica	-	X	-	-	X	X	X
Ecuador	X	-	-	-	X	-	-
México	X	-	-	X	-	-	-
Perú	X	-	-	X	-	-	-

Uruguay	-	X	X	X	-	-	-
Estados Unidos *	-	-	-	-	-	-	-
Canadá	X	-	-	-	X	-	-
India	X	X	-	X	-	-	-
Corea de Sur	X	X	-	X	X	X	X
Singapur	X	X	-	X	X	-	-
España (gral)	X	-	-	-	-	-	-
Alemania	-	X	-	X	-	-	-
Francia	X	X	-	X	-	-	-
Italia	-	-	-	X	-	-	-
Australia	X	-	-	X	-	-	-
Nueva Zelanda	-	X	-	-	-	-	-
Egipto	X	-	-	-	X	-	-

## 5. Análisis de funcionalidades

Como se explicó anteriormente, en este contexto de pandemia se desarrollan soluciones tecnológicas con diferentes propósitos. A continuación, se listan una categorización de funcionalidades que las aplicaciones estudiadas pueden implementar. Los resultados obtenidos se presentan en la tabla 4.

- **Autodiagnóstico:** permiten al usuario responder una serie de preguntas acerca de su estado de salud. A partir de sus respuestas, la aplicación determina si el usuario presenta síntomas compatibles con COVID-19 y brinda recomendaciones. Algunas aplicaciones para ampliar esta funcionalidad pueden acudir a la geolocalización del usuario que realizó el autodiagnóstico para mostrarle centros de salud cercanos, o alertar a las autoridades pertinentes. El objetivo es reconocer zonas de riesgo.
- **Informativa:** brindan información centralizada y oficial relativa a la situación actual de la pandemia en el país, así como también información sobre la pandemia, síntomas de la enfermedad, prevenciones y cuidados. Junto con el autodiagnóstico, es común encontrar también este tipo de funcionalidad en sitios web de gobierno.
- **Gestión interna para la pandemia:** funcionalidades directamente relacionadas con medidas que han adoptado los gobiernos en función de controlar y gestionar el paso de la pandemia en sus territorios, como por ejemplo el control de la movilidad a partir de códigos QR, disposición de vías de comunicación a nivel regional, mención de centros de salud cercanos al usuario, soluciones que implementan técnicas de geofencing de modo de alertar a las personas si se encuentran próximas a una zona de riesgo, vías de denuncia ciudadana, entre otras.
- **Rastreo de contactos:** consiste en utilizar alguna técnica para identificar a las personas contagiadas y quienes estuvieron cerca de un caso positivo, con el fin de aislarla y de esa manera cortar la cadena de contagios. Esta

funcionalidad se puede llevar a cabo mediante uso de tecnología bluetooth, códigos QR, entre otras.

- **Control de ubicación de cuarentena:** utilizan la ubicación de dispositivo para rastrear en qué lugar pasa la cuarentena un usuario permitiendo conocer si la persona en cuarentena está saliendo de ese lugar.
- **Registro de usuarios mediante su identificación nacional:** requieren que el usuario ingrese datos personales para su funcionamiento.

**Tabla 4.** Funcionalidades de las aplicaciones estudiadas.

\* Solicita permiso Bluetooth pero no especifica que la aplicación realice rastreo de contactos.

País	Autodiagnóstico	Informativa	Gestión interna para la pandemia	Contacto Tracing	Control de ubicación de cuarentena	Registro de usuarios mediante su identificación nacional
Argentina	X	X	X	-	-	X
Bolivia	X	X	X	-	-	X
Brasil	X	X	-	-	-	X
Chile	X	X	X	-	X	X
Colombia	X	X	X	X	X	X
Costa Rica	X	X	-	-	-	X
Ecuador	X	X	X	-	-	X
Mexico	X	X	-	*	-	X
Perú	X	X	X	X	-	X
Uruguay	X	X	X	X	-	X
Estados Unidos	X	X	-	-	-	-
Canada	X	X	X	-	-	-
India	X	X	X	X	-	-
Corea de Sur	X	-	X	X	X	X
Singapur	-	-	-	X	-	X
España (gral)	X	X	X	-	X	X
Alemania	-	-	X	X	-	-
Francia	-	-	-	X	-	-
Italia	-	-	-	X	-	-
Australia	-	-	-	X	-	-
Nueva Zelanda	-	-	-	X	-	-
Egipto	X	X	X	-	-	-

## 6. Análisis de resultados obtenidos

### 6.1 Análisis de información básica

En base a lo analizado en la tabla 2, se desprenden los siguientes resultados

- De las 22 aplicaciones analizadas 20 están disponibles para las plataformas Android y iOS. La aplicación de Egipto está sólo disponible en Android mientras que la aplicación utilizada en Estados Unidos sólo está disponible para iOS.
- Del total de las aplicaciones, 7 han sido desarrolladas empleando un enfoque multiplataforma, mientras que las restantes 15 han sido desarrolladas de manera nativa. De las tecnologías de desarrollo multiplataforma empleadas se ha encontrado un conjunto variado de las que se pueden mencionar Ionic, Capacitor, Xamarin, ReactNative, Flutter, Appcelerator y NativeScript.
- De las 7 aplicaciones multiplataforma, 6 pertenecen a países americanos. 5 de ellos son de la región latinoamericana.
- El promedio de tamaño de las aplicaciones en Android es de 20 MB mientras que para las aplicaciones en iOS es 65 MB.
- De las 22 aplicaciones sólo 7 cuentan con un repositorio público donde se encuentra disponible el código fuente de las mismas. Para el caso de Argentina, si bien se conoce dónde estará alojado el código, al momento de escribir este trabajo no se ha publicado. Por otra parte, India sólo publicó el código correspondiente a la aplicación en Android, que representa el 98% de sus usuarios.
- Brasil y Costa Rica han modificado aplicaciones preexistentes en ambas plataformas analizadas, agregándole contenido o funcionalidad referente a la pandemia. Lo mismo sucede con la versión para dispositivos Android de Colombia, no obstante, la versión para iOS fue creada íntegramente.

### 6.2 Análisis de los permisos

El siguiente análisis se realiza sobre 21 de las 22 aplicaciones estudiadas debido a que no se ha podido tener acceso a la aplicación de Estados Unidos.

- El 76% de las aplicaciones analizadas, requieren permiso de ubicación.
- El permiso de Bluetooth es utilizado por más del 50% de las aplicaciones, evidenciando la cantidad de ellas que lo utilizan para rastreo de contactos.
- La cámara es requerida por 9 aplicaciones del total.
- La única aplicación que requiere el acceso al micrófono es la de Uruguay.
- El almacenamiento de datos es solo requerido por 8 aplicaciones.
- El uso del permiso del teléfono es requerido por 5 aplicaciones, principalmente para permitir el contacto con los centros de salud.
- En cuanto al id del dispositivo solo es requerido por las aplicaciones de Colombia, Costa Rica y Corea del Sur

### 6.3 Análisis de funcionalidades

Por último, a partir del análisis de la tabla 4 se puede indicar:

- 16 de las 22 aplicaciones analizadas permiten a los usuarios realizar un autodiagnóstico a través de la carga de síntomas como fiebre, tos seca, dolor de cabeza, dificultad para respirar, falta de aire, etc. La mayoría de las

aplicaciones que presentan esta funcionalidad, se combinan con la georreferenciación del usuario.

- 16 de estas aplicaciones proveen funcionalidad informativa.
- En referencia a la gestión interna para la pandemia, Brasil, Costa Rica, México, Singapur, Francia, Italia, EEUU, Australia y Nueva Zelanda no utilizan las aplicaciones para realizarla. El resto si lo hace. Como por ejemplo la de Chile, que permite hacer denuncias sobre la violación de la cuarentena; en el caso de la de Corea del Sur permite saber dónde realiza la cuarentena un usuario y se los alerta si están saliendo de su zona de aislamiento.
- Se ha podido establecer que el 50% de las aplicaciones analizadas hacen rastreo de contactos a través del uso de bluetooth, detectando la cercanía entre dos dispositivos. Estas son las aplicaciones de Colombia, Corea del Sur, Singapur, Alemania, Francia, Italia y Australia. En el caso de Perú y Uruguay, utilizan la ubicación del dispositivo para informar los posibles contagios. La aplicación de la India además de bluetooth utiliza el GPS, determinando si por su ubicación, el usuario se encuentra en una de las áreas infectadas. En el caso de Nueva Zelanda, el usuario registra su paso por los diferentes lugares a través del escaneo de códigos QR y si coincide con algún lugar de un usuario infectado, se envía una alerta.
- Con respecto a la localización de usuario para realizar control de la cuarentena, solo la implementan Chile, Colombia, Corea del Sur y España.
- Todas las aplicaciones estudiadas de Latinoamérica requieren que el usuario ingrese información personal, para acceder a otras funcionalidades.

## 7. Conclusiones y Trabajo Futuro

Al momento de escritura de este informe, el mundo se encuentra combatiendo la pandemia de COVID-19. Debido a la alta velocidad de transmisión del virus los gobiernos se han visto obligados a tomar drásticas medidas para contenerlo y no saturar sus sistemas de salud.

La tecnología tiene un rol fundamental que permite obtener mejores resultados que en crisis sanitarias pasadas. Un ejemplo del uso de la tecnología en este contexto de pandemia son las aplicaciones móviles, las cuales permiten a los gobiernos ofrecer servicios a la sociedad tales como información sobre la pandemia; cuidados y prevenciones; posibilidad de autodiagnóstico; rastrear casos de contacto estrecho de aquellas personas que han resultado contagiadas con el virus y monitorear aquellas personas que deben cumplir una cuarentena.

En este trabajo se han analizado 22 aplicaciones móviles promovidas por gobiernos de 22 países como herramientas tecnológicas para hacer frente al COVID-19. Del total de países alcanzados por este estudio, 12 pertenecen al continente americano, región más golpeada por la pandemia. El criterio de selección de países se basó en el impacto de la pandemia en cada país y además se han considerado países pioneros en la adopción de la tecnología para combatir al virus. El análisis realizado abarcó aspectos técnicos, permisos requeridos por las aplicaciones y funcionalidades que éstas proveen.

Las aplicaciones de software presentadas en este trabajo corresponden a aquellas propuestas por los gobiernos a nivel nacional, algunos de los países alcanzados por este análisis presentan otros aplicativos en otros niveles de gobierno (municipales, provinciales, entre otros). Por otra parte, el estudio tuvo en cuenta el estado de los aplicativos al 1 de Julio de 2020. Debido a las características propias del

software, dichas aplicaciones pueden continuar evolucionando en el tiempo, incorporando más funcionalidades que podrán ser analizadas en un trabajo futuro.

La mayoría de las aplicaciones estudiadas tienen su versión para ser ejecutada en las plataformas con más presencia en el mercado: Android y iOS. En cuanto al enfoque de desarrollo utilizado, de 22 aplicaciones 7 optaron por un desarrollo multiplataforma, mientras que el resto optó por un desarrollo nativo. Es interesante notar que, de las 7 aplicaciones desarrolladas con enfoque multiplataforma, 5 pertenecen a países latinoamericanos.

Las funcionalidades comunes a la mayoría de las aplicaciones son la de autodiagnóstico e informativa. La primera permite a los usuarios evaluar su estado de salud y detectar síntomas compatibles con la enfermedad. La segunda funcionalidad provee a la sociedad de información fidedigna acerca de la pandemia, métodos de prevención y cuidados, informes sobre el estado actual del virus en dicho país, etc.

La mitad de las aplicaciones estudiadas presentan la funcionalidad de rastreo de contactos, la cual aprovechando tecnologías presentes en los dispositivos móviles permiten conocer cadenas de contagios. Para este propósito, la mayoría utiliza bluetooth.

En función de los permisos requeridos para su funcionamiento, de las 22 aplicaciones estudiadas 16 solicitan al usuario el permiso para acceder a su ubicación, lo que puede generar cierto rechazo, presentando resistencia en confiar esos datos a los diferentes gobiernos.

Como trabajo futuro se plantea analizar la efectividad de este tipo de soluciones tecnológicas para contener el avance de la pandemia. Las aplicaciones aquí analizadas surgen en respuesta a las primeras fases epidemiológicas, donde el esfuerzo debe enfocarse en brindar información de prevención y contener la velocidad de contagios del virus para no saturar los sistemas de salud. Cuando el mundo llegue a una etapa post pandémica, se debe aprovechar el uso de la tecnología para encontrar soluciones a problemas que surgen como consecuencia del paso del virus en cada territorio.

## Referencias

1. World Health Organization. 11 de Marzo 2020. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. Disponible en <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
2. World Health Organization. 29 de Marzo 2020. *Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations*. Disponible en <https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations>
3. Hernández-Orallo, E., Manzoni, P., Calafate, C. T., & Cano, J. C. (2020). Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: the case of COVID-19. IEEE Access.
4. Wilson, C., & Jumbert, M. G. (2018). *The new informatics of pandemic response: humanitarian technology, efficiency, and the subtle retreat of national agency*. Journal of International Humanitarian Action, 3(1), 1-13.
5. Morley, J., Cowls, J., Taddeo, M., & Floridi, L. (2020). *Ethical guidelines for COVID-19 tracing apps*.
6. Elavarasan, R. M., & Pugazhendhi, R. (2020). *Restructured society and environment: A review on potential technological strategies to control the COVID-19 pandemic*. Science of The Total Environment, 138858.

7. Bedford, J., Enria, D., Giesecke, J., Heymann, D. L., Ihekweazu, C., Kobinger, G., ... & Ungchusak, K. (2020). *COVID-19: towards controlling of a pandemic*. *The Lancet*, 395(10229), 1015-1018.
8. Park, S., Choi, G. J., & Ko, H. (2020). *Information technology-based tracing strategy in response to COVID-19 in South Korea—privacy controversies*. *Jama*.
9. Cho, H., Ippolito, D., & Yu, Y. W. (2020). *Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs*. arXiv preprint arXiv:2003.11511.
10. Paul Mozur, Raymond Zhong and Aaron Krolik. 1 de Mayo de 2020. *In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags*. Disponible en <https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html>. Accedido Junio 2020.
11. Fred Sainz, Apple. 10 de Abril de 2020. *Apple and Google partner on COVID-19 contact tracing technology*. Disponible en <https://www.apple.com/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/>
12. Department of Health and Social Care. 18 de Junio de 2020. *Next phase of NHS coronavirus (COVID-19) app announced*. Disponible en <https://www.gov.uk/government/news/next-phase-of-nhs-coronavirus-covid-19-app-announced>. Accedido Junio 2020.
13. <https://www.coviid.me/#Governments>
14. Adrian Kriesch, 15 de Mayo de 2020, *South African app aims to slow spread of COVID-19 in developing nations*. Disponible en <https://www.dw.com/en/south-african-app-aims-to-slow-spread-of-covid-19-in-developing-nations/a-53447346>. Accedido en Junio 2020.
15. Delia, L., Galdamez, N., Thomas, P., Corbalan, L., & Pesado, P. (2015, May). Multi-platform mobile application development analysis. In 2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS) (pp. 181-186). IEEE.
16. Delia, L., Thomas, P., Corbalan, L., Sosa, J. F., Cuitiño, A., Cáseres, G., & Pesado, P. (2018, July). Development approaches for mobile applications: comparative analysis of features. In Science and Information Conference (pp. 470-484). Springer, Cham.
17. Dex2Jar. <https://github.com/pxb1988/dex2jar>
18. Java Decompiler. <http://java-decompiler.github.io/>
19. <https://github.com/argob/cuidar>
20. <https://github.com/lbarroeta/coronavirus-sus/>
21. [https://github.com/nic-delhi/AarogyaSetu\\_Android](https://github.com/nic-delhi/AarogyaSetu_Android)
22. <https://github.com/opentrace-community>
23. <https://github.com/corona-warn-app>
24. <https://github.com/immuni-app>
25. <https://github.com/AU-COVIDSafe>
26. Permisos del sistema. Disponible en <https://developer.android.com/guide/topics/security/permissions>. Accedido Julio 2020.

# A Language for the Specification of the Schema of Spreadsheets for the Materialization of Ontologies

Sergio Alejandro Gómez<sup>1,2</sup> and Pablo Rubén Fillottrani<sup>1,2</sup>

<sup>1</sup>Laboratorio de I+D en Ingeniería de Software y Sistemas de Información (LISSI)  
Departamento de Ciencias e Ingeniería de la Computación  
Universidad Nacional del Sur

San Andrés 800 - Campus Palihue – Bahía Blanca, Buenos Aires, Argentina  
Email: {sag,prf}@cs.uns.edu.ar

<sup>2</sup>Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC-PBA)

**Abstract.** Ontology-based Data Access (OBDA) is concerned with providing end-users and applications with a way to query legacy databases through a high-level ontology that models both the business logic and the underlying data sources, accessed by mappings that define how to express records of the database as ontological assertions. In this research, we are concerned with providing with tools for performing OBDA with relational and non-relational data sources. We developed an OBDA tool that is able to access H2 databases and CSV files allowing the user to explicitly formulate mappings, and populating an ontology that can be saved for later querying. In this paper, we present an extension of our previous work as a language for specifying the schema of the data in a spreadsheet data application. This specification is then used to access the contents of a set of Excel books and express them as a relational database with the ultimate goal of materializing its data as an OWL/RDF ontology. We characterize the syntax and semantics of the language, present a prototypical implementation and report on the performance tests showing that our implementation can handle a workload of Excel tables of the order of ten thousand records.

**Keywords.** Ontology-based data access, Ontologies, Relational databases, Spreadsheets.

## 1 Introduction

Despite their simplicity and ubiquity, spreadsheets are still important because they provide a semi-structured way of representing the information of an organization in a distributed way when there is no formal database; even, many times, despite the existence of a centralized system in the company, informal or operational information not covered by the main system is managed in spreadsheets. Although spreadsheet applications (such as MS Excel, Apache Open Office, or Libre Office) give the possibility of making totalizations and filters, these tools

allow limited functionality and are difficult to integrate with the rest of the organization's information, having to resort to data mining and datawarehousing solutions that are not always available to the layman.

Ontology-based data access [1] is a prominent approach to accessing the content of heterogeneous and legacy databases that has gained relevance in the past years in which the database schema along with the semantics of the business model they are exposed as an OWL ontology and the data as RDF triples in distributed form on the web. These OWL/RDF ontologies can be queried through SPARQL end-points.

In this research, we are interested in studying formal models and novel ways of performing OBDA, with the aim of carrying out concrete implementations. In this sense, in recent times, we have been developing a prototype that allows to export the schema of a relational database in H2 format as an OWL ontology and its relational instance as an RDF graph, also allowing the expression of mappings to define concepts from complex SQL queries [2]. In this paper, we present an extension to our OBDA prototype that allows a user to specify a spreadsheet application using a schema definition language. This language allows a naive user to specify the format of the data in the tables contained in sheets of several books, indicating the orientation of the tables, format of columns and rows, cross-relations between tables and books. This allows the spreadsheets to be interpreted as databases and ultimately being integrated with the rest of the OBDA application. We assume that the reader has a basic knowledge of Description Logics (DL) [3], relational databases and the Web Ontology Language [4].

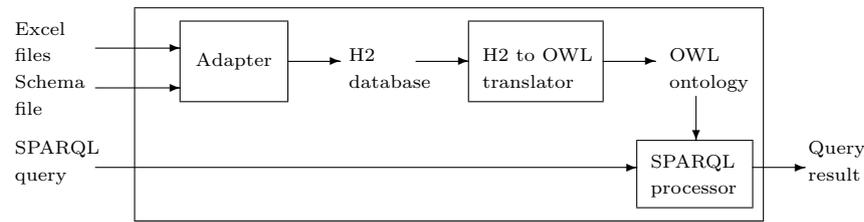
The rest of the paper is structured as follows. In Sect. 2, we present a framework for conceptual modeling of spreadsheets as ontologies. In Sect. 3, we show an empirical evaluation of the performance of the prototype creating tables and ontologies from several Excel files of increasing size. In Sect. 4, we discuss related work. Finally, in Sect. 5, we conclude and foresee future work.

## 2 A Framework for Representing Spreadsheets

Now we present a theoretical framework to represent the data of a spreadsheet application. Later, with this framework, we will define a language to describe the schema of the data. Such schema will be used to access the contents of the spreadsheets, interpret them, generate an SQL script, create and populate an H2 database such script, and then materialize an OWL/RDF ontology with the contents of such a database. This ontology could then be queried via a SPARQL processor (see Fig. 1). We provide the syntax of the data description language in the spreadsheet application using a BNF grammar and give its operational semantics in terms of this framework. We will use a running example throughout the article to illustrate how to use it.

A spreadsheet application data is a set of books. More formally:

**Definition 1.** *An spreadsheet application  $A$  is a pair  $(books, m)$  where  $books$  is a set of books and  $m$  is a map from a unique identifier into an object of the application.*



**Fig. 1.** Architecture of the system

A book is basically a set of sheets along with further information. Formally:

**Definition 2.** A book  $b$  is a tuple  $(id, path, sheets, sheetByID)$  where  $id$  is the identifier of the book,  $path$  is the absolute path of the Excel file defining the book,  $sheets$  is a list of sheets, and  $sheetByID$  is a map from sheet identifier into a sheet.

A sheet is composed by a set of tables. Formally:

**Definition 3.** A sheet  $s$  is a tuple  $(id, name, tables, tableByID, containerBookID)$  where  $id$  is the unique identifier of the sheet,  $name$  is the sheet's name in the container Excel book,  $tables$  is the set of tables contained in this sheet,  $tableByID$  is a map from unique table identifier into a table, and  $containerBookID$  is the identifier of the book containing the sheet.

A table has a header, a set of records, and an orientation (either horizontal or vertical). A cell range defines a rectangle of the data sheet specified by two cell references. Tables can contain references to other tables. Formally:

**Definition 4.** A table  $t$  is a tuple  $(id, className, orientation, initialDataCell, finalDataCell, initialHeaderCell, finalHeaderCell, headerInfo, indexOfKeyField, crossReferences, containerSheetID, containerBookID)$  where  $id$  is the unique identifier of the table,  $className$  is the class in the target ontology defined by the table,  $orientation$  is either vertical or horizontal,  $initialDataCell$  is the top-left corner of the table's data,  $finalDataCell$  is the bottom-right corner of the table's data,  $initialHeaderCell$  is the top-left corner of the table's header,  $finalHeaderCell$  is the bottom-right corner of the table's header,  $headerInfo$  is a map from integer  $i$  into a header datum object  $h_i$ ,  $crossReferences$  is a set of cross-references from this table into other tables,  $containerSheetID$  is the identifier of the sheet containing this table, and  $containerBookID$  is the identifier of the book containing this table. A header datum is a tuple  $(i, name, type)$  where  $i$  is the 1-based index of the header datum in its container map,  $name$  is the name of the field, and  $type$  is the type of the field, that can be one of string, numeric (either integer or real), boolean, or date. A cell has a row (a positive number) and a column (a 1-based positive number). A range is pair  $(c_i, c_f)$  composed of an initial cell  $c_i$  and a final cell  $c_f$ . A cross-reference is a tuple  $(i, t, j)$  where  $i$

is the index of the field in the source table,  $t$  is the identifier of the destination table and  $j$  is the index of the field in the destination table.

## 2.1 Grammar for the Spreadsheet Description Language

We need a language for expressing the elements of this framework. Let us consider the spreadsheet in Fig. 2 containing two tables representing people and their cell phones. We will use that example in order to introduce the elements of our language for describing the schema of the data in the spreadsheet with the goal of materializing an ontology from it that can be queried using SPARQL. We now define the grammar for writing scripts for defining the structure of Excel application data. We discuss each construct by giving its meaning, the BNF grammar that defines its syntax, and an example describing its elements.

PersonID	Name	DateOfBirth	Checked	Weight	Status
1	John	1/1/1981	TRUE	100.5	heavy
2	Mary	2/2/1982	FALSE	60.5	light
3	Paul	3/3/1983	TRUE	80.5	heavy

CellID	1	2	3	4
Brand	Samsung	Apple	Nokia	Samsung
Model	S8	Iphone 11	1100	J7
Owner	1	2	1	2

Fig. 2. A spreadsheet representing people and their cell phones

A script is a sequence of commands and is the start symbol of the grammar:

$$\langle \text{script} \rangle ::= \langle \text{command} \rangle^*$$

There are several available commands to be used in the description of schemas of Excel files.

$$\begin{aligned} \langle \text{command} \rangle ::= & \langle \text{book-declaration} \rangle \mid \langle \text{sheet-declaration} \rangle \mid \langle \text{table-declaration} \rangle \\ & \mid \langle \text{table-header-declaration} \rangle \mid \langle \text{table-data-declaration} \rangle \mid \langle \text{table-field-declaration} \rangle \\ & \mid \langle \text{table-key-field-declaration} \rangle \mid \langle \text{cross-ref-declaration} \rangle \mid \langle \text{comment} \rangle \end{aligned}$$

A book can be declared by giving it an identifier and a path. Identifiers are surrounded by quotation marks and are composed in the usual way.

$$\langle \text{book-declaration} \rangle ::= \text{book } \langle \text{id} \rangle \text{ has-path } \langle \text{path} \rangle$$

$$\langle \text{id} \rangle ::= " \langle \text{identifier} \rangle "$$

$$\langle \text{identifier} \rangle ::= \langle \text{letter} \rangle . (\langle \text{letter} \rangle \mid \langle \text{digit} \rangle)^*$$

$$\langle \text{letter} \rangle ::= \text{a} \mid \text{b} \mid \dots \mid \text{z} \mid \text{A} \mid \text{B} \mid \dots \mid \text{Z}$$

$$\langle \text{digit} \rangle ::= 0 \mid 1 \mid \dots \mid 9$$

$$\langle \text{path} \rangle ::= "\dots \text{windows file path} \dots"$$

*Example 1.* Consider the piece of code that expresses that book  $b_1$  has as its path the Excel file `book1.xlsx` located in the `Escritorio8` subfolder in the desktop folder: `book "b1" has-path "c:/users/sgomez/Desktop/Escritorio8/book1.xlsx"`.

A book has at least one data sheet. Each sheet has an identifier in this schema file, a name in the spreadsheet and it is located in a book.

$\langle \text{sheet-declaration} \rangle ::= \text{sheet } \langle id \rangle \text{ name } \langle id \rangle \text{ in } \langle id \rangle$

*Example 2.* Consider the code: `sheet "s1" name "Data" in "b1"`. It expresses that the spreadsheet  $s_1$  has been named *Data* and it is located in the book  $b_1$ .

Each spreadsheet can have several tables. Each table has an identifier, is contained in a certain spreadsheet, defines a class and has an orientation which either is horizontal or vertical.

$\langle \text{table-declaration} \rangle ::= \text{table } \langle id \rangle \text{ in-sheet } \langle id \rangle \text{ class-name } \langle id \rangle \text{ orientation } \langle \text{orientation-literal} \rangle$

$\langle \text{class-name} \rangle ::= \langle id \rangle$

$\langle \text{orientation-literal} \rangle ::= \text{horizontal} \mid \text{vertical}$

*Example 3.* Consider the commands: `table "t1" in-sheet "s1" class-name "Person" orientation vertical` and `table "t2" in-sheet "s1" class-name "Phone" orientation horizontal`. They define that there are two tables:  $t_1$  and  $t_2$ , which are both located in sheet  $s_1$ . Table  $t_1$  defines a class name *Person* while table  $t_2$  defines a class named *Phone*. The orientation of  $t_1$  is vertical but the orientation of  $t_2$  is horizontal.

Every table definition is composed of header and data sections, with syntax:

$\langle \text{table-header-declaration} \rangle ::= \text{header } \langle id \rangle \text{ range } \langle \text{range-specification} \rangle$

$\langle \text{table-data-declaration} \rangle ::= \text{data } \langle id \rangle \text{ range } \langle \text{range-specification} \rangle$

$\langle \text{range-specification} \rangle ::= \text{" } \langle \text{cell-spec} \rangle \text{: } \langle \text{cell-spec} \rangle \text{"}$

$\langle \text{cell-spec} \rangle ::= \langle \text{letter} \rangle^+ \langle \text{digit} \rangle^+$

*Example 4.* Consider the commands for defining the limits of tables  $t_1$  and  $t_2$ : `header "t1" range "b2:g2", data "t1" range "b3:g5", header "t2" range "b8:b11", and data "t2" range "c8:f11"`.

Fields are declared by specifying the table to which they belong, an index, a name and a type. There is an special field called the key field:

$\langle \text{table-field-declaration} \rangle ::= \text{field } \langle id \rangle \text{ index } \langle \text{positive-integer} \rangle \text{ name } \langle id \rangle \text{ type } \langle \text{type-id} \rangle$

$\langle \text{type-id} \rangle ::= \text{integer} \mid \text{string} \mid \text{date} \mid \text{real}$

$\langle \text{table-key-field-declaration} \rangle ::= \text{key-field } \langle id \rangle \text{ index } \langle \text{positive-integer} \rangle$

$\langle \text{positive-integer} \rangle ::= (1..9) \langle \text{digit} \rangle^*$

*Example 5.* Consider the piece of code for defining the fields of tables  $t_1$  and  $t_2$ :

```
field "t1" index "1" name "PersonID" type integer
field "t1" index "2" name "Name" type string
field "t1" index "3" name "DateOfBirth" type date
field "t1" index "4" name "Checked" type boolean
field "t1" index "5" name "Weight" type real
field "t1" index "6" name "Status" type string
key-field "t1" index "1"
field "t2" index "1" name "CellID" type integer
field "t2" index "2" name "Brand" type string
field "t2" index "3" name "Model" type string
field "t2" index "4" name "Owner" type integer
key-field "t2" index "1"
```

The table  $t_1$  has 6 fields named *PersonID*, *Name*, *DateOfBirth*, *Checked*, *Weight* and *Status* of type integer, date, boolean, real and string, resp. The table  $t_2$  has 4 fields named *CellID* and *Owner* both of type integer, and *Brand* and *Model* of type string. The key field of  $t_1$  is *PersonID* while the key field of  $t_2$  is *CellID*. Notice that no indications are given here if the contents of a cell is either a formula or a value and it is neither necessary. For instance the column *Status* is a formula of the form: =IF(F3>=80, "heavy", "light") indicating that if the weight of the person is greater than or equal to 80 kilograms, the person is considered as heavy, otherwise is deemed as light.

A table can have cross-references to other tables.

$\langle \text{cross-ref-declaration} \rangle ::= \text{cross-ref from } \langle \text{id} \rangle \text{ index } \langle \text{positive-integer} \rangle \text{ into } \langle \text{id} \rangle \text{ index } \langle \text{positive-integer} \rangle$

*Example 6.* The following piece of code defines a cross-reference from field number 4 of table  $t_2$  into field number 1 of table  $t_1$ :

```
cross-ref from "t2" index "4" into "t1" index "1"
```

One-line comments are allowed in our scripting language and they begin with the hashtag character.

$\langle \text{comment} \rangle ::= \# \langle \text{character} \rangle^*$

$\langle \text{character} \rangle ::=$  any Ascii character excluding end of line

## 2.2 Semantics of Spreadsheet Constructors

The semantics of the empty spreadsheet application *create* is  $(\{\}, \{\})$ . The semantics of commands is given in terms of the function **Sem** from sequences of commands by spreadsheet applications into spreadsheet applications. The semantics of a book declaration is as follows:

$\text{Sem}(\text{sheet } "id" \text{ name } "n" \text{ in } "bid", (books, m)) = (books', \{(id, s)\} \cup m)$  where

$$\begin{aligned} books' &= books - \{b\} \cup \{b'\} \\ b &= m(bid) = (bid, p, sheets, sheetByID), \\ b' &= (bid, p, \{s\} \cup sheets, \{(id, s)\} \cup sheetByID) \\ s &= (id, n, \{\}, \{\}, bid) \end{aligned}$$

The semantics of the declaration of a table *id*, in sheet *sid*, determining a class *c*, with orientation *o*, with *n* fields named  $name_1, \dots, name_n$  of types  $t_1, \dots, t_n$ , key field *k*, *m* cross-references from fields  $i_1, \dots, i_m$  into foreign tables  $tid_1, \dots, tid_m$  and foreign fields with indexes  $j_1, \dots, j_m$ , resp., header info in the range  $h_1 : h_2$  and data info in the range  $d_1 : d_2$  is given shown in Fig. 3.

## 2.3 Generation of Databases and Ontologies from Spreadsheets

We now discuss the generation of OWL/RDF ontologies from spreadsheet applications. Given a book with mapping *m* of identifiers into objects, let *t* be a table such that  $t = (id, c, o, d_1, d_2, h_1, h_2, head, k, cross, s)$ , such that  $cross =$

```

Sem(sec, (books, m)) = (books', {(id, t)} ∪ m) where
  sec = (table "id" in-sheet "sid" class-name "c" orientation o
    header "id" range "h1 : h2"
    field "id" index "1" name "name1" type t1
    ...
    field "id" index "n" name "namen" type tn
    key-field "id" index "k"
    data "id" range "d1 : d2"
    cross-ref from "id" index "i1" into "tid1" index "j1"
    ...
    cross-ref from "id" index "im" into "tidm" index "jm")
  s = (sid, name, ts, tableByID, containerBookID) = m(id)
  t = (id, c, o, d1, d2, h1, h2, head, k, cross, sid)
  s' = (sid, name, {t} ∪ ts, {(id, t)} ∪ tableByID, containerBookID)
  books' = books - {b} ∪ {b'}
  b = (bid, p, sheets, sheetByID) = m(containerBookID)
  b' = (bid, p, sheets', sheetByID)
  sheets' = {s} ∪ sheets
  cross = {(i1, tid1, j1), ..., (im, tidm, jm)}
  head = λi.(i, name, ti), with i = 1, ..., n

```

**Fig. 3.** Semantics of table declaration commands

$\{(i_1, tid_1, j_1), \dots, (i_m, tid_m, j_m)\}$ , and  $head = \lambda i.(i, name_i, t_i)$ , with  $i = 1, \dots, n$ . The SQL code in Fig. 4 represents the schema of table  $t$ , where *second* and *sixth* are the projectors of the second and the sixth components of a tuple, resp. Then this SQL code is used to materialize an H2 database, which in turn is used to materialize an OWL/RDF ontology using the methodology described in our previous work [5] and in accordance to the architecture shown in Fig. 1.

```

create table "c" (
  "name1" t1, ..., "namek" tk primary key, ..., "namen" tn,
  foreign key ("second(head(i1))" references "second(m(tid1))" ("second(sixth(m(tid1))(j1))"),
  ..., foreign key ("second(head(im))" references "second(m(tidm))" ("second(sixth(m(tidm))(jm))" );

```

**Fig. 4.** SQL script for creating a generic table  $t$

*Example 7.* The spreadsheet in Fig. 2 is represented by the SQL script in Fig. 5. Then, from this script, a database is created and the ontology materialized from that database has the following DL axioms (that are ultimately serialized as OWL/RDF):  $\text{Person} \sqsubseteq \exists \text{PersonID}, \exists \text{PersonID}^- \sqsubseteq \text{Integer}, \text{Person} \sqsubseteq \exists \text{name}, \exists \text{name}^- \sqsubseteq \text{String}, \text{Person} \sqsubseteq \exists \text{dateOfBirth}, \exists \text{dateOfBirth}^- \sqsubseteq \text{Date}, \text{Person} \sqsubseteq \exists \text{checked}, \exists \text{checked}^- \sqsubseteq \text{Boolean}, \text{Person} \sqsubseteq \exists \text{weight}, \exists \text{weight}^- \sqsubseteq \text{Real}, \text{Person} \sqsubseteq \exists \text{status}, \exists \text{status}^- \sqsubseteq \text{String}, \text{Phone} \sqsubseteq \exists \text{cellID}, \exists \text{cellID}^- \sqsubseteq \text{Integer}, \text{Phone} \sqsubseteq \exists \text{brand}, \exists \text{brand}^- \sqsubseteq \text{String}, \text{Phone} \sqsubseteq \exists \text{model}, \exists \text{model}^- \sqsubseteq \text{String}, \text{Phone} \sqsubseteq \exists \text{owner}, \exists \text{owner}^- \sqsubseteq \text{Integer}, \text{Phone} \sqsubseteq \exists \text{ref\_owner}, \exists \text{ref\_owner}^- \sqsubseteq \text{Person}$ . The assertions for representing the first record of the class **Person** are:  $\text{PersonID}(\text{Person}\#1, 1)$ ,  $\text{name}(\text{Person}\#1, \text{JOHN})$ ,  $\text{dateOfBirth}(\text{Person}\#1, 1981-01-01)$ ,  $\text{checked}(\text{Person}\#1, \text{TRUE})$ ,  $\text{weight}(\text{Person}\#1, 100.5)$ , and  $\text{status}(\text{Person}\#1, \text{HEAVY})$ .

```

create table "Person"(
    "PersonID" int primary key,      "Name" varchar(50), "DateOfBirth" date,
    "Checked" boolean, "Weight" real, "Status" varchar(50) );
create table "Phone"(
    "CellID" int primary key, "Brand" varchar(50), "Model" varchar(50), "Owner" int,
    foreign key ("Owner") references "Person"("PersonID") );
insert into "Person"("PersonID", "Name", "DateOfBirth", "Checked", "Weight", "Status")
    values (1, 'John', '1981-01-01', true, 100.5, 'heavy');
insert into "Person"("PersonID", "Name", "DateOfBirth", "Checked", "Weight", "Status")
    values (2, 'Mary', '1982-02-02', false, 60.5, 'light');
insert into "Person"("PersonID", "Name", "DateOfBirth", "Checked", "Weight", "Status")
    values (3, 'Paul', '1983-03-03', true, 80.5, 'heavy');
insert into "Phone"("CellID", "Brand", "Model", "Owner") values (1, 'Samsung', 'S8', 1);
insert into "Phone"("CellID", "Brand", "Model", "Owner") values (2, 'Apple', 'Iphone 11', 2);
insert into "Phone"("CellID", "Brand", "Model", "Owner") values (3, 'Nokia', '1100', 1);
insert into "Phone"("CellID", "Brand", "Model", "Owner") values (4, 'Samsung', 'J7', 2);

```

**Fig. 5.** SQL code obtained from the spreadsheet in Fig. 2

**Table 1.** Running times for ontology generation from Excel files

Number of records	Excel file size [Megabytes]	Time for loading Excel file [seconds]	Time for creating ontology [seconds]	Size of ontology file [Megabytes]
10	0.012	0.901	0.276	0.115
100	0.033	1.774	0.359	0.910
1,000	0.255	5.825	1.067	8.951
10,000	2.640	29.703	4.253	90.951
100,000	26.742	Out of memory error	–	–

### 3 Experimental Evaluation

We now discuss some of the tests we have performed in order to test how our application handles increasing demands in spreadsheet size. The performance of our system is affected mainly by the fact that Excel records are materialized as RDF triples and also by four factors: (i) the system is implemented in the JAVA programming language; (ii) the database management system that we use as intermediate data representation is H2, (iii) the handling of the global ontology is done via the OWL API [6], and (iv) the access to Excel files is implemented using the Apache POI library [7]. Our tests were conducted on an ASUS notebook having an Intel Core i7, 3.5GHz CPU, 8GB RAM, 1TB HDD, and Windows 10. They involved the creation of databases with a single table extracted from Excel books containing only a sheet with a table containing 100 fields of numeric type filled with an increasing number of records. In Table 1, we can see the times for loading the Excel files and the size of the materialized ontologies. Therefore, we conclude that our application can only handle spreadsheets containing tables with a size of tens of thousands records.

## 4 Related Work

*XLWrap* [8] constitutes an approach for generating RDF graphs of arbitrary complexity from various spreadsheet layouts, including cross tables and tables where data is not aligned in rows. They provide a functionality similar to ours but relying in JSON for the description of data. Our approach features a simpler language aimed towards naive users. *NOR2O* [9] can convert Excel to Scovo and Data Cube Vocabulary but it is no longer maintained. *Excel2rdf*<sup>1</sup> is a Java-based command-line utility that converts Excel files into valid RDF files but as far as we know it is not possible to make precise definitions of the data contained nor export terminologies as done in our proposal. *RDBToOnto*<sup>2</sup> allows to automatically generate fine-tuned OWL ontologies from relational databases. A major feature of this full-fledged tool is the ability to produce structured ontologies with deeper hierarchies by exploiting both the database schema and the stored data. RDBToOnto can be exploited to produce RDF Linked Data. It can also be used to generate highly accurate RDB-to-RDF mapping rules (for D2RQ Server and Triplify). *Spread2RDF*<sup>3</sup> is a converter for complex spreadsheets to RDF and a Ruby-internal DSL for specifying the mapping rules for this conversion. Other solutions to the problem of wrapping Excel files into semantic technologies have migrated from the academic world to the commercial world. For example, *Open Anzo*<sup>4</sup> used to include both an open-source enterprise-featured RDF quad store and a sophisticated service oriented, semantic middleware platform that providing support for multiple users, distributed clients, offline work, real-time notification, named-graph modularization, versioning, access controls, and transactions, giving support to applications based on W3C semantic technology standards like OWL, RDF and SPARQL. This project is no longer available as it has turned into a company named *Cambridge Semantics*<sup>5</sup>. *TopBraid Composer*<sup>6</sup> can convert Excel spreadsheets into instances of an RDF schema. *TabLinker*<sup>7</sup> can convert non-standard Excel spreadsheets to the Data Cube vocabulary. Our work converts the contents of the records in Excel sheets to RDF but also allows to precisely define the schema of the data in OWL.

## 5 Conclusions and Future Work

We have presented a framework for the representation of spreadsheet applications along with a description language of the schema of the data stored in them. Furthermore, we have given a formal specification of the syntax of such a language with a BNF grammar and its formal semantics in terms of the representation framework. We have shown an example of how it is used. We have

<sup>1</sup> <https://github.com/waqarini/excel2rdf>

<sup>2</sup> <https://sourceforge.net/projects/rdbtoonto/>

<sup>3</sup> <https://github.com/marcelotto/spread2rdf>

<sup>4</sup> <https://www.w3.org/2001/sw/wiki/OpenAnzo>

<sup>5</sup> <http://www.cambridgesemantics.com>

<sup>6</sup> <https://www.topquadrant.com/knowledge-assets/faq/tbc/>

<sup>7</sup> <https://github.com/Data2Semantics/TabLinker/wiki>

provided a prototypical implementation, showing how it is integrated into an ontology-based data access system with the aim of publishing such spreadsheets as freely available ontologies on the Semantic Web. We believe that this language provides a valid alternative to more technical options like JSON from which naive users can benefit while providing more control than WYSIWYG-type applications that provide similar functionality. Also, we have carried out experimental tests to determine the workload that our implementation can effectively handle, showing its viability for spreadsheets containing tables with thousands of records.

As part of future work, we are interested in continuing to explore other types of NoSQL database models and thinking about integrating them into our ontology-based data access prototype with the aim of developing novel algorithms and techniques such as virtualization by query-rewriting to provide more flexibility in regards to volatile data than the one offered by the materialization approach.

**Acknowledgments.** This research is funded by Secretaría General de Ciencia y Técnica, Universidad Nacional del Sur, Argentina and by Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC-PBA).

## References

1. Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., Zakharyashev, M.: *Ontology-Based Data Access – A Survey*. In: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*. (2018) 5511–5519
2. Gómez, S.A., Fillottrani, P.R.: *Materialization of OWL ontologies from relational databases – A practical approach*. In Pesado, P., Arroyo, M., eds.: *Computer Science – CACIC 2019 selected papers*, Cham, Springer International Publishing (2020) 285–301
3. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: *An Introduction to Description Logic*. Cambridge University Press (2017)
4. Bao, J., Kendall, E.F., McGuinness, D.L., Patel-Schneider, P.F.: *OWL 2 Web Ontology Language Quick Reference Guide (Second Edition)* W3C Recommendation 11 December 2012 (2012)
5. Gómez, S.A., Fillottrani, P.R.: *Towards a Framework for Ontology-Based Data Access: Materialization of OWL Ontologies from Relational Databases*. In Pesado, P., Aciti, C., eds.: *X Workshop en Innovación en Sistemas de Software (WISS 2018)*, XXIV Congreso Argentino de Ciencias de la Computación CACIC 2018. (2018) 857–866
6. Matentzoglou, N., Palmisano, I.: *An Introduction to the OWL API*. Technical report, The University of Manchester (2016)
7. Minh, N.H.: *How to Read Excel Files in Java using Apache POI* (2019)
8. Langegger, A., Wöß, W.: *XLWrap – Querying and Integrating Arbitrary Spreadsheets with SPARQL*. In: *Proceedings of the 8th International Semantic Web Conference (ISWC2009)*, Washington D.C. LNCS 5823, Springer (2009)
9. Terrazas, B.V., Gomez-Perez, A., Calbimonte, J.P.: *NOR2O: a library for transforming non-ontological resources to ontologies*. In: *ESWC’10*. (2010)

## Genvi: propuesta de un desarrollo tecnológico para abordar la violencia de género

Juan Pablo Quiñones<sup>1</sup>, Viviana Harari<sup>1</sup> and Ivana Harari<sup>1</sup>

<sup>1</sup> Facultad de Informática, Universidad Nacional de La Plata. La Plata, CP 1900, Argentina  
[jpq.1987@gmail.com](mailto:jpq.1987@gmail.com); {vharari, iharari}@info.unlp.edu.ar

**Abstract.** Según Florence Raes, representante de la ONU Mujeres de Paraguay y Argentina, en América Latina 1 de cada 3 mujeres en promedio, ha padecido violencia física y/o sexual en una relación íntima a lo largo de su vida. La violencia de género es una problemática social y cultural que se da a nivel mundial y en diferentes estratos sociales. La cantidad de mujeres que sufren este delito son alarmantes y, año tras año aumentan a pesar de las leyes y programas que se vienen impulsando desde diferentes sectores de la Sociedad. En la mayoría de los casos, el silencio que rodea a la víctima se encuentra condicionado por factores económicos, políticos, legales e institucionales como así también por el conjunto de tradiciones, costumbres y reglas vigentes en cada sociedad. Esta problemática, según especialistas en temas de género, es compleja y multicausal y, por lo tanto, requiere de una respuesta integral y articulada surgida a partir de un abordaje interdisciplinario e interinstitucional. Este artículo presenta una propuesta de una aplicación móvil que permita aportar desde lo tecnológico, en el abordaje de esta problemática, mediante la integración de información útil y servicios, mapas con ubicaciones de entidades cercanas que tratan la violencia contra la mujer y que brinde servicios de asistencia durante el tránsito y movimiento de la mujer.

**Keywords:** Aplicaciones móviles, Flutter, Software libre.

### 1 Introducción

La Violencia contra las Mujeres es definida, según la ley argentina 26.485 artículo 4 [1], en los siguientes términos: “*Se entiende por violencia contra las mujeres toda conducta, acción u omisión, que de manera directa o indirecta, tanto en el ámbito público como en el privado, basada en una relación desigual de poder, afecte su vida, libertad, dignidad, integridad física, psicológica, sexual, económica o patrimonial, como así también su seguridad personal. Quedan comprendidas las perpetradas desde el Estado o por sus agentes. Se considera violencia indirecta, a los efectos de la presente ley, toda conducta, acción u omisión, disposición, criterio o práctica discriminatoria que ponga a la mujer en desventaja con respecto al varón*”

Esta ley nacional, sancionada en el año 2009, junto a otras legislaciones provinciales intentan proveer un marco normativo para el tratamiento de mujeres en situación de violencia y, dotar de mayor institucionalidad a los dispositivos y acciones de atención y prevención.

Pero, a pesar de ello, del surgimiento de programas, recursos y centros de atención, del accionar colectivo mediante movilizaciones y del trabajo en territorios mayoritariamente de ONGs (Organizaciones no gubernamentales), las cifras de femicidios aumentan año tras año, estimándose durante el 2019 una muerte cada 33 horas en la República Argentina [2]. La falta aún de más iniciativas de formación, intervención y aplicación de esta Ley que cumple ya más de 10 años, pueden ser causales de estas estadísticas alarmantes.

En este artículo, se explicará una propuesta de desarrollo para móviles que permita abordar esta problemática de violencia de género, mediante un recurso que centraliza toda la información necesaria para mantener informada a la mujer. El recurso, que centraliza información para la misma, contiene datos de entidades de protección, de asesoramiento y, centros para realizar denuncias. Para todo esto se tiene en cuenta la geolocalización de la mujer. A su vez, que provea servicios para utilizar el celular como dispositivo de acompañamiento, advirtiendo a terceros sobre recorridos y movimientos que la persona damnificada y víctima realizará.

Se explicará entonces, detalles de la herramienta tecnológica que se propone, que permita brindar alternativas de solución a necesidades que aún están desatendidas. Una herramienta que sea simple de utilizar, accesible y colaborativa, que reúna y nucleee toda la información actualizada y que se pueda descargar en el celular para tenerla siempre presente.

## 2 Marco teórico sobre la violencia de género

Se conoce como violencia “*al uso intencional de la fuerza o el poder físico, de hecho o como amenaza, contra uno mismo, otra persona o un grupo o comunidad, que cause o tenga muchas probabilidades de causar lesiones, muerte, daños psicológicos, trastornos del desarrollo o privaciones*” [3].

Si bien las violaciones de los derechos humanos afectan tanto a los hombres como las mujeres, su impacto varía de acuerdo con el sexo de la víctima. Los derechos humanos, considerados como un conjunto de pautas éticas con proyección jurídica, surgen de la necesidad de todos los individuos, de contar con las condiciones esenciales para una vida digna. Pero estos han sido producto de un largo proceso de construcción y cambio a lo largo de los dos últimos siglos, donde se cuestiona la validez universal del androcentrismo y del modelo del hombre occidental. Como consecuencia de esto, se ha reconocido la necesidad de considerar y reconocer las especificidades de los individuos, ya sean de género, de etnia, de edad o de cualquier otra índole. El reconocimiento de esta heterogeneidad no conduce a la fragmentación o atomización de la condición humana, sino que por el contrario, contribuye a una real universalización de los individuos y de los derechos humanos, basada en el principio de la pluralidad y el respeto de las diferencias y de la diversidad [4].

Toda agresión perpetrada contra una mujer tiene alguna característica que permite identificarla como violencia de género. Esto significa que está directamente vinculada a la desigual distribución del poder y a las relaciones asimétricas que se establecen entre varones y mujeres en nuestra sociedad, que perpetúan la desvalorización de lo

femenino y su subordinación a lo masculino. Lo que diferencia a este tipo de violencia de otras formas de agresión y coerción es que el factor de riesgo o de vulnerabilidad es el solo hecho de ser mujer [4].

Las Naciones Unidas definen la violencia contra la mujer como: *“todo acto de violencia de género que resulte, o pueda tener como resultado un daño físico, sexual o psicológico para la mujer, inclusive las amenazas de tales actos, la coacción o la privación arbitraria de libertad, tanto si se producen en la vida pública como en la privada”* [5]

Los elementos claves de esta definición son que el factor de riesgo es justamente el ser mujer, incluye agresiones físicas, psicológicas, sexuales así como amenazas de tales actos. Estas agresiones pueden darse en el ámbito de lo público como privado, y que su objetivo principal es mantener la subordinación de la mujer al hombre.

Entonces, la violencia de género, violencia machista o violencia hacia la mujer son conceptos que denotan la discriminación y desequilibrio de poder entre mujeres y hombres en la sociedad. Puede adoptar diversas formas, lo que permite clasificar el delito, como ser: violación sexual e incesto, acoso sexual en el trabajo y en las instituciones de educación, violencia sexual contra mujeres detenidas o presas, actos de violencia contra las mujeres desarraigadas, tráfico de mujeres y violencia doméstica.

Cada una de estas formas de violencia, presentan subaristas particulares. Por ejemplo, en el caso de violencia doméstica, entre los sucesos que desencadenan la violencia masculina en la pareja que se percibe como falta de subordinación, pueden ser: no obedecer al hombre, contestarle mal, no tener la comida preparada a tiempo, no atender adecuadamente a los hijos o el hogar, preguntarle al hombre por cuestiones de dinero o de sus enamoradas, salir sin el permiso del hombre, negarse a mantener relaciones sexuales con el hombre, sospechas, por parte del hombre, de infidelidad de la mujer. Y, se detectan diversos comportamientos dominantes o manifestaciones frente a esa falta de subordinación percibida por el agresor, además de daños físicos y psicológicos, como aislar a una persona de su familia y amigos, privación de la libertad, amenazas, persecución, control, vigilar sus movimientos y restringir su acceso a la información o asistencia, entre otros.

Es imprescindible abordar la violencia de género contra las mujeres desde una perspectiva educativa que ofrezca posibilidades de cambios culturales en la sociedad, por lo que se requieren profundas modificaciones en esta área. La intervención desde diferentes áreas, incluyendo desde lo tecnológico pueden implicar un aporte significativo para que el cambio social que exige el respeto de los derechos de las mujeres, se lleve a cabo.

### **3 Diagnóstico sobre recursos tecnológicos para abordar la violencia de género**

En respuesta a la problemática de violencia de género, el Estado cuenta con varios programas, recursos y centros de atención que toman iniciativa a través de distintos medios de comunicación. Entre estos, se destacan:

**Línea Nacional 144:** La Línea telefónica Nacional 144 está destinada a brindar información, orientación, asesoramiento y contención para las mujeres en situación de violencia de todo el país, los 365 días del año, las 24 horas, de manera gratuita [6].

**Línea 137:** La línea telefónica, perteneciente al Ministerio de Justicia y DDHH de la Nación, comenzó su experiencia en la Ciudad de Buenos Aires y hoy está siendo extendida a otras provincias. Está destinada para que cualquier persona en situación de urgencia, ya sea la víctima o algún familiar o vecino/a puede llamar al 137. Una vez verificado el llamado, un móvil no identificable con dos oficiales de Policía Federal Argentina se desplazará hasta la escena, acompañados por un/a psicólogo/a y un/a trabajador/a social para dar respuesta inmediata, conteniendo a la víctima y desarmando a la persona agresora que será conducida a la comisaría correspondiente [7].

**Atención integral:** El Ministerio de Justicia y Derechos Humanos de la Nación ofrece atención integral a las víctimas de violencia de género con el objetivo de acompañarlas aportándoles el apoyo jurídico, social y psicológico necesario [8]. En caso que la mujer presentara lesiones sería trasladada a un hospital público. Al mismo tiempo informan al juez con competencia acerca del hecho acontecido. Una oficina de orientación y asistencia le ofrecerá a la víctima las direcciones de instituciones que puedan orientarlas psicoterapéuticamente.

**Centros de Atención Personalizada:** El Instituto Nacional de las Mujeres participó en la sistematización de una Guía de Recursos a nivel federal (Recurso) que se pone a disposición de las personas que requieren atención [9]. Esta publicación da cumplimiento al requerimiento de la Ley 26.485 de Protección Integral para Prevenir, Sancionar y Erradicar la Violencia contra las Mujeres en los Ámbitos en que Desarrollen sus Relaciones Interpersonales respecto de las acciones positivas que debe desarrollar el Consejo Nacional de las Mujeres, entre las que menciona el diseño y publicación de una guía de programas y servicios de asistencia directa (Art. 9º).

**Mecanismos de vigilancia electrónica:** Los mecanismos de vigilancia electrónica utilizados en situaciones de violencia de género, permiten monitorear en tiempo real a los agresores y proteger a las víctimas cuando se sienten amenazadas. Son utilizados en los casos donde los jueces o fiscales lo requieran [10][11].

Por lo general, el mecanismo de vigilancia electrónica está compuesto por dos niveles de seguridad que operan de forma conjunta y sincronizada: **de protección a la víctima y de rastreo del agresor**. En el *nivel de protección*, la víctima dispone de un dispositivo conocido como unidad de rastreo (ver Figura 1), el cual genera una alarma sonora y visual cada vez que se detecta la presencia del agresor dentro de la zona restringida (suele establecerse un perímetro en metros). Existe un centro de monitoreo en tiempo real que recibe todas las alertas enviadas por estas unidades de rastreo. En caso de que la víctima se sienta amenazada, el dispositivo cuenta con un botón antipánico que emitirá una señal de alerta al centro de monitoreo.



**Fig. 1.** Modelo de una unidad de rastreo para proteger a la víctima

En el *nivel de rastreo*, el agresor recibe una tobillera (ver Figura 2) con tecnologías basadas en Sistemas de Posicionamiento Global (en inglés GPS; Global Positioning System) o Radiofrecuencia (en inglés RF; Radio frequency) que actúa como un rastreador, y permite recibir en el centro de monitoreo avisos en forma de mensajes y alertas visuales, audibles o vibratorias si es que viola la restricción de acercarse a la víctima o de ingresar en zonas restringidas por la Justicia o si se considera necesario informar las novedades de ubicación, estado y/o última comunicación. En unidades con prestaciones avanzadas es posible procesar y almacenar una gran cantidad de datos (como por ejemplo horarios y posiciones) en una memoria interna antes de ser descargados al centro de monitoreo.



**Fig. 2.** Modelo de una tobillera electrónica para rastrear al agresor

#### **4 Propuesta de desarrollo**

En líneas generales estas iniciativas gubernamentales, si bien tienen un valor considerable respecto a la atención y actuación en casos de violencia, se podrían potencializar sus funciones y acceso, mediante un recurso en línea que integre y centralice sus servicios e información.

El recurso en línea podría mejorar el acceso al Recurso mostrando la información que maneja (contactos, lugares, horarios y días de atención, entre otros) de una manera geolocalizada, visual y de rápido acceso.

A su vez el usuario podría aportar información extra como opiniones, experiencia de la atención recibida, quejas u otro tipo de intervención, ayudando a la comunidad y aportando a la mejora del Recurso.

Resulta importante que estas mejoras se apliquen a través de la utilización de dispositivos móviles. En la actualidad, los dispositivos móviles han ido ganando terreno en la vida cotidiana de toda persona. Cumpliendo no solo la función de realizar llamadas y enviar mensajes de texto, sino también la de consumir y producir cualquier clase de información desde cualquier lugar y en cualquier momento, lo que permite usar aplicaciones móviles que utilicen información implícitamente adquirida del contexto y modifiquen su comportamiento en base a esta. Estas aplicaciones son

conocidas como “*Aplicaciones sensibles al contexto*” [12]. El beneficio de estas aplicaciones permitirá ofrecer a los usuarios la posibilidad de acceder a un conjunto de información y servicios que requieran, cuando y donde lo deseen. El contexto puede ser usado para determinar qué información y servicios hacer disponibles o brindar a los usuarios en una situación determinada.

### 5.1 Objetivos

Desarrollar un recurso socio-tecnológico móvil, colaborativo y con geolocalización sobre violencia de género, que facilite el acceso integral a servicios e información oficial, aportando nueva funcionalidad.

### 5.2 Aspectos funcionales

Se pretende que esta aplicación móvil cubra las siguientes componentes funcionales:

- **Centralización de la información:** se nuclea toda la información sobre teléfonos de emergencia, datos útiles, leyes, programas, asesoramiento, información contextual a través de la geolocalización y sistema de búsqueda.
- **Recursero:** se geolocalice en un mapa los centros, entidades, ONGs relacionadas con el tema, refugios, instituciones para la prevención, asesoramiento y realización de denuncias. Disponer de información detallada de cada recurso y sus horarios de atención. Se permita la participación de la persona, registrando información sobre cómo fue atendida, su opinión, y relato sobre su experiencia en la atención brindada.
- **Acompañame!:** se ofrezca un acompañamiento virtual durante el recorrido que la persona desee realizar. Esta funcionalidad permite, con permiso previo, notificar a terceras personas acerca del lugar donde va la víctima y tiempos de llegada estimativos.

### 5.3 Arquitectura

Esta aplicación está pensada como un desarrollo móvil, con características colaborativas y de geolocalización. Respetará las normas de accesibilidad para el uso de personas con discapacidad y principios de usabilidad para garantizar una calidad de uso aceptable.

La arquitectura cliente servidor de la aplicación móvil tendrá el siguiente aspecto: la comunicación con los servicios web a utilizar desde la aplicación, se realizará mediante REST [13] debido a que garantiza mayor flexibilidad, presenta una mejor performance, ya que los mensajes son de menor tamaño y, los tiempos de respuesta más cortos. Esto determina una solución más liviana, presenta tecnologías de uso generalizado y se adapta correctamente a las necesidades de conexión de los dispositivos móviles y de otros servicios externos que consuman la información

generada [14]. Por lo tanto, se utilizará servicios RESTful para la conexión entre el back-end (servidor) y el dispositivo (cliente).

#### 5.4 Aspectos técnicos

Las decisiones respecto a las cuestiones técnicas del desarrollo de una versión preliminar, estuvieron condicionadas por la optimización de las capacidades, de los tiempos de latencia para minimizar costos de comunicación y por la manera en que afecten la experiencia de uso.

Considerando el alcance como premisa básica para la elección de la plataforma, se tiene que Android cuenta con el 92.67% de la cuota del mercado en Argentina mientras que iOS con el 6.53% restante [15][16]. Esta diferencia se debe a que Android es un sistema operativo de código abierto y permite que cada fabricante pueda personalizarlo y adaptarlo a sus dispositivos. Esta libertad garantiza que una aplicación se encuentre accesible a una mayor cantidad de dispositivos y de usuarios potenciales.

El hecho de tener que desarrollar una versión de la aplicación para cada plataforma hace que las aplicaciones nativas queden en la peor posición en este punto, aumentando notoriamente el costo de desarrollo, es por eso que se decidió por el desarrollo en Flutter [17], kit de herramientas de interfaz del usuario de Google, por permitir realizar aplicaciones compiladas nativamente, para móvil, web y escritorio desde una única base de código y, garantizar las siguientes cuestiones básicas para la propuesta planteada:

- **Acceso directo a las utilidades del sistema operativo del dispositivo:** permite acceder a los dispositivos de almacenamiento internos y externos, a sensores como GPS y acelerómetro, a la agenda de contactos, al sistema de notificaciones, entre otros, en forma más directa y óptima. Este acceso garantiza una buena experiencia de usuario y que las aplicaciones tengan un rendimiento optimizado.
- **Ejecución en modo offline:** garantiza que la aplicación se encuentre disponible en cualquier momento que desee ser utilizada.
- **Distribución a través de las tiendas de los fabricantes:** permite aumentar la visibilidad de las aplicaciones y obtener actualizaciones de manera automática.
- **Interfaz adaptada al sistema operativo al cual el usuario está acostumbrado:** favorece la usabilidad y la interacción con el usuario.

## 5 Diseño prototípico de la aplicación

Se trabajó sobre el proceso de diseño de la interfaz del usuario, desarrollando bocetos prototípicos, que permitan observar las características funcionales como también aspectos interactivos y de percepción.

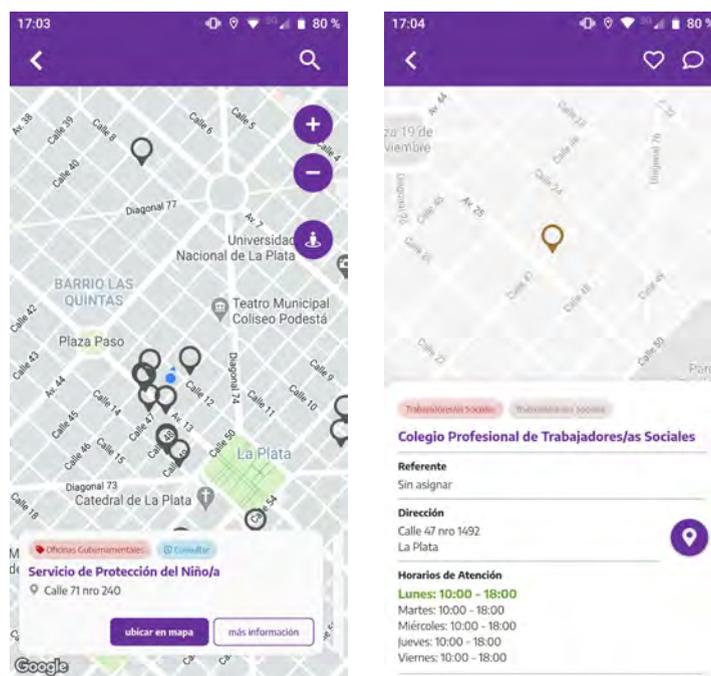
El diseño tiene el siguiente aspecto:

A través de la **pantalla principal** (ver Figura 3) y la vista del **menú lateral de opciones** (ver Figura 4) se podrá realizar llamadas a los principales números de atención, modificar los datos del perfil, realizar ajustes de preferencias de uso y acceder al resto de las funcionalidades:



**Fig. 3.** Vista de la página principal de la aplicación y **Fig. 4.** Vista del menú lateral de opciones.

Desde la actividad de **recursero** (ver Figura 5) se podrá visualizar los centros de atención cercanos a la persona, utilizando el sensor de GPS del dispositivo, y accediendo a la **Información detallada** de cualquiera de estos centros (ver Figura 6), podrá consultar los horarios de atención, llamar a los números telefónicos disponibles, consultar canales alternativos de comunicación (mail, redes sociales, sitios web), realizar reportes que permitan expresar su conformidad o no en la atención recibida, entre otras funcionalidades:



**Fig. 5.** Vista del mapa con el recurso geolocalizado y **Fig. 6.** Vista del detalle de un centro de recurso elegido.

Desde el servicio **Acompañame!** (ver Figuras 7, 8 y 9) se podrá programar un recorrido seguro desde la ubicación en la que se encuentra la persona y hasta un lugar de destino, para ello se deberá indicar el medio de transporte (bici, auto o a pie) para poder medir tiempo, distancia y recorrido sugerido:



**Fig. 7.** Vista inicial para la selección del destino, **Fig. 8.** Vista de la selección del medio de transporte y **Fig. 9.** Vista del recorrido sugerido con la distancia y tiempo estimados

## 6 Revisión de la propuesta

En un marco metodológico de Diseño centrado en el usuario, cada paso fue revisado mediante la convocatoria y participación de mujeres dando lugar a manifestar su opinión respecto a la funcionalidad que se propone brindar en la herramienta como de sus aspectos de diseño.

El prototipo diseñado de alta fidelidad, permitió dar una simulación del producto a desarrollar garantizando experimentar su utilización y navegación antes de su desarrollo definitivo, dando lugar a modificaciones y adecuaciones según las opiniones, críticas o sugerencias realizadas por las mujeres que lo testearon.

Detalles de la evaluación no es objetivo de este artículo, pero es interesante, obtener una retroalimentación sobre el primer impacto que se obtuvo de las personas, en su rol de usuario.

Se realizó la evaluación a 15 mujeres, de las cuales 5 tienen entre 15 a 20 años de edad, 5 entre 21 a 35 años y otras 5, mayores a 36 años.

Luego de utilizar el prototipo de la herramienta leyendo su contenido y probando determinadas funciones, se indagó sobre cuestiones funcionales, como ser: si estaba conforme con la información brindada en la aplicación y su forma de búsqueda, sobre el mapa colaborativo de entidades afines al tema, sobre el servicio de asistencia en la circulación y movimiento de la víctima. Además, se les preguntó sobre cuestiones de interfaz del usuario, como si le resultaba fácil de usar, si la arquitectura de la información le resultaba adecuada y si los aspectos visuales le agradaban.

Se utilizó una escala de Likert [18] de 1 a 5, donde 1 significa estar totalmente en desacuerdo y 5, totalmente de acuerdo.

Los resultados obtenidos arrojaron los siguientes datos:

- Unas 11 mujeres manifestaron estar de acuerdo o totalmente de acuerdo con la información transmitida en la aplicación.
- 13 mujeres estuvieron de acuerdo o totalmente de acuerdo con el mapa del Recursero.
- Todas manifestaron estar de acuerdo o totalmente de acuerdo con el servicio de Acompañame!
- Todas las mujeres coincidieron estar de acuerdo o totalmente de acuerdo con que la herramienta resulta fácil de usar.
- 14 mujeres respondieron estar de acuerdo o totalmente de acuerdo con que la organización de la información le resultó adecuada.
- 12 mujeres manifestaron estar de acuerdo o totalmente de acuerdo con los aspectos visuales.

Entre las opiniones o sugerencias vertidas sobre la funcionalidad, se encuentran:

- Es tan importante tener un lugar centralizado con toda la información. (Emilia, 16 años).
- Qué bien saber que es información formal, verdadera, de fuentes confiables. (Fiorella, 18 años).
- En el Recursero sería ideal mostrar los lugares donde te atienden bien, que haya nivel de satisfacción. porque hay mucho desprecio, descreimiento, hasta inclusive en los lugares para la atención de la mujer.
- No se entiende si en información útil va a estar los números y lugares más importantes o si la información legal, recomendaciones, pasos a seguir, signos a tener en cuenta, prevención, entre otros. (Melina, 21 años)
- Debería existir una opción donde se nuclea información a tener en cuenta como derechos, cómo hacer la denuncia, cómo actuar antes que sea tarde (Carina, 44 años)
- El Recursero además del mapa sería ideal que tenga un listado con los lugares más cercanos a mí.
- La opción de Acompañame! es muy importante, sería ideal que te sugiera caminos o recorridos donde haya cerca entidades del Recursero.
- Acompañame! podría marcarte donde está el agresor? para evitar ese camino. (Alicia, 46 años).
- Se puede guardar los recorridos en Acompañame!?

Con respecto al diseño de la interfaz del usuario:

- Se planteó que quizás para las/los niñas/niños y las/los jóvenes se podría tener una opción con información más directa, como ser: Sabías que? Cómo actuar, Tus derechos, entre otros.
- Mayor contraste de los colores o más grande o gruesa la letra para que se distinga mejor.

Estos aportes de las mujeres que participaron en el testeado del prototipo, constituye una fuente de información muy importante a incorporar en la versión definitiva de la aplicación.

## Conclusiones

La violencia de género es una problemática social y cultural que atraviesa diferentes estratos sociales a nivel mundial. En la Argentina, en el año 2019 se estimó una muerte por femicidio cada 33 horas y, en lo que va del año 2020 ya se registran más de 70 casos. Si bien se ha trabajado desde el estado y las ONG en implementación de programas, recursos, centros de atención, entre otros, con eso no alcanza, basta revisar los datos que arroja la realidad.

A través del desarrollo de la herramienta Genvi se intenta aportar un recurso más, que brinde alternativas de solución a necesidades que aún están desatendidas. Es una herramienta simple de utilizar, accesible y colaborativa puesta en un celular, que permite que la persona que lo utilice pueda no solo contar con información importante de forma más directa y rápida, sino que también pueda sentirse un poco más acompañada al momento de transitar por la vía pública.

Pruebas realizadas sobre un prototipo de alta fidelidad arrojaron buenos resultados y una gran aceptación de la aplicación y su funcionalidad, lo que incentiva aún más para continuar y finalizar el desarrollo para que la misma pueda estar disponible para quien quiera y necesite usarla.

## Referencias

1. Ley 26.485, <https://www.argentina.gob.ar/normativa/nacional/ley-26485-152155>, último acceso 2020/06/01
2. Registro Nacional de Femicidios de la Justicia Argentina (RNFJA), <https://www.csjn.gov.ar/omrecopilacion/omfemicidio/homefemicidio.html>, último acceso 2020/06/01
3. Organización Mundial de la Salud: Informe mundial sobre la violencia y la salud: resumen. (2002)
4. Rico, María Nieves: Violencia de género: un problema de derechos humanos. (1996).
5. Naciones Unidas: <https://www.unwomen.org/es/what-we-do/ending-violence-against-women/global-norms-and-standards>, último acceso 2020/06/01
6. Línea 144, <https://www.argentina.gob.ar/generos/linea-144>, último acceso 2020/06/01
7. Línea 137, <https://www.argentina.gob.ar/justicia/violencia-familiar-sexual>, último acceso 2020/06/01
8. Centro de Acceso a Justicia (CAJ), <https://www.argentina.gob.ar/justicia/afianzar/caj/conoce-y-ejerce-tus-derechos/guia-de-informacion-sobre-violencia-contra-las-mujeres>, 2019/11/19
9. Recurso del Instituto Nacional de las Mujeres, <https://www.argentina.gob.ar/mapamujeres>, último acceso 2019/10/17

10. Tobilleras Mujeres Seguras,  
<http://www.buenosaires.gob.ar/desarrollohumanoyhabitat/mujer/tobilleras-mujeres-seguras>  
, último acceso 2019/11/19
11. Dispositivo de Alerta para Mujeres Agredidas (DAMA),  
[http://servicios.tigre.gob.ar/prensa/3398\\_violencia\\_de\\_genero\\_casi\\_100\\_aprehendidos\\_por\\_sistema\\_sms\\_y\\_dama\\_en\\_tigre.htm](http://servicios.tigre.gob.ar/prensa/3398_violencia_de_genero_casi_100_aprehendidos_por_sistema_sms_y_dama_en_tigre.htm) , último acceso 2019/11/19
12. Baldauf, M., Dustdar, S. and Rosenberg, F.: A survey on context-aware systems. In: Int. J. Ad Hoc and Ubiquitous Computing, Vol. 2, No. 4, pp. 1. (2007)
13. REST, [https://www.ics.uci.edu/~fielding/pubs/dissertation/rest\\_arch\\_style.htm](https://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm), último acceso 2019/12/10
14. Hatem Hamad, Motaz Saad, Ramzi Abed: Performance Evaluation of RESTful Web Services for Mobile Devices. (2010)
15. Android, <https://developer.android.com/>, último acceso 2020/06/20
16. iOS, <https://developer.apple.com/> , último acceso 2020/06/02
17. Flutter, <https://flutter-es.io/> , último acceso 2020/06/02
18. Likert, Rensis: A technique for the measurement of attitudes. (1932)

## Sistema para el registro de factores psicológicos que inciden en los trastornos gastrointestinales funcionales. “Digestivamente”.

Juan Zaragoza <sup>1</sup>, Jorge Ierache <sup>1</sup>, Nicolas Domínguez <sup>1</sup>, Hernan Merlino <sup>1</sup>, Milagros Celleri <sup>2</sup>, Florencia Koutsovitis <sup>2</sup>, Cristian Garay <sup>2</sup>.

Laboratorio de Sistemas de Información Avanzados, Departamento Computación, Facultad de Ingeniería, Universidad de Buenos Aires, Argentina<sup>1</sup>

{Jierache, jzaragoza, ldominguez, hmerlino}@fi.uba.ar

<http://lsia.fi.uba.ar/1>

Facultad de Psicología, Universidad de Buenos Aires, Argentina <sup>2</sup>  
{mcelleri, fkoutsovitis, cristiangularay}@psi.uba.ar

**Resumen.** El trabajo se centra en el desarrollo de un prototipo de aplicación móvil, su sistema de gestión médico y explotación de datos de los pacientes. El desafío del mismo lleva a entender algunos conceptos claves acerca 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.

**Keywords:** sistema de registro; factores psicológicos; intestino irritable; trastornos gastrointestinales funcionales.

### 1 Introducción

Este desarrollo 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. Los trastornos gastrointestinales funcionales son considerados un problema en la Salud Pública, debido a su alta prevalencia y los altos costos que generan, tanto económicos como de recursos en el sistema de salud, siendo los más prevalentes el síndrome de intestino irritable y la dispepsia funcional. Estos trastornos se caracterizan por la ausencia de una fisiopatología clara que explique el cuadro [1]. El síndrome de intestino irritable [2]. 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 [3]. 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 endoscopías 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 [4]. 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 [5]. El tratamiento psicológico de primera elección para el síndrome de intestino irritable es la terapia cognitivo-conductual (TCC) [6] y, en la dispepsia funcional también este tipo de terapia ha demostrado buenos resultados [7]. La TCC supone que los pensamientos, las emociones y las conductas de las personas se encuentran relacionados y se influyen mutuamente. La 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 [8]. En la sección 2 se presenta un resumen de los trabajos relacionados, destacando algunas aplicaciones que contribuyen con la problemática de del Síndrome de Intestino Irritable (SII). En la sección 3 se aborda la propuesta del sistema, en la sección 4 se presenta los resultados del desarrollo realizado que nominamos “Digestivamente”, en la sección 4 las conclusiones y futuras líneas de trabajo e investigación

## **2 Trabajos relacionados con aplicaciones en el contexto de la problemática del Síndrome de Intestino Irritable**

En particular aplicaciones como: FODMAP para la asistencia en dieta baja caloría, ayudan a disminuir los síntomas SII, con un control de los alimentos que se consumen [X9]. La app B.Energy, de ALVUM [X10] 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 ganar en autoconfianza por parte de los pacientes. Otra app es mySymptoms Food Diary & Symptom Trackerpermite [X11], permite hacer seguimiento de los síntomas y de la alimentación, y promete encontrar relaciones a través de un algoritmo de cálculo específico, ya que, de hecho, es un problema muy difícil de resolver porque los síntomas muchas veces no vienen inmediatamente al terminar las comidas, facilita exportar informes para compartirlos con médicos o nutricionistas sino horas después. La aplicación Bowelle realiza el seguimiento de síntomas y los relaciona también con las comidas. [X12], 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, sin embargo esta primera versión no incorpora información de los alimentos que ingiere el paciente.

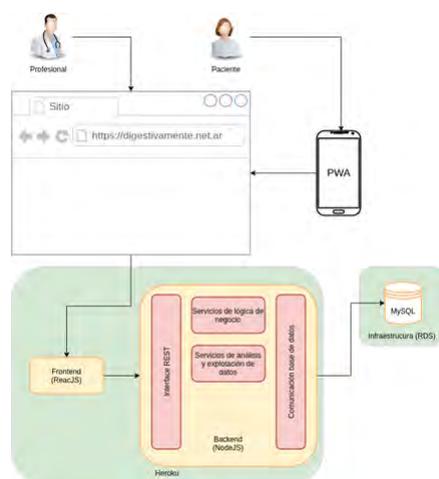
### **3 Objetivo del Sistema**

El principal objetivo de la aplicación es contribuir a mejorar la calidad de vida de las personas que sufren SII, con capacidad de registro de síntomas psicológicos en pacientes bajo tratamiento en un hospital públicos. El Síndrome de Intestino Irritable constituye una patología con alta prevalencia en la consulta en los hospitales y que se vincula con elevados costos económicos. No se conoce un claro mecanismo patofisiológico que explique el cuadro, pero se conocen distintos factores que influyen en su desarrollo y mantenimiento. Entre estos factores, se encuentran los factores psicológicos, principalmente el estrés, la ansiedad y la depresión. La utilización de tecnología y tecnología móvil en salud ha ido incrementando globalmente en los últimos años, y se han comenzado a utilizar distintas aplicaciones para el registro e intervención en diversos problemas clínicos, mostrando resultados alentadores. El objetivo 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, facilitar el acceso y explotación de datos por parte del personal médico. Se trata de un trabajo descriptivo y transversal. En una primera etapa de la prueba piloto, se evaluó el funcionamiento de la aplicación, se realizarán en una segunda etapa encuestas de usabilidad de la aplicación entre los pacientes en la etapa de pruebas. 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.

### **4 Resultados del desarrollo del sistema “Digestivamente”**

El sistema desarrollado denominado “Digestivamente”. permite entre sus principales funcionalidades: a) Registro de los valores diarios de Escala de Bristol de los pacientes, sus sensaciones físicas y emocionales, como así también registros de pensamientos y actividades, b) Ofrece una interfaz que le permite apropiadamente a los pacientes determinar los valores apropiados de la escala de Bristol, c) Registra las emociones que enfrentan los pacientes al finalizar el día, d) Registra los estresores semanales que los pacientes han vivido durante el transcurso de la semana., e) Evalúa los resultados de cada paciente a través de gráficos que muestran las relaciones entre cada una de las variables mencionadas anteriormente por parte de los médicos y facilite 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) Carga y administra los diferentes usuarios: médicos y pacientes, h) Visualiza y analiza la información recolectada para cada uno de los pacientes asistiendo de esta forma a la identificación de las causas de los síntomas encontrados, i) Permite 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 visualizar los datos 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 [13], como entorno en tiempo de ejecución multiplataforma, de código abierto, para la capa del servidor ,c) React React.js [14] biblioteca Javascript de código abierto empleada en el desarrollo de las interfaces de usuario con el objetivo de facilitar el desarrollo de aplicaciones en una sola página; como la principal herramienta durante el desarrollo de nuestra aplicación web, d) Material UI Framework Material UI [15] librería con componentes React se aplicó para desarrollo web rápido y eficiente basado en Material Design, un lenguaje visual que sintetiza los principios clásicos del buen diseño con la innovación de la ciencia y la tecnología, e) npm De sus siglas NPM (Node Package Manager) [16] se seleccionó en razón a que es un gestor de paquetes desarrollado en su totalidad bajo el lenguaje JavaScript por Isaac Schlueter, f) Express.js[17],o Express, framework de la aplicaciones web para Node.js, g) MySQL [18] sistema de gestión de base de datos relacional (RDBMS), h) AWS Amazon Web Services (AWS) [19] plataforma de servicios de nube ,i) Heroku Heroku [20] como plataforma de servicio (PaaS) que permite construir, correr y operar aplicaciones enteramente en la nube, j) db4free.net db4free.net [21], se aplicó en razón a que ofrece un servicio de pruebas para la última versión de un servidor MySQL, k) para crear experiencias de usuario similares a las aplicaciones nativas en dispositivos móviles y de escritorio PWA Una Progressive Web Application o PWA [22], este es un tipo de software web, creado con tecnologías como HTML, CSS y JavaScript. Se encuentra destinado a funcionar en cualquier plataforma que use un navegador web compatible con los estándares. Entre sus

funcionalidades encontramos: trabajar sin conexión, notificaciones push y acceso al hardware del dispositivo. Se presenta en la figura 1 la arquitectura del prototipo junto con los componentes que lo identifican



**Fig. 1** Diagrama la arquitectura del prototipo junto con los componentes que lo identifican

Se muestran a continuación las pantallas de usuario aplicación móvil para cargar un registro diario, en figura 2a el login, como así también el video inicial de guía [23], en figura 2b el inicio de carga de registro escala de Bristol, en figura 2c la carga del registro de sensaciones y pensamientos.

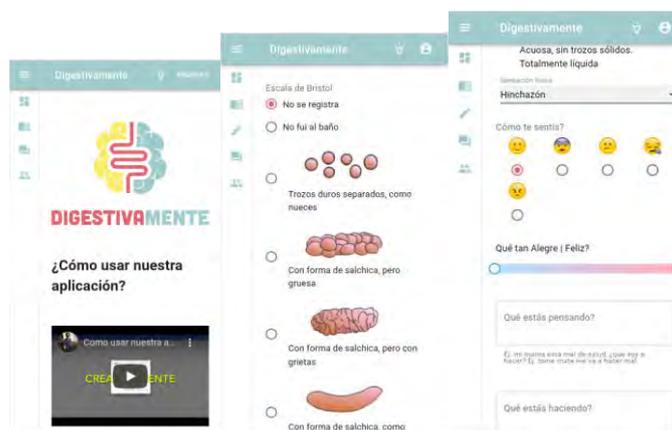


Fig 2a. Dashboard

Fig. 2b. items Bristol

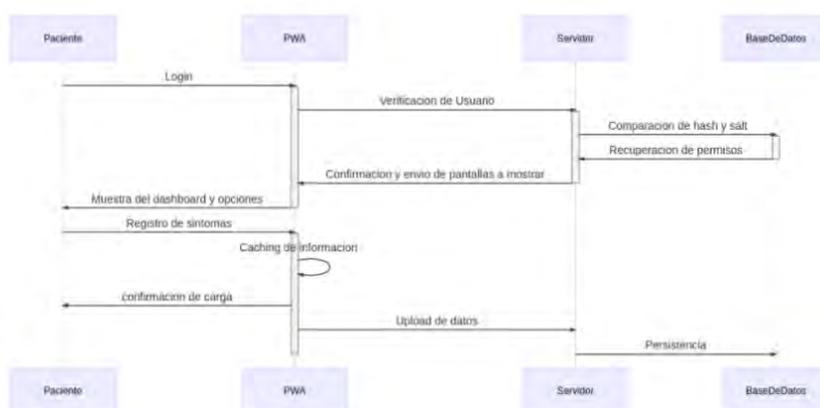
Fig, 2c. sensaciones y pensamientos.

En particular se detalla en la tabla 1 la clasificación de las Emociones y sus emojis empleados en la aplicación.

Emoción	Emoji
Alegre   Feliz	😊
Miedoso/a/   Asustado/a	😱
Preocupado/a   Ansioso/a	😟
Triste   Angustiado/a	😞
Enojado/a	😡

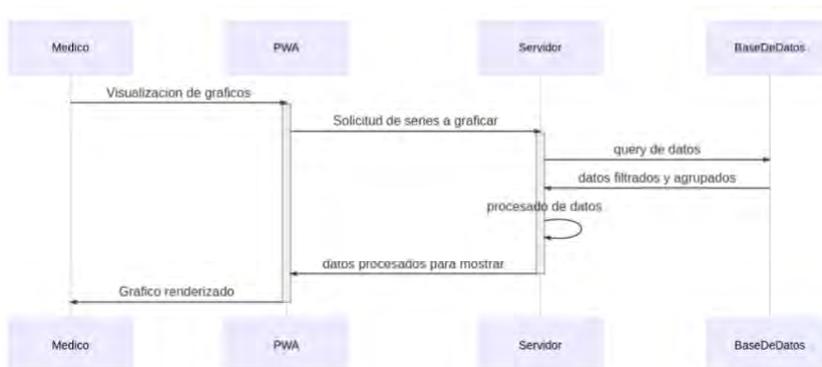
**Tabla 1.** Emociones y sus emojis

En la figura 3 se presenta el diagrama de secuencia correspondiente al registro de datos de paciente.



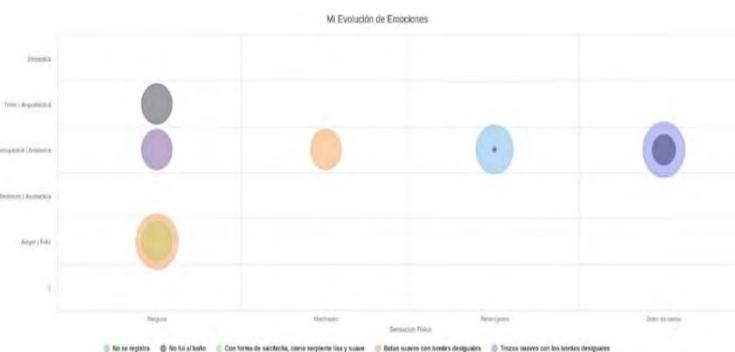
**Fig. 3** Diagrama de secuencia de registros de datos de pacientes

En la figura 4 el diagrama de secuencia correspondiente a la visualización de la información gráfica por parte de los usuarios

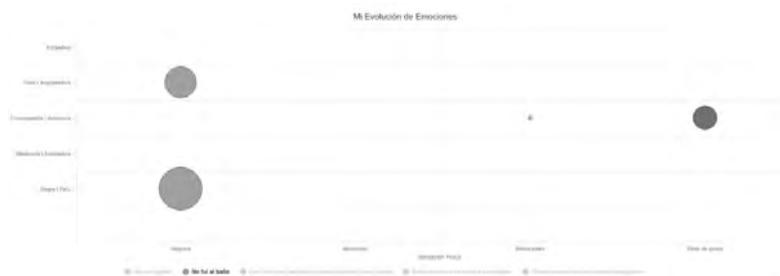


**Fig. 4 Diagrama de secuencia de visualización de gráficos**

Se presenta a continuación resultados generales de uno de los casos de prueba realizado, en este orden en la figura 5. se puede observar la visualización de un seguimiento de los valores de la escala Bristol, las emociones, y síntomas en un periodo de un mes. La correlación de las emociones (eje y) y las sensaciones físicas (eje x) agrupadas por escala de Bristol. El tamaño de los círculos representa la predominancia de una emoción en relación con los síntomas y la escala de Bristol indicadas en color. En figura 6 se muestra para un valor seleccionado de la escala de Bristol, que emociones y sensaciones tuvo el paciente, en particular el estado emocional preocupado y en la condición Bristol no fui al baño, se manifiesta el síntoma de dolor de panza en mayores ocasiones que el de retorcionjes .



**Fig. 5 Evolución de emociones- síntomas de un paciente provisto por la interfaz de consultas del sistema relacionando registro deposiciones según escala Bristol**



**Fig. 6** para un valor particular “no fui al baño” se muestran la emociones y sensaciones físicas de un paciente

#### 4. Conclusiones y futuras líneas de investigación

Las pruebas basales fueron satisfactorias, permitieron mejorar las funcionalidades del prototipo en desarrollo, como así también el desarrollo de nuevas capacidades en orden a la explotación y representación de datos para uso de pacientes en particular para que éstos observen su evolución y situación frente a estresores, como así también que el equipo profesional de la salud pueda realizar un seguimiento detallado y la explotación de datos comparativos en función de diagnósticos y tratamientos.. Las futuras líneas de investigación, se orientan al descubrimiento de patrones y la explotación de reconocimiento automático de imágenes en el contexto de la escala de Bristol. Las líneas de trabajo actuales se concentran en la seguridad con los estándares de calidad necesarios para un ambiente productivo, como así también en las adecuaciones de explotación del sistema y las mejoras usabilidad de la aplicación del paciente, psicólogos y médicos. En el marco de otras de las líneas de trabajo se considerara incorporar el registro de la alimentación del paciente

## Referencias

1. Mahadeva S, Goh KL. Epidemiology of functional dyspepsia: A global perspective. *World J Gastroenterol.* 2006;12(17):2661–6.
2. Brizuela Quintanilla RA. Síndrome de intestino irritable. *Rev Cuba Med Mil.* 1997;26(1):63–8.
3. Sebastián Domingo JJ. The new Rome criteria (IV) of functional digestive disorders in clinical practice. *Med Clin (Barc)* [Internet]. 2017;148(10):464–8. Available from: <http://dx.doi.org/10.1016/j.medcli.2016.12.020>
4. 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.
5. Jia L, Jiang SM, Liu J. Behavioral gastroenterology: An emerging system and new frontier of action. *World J Gastroenterol.* 2017;23(33):6059–64.
6. American Psychological Association. Diagnosis: irritable bowel syndrome, psychological treatments [Internet]. [Consultado 28 may 2020]. Disponible en: <https://www.div12.org/diagnosis/irritable-bowel-syndrome/>
7. 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.
8. Beck J. Terapia cognitiva: conceptos básicos y profundización. Editorial Gedisa; 1995.
9. 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
10. <https://www.alvumhealth.com/2019/02/21/alvum-lanza-b-energy-aplicacion-movil-sindrome-intestino-irritable-sii-colon-irritable/>, [Consultado Sep 2020].
11. [https://colon-irritable.es/5-apps-android-iphone-intestino-colon-irritable/#mySymptoms\\_Food\\_Diary\\_Symptom\\_Tracker](https://colon-irritable.es/5-apps-android-iphone-intestino-colon-irritable/#mySymptoms_Food_Diary_Symptom_Tracker) [Consultado Sep 2020].
12. <https://colon-irritable.es/5-apps-android-iphone-intestino-colon-irritable/#Bowel> [Consultado Sep 2020].
13. NodeJS. Entorno de ejecución para JavaScript construido con el motor de JavaScript V8 de Chrome. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://nodejs.org/>
14. React. Biblioteca de JavaScript para desarrollo de clientes web. [Internet]. [Consultado 29 may 2020]. Disponible en: <https://reactjs.org/>
15. Material UI. Componentes React para un desarrollo web rápido y fácil. [Internet]. [Consultado 29 may 2020]. Disponible en: <https://material-ui.com/>
16. Npm. Sistema de gestión de paquetes por defecto para Node.js. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://www.npmjs.com/>
17. Express. Infraestructura web rápida, minimalista y flexible para Node.js. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://expressjs.com/>
18. MySQL. Sistema de gestión de base de datos relacional. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://www.mysql.com/>
19. AWS. Servicios de computación en la nube pública de Amazon. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://aws.amazon.com/>
20. Heroku. Plataforma como servicio en la nube. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://www.heroku.com/>
21. Db4free.net. Servicio de pruebas de base de datos MySQL. [Internet]. [Consultado 28 may 2020]. Disponible en: <https://www.db4free.net/>
22. PWA. Plataformas web en dispositivos móviles. [Internet]. [Consultado 28 may 2020]. Disponible en: [https://en.wikipedia.org/wiki/Progressive\\_web\\_application](https://en.wikipedia.org/wiki/Progressive_web_application)
23. <https://www.youtube.com/watch?v=82C1D4jZats&feature=youtu.be> [Consultado 29 may 2020].

# Workshop Procesamiento de Señales y Sistemas de Tiempo Real

Coordinadores

Horacio Villagarcia  
Wanza (UNLP)  
Emanuel Frati (UNdeC)  
Jorge Ierache (UM)

## Diseño y desarrollo de un sistema embebido para una trampa pitfall con data logger

Mariano R. Droz<sup>1</sup>, Juan A. Ramos<sup>1</sup>, Pedro D. Benitez<sup>1</sup>, Luz M. Zapata<sup>1</sup>, Beatriz M. Diaz<sup>2</sup>

<sup>1</sup> Facultad de Ciencias de la Alimentación, Universidad Nacional de Entre Ríos, Av. Monseñor Tavella N° 1450, (CP 3200) Concordia, Entre Ríos, República Argentina  
{mariano.droz, juan.ramos, pedrodaniel.benitez, luzmarina.zapata}@uner.edu.ar  
<http://www.fcal.uner.edu.ar>

<sup>2</sup> Instituto Nacional de Tecnología Agropecuaria (INTA), Estación Experimental Agropecuaria Concordia (EEA Concordia), Ruta Provincial 22 y vías del Ferrocarril, (CP 3200) Estación Yuquerí, Concordia, Entre Ríos, República Argentina  
diaz.beatriz@inta.gob.ar  
<https://inta.gob.ar/concordia>

**Abstract.** Las trampas de caída (pitfall) son utilizadas en estudios de biodiversidad de los artrópodos del suelo. Los avances tecnológicos posibilitaron su evolución surgiendo las trampas de caída por tiempo. El objetivo fue diseñar y desarrollar un sistema embebido para un prototipo automatizado de pitfall dotada de un data logger y sensores meteorológicos y edáficos. La metodología incluyó actividades propias del proceso de diseño de software de tiempo real y, complementariamente, utilizó un enfoque iterativo e incremental. El prototipo se conformó por 3 módulos, el Data Logger, el Recolector y el de Alimentación. Se utilizó Arduino como plataforma de desarrollo y mediante 6 ensayos de laboratorio se comprobó que el sistema cumplió los requerimientos funcionales y que su diseño arquitectónico contribuye al mantenimiento y evolución del dispositivo. Los resultados indican la factibilidad de concretar ensayos a campo en diferentes ambientes productivos, contemplando en el futuro la utilización de Minería de Datos.

**Keywords:** Sistema embebido, Trampa de caída (pitfall) por tiempo, Data logger, Sensores, Artrópodos.

### 1 Introducción

Entre los instrumentos utilizados para llevar a cabo estudios de biodiversidad de los artrópodos del suelo, se encuentran las trampas de caída, denominadas también como “pitfall”. El diseño básico de este tipo de trampa consiste en un contenedor enterrado en el suelo con la parte superior a ras de la superficie [1] y su uso se remonta a más de 120 años [2].

Éstas funcionan según el principio de que un artrópodo que se mueve en la superficie simplemente cae en un contenedor abierto (normalmente circular) excavado en el suelo. Los artrópodos generalmente mueren en una solución conservante y se devuelven a un laboratorio para su posterior clasificación e identificación [3]. De este modo, la trampa pitfall genera una estimación de la abundancia-actividad, considerando que la abundancia es un indicador de la actividad de los artrópodos, dando lugar así a su uso para estudios de biodiversidad de la fauna edáfica en diferentes escenarios, naturales o productivos [4, 5].

Producto de la evolución tecnológica de las trampas pitfall convencionales surgieron las trampas de caída por tiempo (time-sorting pitfall trap). Éstas conservan el mismo principio que las tradicionales, con la diferencia principal que el dispositivo recolector o contenedor va cambiando mediante una rotación automática cada cierto intervalo de tiempo previamente programado. Una vez cumplido el ciclo completo de rotación, se accede a la trampa y se retiran los contenedores para ser llevados al laboratorio [6, 7].

Más allá que, actualmente es posible diseñar y desarrollar unidades de registro de datos (data loggers) con hardware abierto y de bajo costo [8, 9], éstas trampas analizadas no utilizaron sensores de luz, de temperatura y humedad relativa ambiente, y de temperatura y humedad del suelo de manera simultánea.

Teniendo en cuenta lo expuesto, investigadores del Grupo Hortícola de la EEA Concordia del INTA plantearon a investigadores de la carrera Ingeniería en Mecatrónica de la Facultad de Ciencias de la Alimentación, la necesidad de avanzar en un desarrollo local adaptado a sus requerimientos.

Por ello, uno de los objetivos fue diseñar y desarrollar un sistema embebido para un prototipo automatizado de trampa de caída por tiempo, que utilice hardware abierto y de bajo costo, dotado de un data logger más completo, con pantalla y mayor cantidad de sensores, el cual permita medir simultáneamente la luz, la temperatura y humedad relativa ambiente, y la temperatura y humedad del suelo. A su vez, otro objetivo consistió en que el sistema cuente con una funcionalidad innovadora, que permita la fácil parametrización de los ensayos en estudios de biodiversidad de artrópodos de suelo y la posibilidad de soportar 2 modos de operación.

## 2 Metodología

Se efectuaron actividades propias de un proceso de diseño de software de tiempo real [10]: selección de la plataforma, identificación de estímulos/respuestas, análisis de temporización, diseño de procesos, diseño de algoritmos, diseño de datos y planificación del proceso.

Además, se efectuaron otras actividades relacionadas a la selección y adquisición de componentes accesorios, estructurales, mecánicos y electrónicos. Sumado a ello, se realizó el diseño e impresión de una placa de circuito impreso (PCB) y piezas 3D.

Teniendo en cuenta que el orden de las actividades en el proceso de diseño de software de tiempo real depende del tipo de sistema a desarrollar, así como de sus requerimientos de proceso y plataforma [10], se decidió utilizar también un enfoque iterativo e incremental [11].

La selección de este abordaje metodológico complementario radicó en que luego del análisis preliminar se arribó a la conclusión que sería práctico dividir el trabajo en partes más pequeñas o miniproyectos. Vale señalar que las iteraciones fueron controladas y planificadas [11], permitiendo en algunas instancias el avance de miniproyectos en paralelo.

### 3 Arquitectura del Hardware

El dispositivo está compuesto por 3 módulos, el *Módulo Data Logger (MDL)*, el *Módulo Recolector (MR)* y el *Módulo de Alimentación (MA)* (Fig. 1). Este diseño modular permite que cada módulo cuente con una funcionalidad específica y definida, estableciendo las interfaces de conexión y comunicación entre los mismos.

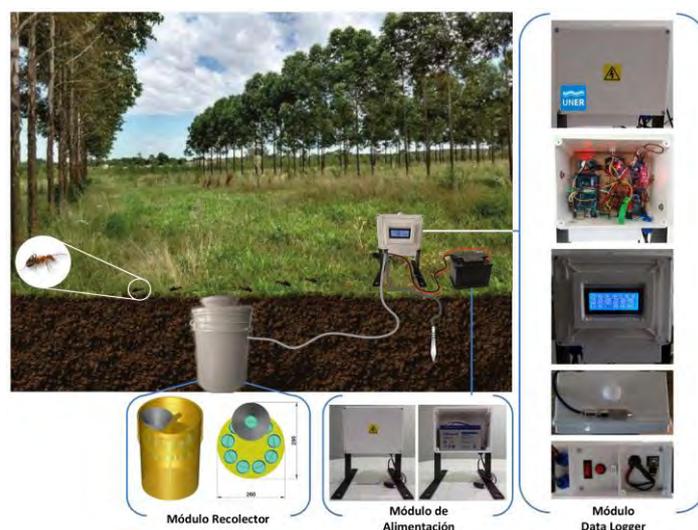


Fig. 1. Arquitectura del Hardware dispuesta en Módulos.

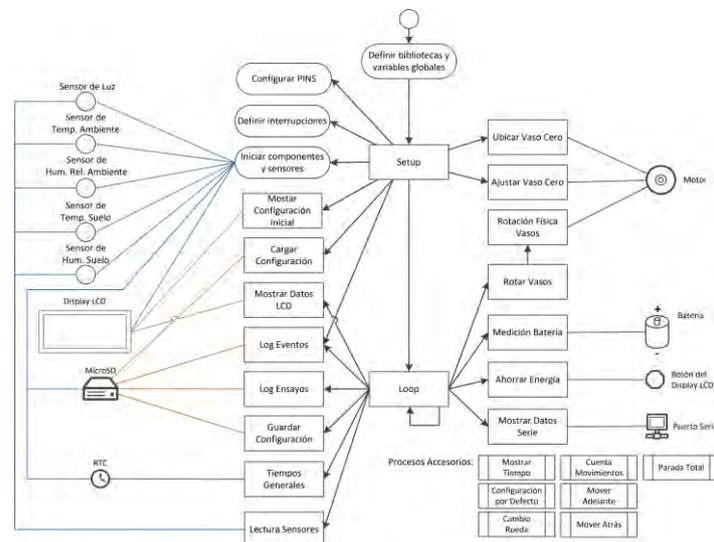
### 4 Sistema embebido

El software se programó utilizando el entorno de desarrollo Arduino IDE 1.8.10, una placa Arduino Mega 2560 Rev. 3 y cuenta con 1760 líneas de código. Para el diseño y desarrollo del sistema embebido se tuvieron en cuenta estos requerimientos funcionales (RF):

- 1) *Parametrizar el ensayo*: Permitir la configuración del ensayo a través de un archivo de configuración que se almacene en la tarjeta MicroSD.
- 2) *Controlar el funcionamiento del dispositivo*: Efectuar un control, al iniciar y durante su funcionamiento, de los principales componentes del dispositivo.

- 3) *Mostrar información para control*: Informar al operador, a través del display LCD, el resultado de los controles iniciales, los parámetros configurados, el modo y estado de funcionamiento.
- 4) *Registrar datos*: Registrar, cada cierto intervalo de tiempo configurado, la fecha, la hora, las mediciones de los sensores y el voltímetro, el número de rotación y el número de recipiente colector (vaso) utilizado.
- 5) *Cambiar el recipiente colector*: Rotar el disco del *Módulo Recolector* cada cierto intervalo de tiempo configurado para utilizar otro recipiente colector.
- 6) *Ahorrar energía*: Cuando el dispositivo no efectúe alguna tarea entrar en modo reposo hasta el próximo evento programado, o bien, hasta que el operador accione el botón de encendido del display LCD.
- 7) *Continuar un ensayo en proceso*: Retomar un ensayo en proceso cuando el mismo haya sido interrumpido.
- 8) *Registrar los principales eventos del sistema*: Llevar una bitácora con la información de los principales eventos del sistema.
- 9) *Soportar 2 modos de operación*: Permitir que el operador, a través del archivo de configuración que se almacene en la tarjeta MicroSD, indique el modo de funcionamiento del dispositivo: *Modo Data Logger* donde se omite el RF N° 5 y se reduce la información proporcionada para el RF N° 3, y el *Modo Completo (Modo Pitfall)* que abarca la totalidad de los RF.

Una vez analizados los requerimientos funcionales y luego del proceso de desarrollo guiado por el enfoque metodológico previamente descrito, se construyó el sistema embebido que se presenta en la Fig. 2.



**Fig. 2.** Arquitectura del sistema embebido (procesos y principales componentes).

En la plataforma Arduino [12] existen 2 procesos fundamentales, *Setup* que se ejecuta sólo una vez al iniciar y *Loop* que se ejecuta indefinidamente hasta que la

placa se apaga. A partir de ello, a continuación se describen cada uno de los principales componentes del código fuente representado en la Fig 2.

- *Definir bibliotecas y variables globales*: Llamada a las bibliotecas necesarias para utilizar los sensores, los componentes (el display LCD, la tarjeta MicroSD y el Real Time Clock RTC), el protocolo I2C (utilizado por el RTC, el display LCD, el sensor de luz y el sensor de temperatura y humedad del suelo) y el modo de ahorro de energía. Además, se efectúa la declaración de las constantes y variables globales, los prototipos de las funciones y los procedimientos.

- *Configurar PINS*: Se declaran los PIN a utilizar y si los mismos serán empleados como entrada o salida.

- *Definir interrupciones*: Se definen las 2 interrupciones utilizadas y el modo de disparo de las mismas. Una interrupción está asociada al botón de encendido y apagado del display LCD y, la otra, al sensor fotoeléctrico dispuesto en el *Módulo Recolector* para controlar los movimientos.

- *Iniciar componentes y sensores*: Se inician todos los sensores y los componentes utilizados. Asimismo, se efectúa un control de los mismos para detectar el estado de su funcionamiento, informando el resultado en el display LCD (Fig. 3). La continuidad del proceso de inicio depende de un resultado satisfactorio de dichos controles. Permite cubrir parcialmente el RF N° 3.



**Fig. 3.** a) Pantalla al *Iniciar componentes y sensores*. b) y c) Pantallas del procedimiento *Mostrar Configuración Inicial*.

- *Cargar Configuración*: Procedimiento que permite cumplir con los RF N° 1, 7 y 9. Lee y procesa el archivo de configuración CONFIG.INI (Tabla 1) desde la tarjeta MicroSD, si la operación es exitosa, las variables de operación se establecen con los valores leídos; si se encuentran errores durante esta operación, se invoca al procedimiento *Configuración por Defecto* para establecer los valores de las principales variables del sistema.

- *Ubicar Vaso Cero*: Procedimiento invocado cuando el modo de operación es completo. Se ocupa de ubicar el vaso cero, es decir, el vaso residual con el que comienza un ensayo hasta que se cumplen las condiciones para el primer movimiento.

- *Ajustar Vaso Cero*: En el modo de operación completo, este procedimiento es invocado una vez que se ha ubicado el vaso cero, se efectúa un movimiento de ajuste para que dicho vaso quede correctamente centrado con respecto al embudo.

- *Mostrar Configuración Inicial*: Este procedimiento se invoca en la etapa final del *Setup* y se encarga de mostrar en el display LCD los valores de las principales variables. De este modo, el operador verifica que el ensayo inicie con los parámetros deseados. Este es uno de los procedimientos que se ocupa del RF N° 3. Las pantallas con esta información se encuentran en la Fig 3.

- *Log Eventos*: Procedimiento cuyo objetivo es dar cumplimiento al RF N° 8. Para ello, escribe los datos en el archivo EVENTOS.CSV residente en la tarjeta MicroSD (Fig. 4a). Es invocado al finalizar el *Setup* para registrar el arranque del dispositivo y en el *Loop* cuando resulta necesario.

- *Log Ensayos*: Procedimiento cuyo objetivo es dar cumplimiento al RF N° 4. Para ello, escribe los datos en el archivo ENSAYOS.CSV residente en la tarjeta MicroSD (Fig. 4b).

#### a) EVENTOS.CSV

```
FECHA;HORA;TIPO;EVENTO
21/5/2020;9:12:24;S;ENCENDIDO
21/5/2020;9:12:24;E;AHORRO POR 225 BLOQUES DE 8 SEGUNDOS
21/5/2020;9:12:44;E;ACTIVO
21/5/2020;9:12:57;D;LCD ON
21/5/2020;9:12:57;E;AHORRO POR 221 BLOQUES DE 8 SEGUNDOS
```

#### b) ENSAYOS.CSV

```
FECHA;HORA;LUZ;TEMP_AMB;HUM_AMB;TEMP_SUELO;HUM_SUELO;V_BAT;PORC_BAT;ROTACION;VASO
21/5/2020;9:48:49;9;20.7000007629;74.0999984741;21.0999984741;83.3480300903;12.67;90;0;0
21/5/2020;10:17:39;20;22.6000003814;67.5000000000;21.2899971008;83.1058959960;12.65;90;1;1
21/5/2020;11:30:50;12;23.3999996185;66.0999984741;22.2799987792;82.0081253051;12.59;80;1;1
21/5/2020;12:44:2;4;24.5000000000;65.3000030517;21.9499969482;84.4952392578;12.53;80;1;1
```

**Fig. 4.** Formato de los archivos EVENTOS.CSV y ENSAYOS.CSV.

- *Ahorrar Energía*: Procedimiento que es ejecutado cuando se dispara la interrupción al presionar el botón de encendido y apagado del display LCD. Establece el valor de una variable para determinar si el dispositivo debe o no activar el ahorro de energía hasta el próximo evento programado. Es fundamental para dar cumplimiento al RF N° 6.

- *Tiempos Generales*: Procedimiento que es invocado en muchas partes del código ya que actualiza la fecha, la hora y variables relacionadas al tiempo, las cuales son muy importantes para el funcionamiento del sistema.

- *Lectura Sensores*: Procedimiento encargado de efectuar las lecturas de los sensores y verificar el funcionamiento de los mismos.

- *Medición Batería*: Procedimiento que mediante un voltímetro efectúa la medición del voltaje de carga de la batería y determina el porcentaje de autonomía restante.

- *Rotar Vasos*: Procedimiento que controla y determina el número de rotación, el vaso actual y el siguiente, verificando que los valores se encuentren de acuerdo a lo programado. Luego de ello, invoca al siguiente procedimiento que efectúa y controla la rotación física de los vasos. Es el principal procedimiento que se ocupa del RF N° 5.

- *Rotación Física Vasos*: Procedimiento que se ocupa de controlar y accionar el funcionamiento del *Módulo Recolector*, efectuando la rotación de los vasos y controlando los movimientos necesarios para que cada vaso quede alineado al embudo.

- *Mostrar Datos LCD*: Proporciona información sobre la fecha, la hora, el modo de operación, los errores encontrados, las mediciones de los sensores y los tiempos de operación. Este es uno de los procedimientos que se ocupa del RF N° 3. Las pantallas con esta información se encuentran en la Fig 5.



**Fig. 5.** Pantallas del procedimiento *Mostrar Datos LCD*.

- *Guardar Configuración*: Procedimiento que hace persistentes en el archivo CONFIG.INI de la tarjeta MicroSD los cambios de los parámetros del ensayo a medida que los mismos son actualizados. Este procedimiento en combinación con *Cargar Configuración* permite dar cumplimiento a los RF N° 7 y 9.

- *Mostrar Datos Serie*: Procedimiento empleado durante el desarrollo del sistema, permite conocer los valores de las diferentes variables y efectuar la corrección de errores y la depuración del sistema.

- *Mostrar Tiempo*: Procedimiento que convierte el tiempo pasado como parámetro a una notación +/-HHH:MM:SS. Es utilizado por el procedimiento *Mostrar Datos LCD* para proporcionar al operador la información de los tiempos en una notación entendible, permitiendo controlar el número de caracteres utilizados en el display LCD.

- *Configuración por Defecto*: Procedimiento que parametriza el ensayo con valores por defecto, es invocado en el *Setup* sólo cuando el procedimiento *Cargar Configuración* ha presentado algún error.

- *Cambio Rueda*: Procedimiento asociado a la interrupción empleada al iniciar el dispositivo durante el *Setup*. Es invocado por la interrupción que se dispara cuando el sensor fotoeléctrico del *Módulo Recolector* detecta un cambio de señal de alto a bajo (FALLING). Funciona de manera complementaria con el procedimiento *Ubicar Vaso Cero*, durante la configuración inicial del dispositivo.

- *Cuenta Movimientos*: Procedimiento ligado a la interrupción que se dispara cuando el sensor fotoeléctrico del *Módulo Recolector* detecta un cambio de señal (CHANGE), luego que el vaso cero ha sido ubicado. Actúa complementariamente con los procedimientos *Ajustar Vaso Cero* y *Rotación Física Vasos*, permitiendo moverse correctamente de un vaso a otro.

- *Mover Adelante*: Procedimiento que activa el motor para que se mueva en un sentido creciente al número de vaso. Es invocado por los procedimientos *Ubicar Vaso Cero* y *Rotación Física Vasos*.

- *Mover Atrás*: Procedimiento que activa el motor para que se mueva en un sentido decreciente al número de vaso. Es invocado por el procedimiento *Ajustar Vaso Cero*.

- *Parada Total*: Procedimiento que detiene el motor. Es invocado por los procedimientos *Ubicar Vaso Cero*, *Ajustar Vaso Cero* y *Rotación Física Vasos*.

Cabe señalar que el RF N° 2 es abordado en el *Setup* al momento de *Iniciar componentes y sensores* y *Cargar Configuración*. Luego, durante el *Loop* en los procedimientos: *Lectura Sensores*, *Log Eventos*, *Log Sensores*, *Guardar Configuración*, *Rotación Física Vasos* y *Tiempos Generales*.

## 5 Ensayos de Laboratorio

Mediante el archivo CONFIG.INI (Tabla 1) el operador configura el ensayo a realizar. A su vez, a medida que el dispositivo se encuentra operativo, se actualiza el valor de ciertas variables del archivo. Más allá que durante el diseño y desarrollo las pruebas del sistema fueron una etapa insoslayable del proceso, es preciso señalar 6 ensayos de laboratorio que evaluaron el sistema bajo diferentes condiciones.

**Tabla 1.** Archivos CONFIG.INI con los parámetros de los ensayos realizados.

<b>Ensayo N° 1:</b> Con ahorro de energía en bloques de 8 s. Sin reinicio. Inicio y fin 02/05/2020.	<b>Ensayo N° 2:</b> Con ahorro de energía en bloques de 1 s. Con reinicio. Inicio 03/05/2020 y fin 04/05/2020.	<b>Ensayo N° 3:</b> Con ahorro de energía en bloques de 1 s. Con reinicio y pausa. Inicio 19/05/2020 y fin 20/05/2020.
TiempoEsperaLog;10 TiempoIntervaloLog;10 TiempoUltimoLog;0 TiempoEsperaRot;10 TiempoIntervaloRot;10 TiempoUltimaRot;0 TiempoInicioRot;0 TiempoFinRot;0 vaso;0 rotacion;0 cant_vasos;8 cant_max_rotacion;3 estadoRot;1	TiempoEsperaLog;30 TiempoIntervaloLog;1800 TiempoUltimoLog;0 TiempoEsperaRot;60 TiempoIntervaloRot;3600 TiempoUltimaRot;0 TiempoInicioRot;0 TiempoFinRot;0 vaso;0 rotacion;0 cant_vasos;8 cant_max_rotacion;3 estadoRot;1	TiempoEsperaLog;1800 TiempoIntervaloLog;3600 TiempoUltimoLog;0 TiempoEsperaRot;3600 TiempoIntervaloRot;10800 TiempoUltimaRot;0 TiempoInicioRot;0 TiempoFinRot;0 vaso;0 rotacion;0 cant_vasos;8 cant_max_rotacion;1 estadoRot;1 ModoOperacion;1
<b>Ensayo N° 4:</b> Con ahorro de energía en bloques de 8 s. Con reinicio y pausa. Inicio 21/05/2020 y fin 22/05/2020.	<b>Ensayo N° 5:</b> Con ahorro de energía en bloques de 1 s. Sin reinicio. Inicio 30/05/2020 y fin 31/05/2020.	<b>Ensayo N° 6:</b> Sin ahorro de energía en bloques de 1 s. Sin reinicio. Inicio 23/06/2020 y fin 24/06/2020.
TiempoEsperaLog;1800 TiempoIntervaloLog;3600 TiempoUltimoLog;0 TiempoEsperaRot;3600 TiempoIntervaloRot;10800 TiempoUltimaRot;0 TiempoInicioRot;0 TiempoFinRot;0 vaso;0 rotacion;0 cant_vasos;8 cant_max_rotacion;1 estadoRot;1 ModoOperacion;1	TiempoEsperaLog;1800 TiempoIntervaloLog;1800 TiempoUltimoLog;0 TiempoEsperaRot;3600 TiempoIntervaloRot;3600 TiempoUltimaRot;0 TiempoInicioRot;0 TiempoFinRot;0 vaso;0 rotacion;0 cant_vasos;8 cant_max_rotacion;1 estadoRot;1 ModoOperacion;2	TiempoEsperaLog;1800 TiempoIntervaloLog;3600 TiempoUltimoLog;0 TiempoEsperaRot;3600 TiempoIntervaloRot;10800 TiempoUltimaRot;0 TiempoInicioRot;0 TiempoFinRot;0 vaso;0 rotacion;0 cant_vasos;8 cant_max_rotacion;1 estadoRot;1 ModoOperacion;1

Las variables de este archivo cumplen las siguientes funciones. *TiempoEsperaLog*: Segundos a esperar para registrar los datos luego del arranque del dispositivo. *TiempoIntervaloLog*: Intervalo en segundos entre cada registro en el archivo ENSAYOS.CSV. *TiempoUltimoLog*: Tiempo de la última escritura en el archivo ENSAYOS.CSV. *TiempoEsperaRot*: Segundos a esperar para rotar los vasos luego del arranque del dispositivo. *TiempoIntervaloRot*: Segundos entre cada rotación de los vasos. *TiempoUltimaRot*: Tiempo de la última rotación de los vasos.

*TiempoInicioRot*: Tiempo en que se inició la rotación de los vasos. *TiempoFinRot*: Tiempo en que finalizó la rotación de los vasos. *vaso*: Número de vaso actualmente utilizado. *rotacion*: Número de rotación en curso. *cant\_vasos*: Cantidad de vasos del dispositivo. *cant\_max\_rotacion*: Número de rotaciones (vueltas) del ensayo. *estadoRot*: Estado de la rotación (1 - ESPERA, 2 - ROTANDO, 3 - FIN, 4 - DETENIDO). *ModoOperacion*: 1 - Modo *Pitfall*, otro valor Modo *Data Logger*.

Durante la ejecución del sistema, las siguientes variables son actualizadas: *TiempoUltimoLog*, *TiempoUltimaRot*, *TiempoInicioRot*, *TiempoFinRot*, *vaso*, *rotacion* y *estadoRot*. En el modo de operación *Data Logger* sólo se consideran las variables *TiempoEsperaLog*, *TiempoIntervaloLog*, *TiempoUltimoLog* y *ModoOperacion*, las demás variables no son tenidas en cuenta.

## 6 Resultados

El *Ensayo N° 1* permitió comprobar el funcionamiento correcto del *MR* en combinación con el *MDL*, allí se detectaron ajustes a realizar en el procedimiento *LogEventos*. El *Ensayo N° 2* posibilitó determinar que la autonomía de la batería (12 Volts 7 Ah) no superó las 18 h, siendo necesaria una recarga o un remplazo de la misma durante su ejecución.

A partir del *Ensayo N° 3* se incorporó el parámetro *ModoOperacion* y contribuyó a comprobar que más allá de ampliar los períodos de tiempo, la autonomía de la batería continuaba sin superar las 18 h; por otra parte, cuando se interrumpía un ensayo, el período de tiempo de permanencia de un vaso comenzaba desde 0. Así, el *Ensayo N° 4* confirmó lo analizado en los 2 ensayos previos respecto a la autonomía de la batería, también permitió establecer que la autonomía no estaba relacionada respecto a la utilización de bloques de ahorro de energía de 1 s u 8 s, y posibilitó verificar que se subsanó el inconveniente detectado en el ensayo anterior, el cual ante un reinicio el período de tiempo de permanencia de un vaso comenzaba desde 0.

Mediante el *Ensayo N° 5* se comprobó el correcto funcionamiento del sistema en el modo de operación *Data Logger*. A su vez, terminó de comprobar algo detectado en los ensayos previos, relacionado a que el uso reiterado de la función *LowPower.powerDown* de la biblioteca *LowPower.h*, si bien contribuye a una mayor autonomía de la batería, tiene como desventaja que introduce un margen de tiempo superior entre 6 y 7 min en los intervalos temporales de rotación y registro de datos. En tanto, en el *Ensayo N° 6* se corroboró que sin la funcionalidad de ahorro de energía el sistema sólo cuenta con una autonomía de 15 h; no obstante, los intervalos temporales de rotación y registro de datos son más precisos, con un margen de +/- 6 s.

## 7 Conclusiones y Trabajo Futuro

Ha sido posible diseñar y desarrollar un sistema embebido para un prototipo automatizado de trampa de caída por tiempo, dotado de un data logger parametrizable y más completo, con sensores de luz, temperatura y humedad relativa ambiente, y

temperatura y humedad del suelo. Los ensayos fueron satisfactorios y permitieron la evolución del sistema, cumpliendo los requerimientos funcionales. El diseño arquitectónico contribuye al mantenimiento y evolución del dispositivo, resultando factible su utilización en estudios de biodiversidad de artrópodos de suelo.

El enfoque de ahorro de energía utilizado si bien contribuye a preservar la autonomía de la batería, también conlleva a la pérdida de exactitud en los tiempos. En consecuencia, este aspecto tendrá que ser revisado.

Como instancia posterior, se tiene previsto dotar al *Módulo de Alimentación* de una batería de mayor capacidad para prolongar la autonomía, a partir de ello, realizar ensayos de campo en diferentes ambientes productivos de la EEA Concordia del INTA, lo cual abre la posibilidad de efectuar comparaciones con dispositivos similares previamente desarrollados y, a su vez, realizar investigaciones de artrópodos de suelo mediante un abordaje novedoso que utilice herramientas estadísticas y Minería de Datos.

## Referencias

1. Hohbein, R.R., Conway, C.J.: Pitfall traps: A review of methods for estimating arthropod abundance. *Wildl. Soc. Bull.* 42, 597-606 (2018). <https://doi.org/10.1002/wsb.928>
2. Dahl, F.: Vergleichende Untersuchungen Über die Lebensweise wirbelloser Aasfresser. *Sitzungsberichte – K. Preuss. Akad. der wissenschaften.* 17-30 (1896)
3. Greg Sherley and Ian Stringer: DOCCM-248862 Invertebrates: pitfall trapping v1. 0. Dep. Conserv. Te Papa Atawhai, Wellingt. 1-30 (2016)
4. Brown, G.R., Matthews, I.M.: A review of extensive variation in the design of pitfall traps and a proposal for a standard pitfall trap design for monitoring ground-active arthropod biodiversity. *Ecol. Evol.* 6, 3953-3964 (2016). <https://doi.org/10.1002/ece3.2176>
5. Knapp, M., Saska, P., Knappová, J., Vonička, P., Moravec, P., Kúrka, A., Anděl, P.: The habitat-specific effects of highway proximity on ground-dwelling arthropods: Implications for biodiversity conservation. *Biol. Conserv.* 164, 22-29 (2013). <https://doi.org/10.1016/j.biocon.2013.04.012>
6. Buchholz, S.: Design of a time-sorting pitfall trap for surface-active arthropods. *Entomol. Exp. Appl.* 133, 100-103 (2009). <https://doi.org/10.1111/j.1570-7458.2009.00902.x>
7. McMunn, M.S.: A time-sorting pitfall trap and temperature datalogger for the sampling of surface-active arthropods. *HardwareX.* 1, 38-45 (2017). <https://doi.org/10.1016/j.ohx.2017.02.001>
8. Wickert, A.D., Sandell, C.T., Schulz, B., Ng, G.C.: Open-source Arduino-compatible data loggers designed for field research. 2065-2076 (2019)
9. Baker, E.: Open source data logger for low-cost environmental monitoring. *Biodivers. Data J.* 2, (2014). <https://doi.org/10.3897/BDJ.2.e1059>
10. Sommerville, I.: Software Embebido. En: *Ingeniería de Software.* pp. 537-564. Pearson Educación S.A. (2011)
11. Jacobson, I., Booch, G., Rumbaugh, J.: *El Proceso Unificado de Desarrollo de Software.* (2000)
12. Arduino AG: What is Arduino?, <https://www.arduino.cc/en/Guide/Introduction>

## Dispositivo GPS para monitoreo del rendimiento deportivo

Luisina Santos<sup>1</sup>, Marcelo Guiguet<sup>2</sup>, Pablo Luengo<sup>3</sup>, Mónica Sarobe, Eduardo Alvarez, Gustavo Useglio, Federico Gómez, Matías Capelli, Carlos Di Cicco

Instituto de Investigación y Transferencia en Tecnología (ITT)<sup>4</sup>  
Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)  
Escuela de Tecnología (ET)  
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

{luisina.santos, marcelo.guiguet, pablo.luengo, monica.sarobe, eduardo.alvarez, gustavo.useglio, federico.gomez, matias.capelli, carlos.dicicco}@itt.unnoba.edu.ar

**Abstract.** El presente trabajo se enmarca en un proyecto que utiliza la tecnología del GPS aplicada al deporte para generar un sistema de entrenamiento que permita el monitoreo de ciertos aspectos de interés del rendimiento físico del atleta en su contexto real. Facilitando así la superación deportiva mediante el análisis científico de los datos obtenidos. El propósito es desarrollar un prototipo gps económico que permita recolectar los datos necesarios para calcular los parámetros de rendimiento físico de un atleta, enfocándose en una rápida transición a un equipo final de fácil fabricación nacional y accesible para instituciones deportivas tanto amateurs como profesionales. El trabajo se desarrolla en el Instituto de Investigación y Transferencia de Tecnología de la UNNOBA, en el marco del proyecto de investigación denominado “Informática y Tecnologías Emergentes”, contó con la participación de la empresa Silamberts S.R.L. y se obtuvo financiación de la Fundación Dr. Manuel Sadosky.

**Keywords:** Global Positioning System, IMU, Rendimiento Deportivo, Datalogger.

### 1 Introducción

Este trabajo, utiliza la tecnología GPS y la aplica al deporte para desarrollar un sistema de entrenamiento para atletas de diversas disciplinas que permita monitorear los aspectos esenciales del rendimiento físico de los mismos. El antecedente de mayor relevancia es una tecnología australiana diseñada en 1990 por el equipo de investigadores del Centro de Investigación Cooperativa a cargo de Igor van de

---

<sup>1</sup> Becario ITT

<sup>2</sup> Becario CIN

<sup>3</sup> Docente Investigador - ITT

<sup>4</sup> ITT – Centro Asociado CIC

Griendt y Shaun Holthouse. En sus estudios, realizados junto al Instituto Australiano del Deporte para el análisis del rendimiento físico de atletas por medio de evidencia científica, Griendt y Holthouse encontraron que si bien el monitoreo en laboratorio permitía tener una idea detallada de las exigencias, los atletas no realizaban los mismos esfuerzos durante una competencia. Para trasladar la experiencia al campo real, desarrollaron sensores portátiles que, a fines del año 2006, comenzaron a ser comercializados bajo el nombre de Catapult [1] y prontamente se convirtieron en líderes mundiales del análisis atlético.

Actualmente, en Argentina sólo los clubes más acaudalados pueden acceder al equipamiento Catapult. Quienes no lo hacen utilizan un GPS que recolecta datos simples como distancias, tiempos, intensidades y frecuencias. A partir de ellos el preparador físico elabora informes que sirven para el diagnóstico y la planificación del entrenamiento individual según, por ejemplo, el período y la duración de la competencia. El desafío que aquí se propone es la elaboración de una tecnología de bajo costo ubicada a nivel medio entre el GPS estándar y el avanzado como el de Catapult. Esta tecnología busca brindar a los clubes locales una herramienta económicamente accesible que facilitará el diagnóstico y la planificación del entrenamiento de deportistas profesionales y amateurs.

### **1.1 Consideraciones de la utilización de gps deportivos**

Por su peso y tamaño reducido y el ahorro de tiempo que supone el análisis automatizado de múltiples atletas y el feedback casi inmediato, estos dispositivos se convierten en una herramienta ampliamente superadora de técnicas de monitorización con carga externa [2][3]. Si bien aún presentan ciertas limitaciones en lo que hace a la fiabilidad y a la validez en la medición de carreras de alta intensidad o distancias reducidas, la inclusión de variables sin referencia exclusiva a la locomoción -tales como giros, saltos, cambios de dirección o impactos- es un valor añadido, mientras que los softwares comercializados son cada vez más potentes y flexibles en lo que al tratamiento de los datos se refiere [4].

El elevadísimo número de niveles y variables configurables a posteriori, obliga a filtrar y optimizar datos, incorporando indicadores globales capaces de recoger la máxima cantidad de información de la forma más sencilla, facilitando la programación, el control y el seguimiento de la demanda física y táctica de los deportistas. Acompañado del estudio de variables fisiológicas y del análisis de video, el seguimiento por GPS puede propiciar el abordaje holístico del rendimiento de jugadores y equipos [5].

De todas las variables con las que cuentan los equipos de GPS profesionales, en esta primera instancia se determinó centrarse en aquellas que los especialistas consideran de mayor relevancia descartando aquellas que proporcionan detalles más específicos. Así, de un encuentro con especialistas de dos clubes de la ciudad Junín (Bs As) se estableció que resultan ser cinco las fundamentales para el trabajo de todo preparador físico o director técnico. Estas variables son la posición en cancha, los saltos y tres rangos de desplazamientos: de 0 a 3 m/s, de 3 a 6 m/s y de 6 a 12 m/s.

## 2 Contexto

Como ya se mencionó en este documento el presente trabajo se desarrolla en el marco del proyecto de investigación denominado “Informática y Tecnologías Emergentes”. Allí mismo un grupo de investigadores ha trabajado en la construcción de un prototipo de GPS aplicado al deporte. Actualmente se continúa con el desarrollo de versiones más avanzadas. El equipo de trabajo, se desempeña en el ITT y ha establecido vínculos con clubes de la región para continuar con las pruebas de campo que se vienen realizando.

## 3 Construcción del prototipo

En la fase preliminar del trabajo se desarrolló un prototipo GPS con un Arduino Nano [6] y los módulos: GY-GPS6MV2 [7] (módulo GPS con antena cerámica), MPU9250 [8] (módulo IMU con acelerómetro de 3 ejes, giroscopio de 3 ejes y magnetómetro de 3 ejes), lector microSD [9] y RTC DS3231 (módulo de reloj en tiempo real) [10]. Este prototipo inicial posibilitó llevar a cabo las primeras pruebas, de las cuales se pudo determinar que la idea de construir un equipo de monitoreo de rendimiento deportivo de bajo costo era un desafío claramente viable.

Si bien esta versión del prototipo permitió obtener buenos resultados aún quedaban cuestiones por resolver, siendo el tamaño y la autonomía de la fuente de alimentación los temas más críticos. Para obtener un prototipo completamente funcional, el presente trabajo se ve dividido en dos etapas de mejora bien diferenciadas. Por un lado, se da enfoque a la transmisión de datos desde el prototipo al ordenador donde la información será procesada a través de una App Web desarrollada específicamente para el proyecto. Por otra parte, se refleja el proceso realizado para la construcción de un nuevo prototipo que cuenta con las mismas características que la versión anterior, pero esta vez sobre una placa de circuito impreso diseñada específicamente para el problema.

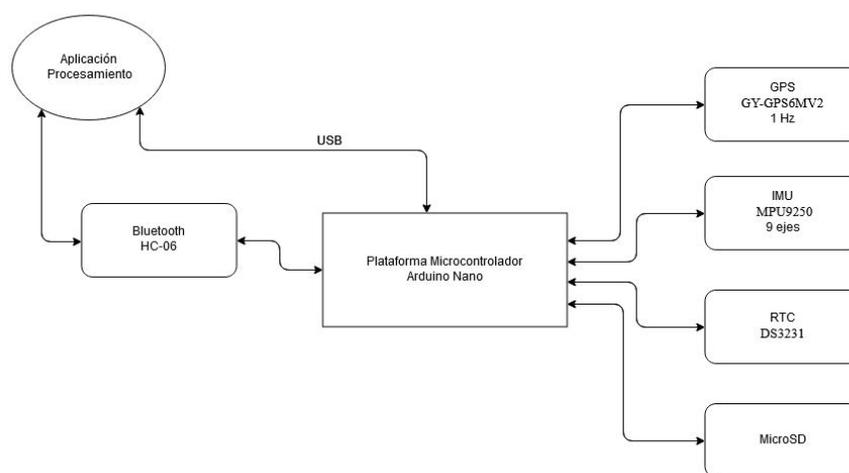
El objetivo es entonces reducir el tamaño del prototipo final, brindarle autonomía energética y, además diseñar e implementar un protocolo de comandos para comunicación que permita la transmisión de datos de forma segura y sin errores.

Con el objetivo de realizar una interconexión de los módulos mencionados anteriormente, disminuir el tamaño final del equipo y evitar errores por falso contacto de cableado se lleva a cabo el diseño de una placa de circuito impreso.

Para aprovechar al máximo el espacio de utilización de la placa y añadir funcionalidad al dispositivo desarrollado, se agrega un módulo que permite la conectividad vía Bluetooth (hc-06 [11]). La fuente de alimentación es una celda de litio de 3.7v modelo 18650. En la figura 1 se puede visualizar el diagrama de bloques del prototipo.

Los siguientes componentes se utilizan para el sistema de alimentación de la placa:

- TP4056 cargador con protección incluida
- Elevador de tensión a 5v
- 1 batería litio 18650
- Selector de alimentación construido con diodos



**Fig. 1.** Diagrama de bloques

### 3.1 Consideraciones en el firmware para el almacenamiento de datos

Los datos de los módulos GPS e IMU se almacenan en la memoria microSD en dos archivos diferenciados:

- G: Contiene los datos extraídos a partir del módulo GY-GPS6MV2
- I: Contiene los datos extraídos a partir del módulo MPU-9250

Se genera entonces, una trama completa para cada archivo definida de la siguiente manera:

G = <contador, fecha, latitud, longitud, cant. de satélites activos>

I = <contador, fecha, ax, ay, az, gx, gy, gz>

\* fecha en formato de tiempo unix.

Debido a la capacidad reducida de la memoria de la plataforma Arduino Nano resultó necesario aplicar optimizaciones en la librerías intervinientes, principalmente en lo referido a la reducción de código y variables que no eran indispensables.

Para optimizar las funciones encargadas de la captura de los datos de los sensores IMU y GPS para almacenarlos en una memoria microSD se tuvieron en cuenta los siguientes factores:

- Código innecesario
- Variables evitables
- Simplificación de operaciones
- Reducción de funcionalidades no prioritarias

## 4 Desarrollo del protocolo de comunicación

El objetivo de esta fase del trabajo es el desarrollo de un protocolo que permita la transmisión de datos con una tasa de error cero. Esto es, cada línea enviada desde el dispositivo hacia el ordenador será retransmitida cuantas veces sea necesario para asegurar la totalidad de los datos en el receptor.

El protocolo de comunicación a desarrollar estará basado en comandos de petición/respuesta donde el ordenador será el componente activo: Se encargará de solicitar al dispositivo cada línea de datos almacenada y en caso de detectar algún error en la trama (fallo de checksum) se procederá a la retransmisión.

También es necesario realizar un control de la conexión para reconectar automáticamente en caso de ser requerido.

Conociendo el formato que tiene la trama que se quiere transmitir y haciendo foco en la numeración incremental línea a línea que ambos archivos poseen en común, se espera que el protocolo la utilice para reconocer y controlar qué dato se está enviando en forma de ACK.

### 4.1 Diseño

En función del estado momentáneo en el que se encuentra el dispositivo se comporta como se especifica en la siguiente tabla.

**Tabla 1.** Flujo de ejecución del protocolo de comandos

Estado	Descripción
Inicio	El ordenador da inicio al protocolo a través de un comando enviado por el usuario.
Trama correcta	Si el ordenador recibe la trama completa y sin errores, entonces devuelve un ACK con el último número de línea almacenado.
Trama incorrecta	Si se produce un error en la transmisión, el ACK no coincide con el esperado por el dispositivo. Pasa a retransmisión.

Retransmisión	Si el ACK esperado por el dispositivo no coincide con el recibido por el ordenador, entonces debe retransmitir la última trama enviada.
Time-Out	Si en un tiempo específico el ordenador no recibe ninguna trama se asume una pérdida de conexión. Pasa a reconexión.
Reconexión	Se restablece la conexión sin perder el ACK de la última línea recibida correctamente y se envían los datos desde la pérdida en adelante.
Fin	El dispositivo reconoce que la totalidad de las líneas (n) se enviaron de forma correcta y finaliza la comunicación.

---

Se diagrama a continuación la secuencia ideal al momento de establecer la comunicación y realizar la transferencia de datos:

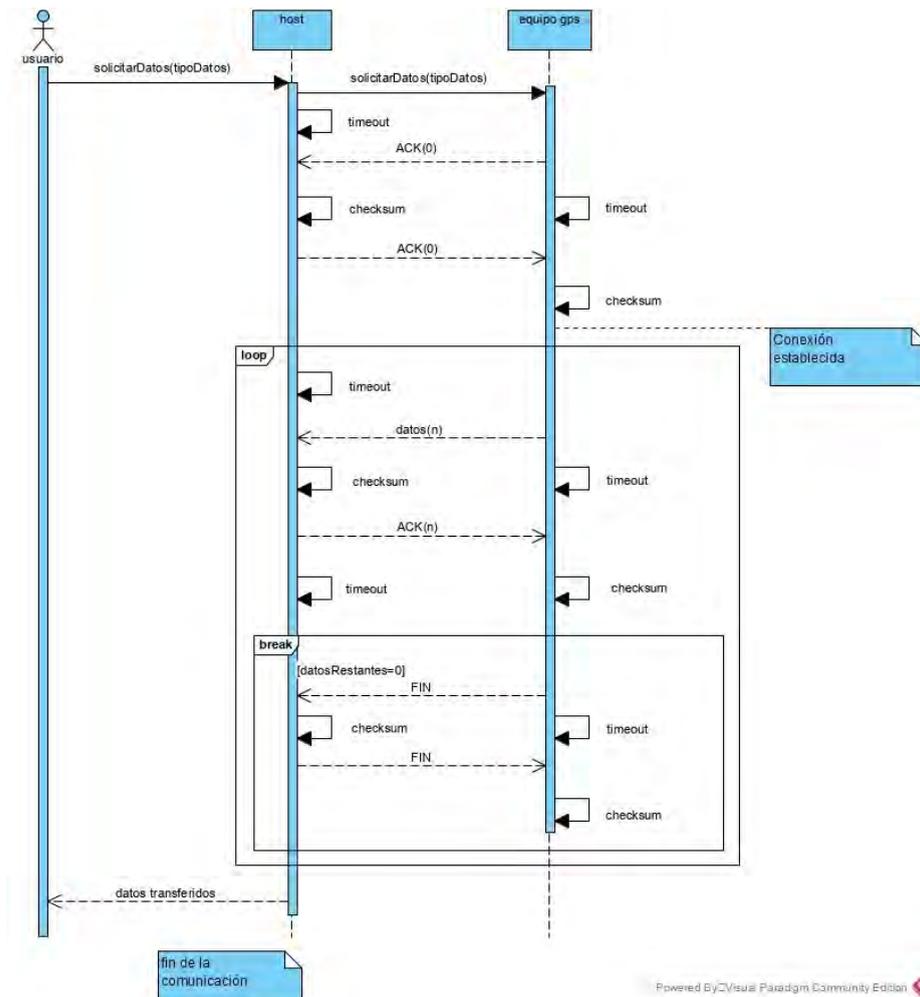


Fig. 2. Diagrama de secuencia ideal

## 4.2 Implementación

La implementación del diseño propuesto consiste principalmente en dos desarrollos que funcionen en conjunto. Por un lado, es necesaria una aplicación de escritorio que envíe los comandos definidos y reciba las tramas correspondientes a través del puerto Serial (vía USB o Bluetooth).

Por el otro, será necesaria la implementación de un algoritmo que le permita al microcontrolador Arduino conectarse al ordenador y enviar las tramas solicitadas por éste (Firmware).

## 5 Pruebas

Con el fin de llevar a cabo las correspondientes pruebas de campo se optó por montar el dispositivo en un contenedor plástico que se adapte a un corpiño deportivo.

El equipo de Fútbol femenino de la UNNOBA, en la ciudad de Junín, puso a prueba el prototipo durante un entrenamiento y se comprobó el correcto funcionamiento del módulo de registración. Además, se realizaron varias pruebas de transferencia de archivos a través del puerto USB. Las mismas reflejan una tasa de error cero con un tiempo de transferencia de 480 tramas por minuto aproximadamente.

## 6 Conclusión y trabajo a futuro

Los resultados positivos de las pruebas realizadas hasta el momento ratifican la viabilidad del desafío propuesto de construir un equipo accesible económicamente para monitorear el rendimiento deportivo. En esta primera etapa se logró construir un prototipo de dispositivo gps que es capaz de registrar el movimiento del atleta, transmitir los datos a una computadora para su posterior procesamiento y ser de un tamaño que permite su utilización en un corpiño deportivo.

Estos resultados son alentadores para continuar trabajando en pos de incorporar nuevas funcionalidades y superar las limitaciones que se presentan en esta versión, siendo una de las principales la reducida memoria que tiene la plataforma Arduino Nano, situación que conduce a la necesidad de simplificar aspectos del programa no permitiendo desarrollar un firmware óptimo que pueda aprovechar todas las funcionalidades que pueden brindar los módulos utilizados.

Por lo tanto, aún queda un largo camino de mejoras completamente viables que en la actualidad se encuentran siendo evaluadas. En primer lugar resulta necesario focalizar el trabajo futuro en la incorporación de plataformas de microcontroladores de 32 bit, con especial atención en el tamaño físico de las placas y asignando mayor relevancia a las que incorporan funcionalidades adicionales como socket microSD o comunicaciones inalámbricas en la misma plaqueta, lo que permitiría en principio una reducción del tamaño final del equipo. El objetivo de reducir el tamaño físico del equipo es fundamental para permitir un uso ergonómico y seguro por parte de los atletas. Además, se busca la incorporación de un sensor cardíaco que mida el pulso del atleta y brinde información que sirva como apoyo para que un especialista médico realice el correspondiente diagnóstico.

Otro punto a desarrollar en trabajos futuros es la definición de métodos estandarizados para la evaluación objetiva de los resultados provistos por el equipo, buscando reducir la variabilidad de los resultados por las condiciones particulares de cada prueba y generando opciones válidas de comparación entre cada una de las pruebas realizadas. Esto es fundamental para poder seguir avanzado, ya que de lo

contrario no hay una retroalimentación certera del funcionamiento del equipo y no se tendrá conocimiento de los ajustes que realmente haya que hacerle al equipo.

## 7 Formación de recursos humanos

En esta línea de I/D se han obtenido y se encuentran desarrollando actualmente dos becas de iniciación a la investigación. Asimismo, se desarrollaron dos tesinas de grado, todas ellas dirigidas por miembros de este proyecto.

## Referencias

1. Catapult: About Catapult. Disponible en: <https://www.catapultsports.com/about>
2. Técnicas de análisis y métodos para el incremento de la precisión posicional relativa utilizando receptores GPS monofrecuencia (2017) Alberto Eduardo Riba, Nelson Acosta, Juan Manuel Toloza, Fernando Emmanuel Frati, Carlos Kornuta.
3. Procesamiento de sentencias NMEA-0183 para el análisis de la geometría satelital utilizando receptores GPS de bajo costo (2016) Alberto Eduardo Riba, Jorge Damián Tejada, Nelson Acosta, Juan Manuel Toloza.
4. Mejora de la precisión posicional utilizando receptores GPS de bajo costo (2016) Alberto Eduardo Riba, Nelson Acosta, Juan Manuel Toloza, Fernando Emmanuel Frati, Jorge Tejada, Carlos Kornuta.
5. CASTELLANO, Julen y CASAMICHANA, David. (2014). Deporte con dispositivos de posicionamiento global (GPS): Aplicaciones y limitaciones, UAB: Revista de Psicología del Deporte 23, n° 2, pp. 355-364
6. Datasheet Arduino NANO. Disponible en: <http://www.farnell.com/datasheets/1682238.pdf>
7. U-blox. Datasheet GY-GPS6MV2. Disponible en: [https://www.u-blox.com/sites/default/files/products/documents/NEO-6\\_DataSheet\\_%28GPS\\_G6-HW-09005%29.pdf](https://www.u-blox.com/sites/default/files/products/documents/NEO-6_DataSheet_%28GPS_G6-HW-09005%29.pdf)
8. Iven Sense. MPU-9250 Product Specification. Disponible en: <https://www.invensense.com/wp-content/uploads/2015/02/PS-MPU-9250A-01-v1.1.pdf>
9. (2012, Abril). Micro SD Datasheet, [http://www.supertalent.com/datasheets/5\\_112.pdf](http://www.supertalent.com/datasheets/5_112.pdf)
10. Maxim Integrated Products. (2015). Datasheet RTC DS3231. Disponible en: [https://components101.com/sites/default/files/component\\_datasheet/DS3231%20Datasheet.pdf](https://components101.com/sites/default/files/component_datasheet/DS3231%20Datasheet.pdf)
11. Guangzhou HC Information Technology Co., Ltd. Datasheet HC-06. Disponible en: <https://www.olimex.com/Products/Components/RF/BLUETOOTH-SERIAL-HC-06/resources/hc06.pdf>

## Simplifying concurrency and monitoring on Arduino for Internet of Things

Ricardo Moran<sup>1,2</sup>, Matías Teragni<sup>1</sup>, Gonzalo Zabala<sup>1</sup>

<sup>1</sup>Universidad Abierta Interamericana, Centro de Altos Estudios en Tecnología Informática, Ciudad Autónoma de Buenos Aires, República Argentina

<sup>2</sup>Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Calle 526 e/ 10 y 11, La Plata, Buenos Aires, República Argentina

{Ricardo.Moran, Matias.Teragni, Gonzalo.Zabala}@uai.edu.ar

**Abstract.** The Internet of Things (IoT) presents several challenges and opportunities to improve people's lives. Experts agree on the importance of involving the community in the process of defining and creating IoT in order to succeed. Platforms like Arduino make it simple for non-technical people to build IoT devices. However, they also present difficulties that complicate their adoption and limit their reach. In this paper, we focus on the Arduino language and its limited support for concurrency and monitoring, which we deem essential for the IoT. We explore the existing solutions offered by the Arduino ecosystem and analyze their strengths and weaknesses. Finally, we propose an alternative solution based on a high-level programming language designed to tackle these issues with the help of an embedded virtual machine.

**Keywords:** Internet of Things, cloud computing, programming language, virtual machine, Arduino, concurrency, monitoring

### 1 Introduction

In the last decade, the development of networks of smart devices capable of sensing their environment, connect to the Internet, and publish data to cloud servers has led to a concept known as the Internet of Things (IoT). It is estimated that between 2008 and 2009 the number of “things” connected to the internet has surpassed the number of people [1]. While this evolution of the Internet architecture presents privacy and security risks [2] it also has the potential to improve people's lives [1].

In order to face these challenges in a satisfactory way several experts agree that it is important to reach out and involve more people in the IoT creation [3]. The Maker culture (a technology-based extension of the Do-It-Yourself movement [4]) could play an essential role in involving the society, in part thanks to the rise in popularity of platforms such as Arduino that make relatively easy and inexpensive to build IoT devices.

Arduino is a microcontroller board that has become one of the most popular platforms for building electronic projects, especially among hobbyists, artists, designers, and people just starting with electronics. Apart from being open-source hardware, one of the reasons for its popularity is its software library and integrated development environment (IDE) that provides an abstraction layer over the hardware details, making it possible to build interesting projects without a complete understanding of more advanced microcontroller concepts such as interrupts, ports, registers, timers, and such. At the same time, this abstraction layer can be bypassed to access advanced features if the user needs them making Arduino a platform suitable for both beginners and experts, allowing non-technical users to participate in the IoT phenomenon.

However, there are several aspects in which the Arduino language lacks proper support for the needs of an IoT project, introducing extra complexity that can overwhelm a beginner while building any kind of interesting system.

The first issue we identify is the lack of support for concurrency in the language itself. Even moderately complex problems require some sort of simultaneous task execution. Almost all IoT devices need to be able to perform some combination of the following tasks: communicate with other devices; read the value of several sensors; make decisions based on these values; connect to the Internet and publish their sensor data to the cloud [5]. Usually these tasks need to be done simultaneously but since the Arduino language does not provide any concurrency support it is left to the programmer to implement an arbitrarily complex scheduler in order to avoid interfering one task with another.

The second problem that Arduino presents is its complete lack for remote monitoring, a requirement for almost every IoT project. The lack of a standard protocol forces the programmer to design a different ad-hoc protocol for each project, which limits code reusability and is hard for non-technical users to implement properly. Other solutions are available in the form of open-source libraries but most of them suffer from issues that cannot be dismissed.

In this paper we will discuss the existing solutions, their limitations, and propose an alternative approach based on the use of a high-level programming language especially designed for these use cases that can be executed by a small virtual machine running in the microcontroller.

## 2 Related work

Regarding concurrency support, although the Arduino language does not provide any concurrency abstraction by default, third-party libraries are available that attempt to address this issue.

The most sophisticated libraries we surveyed are implementations of a real-time operating system for microcontrollers known as FreeRTOS [6] [7]. These libraries offer prioritized, preemptive multitasking. However, they are harder to use than the alternatives, requiring the user to specify several configuration parameters, which can be challenging for a novice programmer. Most libraries take a simpler approach, which is to support cooperative multitasking. These libraries differ in their implementation

details and, thus, the programmer needs to understand the tradeoffs of each library in order to correctly predict the behavior of the program. Some libraries, like `ArduinoProcessScheduler` [8] and `Task` [9], force the user into the object-oriented paradigm and require the definition of classes for each independent process. Others, like `Automaton` [10] and `Yet Another State Machine` [11], are event-driven and focus on helping the user to create state machines. Most libraries, however, simply allow the user to specify which procedures should run concurrently by adding them to a global scheduler's list or using macros to define them. `ArduinoThread` [12] and `everytime` [13] take another approach worth mentioning, they do not implement truly independent tasks, but instead provide facilities to schedule the execution of procedures at desired intervals. Each time a task is executed it will run to completion. The main difference between the two is that `everytime` is fully based on compile-time macros, which makes it impossible to dynamically create new processes. While all these libraries and frameworks have their own tradeoffs, they all exhibit the same flaw: they require the user to understand the execution model provided by the library to use it effectively. If the programmer fails to understand the strengths and limitations of the library, they may be punished with potentially hard to debug errors [14]. In the case of the cooperative multitasking libraries, the limitation is evident: the user needs to be extra careful not to use blocking code (like the "delay" function or a long-running loop) or risk interfering with the execution of other tasks. This can be exacerbated by the fact that most third-party libraries, tutorials, and code examples on the web assume total control of the CPU and, thus, may use blocking functions that are not compatible with this programming style. For an experienced programmer this could represent a minor problem but for a beginner it could mean an unsurpassable challenge.

Alternative programming languages for the Arduino platform address the concurrency problem in different ways. We found several implementations of high-level languages and virtual machines for the Arduino platform. Most of them are based on preexisting general-purpose programming languages such as Java [15], Scheme [16] or Python [17]. In these cases, support for concurrency depends on the implementation and is tied to the mechanisms already existing in the language. A more noteworthy example is the `Transterpreter` project [18], a virtual machine explicitly designed to exploit concurrency on embedded systems. This virtual machine runs `occam-pi` programs on several platforms, one of which is Arduino [19]. `Occam-pi` is a variant of the `occam` programming language [20], especially designed to write concurrent programs based on communicating sequential processes (CSP) process algebra [21]. `Occam-pi` has a rich set of runtime libraries that provide functions for interacting with Arduino features such as the serial port, PWM and TWI. Regarding performance, the execution of bytecodes has been reported to be 100 to 1000 times slower than the execution of native code.

Regarding monitoring and communication capabilities, the current options are limited. All Arduino boards have native support for serial communication and the Arduino language includes several functions to allow the user to read or write from the `Serial`. However, writing a custom protocol on top of the `Serial` can be challenging for beginners. For this reason, a popular choice for Arduino developers is the `Firmata` library [22].

`Firmata` is a standard protocol for communicating the Arduino with software on a host computer. It was designed to allow users to write custom firmware without having

to create their own protocol. Client libraries for several languages are available online and the host implementation comes already bundled as one of the default libraries in the Arduino IDE. The protocol is extensible, and it can be used to communicate through the serial port, Wi-Fi, ethernet, and Bluetooth. The usefulness of Firmata has made it a very popular library but it has its drawbacks. On the one hand, the protocol details might not match exactly the requirements of the project, in which case the user is left with no alternative but to dive in the Firmata code and adapt it to suit its needs. This might be challenging for a beginner. On the other hand, if the protocol satisfies the project requirements then the simplest option is to upload the StandardFirmata that comes as an example with the library. The StandardFirmata was designed to include as much functionality as possible into a single firmware, but doing so leaves little room for any custom code the user might wish to add to the sketch (on Arduino UNO StandardFirmata uses 38% of program storage and 52% of dynamic memory) and adding new functionality to the StandardFirmata sketch can be complicated. Another option is to import Firmata into the sketch as a library, in which case the user would have to manually implement the functions to handle the relevant messages provided by the Firmata protocol.

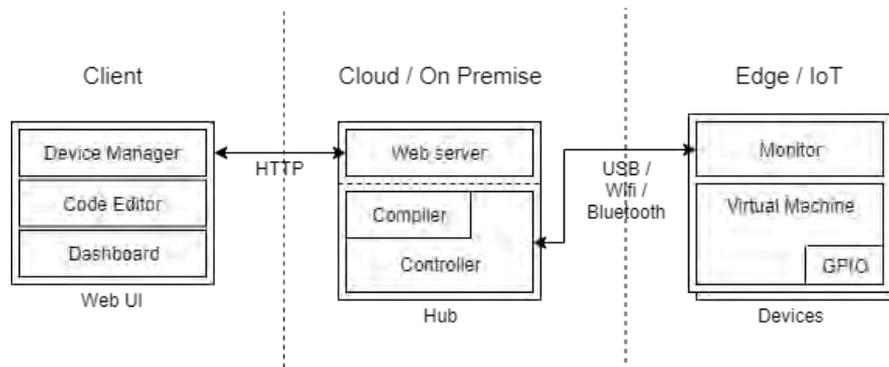
A popular alternative is to use an IoT cloud solution. In this case, the user only needs to use a library provided by the platform it wishes to use and the library will take care of all the communication details. A survey of existing solutions in this space shows a variety of platforms offering services like real time data capture, data visualization and device management related tasks through remote cloud servers while implying “pay-as-you-go” notion [5]. While these platforms simplify the development of IoT systems, they require the user to write code to interact with their cloud services. In most cases, this code is relatively straightforward as it focuses on defining what information is sent to the cloud and what to do with the information received, but the behavior of those libraries is usually opaque and too complex, forcing the programmer to call periodically to certain functions that actually perform the synchronization. Of all the existing solutions we have focused on Thingier.io, the results of our analysis are shown in the sub-section below.

### **3 Proposed solution**

We propose the implementation of a domain-specific language (DSL) supported by a virtual machine running on the Arduino. The nature of a DSL ensures that the logic usually present in IoT solutions can be expressed concisely and the abstraction layer it provides guarantees that the programmer cannot unintentionally interfere with the concurrent task execution. We call this language UziScript and we believe it would be a suitable alternative to the Arduino language for IoT.

## 4 Implementation

### 4.1 General architecture



**Fig 1** Architecture diagram of the UziScript toolchain and virtual machine

The architecture of the proposed solution consists of three distinct components: a firmware for the Arduino devices, a hub that handles the communication with the devices, and client tools that allow users to manage the entire system.

The firmware is a regular Arduino sketch that contains the language runtime, responsible for executing the programs as well as monitoring the state of the device. The hub, connected directly or indirectly to the Arduino boards, contains all the compilation tools and a web server that presents a REST API interface. The client tools are web applications that, through the hub server, allow to program, monitor and control all the devices in the system from any computer or phone that can reach the Hub server.

This architecture has several benefits.

1. **Flexibility:** the client tools, being a web app, could be used from any device, using a web browser or installing it as a native app; the hub could be deployed on the cloud or on premise, depending on the needs of the user.
2. **Portability:** the current firmware supports different Arduino boards (including UNO, Micro, Nano, MEGA 2560, and Yun) and we have also received reports of it working successfully on other compatible boards such as DuinoBot [23], Educabot [24], and TotemDUINO [25].
3. **Interactivity:** the runtime allows to compile and upload new programs at a fraction of the time required to compile an Arduino Sketch, this allows for a more interactive programming style that encourages experimentation and learning.
4. **Scalability:** a single hub could handle multiple devices and multiple hub instances could be deployed on a single system.

In the following sections we will explain some of the design choices we consider relevant for the IoT. For a detailed description of the implementation see [26].

**Programming language.** The UziScript programming language was designed to allow non-expert programmers to express concurrent tasks and facilitate monitoring the IoT device. Its syntax is based on C, which is familiar to most programmers including Arduino developers. In the following sections we will present some code examples.

To support concurrency, we added the “`task`” keyword, which represents behavior that can be executed periodically at a configurable rate. The scheduling and execution of each task is performed automatically by the runtime and the language allows the user to start, stop, pause, or resume any given task. Each task execution is independent.

To support monitoring, the runtime automatically keeps track of all the global variables in the program as well as the value of each pin in the Arduino. Thus, to publish some value to the outside world, the user only has to assign it to a global variable declared using the “`var`” keyword. To check and manipulate the value of either a digital or analog pin, the user does not have to write any code at all because they are monitored automatically.

**Firmware.** The firmware contains the virtual machine responsible for user program execution and a monitor program that allows it to interact with the hub controller. Periodically, this monitor program will send the status of the Arduino and receive commands, allowing the hub to fully control the virtual machine, including directly manipulating the variables and the pin state, debug the current user program, or download a new one.

**Monitoring tools.** The dashboard is a web tool that allows to monitor and control any global in the program or any pin on the Arduino device. In order to simplify this task for the user the dashboard allows to configure a blank canvas with a set of configurable widgets. Depending on the type of widget, it allows to either display or control a value.

## 5 Validation

To validate the effectiveness of the proposed approach we built an IoT greenhouse control system with the following requirements:

1. It can have one or more Arduino devices.
2. It should have a set of sensors to monitor the internal greenhouse conditions as well as a set of actuators that will allow to control said conditions.
3. It must have a dashboard that allows the user to check and control the entire system from a web browser.

We will compare three different implementations of the system. One will be written in the UziScript programming language. For the second we will use a project we built in 2017 using the Arduino language and the thinger.io cloud solution [27]. And for the third we will use the irrigation control system described in Donald Norris book “The Internet of Things: Do-It-Yourself at Home Projects for Arduino, Raspberry Pi and BeagleBone Black” [28].

Before we begin with the analysis it is important to note the differences between these three implementations. The following table shows the needed resources and functionality provided by each system.

	UziScript	Thingier.io	Norris
Microcontrollers required	1	1	2
Sensors	Temperature x2 Door open Ambient light	Temperature x2 Door open Ambient light	Moisture
Actuators	Fans Irrigation	Fans Irrigation	Irrigation

The entire code for the thingier.io is available at GitHub [27], Norris' code can be found in his book [28], and the UziScript program is displayed below.

```

import stepper from 'Stepper.uzi' {
  p0 = D3; p1 = D4; p2 = D5; p3 = D6; steps = 48;
}
import thermistor1 from 'Thermistor.uzi' { pin = A1; }
import thermistor2 from 'Thermistor.uzi' { pin = A2; }

var stepping; var temp1; var temp2; var n_readings = 10;

task control() running {
  "The stepping flag can be modified from the dashboard"
  if stepping { stepper.step(30); }
}

task sensors() running {
  var t1; var t2;
  repeat n_readings {
    t1 = t1 + thermistor1.readDegrees();
    t2 = t2 + thermistor2.readDegrees();
    delayMs(1);
  }
  temp1 = t1 / n_readings;
  temp2 = t2 / n_readings;
}

```

We will compare the following aspects of the solution: concurrency, monitoring, and code size:

**Concurrency.** The UziScript version uses concurrent tasks to group the different responsibilities of the program. This makes the code easier to understand and modify because it decouples the code to control the stepper motor from the code that checks the temperature sensor. In contrast, the sequential nature of the thingier.io version conflates the different tasks into one sequential procedure. The negative effects of this limitation can be seen in the code that gathers the temperature sensor readings on each tick. In Norris' version the separation of concerns into independent tasks is done using two Arduinos instead of just one. His implementation uses one Arduino for the control

system and the other for the monitoring, and connects the two using a radio module. This provides true parallelism but makes the system more complex, expensive, and adds an extra point of failure, since the communication between the Arduinos can be interrupted.

**Monitoring.** While Norris and thinger.io versions have a large portion of the code related to the monitoring aspect of the solution, the UziScript version handles this almost with no extra code. Since the UziScript runtime handles the monitoring transparently, all the pins and global variables are automatically published to the host computer. This allows the user to see and change any value from the web dashboard without requiring extra code in the device to handle the communication.

**Code size.** Norris version uses two Arduino boards, so we need to consider both sketches (143 LOC + 18 LOC = 148 LOC). The thinger.io version takes advantage of several libraries provided by the cloud platform, which allowed us to write much less code (118 LOC). In contrast, the UziScript version can fit in a single page (23 LOC). Although we understand lines of code is not an optimal measure of the size of a program (especially when comparing different languages) the difference is worth noting and shows one of the benefits of using a DSL.

## 6 Conclusions and future work

In this paper we have analyzed the limited support for concurrency and monitoring in the Arduino language. We believe these issues limit the potential and reach of the Arduino regarding IoT.

We have explored some of the different solutions currently available as third-party libraries for the Arduino platform and we have proposed a solution based on a virtual machine and high-level language called UziScript.

We have described the implementation of said language, explaining how it solves the identified issues, and we validated its effectiveness by using it to build a small IoT system and comparing it to alternative solutions. The UziScript language presents benefits in code size, concurrency support, and monitoring capabilities. However, of all the implementations we have analyzed, the IoT cloud solution provides more functionality out of the box. This is not a surprise, being a mature commercial product, the cloud solution allows users to build custom dashboards using a simple web interface with support for multiple widgets for different sensors and actuators. The UziScript dashboard, on the other hand, offers limited visualization options and the virtual machine supports only a small number of sensors and actuators.

Despite its current limitations, we believe the implementation works as a successful proof of concept for the proposed architecture and we encourage anyone interested to proceed further with its development.

There are still a lot of improvements to be made. A major issue we found is performance. Our early measurements show the UziScript virtual machine to be 6 to 10 times slower than native Arduino code. While this performance might be good enough

for some applications, we believe there is still room for improvement both in the runtime and compiler.

Another potential problem is the restriction on program size. Considering that most Arduino boards have limited memory, it is desirable that the virtual machine code and the user programs are as compact as possible.

Finally, we are working towards polishing the client tools to make them more suitable for the IoT. In particular, the dashboard should allow the user to use different visualization strategies to display the data from the devices.

Although it is not finished yet we believe the proposed approach is promising and we hope to explore it further in the future.

## References

- [1] D. Evans, *The Internet of Things: How the Next Evolution of the Internet Is Changing Everything*, Cisco Internet Business Solutions Group (IBSG), 2011.
- [2] S. Sicari, A. Rizzardi, L. Grieco y A. Coen-Porisini, «Security, privacy and trust in Internet of Things: The road ahead,» *Computer Networks*, vol. 76, pp. 146-164, 2015.
- [3] D. De Roeck, K. Slegers, J. Criel, M. Godon, L. Claeys, K. Kilpi y A. Jacobs, «I would DiYSE for it!: a manifesto for do-it-yourself internet-of-things creation,» *NordiCHI '12 Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, pp. 170-179, 2012.
- [4] J. Tanenbaum, A. Williams, A. Desjardins y K. Tanenbaum, «Democratizing technology: pleasure, utility and expressiveness in DIY and maker practice,» *CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2603-2612, 2013.
- [5] P. P. Ray, «A survey on Internet of Things architectures,» *Journal of King Saud University - Computer and Information Sciences*, vol. 30, n° 3, pp. 291-319, 2018.
- [6] P. Stevens, «feilipu/Arduino\_FreeRTOS\_Library,» Github, [En línea]. Available: [https://github.com/feilipu/Arduino\\_FreeRTOS\\_Library](https://github.com/feilipu/Arduino_FreeRTOS_Library). [Último acceso: 13 12 2019].
- [7] Floessie, «Floessie/frt,» Github, [En línea]. Available: <https://github.com/Floessie/frt>. [Último acceso: 13 12 2019].
- [8] A. Wisner, «wizard97/ArduinoProcessScheduler: An Arduino object oriented process scheduler designed to replace them all,» 15 January 2017. [En línea]. Available: <https://github.com/wizard97/ArduinoProcessScheduler>. [Último acceso: 23 July 2017].
- [9] M. Miller, «Makuna/Task,» Github, [En línea]. Available: <https://github.com/Makuna/Task>. [Último acceso: 13 12 2019].
- [10] Tinkerspy, «tinkerspy/Automaton,» Github, [En línea]. Available: <https://github.com/tinkerspy/Automaton>. [Último acceso: 13 12 2019].
- [11] bricofoy, «bricofoy/yasm/,» Github, [En línea]. Available: <https://github.com/bricofoy/yasm/>. [Último acceso: 13 12 2019].

- [12] I. Seidel, «ivanseidel/ArduinoThread: A simple way to run Threads on Arduino,» 15 May 2017. [En línea]. Available: <https://github.com/ivanseidel/ArduinoThread>. [Último acceso: 23 July 2017].
- [13] K. Fessel, «fesselk/everytime: A easy to use library for periodic code execution.,» 2 February 2017. [En línea]. Available: <https://github.com/fesselk/everytime>. [Último acceso: 23 July 2017].
- [14] O. Meerbaum-Salant, M. Armoni y M. Ben-Ari, «Habits of programming in scratch,» ITiCSE '11 Proceedings of the 16th annual joint conference on Innovation and technology in computer science education, pp. 168-172, 2011.
- [15] G. Bob, «HaikuVM: a small JAVA VM for microcontrollers,» 2017. [En línea]. Available: <http://haiku-vm.sourceforge.net/>. [Último acceso: 15 Junio 2017].
- [16] R. Suchocki y S. Kalvala, «Microscheme: Functional programming for the Arduino,» de Scheme and Functional Programming Workshop, Washington, D.C., 2014.
- [17] «PyMite - Python Wiki,» 2014. [En línea]. Available: <https://wiki.python.org/moin/PyMite>. [Último acceso: 15 Junio 2017].
- [18] C. L. Jacobsen y M. C. Jadud, «The Transterpreter: A Transputer Interpreter,» Communicating Process Architectures 2004, vol. 62, pp. 182-196, 2004.
- [19] C. L. Jacobsen, M. C. Jadud, O. Kilic y A. T. Sampson, «Concurrent event-driven programming in occam- $\pi$  for the Arduino,» Concurrent Systems Engineering Series, vol. 68, pp. 177-193, 2011.
- [20] M. Elizabeth y C. Hull, «Occam-A programming language for multiprocessor systems,» Computer Languages, vol. 12, n° 1, pp. 27-37, 1987.
- [21] A. W. Roscoe y C. A. R. Hoare, «The laws of Occam programming,» Theoretical Computer Science, vol. 60, n° 2, pp. 177 - 229, 1988 .
- [22] H.-C. Steiner, «Firmata: Towards Making Microcontrollers Act Like Extensions of the Computer,» NIME, pp. 125-130, 2009.
- [23] A. Rojas, «Reporte Robótica Educativa,» Universidad Nacional de La Pampa (UNLPam), 2017.
- [24] «Educabot,» [En línea]. Available: <https://educabot.org/>. [Último acceso: 13 12 2019].
- [25] Totem, «TotemDUINO | Totemmaker.net,» Totemmaker.net, [En línea]. Available: <https://totemmaker.net/product/totemduino-arduino/>. [Último acceso: 13 12 2019].
- [26] R. Moran, M. Teragni y G. Zabala, «A Concurrent Programming Language for Arduino and Educational Robotics,» de XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017), 2017.
- [27] CAETI GIRA, «GIRA/IoTGreenhouse,» Github, [En línea]. Available: <https://github.com/GIRA/IoTGreenhouse>. [Último acceso: 13 12 2019].
- [28] D. Norris, The Internet of Things: Do-It-Yourself at Home Projects for Arduino, Raspberry Pi and BeagleBone Black, McGraw-Hill Education TAB, 2015.

## Control de cámara de germinación hidropónica mediante IoT, broker Mosquitto y placas SBC en invernaderos de la provincia de Misiones

Marcelo Marinelli, Myriam Kurtz, Rubén R. Urquijo, Guillermo Wurm.  
Instituto de Biotecnología de Misiones "María Ebe Reca". Campus U.Na.M. Ruta  
12 Km 7,5, Miguel Lanús, Posadas, Misiones, Argentina.  
{marcelo, myriam, chinourquijo, guillermow77}@gmail.com

**Resumen.** Se presenta la implementación de un sistema de control automático de sistemas hidropónicos en invernaderos de la provincia de Misiones. Para este propósito se utilizó tecnología IoT, hardware de bajo costo con placas tipo SBC Raspberry Pi y NodeMCU. En la placa Raspberry Pi se instaló un *broker* Mosquitto, el cual controla paquetes de datos MQTT, éste es un protocolo específico para IoT, ideal para trabajar con bajo ancho de banda y provee un sistema de seguridad con tres calidades de servicio. Con el programa NODE-RED se programó la administración de cuatro nodos NodeMCU que se encargan de administrar la lógica del control automático de los sistemas hidropónicos. También en NODE-RED se realiza la presentación de datos en un tablero de control, acceso a redes sociales, a repositorios virtuales y se envían los datos a una web de ThingSpeak para su visualización en tiempo real.

**Palabras claves:** IoT, Raspberry Pi, NodeMCU, MQTT, NODE-RED.

### 1. Introducción

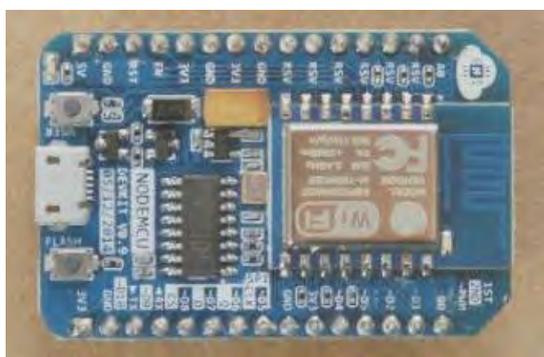
Los cultivos hidropónicos son cada vez más utilizados por sus propiedades sanitarias debido a la no contaminación por contacto con el suelo, como los cultivos tradicionales, esta característica evita que los cultivos estén en contacto con bacterias, parásitos u otro tipo de contaminación. Esta clase de cultivos es conveniente realizarlo en dos etapas: la primera, es la germinación de las semillas hasta que se desarrolle un plantín de determinada característica, según el tipo de especie, para ello se desarrolló en [1] una cámara de germinación con sistema de control difuso. Luego, en la segunda etapa, se trasplantan los platines a otro contenedor que puede ser por sistema de raíz flotante o por sistema NFT (*Nutrient Film Technique*) [2], [3] para su desarrollo definitivo. Los emprendimientos de hidroponía de la provincia de Misiones se realizan en invernaderos para controlar las variables ambientales como humedad, temperatura y radiación solar.

En este trabajo se presenta la implementación de un sistema de control mediante IoT (*Internet of Things*) y dispositivos embebidos de bajo costo del tipo SBC (*Single Board Computer*).

Se utiliza un acceso a internet con router WiFi que provee de una cobertura en toda la superficie de trabajo de un invernadero tipo, que es de aproximadamente 1500 m<sup>2</sup>. De esta forma se vinculan dispositivos inalámbricos de sensado y control de variables en distintos puntos del invernadero.

### 1.1 Hardware de control

Se realiza el sensado y control de dos subsistemas de cultivos hidropónicos: en las cámaras de germinación y en los sistemas NTF. Para esto se utilizan módulos NodeMCU [4] que son dispositivos de código abierto desarrollado para IoT basado en el circuito integrado ESP8266 de ESPRESSIF [5],[6], es un firmware basado en el lenguaje multiparadigma denominado Lua (Luna es portugués). Esta placa de desarrollo tiene conexión a WiFi, 13 puertos GPIO y una entrada analógica entre otras. Existen distintos fabricantes, ya que es de código abierto. Permite protocolos SPI, I<sup>2</sup>C, I<sup>2</sup>S con DMA, UART únicamente para transmisión que puede habilitarse a través del pin GPIO2, tiene dos leds indicadores y una fuente convertidora de continua a continua de 5 V a 3.3 V. El WiFi es compatible con el protocolo 802.11 b/g/n y soporta de autenticación WEP y WPA/WPA2 (Figura 1).



**Figura 1.** Esquema del NodeMCU[4].

Para la programación del módulo se utilizó el IDE Arduino, configurado con la placa NodeMCU 1.0 (ESP-12E Module) como indica la Figura 2.

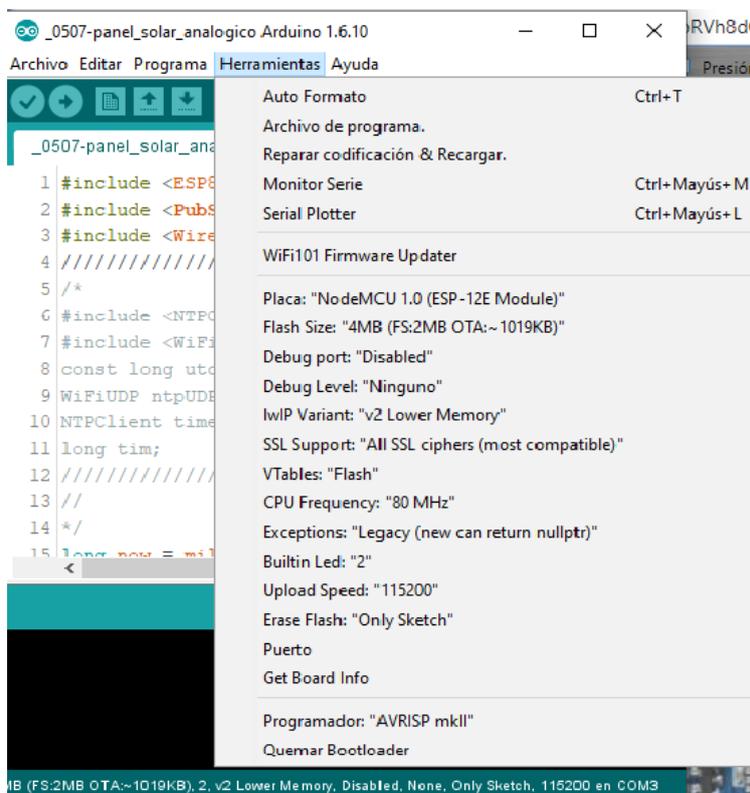


Figura 2. Captura de pantalla de la configuración del IDE Arduino.

## 1.2 Placa Raspberry Pi

Es una placa del tipo SBC (*Single Board Computer*), de bajo costo y soporta distintos sistemas operativo: Linux ARM (Debian, Fedora, Arch Linux), RISC OS2, Windows 10 IoT Core, entre otros. Su tamaño es de 85 mm de longitud por 56 mm de ancho. Requiere alimentación de 5 V y tiene una salida HDMI, una de audio y video compuesto, 4 puertos USB y un conector Ethernet. Tiene una CPU ARM1176JZF-S (armv6k) a 700 MHz, GPU Broadcom VideoCore IV3, memoria RAM de 512 MBytes, capacidad de almacenamiento por medio de Tarjeta SD o SDHC (Figura 3 [6]).

En el conector de expansión posee puertos GPIO (*General Purpose Input/Output*), los cuales representan la interfaz entre la Raspberry Pi y el exterior [7], en total son 17 puertos programables como entrada o salida. Los puertos utilizados son GPIO 11, 12, para sensores de humedad y temperatura y 25 como entrada digital para la obtención del sensor de nivel de nutrientes y el puerto 18 como salida para activar la bomba.

## 1.3 Sensores

Como sensores de humedad y temperatura se utilizó el DHT22, que posee un rango de operación de 0 a 100% HR y de -40 a 80°C y cuenta con una precisión de humedad de 2% RH y de temperatura de 0,5% (Figura 4. a.).

Para medir la temperatura de la solución de nutrientes se empleó el DS18B20 (Figura 4.b.), que posee encapsulado metálico y puede sumergirse sin alterar la

medición. Para el sensado de nivel, fue instalado un sensor con flotante y actuador magnético (Figura 4.c.).



Figura 3. Raspberry Pi 2 modelo B [6].

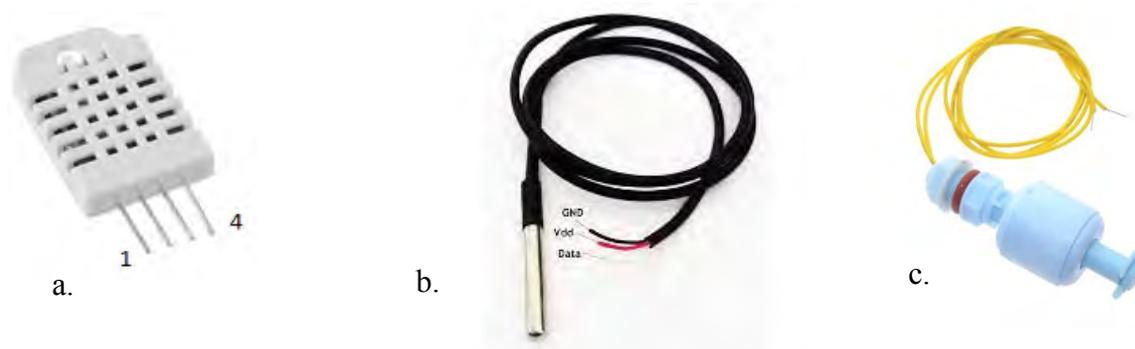


Figura 4. Sensores: a) DHT22 b) DS18B20 c) Nivel magnético.

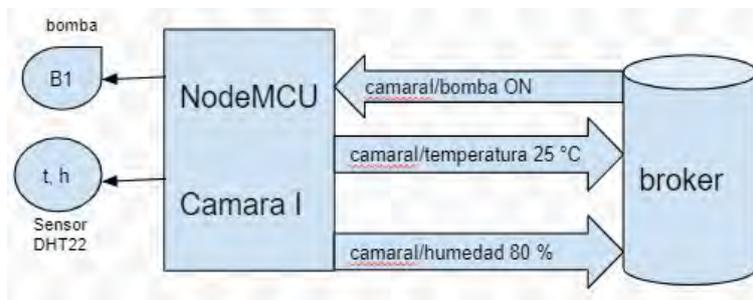
## 2. Protocolo MQTT (Message Queue Telemetry Transport)

Es un protocolo, desarrollado por IBM y Arcom en 1999, de mensajes del tipo publicación/subscripción para trabajar con redes de bajo ancho de banda y alta latencia, lo que lo hace ideal para trabajo en entornos “*machine-to-machine*” (M2M) o “*Internet of Things*” [8]. Debido a sus características, este protocolo es ideal para entrega de paquetes de telemetría en forma confiable, en Versiones 3.1 o superior puede usarse nombre de usuario y contraseña en paquetes MQTT [9].

Los mensajes constan de tres partes: el encabezado fijo de 2 bytes, el encabezado variable de 4 bytes que puede no estar y la carga útil que puede tener hasta un máximo de 256 Mb. Tiene tres opciones de calidad de servicio [10], [11]: QoS 0: At most once, QoS 1: At least once y QoS 2: Exactly once.

Para la implementación del protocolo se necesita un “*broker*”, este administra la comunicación entre los dispositivos que se suscriben para recibir mensajes y los que publican mensajes, cada uno de estos se identifican por un “*topic*”.

En el ejemplo de la Figura 5 se utiliza el tópico camara1 para recibir y transmitir mensajes entre el cliente, en este caso cámara I y el *broker*.



**Figura 5.** Esquema de intercambio de mensajes MQTT.

A nivel de protocolo, para iniciar la comunicación de un cliente con el *broker*, este envía un paquete de control denominado CONNECT, el *broker* contesta con un paquete CONNACK, así se establece un enlace TCP/IP utilizando el puerto 1883 (si utiliza TLS el 8883). El enlace se mantiene hasta la desconexión del cliente. Luego el cliente puede enviar un paquete PUBLISH que contiene el tópico y el contenido. En nuestro ejemplo la cámara I se suscribe con el tópico *camaraI* y envía un mensaje con el valor de la variable humedad, luego la variable temperatura y recibe del *broker* el mensaje para encender la bomba B.

*camaraI/humedad 85%*

*camaraI/temperatura 20°C*

El *broker* también puede enviar mensajes al cliente mediante el mensaje PUBLISH, por ejemplo si pretendemos enviar un mensaje para activar la bomba de la cámara I debe ser:

*camaraI/bomba ON*

### 3. Broker Eclipse Mosquitto

Mosquitto forma parte de la Fundación Eclipse, es un proyecto cuyo sitio web es [iot.eclipse.org](http://iot.eclipse.org) [12], [13]. Es un agente de mensajes de código abierto que implementa las versiones 5.0, 3.1.1 y 3.1 del protocolo MQTT.

En este proyecto se utilizó el *broker* Mosquitto instalado en una placa Raspberry Pi. Como está disponible en el repositorio, se instala con el comando:

```
sudo apt-get install mosquitto
```

En caso contrario, puede descargarse del repositorio: <https://mosquitto.org/blog/2013/01/mosquitto-debian-repository/>.

Una vez instalado en la Raspberry, Mosquitto se puede probar por consola los siguientes comandos: *mosquitto\_sub* y *mosquitto\_pub*, el primero permite suscribirse a un topic del *broker*.

```
mosquitto_sub -h BROKER -t TOPIC (ej. mosquitto_sub -h localhost -t camaraI)
```

El segundo permite publicar mensajes en el *broker* sobre un *topic* determinado y los suscriptores del mismo *topic* lo van a recibir.

**mosquitto\_pub -h BROKER -t TOPIC -m MENSAJE** (ej.  
 mosquitto\_pub -h localhost -t camarai -m "temperatura = 25 °C")

En la Figura 6 se aprecian la captura de pantalla del ejemplo anterior.

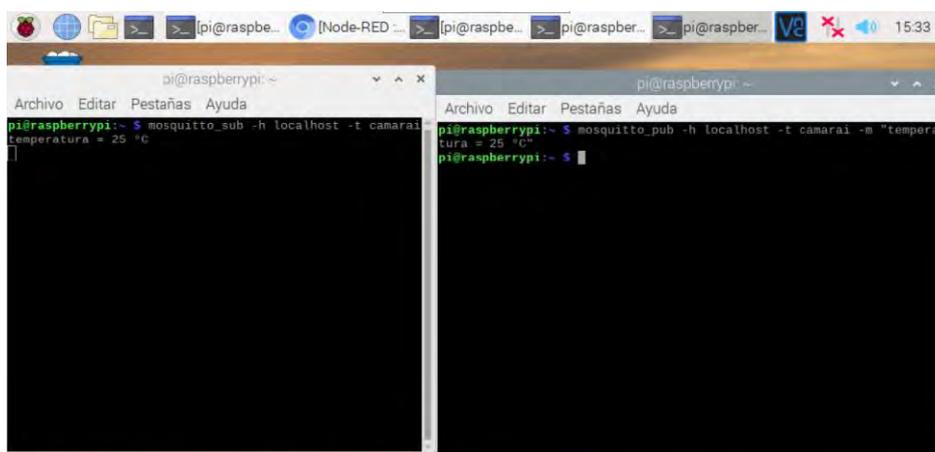


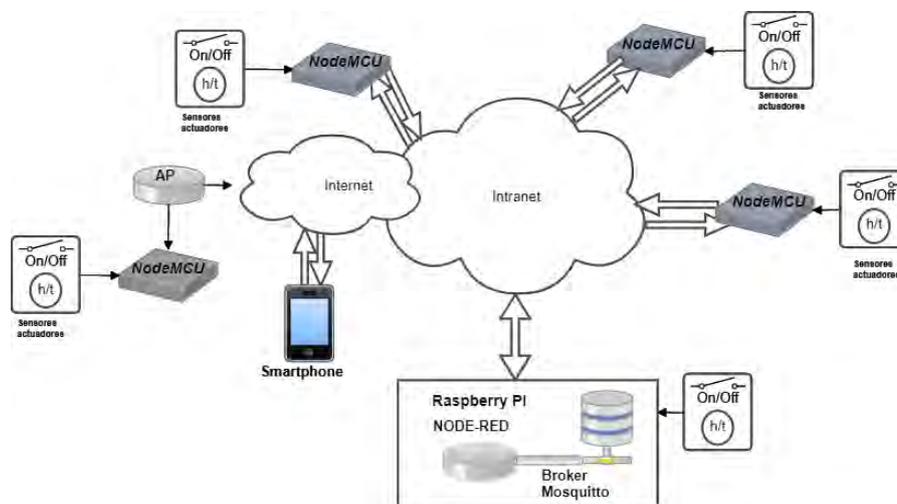
Figura 6. Captura de pantalla del ejemplo.

#### 4. Topología del sistema de control y telemetría de un invernadero en la provincia de Misiones

Debido a las características del invernadero es necesario controlar dos cámaras de germinación, un sistema NFT y los servos de apertura y cierre de las ventanas para recirculación de aire desde el exterior y un sistema de cobertura (media sombra) para disminuir la radiación solar en horas donde el sol se encuentra perpendicular al invernadero Figura 7.

Para el software de control se utilizó NODE-RED, que es una herramienta de programación basada en web para IoT que permite manejar hardware utilizando nodos específicos para cada dispositivo. Es de código abierto, desarrollado por IBM y trabaja con nodos que representan APIs de web, servicios de redes sociales o dispositivos de hardware, se unen mediante conexiones que representan el flujo de datos [13].

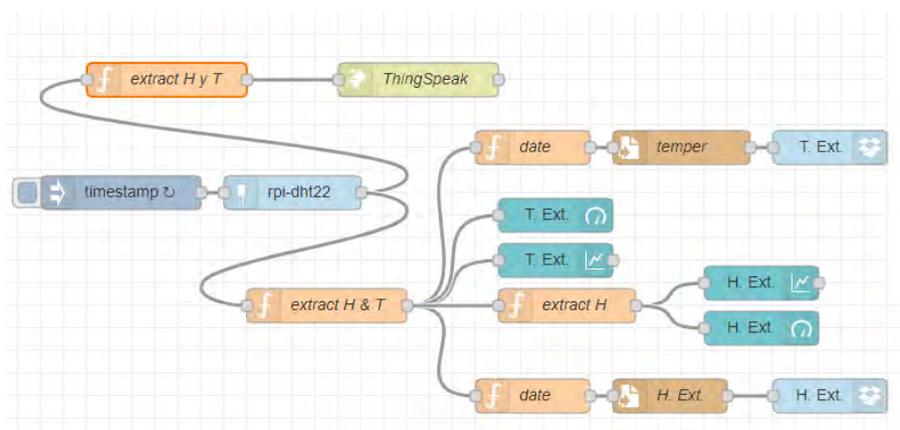
En la versión utilizada del sistema operativo Raspbian usado para la Raspberry Pi ya viene instalado. Posee una interface web del tipo tablero de control en donde se pueden representar los datos y controles en forma gráfica.



**Figura 7.** Topología del sistema de control de invernadero.

La placa Raspberry Pi, en la que se instaló el *broker* Mosquitto, también tiene las siguientes funciones:

- 1) Controla una de las cámaras de germinación mediante un sensor de humedad y temperatura DHT22 (Figura 8), un sensor de temperatura de nutrientes DS18B20 (Figura 9), sensor de nivel de nutrientes y control de la bomba de circulación.
- 2) Almacena los datos en archivos de texto plano.
- 3) Envía a un disco virtual Dropbox los archivos de datos de todas las variables con la fecha y la hora de sensado.
- 4) Envía a ThingSpeak (ver el apartado 4.1) los datos para almacenamiento y visualización en la web.
- 5) Envía los datos a la interfaz Dashboard para visualizar los datos en formato tipo Gage y Chart.



**Figura 8.** Control del sensor DHT22 de la cámara I.

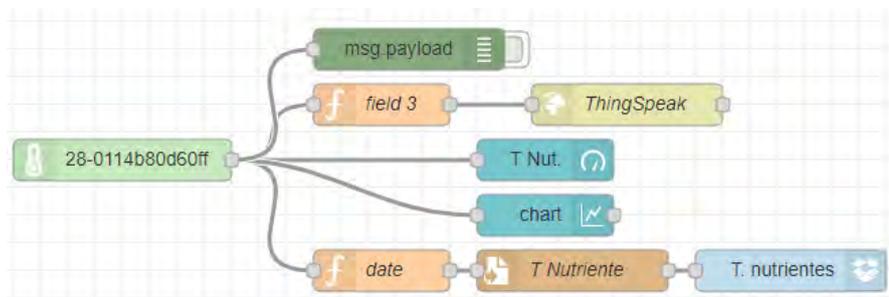


Figura 9. Control del sensor ds18b20 de la cámara I.

#### 4.1 ThingSpeak

Es una plataforma web desarrollada para almacenamiento de datos en la nube para IoT. También permite realizar distintos tratamientos de datos con Matlab. Puede accederse en forma gratuita en <https://thingspeak.com/>.

Para enviar los datos a un canal de ThingSpeak debe programarse un nodo función con los datos del canal y un token que proporciona la cuenta. En el siguiente ejemplo, se envían los datos que llegan al nodo proveniente de un sensor y se reemplazan las X por el token. Luego se conecta con un nodo HTTP para que realice la función GET y envíe el dato correspondiente.

```
var field3 = msg.payload;
msg.url="https://api.thingspeak.com/update?api_key=XXXXXXXXXXXX&
field3="+field3;
return msg;
```

La representación de los datos se realiza en tiempo real, puede apreciarse en la Figura 10.

#### 4.2 Nodos NodeMCU

Cada nodo "NodeMCU" de la Figura 7, está compuesto por los sensores descritos anteriormente con la diferencia que el conexionado se hace por los puertos GPIO del dispositivo denominados D4 con la librería `DallasTemperature.h` para el DS18b20, D3 para el DHT22 con la librería `DHT.h`, D5 para el sensor de nivel de nutrientes y D6 para el control de la bomba.

Para la conexión WiFi, se programan los datos del AP: nombre del ESID y la password, para la conexión con el *broker* se indica la dirección IP del mismo. De esta manera y mediante la librería `PubSubClient.h`, se pueden enviar y recibir paquetes MQTT:

```
client.subscribe("invernadero/camaraiii");

client.publish("invernadero/camaraiii", temperatura)
```



Figura 10. Captura de pantalla de los gráficos de ThingSpeak.

### 4.3 Control automático y mensajería NODE-RED

El *broker* recibe los datos, de los nodos NodeMCU, de cada sector del invernadero (Figura 11), los almacena en archivo de texto plano y en la nube con Dropbox, los presenta en forma gráfica en *Dashboard* e ingresan a un nodo de control que, en función de los datos obtenidos, produce señales de control que por medio de los nodos de salida MQTT envían al respectivo NodeMCU. Con respecto al nivel, la señal recibida es del tipo binaria y en caso de que el nivel esté bajo, envía un mail al administrador y activa señales sonoras y de texto en el tablero de control.

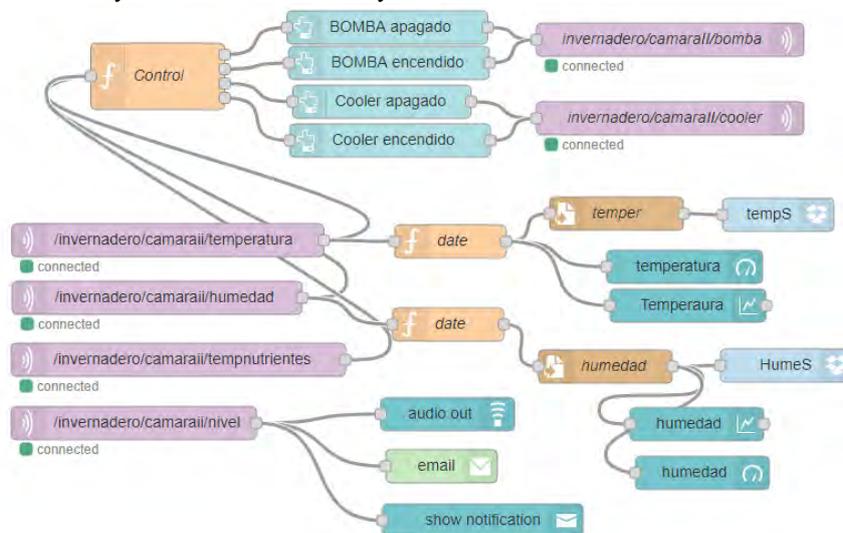


Figura 11. Control de la cámara II con los datos recibidos por MQTT desde un NodeMCU.

## 5. Conclusiones

Se realizó la implementación de un sistema de control para invernaderos de la provincia de Misiones utilizando placas SBC de bajo costo Raspberry Pi y prototipos NodeMCU basados en ESP8266 conectados a un WiFi, utilizando protocolo MQTT con un *broker* Mosquitto. Se diseñó la conexión de los sensores a cada placa y se programó el software embebido en cada caso. Mediante NODE-RED se administran los datos representándolos en un tablero de control y se los almacena en archivos locales y en la nube, utilizando Dropbox y ThingSpeak.

En la topología se aprecia un acceso a un proveedor de Internet, lo que permite visualizar el tablero de control desde un Smartphone o una PC.

## 6. Referencias

1. Marinelli, M., Acosta, N., Toloza, J. M., & Kornuta, C. (2017). Fuzzy Control of a Germination Chamber. *Journal of Computer Science and Technology*, 17(01), 74-78.
2. Hernández, C. J., & Hernández, J. L. (2005). Valoración productiva de lechuga hidropónica con la técnica de película de nutrientes (nft). *Naturaleza y Desarrollo*, 3, 1, 11-16.
3. Garzón López, S. S., & Arévalo de Gauggel, G. (2006). *Evaluación del rendimiento de tres variedades de lechuga bajo el sistema NFT (nutrient film technique) de hidroponía con dos soluciones de nutrientes* (No. T2244). ESCUELA AGRÍCOLA PANAMERICANA.
4. NodeMCU Connect Things EASY, [http://www.nodemcu.com/index\\_en.html](http://www.nodemcu.com/index_en.html) [Ultimo acceso 8/6/2020].
5. Espressif System ESP8266 Series, <https://www.espressif.com/> [acceso 8/6/2020].
6. Raspberry Pi Model 2B <https://www.raspberrypi.org> [acceso 10/6/2020].
7. 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.
8. Ceja, J., Renteira, R., Ruelas, R., & Ochoa, G. (2017). Módulo ESP8266 y sus aplicaciones en el internet de las cosas. *Revista de Ingeniería Eléctrica*, 24-36.
9. Protocolo MQTT (Message Queue Telemetry Transport), <http://mqtt.org> [acceso 6/6/2020].
10. Defossé, N., López, R. A., Gómez, M. E., Konstantinoff, P., Wahler, S. P., Castro, L., & Harris, G. (2017, August). Implementación de middleware publicador/subscriptor para aplicaciones web de monitoreo. In *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*.
11. Naik, N. (2017, October). Choice of effective messaging protocols for IoT systems: MQTT, CoAP, AMQP and HTTP. In *2017 IEEE international systems engineering symposium (ISSE)* (pp. 1-7). IEEE.
12. Eclipse Mosquitto™. An open source MQTT broker, <https://mosquitto.org> [Ultimo acceso 10/9/2020].
13. Open Source for IoT, <https://iot.eclipse.org>, [Ultimo acceso 10/9/2020].
14. Marinelli, M., & Urquijo, R. (2017). Sistema de Control de una cámara de germinación hidropónica con IoT. In *XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017)*.
15. Chanthakit, S., & Rattanapoka, C. (2018, July). Mqtt based air quality monitoring system using node MCU and node-red. In *2018 Seventh ICT International Student Project Conference (ICT-ISPC)* (pp. 1-5). IEEE.

# Análisis de una plataforma de simulación para Cloud Computing. Un caso de estudio

Tomás Rosales<sup>1</sup>, Julián Spinelli<sup>1</sup>, Marcos Di Nardo<sup>1</sup>, Román Bond<sup>2</sup>, Daniel Rosatto<sup>2</sup>, Diego Encinas<sup>1,2,3</sup>, Fernando Romero<sup>1,3</sup>

<sup>1</sup>Facultad de Informática, Universidad Nacional de La Plata. Buenos Aires, Argentina

<sup>2</sup>Proyecto de Investigación SimHPC, Programa TICAPPS, Universidad Nacional Arturo Jauretche. Buenos Aires, Argentina

<sup>3</sup>Instituto de Investigación en Informática (III-LIDI). Facultad de Informática, Universidad Nacional de La Plata - Centro Asociado CIC. Buenos Aires, Argentina  
tomasjoaquinrosales@gmail.com, julspni@gmail.com, mardinar@gmail.com, rbond@unaj.edu.ar, danielrosatto@gmail.com, dencinas@lidi.info.unlp.edu.ar, fromero@lidi.info.unlp.edu.ar

**Resumen** Se presenta una aproximación a la construcción de simuladores en entornos de Cloud Computing por medio de un framework específico. Con este fin, se hace uso de un sistema real desplegado en Amazon Web Services para obtener conjuntos de datos de salida relacionados a la ejecución de tareas en forma paralela. Y por medio de la plataforma de simulación CloudSim se desarrolla una simulación. Finalmente, se procede a la verificación y validación del simulador utilizando los datos de los sistemas físicos y simulados.

**Palabras Claves:** Simulación, Cloud Computing, CloudSim.

## 1. Introducción

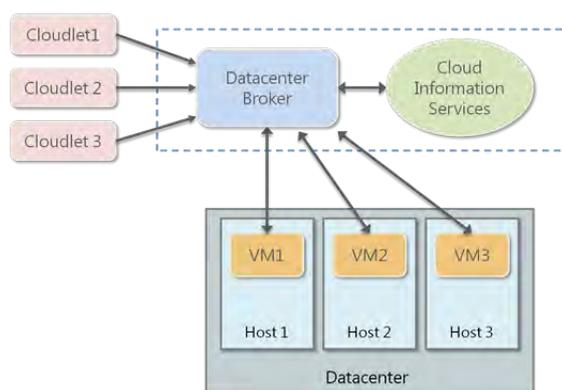
Con el surgimiento de la idea de computación en nube (Cloud Computing) se han abierto nuevas posibilidades para todo el sector IT. Este paradigma está orientado a ofrecer recursos, aplicaciones y plataformas como servicios por Internet, con modelos de precios basados en el consumo utilizado [1]. Sus inicios se deben a proveedores de Internet de gran escala, que construyeron su propia infraestructura Cloud. La idea inicial, fue alquilar recursos ociosos de su propia infraestructura para usuarios que prefirieran utilizar recursos pagando exclusivamente bajo demanda.

Para poder maximizar los beneficios de este modelo computacional complejo, es necesario contar con herramientas que permitan comprender el funcionamiento del paradigma. Por ejemplo, es posible simular el procesamiento de una carga determinada por parte de una aplicación de comercio electrónico y estimar cuánto costaría su ejecución en diferentes proveedores de servicios de Cloud [2]. Las simulaciones, además, son esenciales para intentar prever el comportamiento de la aplicación y para estimar costos, antes de ubicar la aplicación en una determinada plataforma de operación.

En este trabajo se utiliza CloudSim para ejecutar simulaciones de escenarios en Cloud Computing. Cloudsim basa su funcionamiento en un conjunto de herramientas de simulación extensibles que permite el modelado y simulación de sistemas de computación en la nube. El objetivo del presente trabajo es configurar Cloudsim para reflejar las características ofrecidas por un conocido proveedor de servicios (AWS, Amazon Web Services) [3]. Luego se miden los tiempos de respuestas de ambos entornos para una tarea determinada y así obtener conclusiones generales que serán especialmente útiles al momento de determinar la configuración más conveniente para desplegar la aplicación.

## 2. CloudSim

CloudSim [4] es un framework que permite modelar y simular diferentes arquitecturas de Cloud Computing. Mediante esta herramienta es posible representar cada uno de los componentes intervinientes en la nube a través de clases de Java. A partir de dichas clases se modelan las diferentes entidades, como Datacenters, Brokers, Hosts, Máquinas Virtuales, Tareas (Cloudlets), Servicios y Redes, entre otras, cada una con parámetros propios que las caracterizan [5] como se puede ver en la *Figura 1*.



**Figura 1.** Principales entidades en CloudSim.

El escenario de simulación más básico a configurar y ejecutar, consiste en un Data Center con la administración de un broker que asigna los recursos disponibles [6]. El objetivo es procesar todas las instancias de Cloudlets.

### 2.1. Cloudlet

Esta clase representa las tareas que deben ejecutarse en la infraestructura de la nube para su procesamiento. [7]. Cada cloudlet tiene una sobrecarga de

instrucciones y transferencia de datos que debe realizar durante su ciclo de vida. En toda simulación de CloudSim los cloudlets interactúan con el broker y los datacenters con el objetivo que cada Virtual Machine (instancia) sea capaz de procesarlo, según su algoritmo de asignación. En el escenario de simulación que se propone para este trabajo, los parámetros de cada cloudlet son los que se modifican para poder reflejar la carga de trabajo de una tarea determinada.

```

1 // Fifth step: Create one Cloudlet
2 cloudletList = new ArrayList<Cloudlet>();
3
4 // Cloudlet properties
5 int id = 0;
6 long length = 400000;
7 long fileSize = 300;
8 long outputSize = 300;
9 UtilizationModel utilizationModel =
10     new UtilizationModelFull();
11 Cloudlet cloudlet = new Cloudlet(id, length, pesNumber,
12     fileSize, outputSize, utilizationModel,
13     utilizationModel, utilizationModel);

```

A continuación se describen cada uno de los parámetros que describen a un cloudlet:

1. id: identificador único. Importante para el motor de cloudsim, ya que lleva el control de qué cloudlets ejecutó y cuáles no.
2. length: el tamaño del cloudlet en MiB.
3. fileSize: el tamaño de archivo del cloudlet antes de enviarlo a un Datacenter para su procesamiento.
4. outputSize: el tamaño de archivo del cloudlet después de su ejecución en un Datacenter.
5. pesNumber: hace referencia a la unidad de procesamiento. Se define en términos de millones de instrucciones por segundo (MIPS).

Los últimos 3 parámetros hacen referencia al modelo de utilización de la cpu, la memoria ram y del ancho de banda (red), respectivamente. A su vez, las variables que hacen referencia al tamaño de archivo del cloudlet están en bytes.

### 3. AWS

Amazon Web Services (AWS) es una colección de servicios de computación que en conjunto forman una plataforma de computación en la nube, ofrecidas a través de Internet por Amazon.com. Aplicaciones populares como Dropbox, Foursquare, HootSuite utilizan en su infraestructura servicios de AWS. De las ofertas internacionales, es una de las más importantes de computación en la nube considerado como un pionero en este campo que compite directamente con servicios como Microsoft Azure [8] y Google Cloud Platform [9]. AWS ofrece herramientas en diferentes categorías: base de datos, contenedores, robótica y machine learning, entre otras. En este trabajo se utiliza el servicio de cloud computing Amazon EC2 (Elastic Compute Cloud) de la capa gratuita de AWS que

proporciona capacidad cómputo en la nube y todo lo necesario para la creación de instancias, el mantenimiento y el escalamiento de las mismas.

#### 4. Benchmark

Se propuso generar un benchmark propio analizando distintos parámetros, para luego poder relacionarlos y representar a los Cloudlets del simulador. Durante el proceso de investigación de la herramienta CloudSim la realización de dos experimentos permitieron encontrar el parámetro adecuado a modificar.

En el primer experimento se intentó utilizar la propiedad *FileSize*. En un principio, se pensó que al variar este parámetro daría como resultado diferentes tiempos de respuesta de finalización de la tarea que realiza el simulador. En las diferentes pruebas se obtuvo el mismo tiempo de respuesta, con lo que se concluyó que la propiedad *FileSize* no es de utilidad para el fin de este trabajo. En el segundo experimento se utilizó la propiedad *Length*. Como resultado, los tiempos de respuesta de los cloudlets varían proporcionalmente con la cantidad de instrucciones que se indican en este parámetro.

En consecuencia se generó un benchmark para conseguir una forma de obtener cantidades de instrucciones de diferentes tareas y así variar la propiedad *Length*, con lo cual se lograran tiempos de respuesta diferentes y así poder realizar un análisis con respecto a los tiempos de respuesta en AWS. Teniendo en cuenta los experimentos desarrollados anteriormente, se propone generar un programa en lenguaje de programación C que represente la multiplicación de una matriz por otra, es decir:

$$R[n * n] = A[n * n] * B[n * n] \quad (1)$$

Este programa es utilizado como referencia para comparar en AWS y CloudSim. En la siguiente sección, está el detalle de cómo se establece la relación entre el entorno simulado y el ejecutado en la nube de AWS.

```

1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <sys/time.h>
4  #include "matrices.h"
5
6  int *A, *Ai, *R;
7  double timetick;
8
9  int main(int argc, const char* argv[]){
10     if (argc < 2){
11         printf("Falta un parametro.\n");
12         printf("1. Dimension: 512, 1024 y 2048\n");
13         return 0;
14     }
15     N = atoi(argv[1]);
16     int i, j, k;
17     double timetick;
18
19     A = declaration_matrix(A, N);
20     Ai = declaration_matrix(Ai, N);

```

```

21     R = declaration_matrix(R, N);
22     init_matrix(A, R, Ai, N);
23     multi_matrix(R, A, Ai, N);
24     print_matrix(R, N, 0);
25     free_matrix(A, R, Ai);
26 }

```

#### 4.1. Construcción del Benchmark

Por lo explicado anteriormente, el benchmark está definido en 5 módulos que deben mantener relación con CloudSim. Cada módulo corresponde a un cloudlet y tienen la característica que cada uno de ellos es incremental al anterior. Es decir que los módulos van incorporando funcionalidades, como se puede ver en el **Cuadro 1**, y de esta forma se logra que la propiedad **Length** del cloudlet varíe en cada uno de ellos. Los 5 módulos resuelven el problema planteado anteriormente de resolución de dos matrices de NxN.

Cuadro 1.

Módulo	Métodos Invocados
1	<i>Declaración de matriz A</i>
2	<i>Declaración de matriz A y Ai</i>
3	<i>Declaración de matriz A, Ai y R</i>
4	<i>Declaración de matriz A, Ai y R</i> <i>Inicialización de Matriz A, Ai y R</i>
5	<i>Declaración de matriz A, Ai y R</i> <i>Inicialización de Matriz A, Ai y R</i> <i>Multiplicación de matriz A y Ai almacenamiento de resultado en matriz R</i> <i>Liberación de memoria de matrices A, Ai y R.</i>

Utilizando los módulos del **Cuadro 1** se midió únicamente la cantidad de instrucciones de procesamiento que implementan cada uno. Para ello, en primer lugar se trató que el código no posea referencias a operaciones de llamadas al sistema operativo o de entrada/salida. Luego, se obtuvo mediante el Debug de CodeBlocks la cantidad de instrucciones de los módulos. Posteriormente, se midió el crecimiento de los mismos al incrementar el tamaño de las matrices multiplicadas, observando con especial atención las estructuras de repetición **for(int i=0, i<=N, i++)**. A modo de ejemplo, se explica el módulo 1. El módulo 1 está compuesto por la siguiente funcionalidad:

```

1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <sys/time.h>
4  #include "matrices.h"
5  #define N 512
6
7  int *A, *Ai, *R;
8

```

```

9  int main(int argc, const char* argv[]){
10     A = declaration_matrix(A, N);
11 }

```

Para poder contar la cantidad de instrucciones de cada módulo, se utilizó el programa de computadora CodeBlocks, el cuál tiene incorporada las siguientes herramientas para el análisis del código:

1. Disassembly
2. Watches
3. Memory dump

Estas herramientas permiten observar las instrucciones de máquina generadas por el compilador y poder realizar un conteo de las mismas. Utilizando el debugger de CodeBlocks se obtuvo para el módulo 1 lo siguiente:

```

TAMAÑO DEL MAIN: 13 Inst
TAMAÑO DE FUNCIONES: 13 Inst
TAMAÑO TOTAL DE CODIGO: 26 Inst

```

Este proceso se replica para los módulos siguientes teniendo en cuenta el tamaño de la matriz a multiplicar. Es decir, se mide la cantidad de instrucciones de cada módulo para  $N = 512$ ,  $N = 1024$  y  $N=2048$ . Una vez obtenida la cantidad de instrucciones de cada módulo, en el entorno de CloudSim, se puede traducir como la cantidad de instrucciones de cada cloudlet dado en millones de instrucciones (MIPs). Se modifica la escala con la que se está trabajando, por ende, si se necesita modelar 10 instrucciones en Cloudsim se cargan como si fuesen 10 MIPs.

## 5. Ejecución de Benchmark y Mediciones

### 5.1. Mediciones de Tiempo de Respuesta en AWS

Se ejecutan los módulos definidos en el el Cloud AWS y se simulan las tareas que representan a los mismo módulos en CloudSim. En ambos casos se miden los tiempos de respuesta obtenidos a fin de poder analizar los resultados y llegar a conclusiones.

### 5.2. Configuración del Entorno en CloudSim

Para llevar a cabo la simulación es necesario configurar las propiedades en CloudSim lo más cercano al entorno real de AWS, en donde se realizaron las mediciones de tiempo de repuesta. Para esto, se configuró de la siguiente manera la máquina virtual del simulador:

```

1  // VM description
2  int vmid = 0;
3  int mips = 480; // Deberia configurarse en 4800
4  long size = 10000; // image size (MB)
5  int ram = 512; // vm memory (MB)
6  long bw = 1000;
7  int pesNumber = 1; // number of cpus

```

De la instancia de AWS se obtuvieron algunos datos de importancia para la simulación como se puede ver en la **Figura 2**.

```

ubuntu@ip-172-31-27-75:~$ cat /proc/cpuinfo
processor       : 0
vendor_id     : GenuineIntel
cpu family    : 6
model         : 53
model name    : Intel(R) Xeon(R) CPU E5-2676 v4 @ 2.40GHz
stepping     : 2
microcode    : 0x2c
cpu MHz      : 2400.079
cache size   : 30720 KB
physical_id  : 0
siblings     : 1
core_id     : 0
cpu cores   : 1
apicid      : 0
initial apicid : 0
fpu         : yes
fpu_exception : yes
cpuid level : 13
wp         : yes
flags       : fpu vme de pse tsc mtr par mce cmf sstt pge mca
             mmx syscall nx sse3 ssse3 fxsr sse4_1 sse4_2 x2apic
             movbe bmi1 avx2 smep bmi2 rtmr rcpuid mwaitx
             hle
bogomips    : cpu_maltdown_spectre_v1 spectre_v2 spec_store_bypass l1tf
             : #900.15

```

**Figura 2.** Especificaciones técnicas de la máquina virtual en AWS.

CloudSim soporta un máximo de 1000 MIPS. Como se logró determinar que los tiempos de respuesta se calculan con la relación directa entre la cantidad de instrucciones y los MIPS del procesador, se asignó el valor 480 (valor relacionado con los BogoMips) sabiendo que los resultados obtenidos serán 10 veces mayores. Por otro lado, el tiempo con el que se resuelven las tareas en el simulador para el caso de cloudlets limitados por tareas cortas y con un valor muy alto de MIPS, es menor a 1. Este valor coincide con el tiempo entre SimEvents por defecto que establece cuál es el muestreo entre eventos. Debido a esto, se tuvo que reconfigurar este parámetro para obtener mayor resolución en la escala temporal al medir tiempos de tareas muy pequeñas.

### 5.3. Configuración del Cloudlet

Se modifica el parámetro “length” por la cantidad de instrucciones obtenidas de los cálculos realizados en MIPS como se explicó anteriormente.

```

1 // Cloudlet properties
2 int id = 0;
3 long length = 26;
4 long fileSize = 20;
5 long outputSize = 20;
6 UtilizationModel utilizationModel =
7     new UtilizationModelFull();

```

Se ejecutó la simulación en CloudSim y se obtuvo lo siguiente salida del simulador:

```

===== OUTPUT =====
Cloudlet ID   STATUS   Start Time   Finish Time
             0   SUCCESS      0,0001       0,0543

```

Los tiempos de respuesta obtenidos deben ser ajustados para contemplar la limitación con los MIPs del procesador y a su vez, considerar la cantidad de instrucciones como si fueran millones de instrucciones, entonces para el ejemplo mostrado se obtiene lo siguiente:

$$Respuesta[s] = \frac{0,0542}{1,000,000 \times 10} = 0,0000000542 \quad (2)$$

Este proceso se replicó para todos los cloudlets que fueron generados.

#### 5.4. Configuración y Ejecución en Cloud AWS

Se creó una instancia en alguna de las zonas de disponibilidad de AWS y se realizaron 5 ejecuciones de cada módulo y se obtuvieron diferentes tiempos de respuesta. Para las conclusiones se toman los tiempos promedios. Por ejemplo para los módulos con N=512 se obtuvo los siguientes datos que se pueden ver en el *Cuadro 2*.

**Cuadro 2.** Tiempos de respuesta para matriz de N = 512.

Ejecuciones	Módulo 1	Módulo 2	Módulo 3	Módulo 4	Módulo 5
<b>T1</b>	0,000035	0,000038	0,000041	0,007840	0,664568
<b>T2</b>	0,000033	0,000040	0,000041	0,007368	0,662680
<b>T3</b>	0,000031	0,000036	0,000039	0,007665	0,663200
<b>T4</b>	0,000035	0,000038	0,000038	0,007463	0,669473
<b>T5</b>	0,000029	0,000033	0,000042	0,007523	0,668542
<b>Promedio</b>	<b>0,0000326</b>	<b>0,000037</b>	<b>0,0000402</b>	<b>0,0075718</b>	<b>0,6656926</b>

## 6. Presentación de Resultados

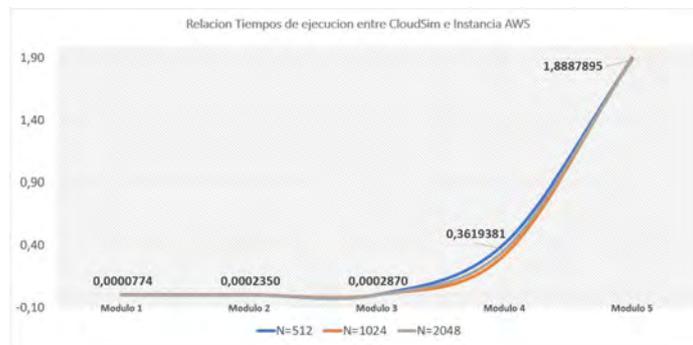
A continuación, se presentan los resultados obtenidos, de forma tabulada como se puede ver en el *Cuadro 3*.

En las imágenes siguientes se grafica la relación entre los tiempos de ejecución en CloudSim y en AWS para todas las tareas implementadas por cada uno de los módulos, generando curvas distintas según el tamaño de las tareas. Ésto se puede ver en la *Figura 3*.

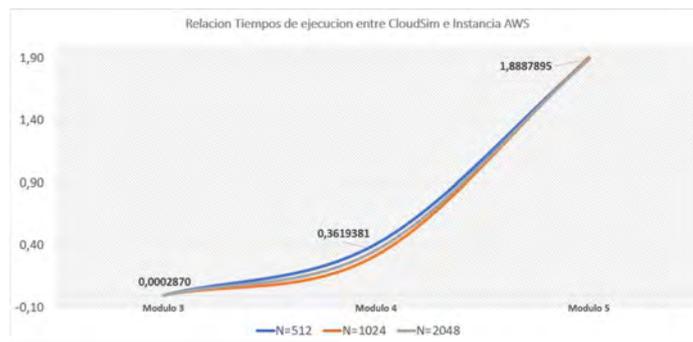
Luego, debido a que los dos primeros módulos incluyen bajo nivel de procesamiento en comparación a los siguientes, se hizo foco en los posteriores para una mejor apreciación de los resultados como se vé en la *Figura 4*.

**Cuadro 3.** Comparación de tiempos de respuestas en CloudSim y Clous AWS

Módulos	#Instrucciones	MI	Tiempo de Respuesta en CloudSim	Tiempo de Respuesta en Cloud AWS	Relación m/AWS	CloudSi-
1 - N=512	26	0,000026	0,00000000542	0,0000326		0,0001662
2 - N=512	44	0,000044	0,00000000917	0,000037		0,0002477
3 - N=512	62	0,000062	0,00000001292	0,0000402		0,0003213
4 - N=512	14942351	14,942351	0,00311298979	0,0075718		0,4111294
5 - N=512	6054740197	6054,740197	1,261404208	0,6656926		1,8948749
1 - N=1024	13	0,000013	0,00000000271	0,000036		0,0000752
2 - N=1024	44	0,000044	0,00000000917	0,000038		0,0002412
3 - N=1024	62	0,000062	0,00000001292	0,000065		0,0001987
4 - N=1024	59768975	59,768975	0,01245186979	0,038918		0,3199514
5 - N=1024	48378151141	48378,15114	10,07878149	5,292631		1,9043046
1 - N=2048	13	0,000013	0,00000000271	0,000035		0,0000774
2 - N=2048	44	0,000044	0,00000000917	0,000039		0,000235
3 - N=2048	62	0,000062	0,00000001292	0,000045		0,000287
4 - N=2048	239075471	239,075471	0,04980738979	0,137613		0,3619381
5 - N=2048	386786132212	386786,1322	80,58044421	42,66248		1,8887895



**Figura 3.** Relación entre el tiempo de ejecución CloudSim y el tiempo de ejecución AWS.



**Figura 4.** Relación entre el tiempo de ejecución CloudSim y el tiempo de ejecución AWS en tareas largas.

## 7. Conclusiones

Se desarrollan las conclusiones en base a las pruebas y análisis realizados. A nivel general, se revela la importancia de contar con una herramienta para modelar arquitecturas de Cloud Computing permitiendo parametrizar una vasta gama de componentes y servicios existentes. Luego, en el caso particular de las pruebas realizadas tomando como parámetro los tiempos de ejecución, se comprueba que, la relación entre los tiempos de respuesta en AWS y en el simulador CloudSim sigue un comportamiento lineal. Aunque, al tratarse de un experimento acotado, si se extiende el rango de pruebas incrementando tamaños seguramente la curva tienda a un valor máximo y se mantenga en ese valor. Además, se comprueba que el valor de la relación entre los tiempos de respuesta en el Cloud AWS y en el simulador CloudSim depende de la tarea realizada.

A partir de las dos últimas conclusiones particulares, se deriva que para poder estimar el tiempo de procesamiento de una tarea específica (y por ende su costo asociado), es imperativo poder determinar una muestra mínima de esa tarea (ejemplo: para renderizar una imagen de 200 MPíxeles sería necesario estimar las instrucciones del renderizado de una imagen de 1 MPíxel). Luego simular la ejecución en CloudSim, realizar esa tarea en un Cloud real (ej: AWS) para determinar la relación de tiempos. Es decir, realizar una calibración del simulador generado y de esta manera se podrían predecir resultados para tareas de mayores tamaños.

## Referencias

1. Buyya, R., Broberg, J., Gościński, A. (Eds.). Cloud computing: principles and paradigms. Hoboken, N.J: Wiley. (2011)
2. Alberto Crespo Guzmán. Um Estudo Da Simulação de Serviços da Amazon EC2 Utilizando o Simulador CloudSim. (2011)
3. Amazon Web Services, <https://aws.amazon.com/es/what-is-aws/>
4. Melbourne Clouds Labs, <http://www.cloudbus.org/cloudsim/>
5. Patiño, Paláez, Granda, Segovia. Diseño y Simulación de un Data Center Cloud Computing que cumpla con la norma PCI-DSS. (2019)
6. Rosatto, D.; Bond, R.; Encinas, D.; Morales, M. “Modelado y simulación de arquitecturas de cloud computing con cloudsim: comunicación entre entidades”. XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). La Plata, Argentina.
7. Calheiros, R., Ranjan, R., Buyya, R. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Journal Software-Practice Experience. Volume 41, John Wiley Sons. (2011)
8. Microsoft Azure, <https://azure.microsoft.com/es-es/overview/>
9. Google Cloud Platform, <https://cloud.google.com/>

# Administración del Tiempo Ocioso Mediante Slack Stealing en FreeRTOS

Francisco E. Páez<sup>1</sup>, José M. Urriza<sup>1</sup>, and Javier D. Orozco<sup>2,3</sup>

<sup>1</sup> Depto. de Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco, Sede Puerto Madryn, Argentina

<sup>2</sup> Depto. de Ingeniería Eléctrica y Computadoras, Universidad Nacional del Sur, Bahía Blanca, Argentina

<sup>3</sup> CONICET

fpaez@unpata.edu.ar, josemurriza@gmail.com

**Resumen** La planificación eficiente de conjuntos de tareas con requerimientos heterogéneos, sin perder la predictibilidad de ejecución de las tareas de tiempo real, es un área que está tomando relevancia en la actualidad. Para lograr este objetivo, se requiere una correcta administración del tiempo ocioso disponible. En este trabajo se presenta una implementación que permite planificar estos conjuntos en FreeRTOS utilizando Slack Stealing.

**Palabras claves:** RTS · SS · RM · Planificación · RTOS

## 1. Introducción

En la actualidad existe una creciente necesidad de integrar, en un mismo sistema embebido, tareas con requerimientos de tiempo real *críticos* (para las que se debe garantizar el cumplimiento de sus restricciones temporales) junto con tareas sin requerimientos temporales estrictos, que deben cumplir con algún tipo de *calidad de servicio* (atención prioritaria, robustez, tolerancia a los fallos, etc.). El sistema debe poder planificar estos dos conjuntos de manera eficiente, brindando una calidad de servicio aceptable sin comprometer los requerimientos de tiempo real. Esto requiere utilizar el tiempo ocioso que dejan las tareas críticas para planificar el resto de las tareas. Diversos métodos han sido propuestos a tal fin y la planificación de estos conjuntos heterogéneos es una importante área de investigación [1].

Este trabajo presenta una implementación de la técnica de *Slack Stealing* (SS) para administrar el tiempo ocioso en el *Sistema Operativo de Tiempo Real* (SOTR) FreeRTOS<sup>4</sup>. Este es un SOTR de código abierto para dispositivos embebidos, con soporte para *Sistema de Tiempo Real* (STR) duros. Está desarrollado en lenguaje C, es de pequeño tamaño, modular y con bajos requerimientos de recursos. A la fecha, soporta más de 33 arquitecturas y es patrocinado por Amazon para su uso en sistemas IoT (*Internet of Things*).

<sup>4</sup> <https://www.freertos.org>

El desarrollo presentado se basa en [2], introduciendo mejoras en su diseño y aprovechando nuevas funcionalidades provistar por FreeRTOS en sus versiones más recientes.

A continuación se presenta una breve introducción a los *STR* y a la técnica de *SS*. En la sección 2 se detalla el diseño e implementación. La sección 3 presenta las pruebas realizadas. Las conclusiones y trabajos futuros se discuten en la sección 4.

### 1.1. Sistemas de Tiempo Real y Slack Stealing

En un *STR* los resultados, además de ser correctos aritmética y lógicamente, deben producirse antes de un determinado tiempo, denominado vencimiento [3]. Si no se admite la pérdida de ningún vencimiento, el *STR* es *duro* o *crítico* y si tolera la pérdida de algunos se lo denomina *blando*. Si existe una cota máxima de pérdidas se dice que es de tipo *firme*.

Como en un *STR crítico* la pérdida de un vencimiento puede tener consecuencias graves (pérdida de vidas, daños materiales, al medio ambiente, etc.), durante su diseño se garantiza que cada tarea cumpla con su vencimiento, mediante *tests de planificabilidad*. Un *STR* que cumple con estos tests se denomina *planificable*. Las primeras contribuciones al respecto fueron realizadas en [4], donde se demostró que cuando todas las tareas solicitan ejecución simultánea ocurre el *peor instante de carga* de un sistema monorecurso (*instante crítico*). Si el sistema es planificable en este instante, lo es en cualquier otro.

El conjunto de reglas que determina cuál tarea ejecutar en un instante dado, se denomina *algoritmo de planificación*, que puede ser estático o dinámico [5]. En los algoritmos dinámicos la prioridad de cada tarea puede modificarse en tiempo de ejecución (*prioridades dinámicas*) o permanecer invariante (*prioridades fijas*). Los algoritmos de planificación dinámicos por prioridades fijas más utilizados son *Rate Monotonic (RM)* [4] y *Deadline Monotonic (DM)* [6].

En un *STR heterogéneo*, el conjunto de tareas del sistema se divide en dos conjuntos bien definidos:

- *Tareas de Tiempo Real (TTRs)*: tareas periódicas con requerimientos de tiempo real duro.
- *Tareas de No-Tiempo Real (TNTRs)*: tareas del sistema que no cuentan con requerimientos de tiempo real estrictos, pero sí deben cumplir con algún otro requerimiento de calidad de servicio.

En sistemas no-saturados, los métodos de *SS* permiten en un instante dado identificar y adelantar parte del tiempo ocioso [7], al que se denomina *Slack Disponible (SD) del sistema en el instante t (SD(t))*. Este tiempo se puede aprovechar para ejecutar las *TNTRs*, retrasando la ejecución de las *TTRs*, sin comprometer su planificabilidad. Esta técnica permite un mejor aprovechamiento del tiempo ocioso que el uso de servidores [8], aunque su complejidad de implementación es mayor.

Varias implementaciones de *SS* han sido propuestas en [9]-[14], las cuales realizan el cálculo del *SD(t)* en tiempo de ejecución o en tiempo de inicialización, y

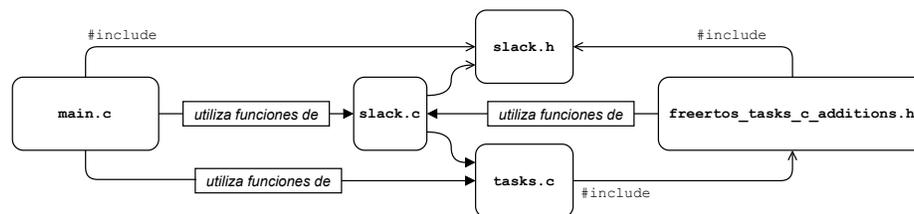
de manera exacta o aproximada. Por otro lado, existen trabajos previos de implementación de estos métodos en un *SOTR*. En [15] se implementa una variante del algoritmo de *SS* aproximado [13] en MaRTE OS<sup>5</sup> [16]. En [17] se presenta una implementación del método de cálculo exacto [14] también sobre MaRTE OS. Una implementación de un método de cálculo aproximado en LejosRT se desarrolla en [18].

## 2. Implementación del Soporte de SS

Se buscó reducir las modificaciones necesarias al núcleo de FreeRTOS en [2], moviendo la implementación de todas las funciones que no requieran utilizar de manera directa variables internas del núcleo a una librería externa.

A partir de la versión 10, FreeRTOS provee una macro<sup>6</sup> para agregar funcionalidad definida por el usuario en el núcleo. Si esta macro es definida con el valor 1, entonces un archivo de nombre `freertos_task_c_additions.h` es incluido al compilar el *kernel*. Las funciones auxiliares que requieren acceder a estructuras de datos internas del núcleo son implementadas de esta manera. Las versiones modificadas de algunas funciones del núcleo también se incluyen en este archivo y se utiliza la opción `wrap` del compilador GCC para reemplazar las originales al enlazar (*linking*) la aplicación con el *kernel*.

La figura 1 presenta como se relacionan los nuevos componentes. El archivo fuente `main.c` contiene el código de la aplicación (creación de las tareas y la lógica de cada una), que invoca tanto funciones de FreeRTOS (archivo `tasks.c`) como de la librería de *SS* (archivos `slack.c` y `slack.h`). Las funciones del núcleo modificadas se encuentran en el archivo `freertos_tasks_c_additions.h`.



**Figura 1.** Esquema de la implementación en FreeRTOS.

Notar que la librería de *SS* y las modificaciones al núcleo son independientes del *hardware*. Por lo tanto, el soporte de *SS* puede utilizarse en cualquier plataforma para la que FreeRTOS este disponible.

A continuación se describen las modificaciones, nuevas funciones y estructuras agregadas al *kernel*.

<sup>5</sup> <https://martel.unican.es>

<sup>6</sup> `configINCLUDE_FREERTOS_TASK_C_ADDITIONS_H` en el archivo `FreeRTOSConfig.h`

## 2.1. Atributos de las tareas

Se considera que las tareas que conforman el sistema se dividen en dos conjuntos:

- *TTRs*: modeladas a partir del modelo propuesto en [4]. Cada tarea genera una serie infinita de instancias, siendo  $j_{k,i}$  la  $k$ -ésima instancia de la *TTR*  $i$ . Cada *TTR* se caracteriza mediante su periodo ( $T_i$ ), su vencimiento relativo ( $D_i \leq T_i$ ), su *peor caso de tiempo de ejecución* ( $C_i$ ), su *peor caso de tiempo de respuesta* ( $R_i$ ), el *SD* en el instante crítico ( $SD_i(0)$ ), el *SD* en el instante actual ( $SD_i(t_c)$ ), el tiempo ejecutado por la instancia actual ( $c_i(t_c)$ ) y un contador de instanciaciones. Notar que  $SD(t_c) = \min SD_i(t_c)$  con  $i = 1, 2, \dots, n$ .
- *TNTRs*: Debido a su heterogeneidad no cuentan con un modelo específico asociado.

Los atributos del modelo de tareas son agrupados en una estructura de datos denominada `SsTCB_t`, que se agrega al *Task Control Block* (*TCB*) de la tarea mediante la funcionalidad *Thread Local Storage* (*TLS*) de FreeRTOS. Esta estructura y otras definiciones se encuentran en el archivo `sLack.h`.

## 2.2. Colas de tareas adicionales

Se agregaron tres nuevas listas de tareas, implementadas mediante el tipo de datos `List_t` de FreeRTOS. Cada elemento de la lista contiene un atributo `xItemValue`, utilizado para ordenar la lista, y un puntero `pvOwner` que hace referencia al *TCB* de la tarea:

- `xSsTaskList`: Reúne todas las *TTR*, sin importar el estado, para simplificar el cálculo del *SD*.
- `xDeadlineTaskList`: Registra los vencimientos absolutos de todas las *TTR*, y se mantiene ordenada por el vencimiento más próximo.
- `xSlackDelayedTaskList`: Contiene las *TNTR* suspendidas por falta de *SD*. Estas tareas continuarán su ejecución cuando  $SD(t_c) > SD_{min}$ .

## 2.3. Modificaciones al núcleo

El cálculo del *SD* se realiza antes del inicio del planificador de FreeRTOS y al finalizar la ejecución de cada instancia de una *TTR*. Además, los contadores de *SD* deben ser actualizados en cada *tick* de reloj. Para lograr esto, se modificó la implementación de las siguientes funciones del núcleo de FreeRTOS:

- `vTaskDelayUntil()`: Bloquea la ejecución de una tarea hasta un instante absoluto, permitiendo implementar tareas periódicas. Se modificó la función para que calcule el *SD* y actualice los atributos correspondientes de la tarea.

- `xTaskIncrementTick()`: Procesa la interrupción de reloj (*tick*). Desde esta función FreeRTOS realiza la activación de las tareas periódicas, controla *timeouts*, etc., y genera de ser necesario un cambio de contexto. Se agregó el control de vencimientos, la actualización de contadores de tiempo de ejecución ( $c_i(t_c)$ ) y de  $SD$ . También se agregó la suspensión de las *TNTR* cuando el  $SD$  sea menor al límite indicado por la aplicación ( $SD(t_c) < SD_{min}$ ).

Desde `vTaskDelayUntil()` se calcula el  $SD$  de la tarea invocando la función `vTaskCalculateSlack()` (implementada en el `slack.c`). Como la tarea ya consumió parte del tiempo de computo del *time slice*, se calcula el  $SD$  en el instante  $t_c + 1$  ( $SD_i(t_c + 1)$ ). Luego, si  $C_i - c_i(t_c) > 0$ , agrega el tiempo ganado a los contadores  $SD_k(t_c)$  de todas las tareas  $k$  de menor prioridad y actualiza  $SD(t_c)$ . Si  $SD(t_c) > SD_{min}$  las *TNTR* bloqueadas son puestas en la cola de tareas listas para ejecutar.

Cuando `xTaskIncrementTick()` mueve una *TTR* a la cola de tareas listas, si esta fue bloqueada por una invocación a `vTaskDelayUntil()` se incrementa su contador de instancias en uno.

Al procesar cada *tick*, `xTaskIncrementTick()` incrementa el contador  $c_i(t_c)$  de la tarea en ejecución. Notar que la ejecución de una instancia en general emplea una fracción de su último *time slice*. Por lo tanto el valor de  $c_i(t_c)$  es, en el peor caso, mayor en un *tick* que el tiempo efectivo de ejecución. Cuando se invoca la función `vTaskDelayUntil()`, se asigna cero a  $c_i(t_c)$ . Por defecto, en FreeRTOS la interrupción de reloj es cada 1 ms.

Además, `xTaskIncrementTick()` resta un *tick* a todos los contadores  $SD_i(t_c)$  si se estuviera ejecutando una *TNTR* o la tarea inactiva (*idle*). Si se estuviera ejecutando una *TTR*, resta los contadores de todas las *TTR* de mayor prioridad. Una vez que los contadores están actualizados, recalcula  $SD(t_c)$ . Luego, si  $SD(t_c) < SD_{min}$  cualquier *TNTR* que estuviera ejecutando o lista para ejecutar, es suspendida y puesta en la cola `xSlackDelayedTaskList`.

Para actualizar los contadores de  $SD$  se utilizan las siguientes funciones auxiliares (implementadas en `slack.c`):

- `vSlackDecrementTasksSlack(i, n)`. Resta  $n$  *ticks* a los contadores de  $SD$  de las *TTR* con prioridad mayor a  $i$ .
- `vSlackDecrementAllTasksSlack(n)`. Resta  $n$  *ticks* a los contadores de  $SD$  de todas las *TTR*.
- `vSlackGainSlack(i, n)`. Suma  $n$  *ticks* a los contadores de  $SD$  de las *TTR* con prioridad menor a  $i$ .
- `vSlackUpdateAvailableSlack()`. Actualiza  $SD(t_c)$ .

El cada *tick* de reloj se verifica si el vencimiento de la primer *TTR* en `xDeadlineTaskList` es mayor al tiempo actual. Por cada tarea que pierda su vencimiento se invoca la función `vApplicationDeadlineMissedHook()`. La función `vTaskDelayUntil()` actualiza el vencimiento absoluto de la tarea y reordena la lista de vencimientos.

Las siguientes funciones se agregan al núcleo, ya que requieren interactuar directamente con la cola de tareas listas para ejecutar:

- `vTaskSlackSuspend()`. Mueve todas las *TNTR* de la cola de tareas listas para ejecutar a la lista de tareas `xSlackDelayedTaskList` (ver sección 2.2).
- `vTaskSlackResume()`. Mueve todas las *TNTR* que se encuentren en la lista `xSlackDelayedTaskList` a la cola de tareas listas para ejecutar.

## 2.4. Interfaz para el desarrollador

Las siguientes funciones forman la interfaz para el desarrollador:

- `vSlackSystemSetup()`: inicializa las colas de tareas adicionales (sección 2.2). Debe invocarse antes de especificar los parámetros de las tareas.
- `vTaskSetParams()`: permite indicar los atributos adicionales de una tarea, tal como su tipo (*TTR* o *TNTR*), período, vencimiento, etc. Si es una *TTR*, la agrega a `xSsTaskList` y su vencimiento inicial ( $D_i$ ) a `xDeadlineTaskList`.
- `vSlackSchedulerSetup()`: realiza la verificación de planificabilidad, el cálculo del peor caso de tiempo de respuesta de las tareas y el cálculo del *SD* en el instante crítico. Debe invocarse antes de iniciar el planificador de FreeRTOS.

Si  $R_i \leq D_i$  para todas las *TTR* el *STR* es planificable. Caso contrario, se invoca la función `vApplicationNotSchedulable()`<sup>7</sup>. La evaluación de planificabilidad no tiene en cuenta el costo del cambio de contexto y se realiza cómo validación del modelo de tareas.

## 2.5. Planificación mediante Slack Stealing

El planificador de FreeRTOS emplea una política apropiativa *First In, First Out (FIFO)* con prioridades [19] y garantiza que siempre ejecutará la tarea de mayor prioridad lista para ejecutar. La cola de tareas listas se implementa como un arreglo de listas. Dadas  $M$  prioridades, la posición cero del arreglo contiene la cola de tareas de menor prioridad y la posición  $M - 1$  la de máxima. Las tareas dentro de cada lista se planifican mediante *FIFO* o *Round Robin (RR)* según la configuración.

El rango de prioridades  $(M, M - n]$  se reserva para las *TNTR* y sólo estarán en la cola de tareas listas si  $SD(t_c) > SD_{min}$ . De esta manera, FreeRTOS les da preferencia hasta que las mismas finalicen, se bloqueen o agoten el *SD*. El rango de prioridades  $(M - n, 1]$  es para las *TTRs*, con sólo una *TTR* por nivel. Las tareas con prioridad cero son ejecutadas cuando no existan *TTRs* o *TNTRs* listas para ejecutar y se excluyen del cálculo de *SD*. Notar que un  $SD(t_c) < 1$  no es detectado. Por lo tanto, las *TNTR* pendientes de ejecución no son puestas inmediatamente en la cola de tareas listas y una *TTR* u otra tarea pueden ejecutarse hasta el siguiente *tick*.

La Figura 2 presenta un ejemplo de organización de la cola de tareas listas. Si  $SD(t_c) = 0$  las *TNTRs* no estarían presentes.

<sup>7</sup> Que debe implementar el desarrollador.

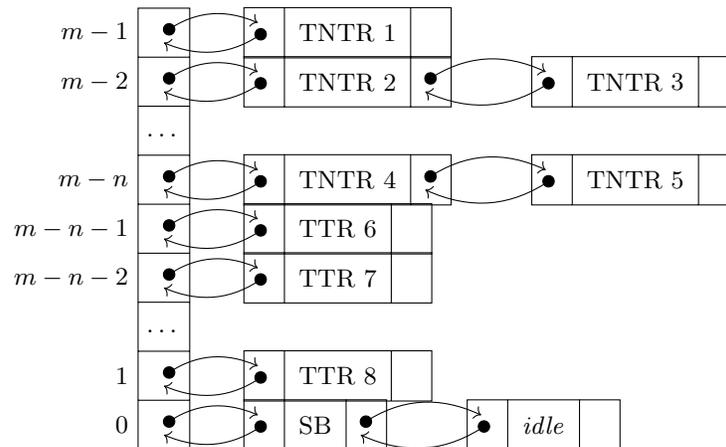


Figura 2. Cola de tareas listas organizada para para planificación mediante SS.

### 3. Resultados Experimentales

#### 3.1. Ejemplo de Ejecución

Se presenta una traza de ejemplo sobre en una placa *mbed* LPC1768 y FreeRTOS v10.1.3. El sistema consta de 4 *TTR*,  $\{(1, 3, 3), (1, 4, 4), (1, 6, 6), (1, 12, 12)\}$  con tiempos expresados en segundos, y dos *TNTR*, TA1 (de máxima prioridad) y TA2, con un tiempo de ejecución aleatorio no mayor a los 2000 ms.

La figura 3 muestra la traza generada por Tracealyzer<sup>8</sup> v3.1.2. Se observa cómo la ejecución de la primer instancia de TA2 es desalojada por TA1 y luego al agotarse el *SD* (en este caso  $SD_{min} = 1$ ). FreeRTOS ejecuta entonces la *TTR* T1, que como ejecuta un tiempo menor a su peor caso de ejecución genera *SD*. Esto permite ejecutar nuevamente a TA2. Se puede ver también cómo en el instante  $t = 12$  todas las *TTR* son desplazadas por la ejecución de otra instancia de TA2. En la salida a través del puerto serial cada línea indica la tarea en ejecución, si inicia (S) o finaliza (E), el valor del reloj, el *SD* del sistema, el *SD* de las *TTR* T1 a T4 y el tiempo ejecutado (en *ticks*) por la tarea.

#### 3.2. Costo del Cambio de Contexto

A continuación se evalúa el *Costo Computacional (CC)* temporal del cambio de contexto, al ejecutar `vTaskDelayUntil()`. Se contabilizó el número de ciclos de CPU para las primeras 30 ejecuciones de cada *TTR*. Se empleó FreeRTOS v10.3.1, sobre una placa *mbed* LPC1768, con el método de *SS* exacto [14]. La duración del *tick* se configuró en 1 ms y se empleó el mecanismo de cambio de contexto optimizado para Cortex-M3 de FreeRTOS. El número de ciclos de CPU se obtuvo con el contador *Clock Cycle Counter (CYCCNT)*.

<sup>8</sup> <https://percepio.com/tracealyzer/>

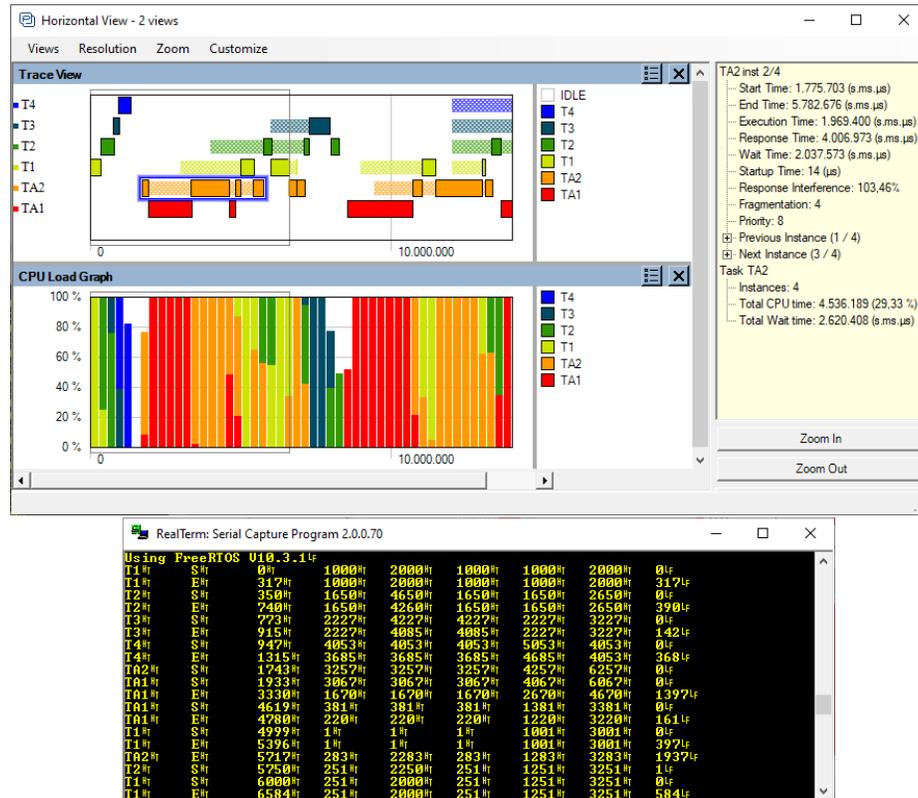
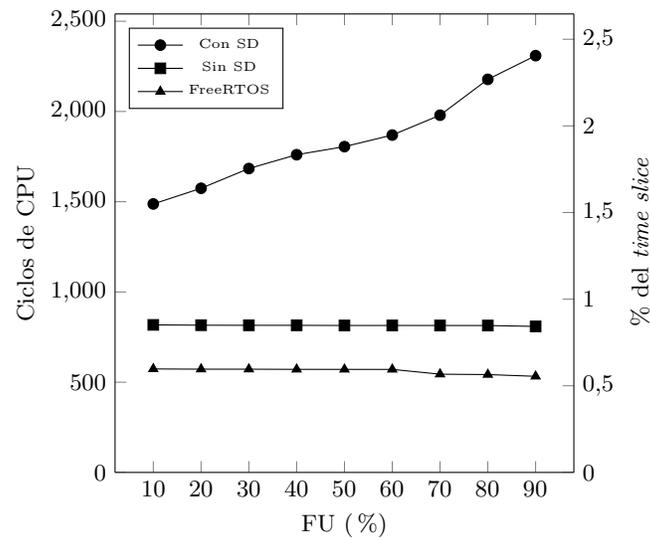


Figura 3. Traza de ejecución mediante Tracealyzer y salida serial.

Se generaron 1000 *STR* de 10 tareas, para cada *Factor de Utilización (FU)* del 10% al 90%, en intervalos de 10%. Los parámetros *T* y *C* de las *TTR* se distribuyeron uniformemente, entre 25 y 1000 *ticks*. Se evaluó el *CC* temporal del cambio de contexto de FreeRTOS sin modificaciones y se comparó con las siguientes configuraciones:

- Sólo actualizando los contadores de *SD* de cada tarea con el valor calculado en el instante crítico, para determinar el costo introducido independientemente del método de *SS*.
- Realizando el cálculo del *SD* en tiempo de ejecución.

La figura 4 presenta los resultados. El costo del cambio de contexto sin el cálculo del *SD* se mantiene constante y dentro del mismo orden que FreeRTOS sin modificaciones. Por otro lado, al realizar el cálculo del *SD* el costo crece con el *FU* del sistema, de manera similar al presentado en [14]. Dado que la interrupción de reloj ocurre cada 1 ms, el *time slice* abarca  $\approx 96000$  ciclos de CPU. Por lo tanto el *CC* temporal del cambio de contexto es, en el peor caso promedio,  $\approx 2,5\%$  del *time slice*.



**Figura 4.** Costo promedio en ciclos de CPU del cambio de contexto al finalizar la instancia de una TTR.

## 4. Conclusiones

Una administración eficiente del tiempo ocioso puede facilitar el diseño e implementación de soluciones que satisfagan los requerimientos funcionales y temporales de sistemas con conjuntos heterogéneos de tareas. Aunque el uso de un *SOTR* facilita cumplir con los requerimientos temporales en muchos casos, no siempre posibilita planificar conjuntos heterogéneos de tareas de manera eficiente, dadas las pocas opciones de políticas de planificación que ofrecen. En este trabajo se presentó una implementación de administración del tiempo ocioso exacto, en un *SOTR* de amplio uso. El diseño propuesto favorece la modularidad y requiere de mínimas modificaciones al *SOTR*.

El código fuente está disponible para su descarga en GitHub<sup>9</sup>, con ejemplos de ejecución en las placas *mbed* LPC1768, FRDM-K64F y EDU-CIAA-NXP.

En trabajos futuros se evaluará el costo computacional real, tanto temporal como espacial, de diversos métodos de *SS*, y el uso del *SD* para cumplir requerimientos como tolerancia a fallas, ahorro de energía, etc.

## Referencias

- [1] A. Burns y R. I. Davis, «A Survey of Research into Mixed Criticality Systems», *ACM Comput. Surv.*, vol. 50, n.º 6, 82:1-82:37, 2017.

<sup>9</sup> <https://github.com/unpsjb-rtsg/slack-freertos>

- [2] F. Páez, J. M. Urriza, R. Cayssials y J. D. Orozco, «Métodos de Slack Stealing en FreeRTOS», en *44 Jornadas Argentinas de Informática (JAIIO)*, JAIIO, ed., SADIO, 2015.
- [3] J. A. Stankovic, «Misconceptions About Real-Time Computing: A Serious Problem for Next-Generation Systems», *Computer*, vol. 21, n.º 10, págs. 10-19, 1988.
- [4] C. L. Liu y J. W. Layland, «Scheduling Algorithms for Multiprogramming in a Hard-Real-Time Environment», *J. ACM*, vol. 20, n.º 1, págs. 46-61, 1973.
- [5] A. Burns, «Scheduling hard real-time systems: a review», *Software Engineering Journal*, vol. 6, n.º 3, págs. 116-128, 1991.
- [6] J. Y. Leung y J. Whitehead, «On the complexity of fixed-priority scheduling of periodic, real-time tasks», *Perform. Eval.*, vol. 2, n.º 4, págs. 237-250, 1982.
- [7] S. Ramos-Thuel y J. P. Lehoczky, «On-line scheduling of hard deadline aperiodic tasks in fixed-priority systems», en *Proceedings of the Real-Time Systems Symposium. Raleigh-Durham, NC, December 1993*, 1993, págs. 160-171.
- [8] T.-S. Tia, J. W.-S. Liu y M. Shankar, «Algorithms and Optimality of Scheduling Soft Aperiodic Requests in Fixed-priority Preemptive Systems», *Real-Time Syst.*, vol. 10, n.º 1, págs. 23-43, 1996.
- [9] R. I. Davis, K. Tindell y A. Burns, «Scheduling slack time in fixed priority preemptive systems», en *Proceedings of the Real-Time Systems Symposium. Raleigh-Durham, NC, December 1993*, 1993, págs. 222-231.
- [10] R. M. Santos, J. M. Urriza, J. Santos y J. Orozco, «New methods for redistributing slack time in real-time systems: applications and comparative evaluations», *Journal of Systems and Software*, vol. 69, n.º 1-2, págs. 115-128, 2004.
- [11] R. C. José Manuel Urriza Javier D. Orozco, «Fast Slack Stealing methods for Embedded Real Time Systems», 2005.
- [12] C. Lin y S. A. Brandt, «Improving Soft Real-Time Performance through Better Slack Reclaiming», en *Proceedings of the 26th IEEE Real-Time Systems Symposium (RTSS 2005), 6-8 December 2005, Miami, FL, USA*, IEEE Computer Society, 2005, págs. 410-421.
- [13] R. I. Davis, *Approximate slack stealing algorithms for fixed priority pre-emptive systems*. University of York, Department of Computer Science, 1993.
- [14] J. Urriza, F. Paez, R. Cayssials, J. Orozco y L. Schorb, «Low cost slack stealing method for RM/DM», *International Review on Computers and Software*, vol. 5, n.º 6, págs. 660-667, 2010.
- [15] A. R. E. Minguet, «.E<sup>x</sup>tensiones al Lenguaje Ada y a los Servicios POSIX para Planificación en Sistemas de Tiempo Real Estricto», Tesis doct., Universidad Politécnica de Valencia, 2003.
- [16] M. A. Rivas y M. G. Harbour, «MaRTE OS: An Ada Kernel for Real-Time Embedded Applications», en *Reliable Software Technologies: Ada Europe 2001, 6th Ade-Europe International Conference Leuven, Belgium, May 14-18, 2001, Proceedings*, D. Craeynest y A. Strohmeier, eds., vol. 2043, Springer, 2001, págs. 305-316.
- [17] L. A. Díaz, F. E. Páez, J. M. Urriza, J. D. Orozco y R. Cayssials, «Implementación de un Método de Slack Stealing en el Kernel de MaRTE OS», en *Proceeding XLIII Jornadas Argentinas de Informática e Investigación Operativa (43 JAIIO) - III Argentine Symposium on Industrial Informatics (SII)*, 2014, págs. 13-24.
- [18] S. Midonnet, D. Masson y R. Lassalle, «Slack-Time Computation for Temporal Robustness in Embedded Systems», *Embedded Systems Letters*, vol. 2, n.º 4, págs. 119-122, 2010.
- [19] C. Svec, «FreeRTOS», en *The Architecture of Open Source Applications, Volume II: Structure, Scale, and a Few More Fearless Hacks*, lulu.com, 2012.

# Análisis Comparativo de Implementaciones HLS de Filtro Sobel en SoC

Roberto Millon<sup>1</sup> , Enzo Rucci<sup>2</sup> , and Emmanuel Frati<sup>1</sup> 

<sup>1</sup> Departamento de Ciencias Básicas y Tecnológicas, UNdeC  
Chilecito (5360), La Rioja, Argentina  
{rmillon,efrati}@undec.edu.ar

<sup>2</sup> III-LIDI, Facultad de Informática, UNLP – CeAs CICPBA.  
La Plata (1900), Bs As, Argentina  
erucci@lidi.info.unlp.edu.ar

**Resumen** El filtro Sobel se ha consolidado como una alternativa popular para detectar bordes en imágenes por su simpleza y baja sensibilidad al ruido. Como consecuencia, Sobel ha sido ampliamente estudiado, existiendo numerosas variantes que ofrecen distintos compromisos entre costo computacional y calidad de filtrado. En este trabajo se sintetizaron tres variantes del filtro Sobel en alto nivel sobre una plataforma SoC y se analizó consumo de recursos, tiempo de procesamiento y calidad de filtrado de cada una de ellas. Los resultados obtenidos muestran que la velocidad de cómputo no presenta variaciones significativas entre las tres variantes mientras que la demanda de recursos sí lo hace, aunque no constituye una restricción de diseño. Por último, en cuanto a la calidad del filtrado, si bien las imágenes producidas son similares, presentan diferencias en los bordes detectados.

**Keywords:** Detección de bordes · Sobel · SoC · Zybo · HLS

## 1. Introducción

Una imagen digital se puede definir como un conjunto de puntos o píxeles dispuestos en una matriz de dos dimensiones ( $x$  e  $y$ ). El procesamiento digital de imágenes consiste en la aplicación de técnicas y algoritmos computacionales complejos sobre una imagen para extraer información de ella o mejorar sus características [15]. Entre estos algoritmos, se pueden mencionar los relacionados a detección de bordes o contornos, ampliamente utilizados como etapa previa a otros algoritmos de procesamiento [12]. Las técnicas de detección de bordes identifican zonas en la imagen donde los píxeles presentan cambios abruptos en intensidad o niveles de grises, permitiendo segmentar una imagen en regiones discontinuas [3]. Como resultado, se reduce significativamente la cantidad de datos en la imagen sin alterar sus propiedades estructurales [16].

Al momento de detectar contornos, no existe un único método. Los operadores Laplacianos (o de segundo orden) permiten detectar bordes en cualquier orientación pero tienen una alta tasa de detección de falsos positivos y son muy

2 Roberto Millon , Enzo Rucci , and Emmanuel Frati 

sensibles al ruido. Por otro lado, se encuentran los algoritmos basados en gradientes (o de primer orden), que se pueden considerar menos sensibles al ruido respecto a los operadores Laplacianos. Sin embargo, no detectan bordes diagonales y generan contornos muy gruesos. Dentro de los algoritmos basados en gradiente, se destaca el operador Sobel por tener mejor inmunidad al ruido respecto a otras alternativas como el operador Roberts-Cross o el Prewitt [16]. Como consecuencia, el algoritmo Sobel ha sido ampliamente estudiado, existiendo numerosas variantes que ofrecen distintos compromisos entre costo computacional y calidad de filtrado.

Respecto a las plataformas hardware elegidas para procesamiento de imágenes, las FPGAs se han destacado por ser dispositivos de bajo consumo energético y alta productividad [6]. La disponibilidad de grandes bancos de memorias internas en estas arquitecturas permiten accesos en forma paralela para ejecutar funciones en pocos ciclos de reloj, a diferencia de otras tecnologías como los procesadores convencionales que operan en forma secuencial y requieren una gran cantidad de ciclos de reloj para resolver las tareas [5].

En este trabajo se presenta un análisis comparativo de tres variantes del filtro Sobel sintetizadas en lenguaje de alto nivel para procesar imágenes en una plataforma System-on-Chip (SoC) ZYBO. El análisis realizado considera uso de recursos, tiempo de procesamiento y calidad del filtrado de cada variante Sobel, y puede resultar útil al momento de tener que elegir una de ellas.

El resto del documento se organiza de la forma siguiente. En la Sección 2 se presentan los antecedentes del presente trabajo mientras que en la Sección 3 se describe la implementación propuesta. A continuación, en la Sección 4 se presentan los resultados experimentales y, finalmente, la Sección 5 resume las conclusiones y posibles trabajos futuros.

## 2. Antecedentes

### 2.1. Filtro de Detección de Bordes Sobel

El filtro Sobel es un método de detección de contornos que se aplica a imágenes en escala de grises [12] y que detecta sus componentes de alta frecuencia mediante la derivada primera. Un componente de alta frecuencia se genera por cambios abruptos de intensidad en píxeles contiguos de la imagen, los cuales se corresponden con un borde. El vector gradiente detecta esos cambios por medio de las derivadas parciales en los ejes horizontal y vertical. La Ecuación 1 define al vector gradiente por sus componentes horizontal  $G_x$  y vertical  $G_y$  [5].

$$\nabla f = \begin{pmatrix} G_x \\ G_y \end{pmatrix} \quad (1)$$

El módulo o magnitud del vector gradiente se denomina fuerza de borde y expresa la tasa de cambio por unidad de distancia. De este modo, una mayor variación de intensidad en píxeles vecinos genera un aumento en la magnitud del gradiente que se computa como un borde. La magnitud del gradiente se define por la Ecuación 2.

$$|\nabla f| = \sqrt{G_x^2 + G_y^2} \quad (2)$$

Las implementaciones del filtro Sobel emplean dos máscaras de convolución  $M_x$  y  $M_y$  para obtener las componentes horizontal y vertical del vector gradiente, respectivamente. Las máscaras de convolución se muestran en la Fig. 1.

$$M_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad M_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Figura 1: Máscaras de convolución Sobel

Desde un punto de vista matemático, las componentes del gradiente se obtienen de la multiplicación de cada máscara con la imagen. El proceso se realiza de izquierda a derecha y de arriba hacia abajo hasta recorrer la imagen completa. En la Fig. 2 se puede observar el proceso de convolución entre una imagen y una máscara del filtro Sobel.

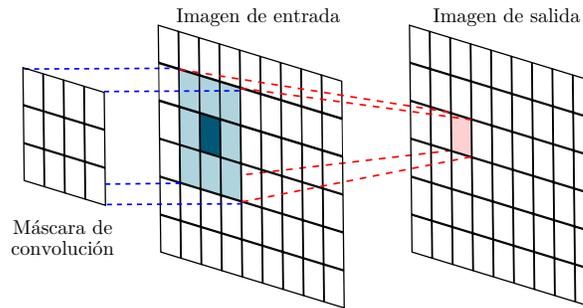


Figura 2: Aplicación del filtro Sobel a una imagen

### Variantes de Sobel.

Existen distintas alternativas del filtro Sobel que tienen por objetivo variar el costo computacional y/o la calidad de imagen obtenida. Un método popular para disminuir la carga computacional consiste en aproximar la magnitud del gradiente por la suma de los valores absolutos de sus componentes, evitando así recurrir a operaciones en punto flotante. Mediante esta alternativa, varios trabajos han logrado reducir el tiempo de procesamiento [5, 14, 18]. En la Ecuación 3 se presenta la fórmula de aproximación para el cálculo de la magnitud del vector gradiente.

$$|\nabla f| = |G_x| + |G_y| \quad (3)$$

Por otra parte, también existen implementaciones arbitrarias de Sobel, como la usada por el editor de imágenes GIMP<sup>3</sup> o la librería de visión por computadora OpenCV<sup>4</sup>. GIMP computa la magnitud del gradiente empleando la fórmula

<sup>3</sup> GNU Image Manipulation Program (GIMP). <https://www.gimp.org/>

<sup>4</sup> Open Source Computer Vision Library (OpenCV). <https://opencv.org/>

4 Roberto Millon , Enzo Rucci , and Emmanuel Frati 

convencional (expresada en la Ecuación 2) pero además dividiendo su resultado por un número real fijo (5.66). Por su parte, OpenCV emplea un enfoque similar al de GIMP, sólo que usando la fórmula de aproximación (expresada en la Ecuación 3) y dividiendo por un número entero fijo (2).

## 2.2. Síntesis de Alto Nivel en FPGAs

Hasta principios de siglo, el modelo de programación en FPGA se centraba en descripciones hardware a nivel de transferencia entre registros (RTL). Las implementaciones en RTL requieren un gran conocimiento de la arquitectura de esta tecnología, además de elevados tiempos de desarrollo y pruebas. Esto limitaba el uso de las FPGAs al no adecuarse a los acotados tiempos de mercado [1].

Numerosas propuestas académicas y comerciales [11], entre ellas las herramientas de síntesis de alto nivel (HLS) de Xilinx, han elevado el nivel de abstracción en el diseño de circuitos al utilizar lenguajes como C, C++, OpenCL, entre otros. Esto permite acelerar los tiempos de desarrollo y reducir las diferencias en los modelos de programación entre procesadores y FPGAs. No obstante, resulta necesario el uso de directivas de optimización específicas para lograr buenas prestaciones en los diseños HLS [17].

Vivado HLS es la herramienta de Xilinx para síntesis de alto nivel en FPGA, la cual transforma un diseño escrito en C, C++ o System C en una implementación RTL [2]. Los diseños RTL son exportadas como bloques de propiedad intelectual (IP) para ser utilizados por otras herramientas de Xilinx. Vivado Design Suite (VDS) emplea un entorno gráfico para utilizar los bloques IP.

## 2.3. Estado del Arte

En el ámbito de las FPGAs, existen numerosas implementaciones HLS del operador Sobel. En [10] se analizan distintas técnicas de optimización para síntesis de alto nivel, eligiendo al algoritmo Sobel como caso de estudio. Al igual que el trabajo anterior, en [7] se usa a Sobel como caso de estudio para comparar y evaluar metodologías de diseño en FPGAs. A diferencia de los anteriores, en [13] se analiza el consumo energético de una implementación Sobel en una plataforma SoC. Por último, varios trabajos han desarrollado soluciones del filtro Sobel para detectar bordes en tiempo real usando diferentes tecnologías y plataformas [4, 8, 19–21]. Sin embargo, hasta donde llega el conocimiento de los autores, no existen trabajos que realicen un análisis comparativo de distintas variantes de este operador en cuanto a costo computacional y calidad del filtrado.

## 3. Implementación

Para el análisis comparativo en el presente trabajo se desarrollaron tres versiones del operador Sobel, las cuales toman como base la implementación [9]. Todas las versiones tienen en común dos estructuras de memoria denominadas *buffer de línea* y *ventana deslizante* [20]. Los buffers de línea son arreglos

unidimensionales que almacenan filas completas de la imagen, con el objeto de mantener el contexto necesario para procesar cada píxel. Estas estructuras de memoria permiten procesar una imagen completa sin necesidad de almacenar todos sus píxeles a la vez, constituyendo un ahorro considerable de memoria. El operador Sobel utiliza tres buffers de línea, ya que se requiere el mismo número de filas al mismo tiempo para el filtrado.

Por su parte, la ventana deslizante es un arreglo bidimensional que contiene el conjunto de píxeles a los cuales se le aplicará las máscaras de convolución. La ventana deslizante se desplaza de izquierda a derecha por los buffers de línea y tiene dimensión  $3 \times 3$ . En la Fig. 3 se puede observar la aplicación de la convolución en Sobel, destacándose en color amarillo los buffer de línea y en color rojo la ventana deslizante.

Las diferencias entre las variantes del operador Sobel se encuentran en la fórmula para computar la magnitud del gradiente. Se distinguen tres versiones:

1. *RMS*: utiliza la fórmula convencional definida en la Ecuación 2.
2. *ABS*: utiliza la fórmula aproximada definida en la Ecuación 3.
3. *GIMP*: utiliza la fórmula convencional definida en la Ecuación 2 pero dividiendo su resultado por un número real fijo (5.66).

Independientemente de la variante Sobel utilizada, es necesario incorporar otro bloque al sistema que permita realizar pruebas reales en la plataforma SoC ZYBO. El sistema completo está formado por un bloque DMA que accede a la memoria microSD para la lectura de imágenes en formato BMP y envía el flujo de píxeles al núcleo de procesamiento de imagen. Luego del filtrado, el flujo de píxeles es devuelto al DMA para su posterior almacenamiento en memoria microSD. La configuración y habilitación de cada módulo del sistema es realizada por el procesador. Todos los módulos fueron integrados usando la herramienta Vivado 2019.1. En la Fig. 4 se observa un diagrama en bloques del sistema completo.

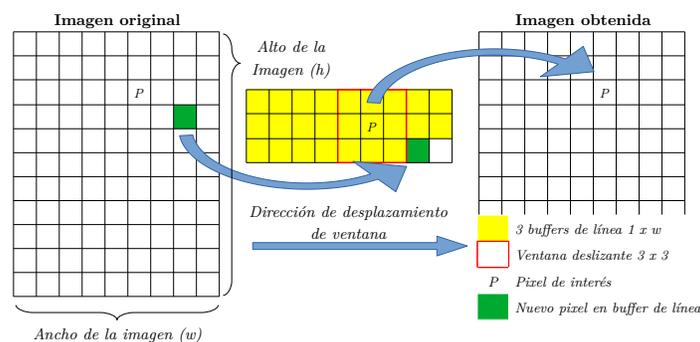


Figura 3: Aplicación de convolución

6 Roberto Millon , Enzo Rucci , and Emmanuel Frati 

El núcleo de procesamiento está formado por tres bloques sintetizados y validados en alto nivel con la herramienta Vivado HLS versión 2019.1 de Xilinx. El primer bloque (RGB2Gray) convierte imágenes a color en escala de grises al promediar el valor de los tres canales (RGB) en cada píxel. El segundo bloque (FiltroSobel) es el responsable de aplicar el operador Sobel para detectar contornos en alguna de sus tres variantes. Por último, se incorpora un bloque adicional (U8toU32) que concatena cuatro caracteres de 8 bits con el fin de sincronizar las comunicaciones con el DMA (ya que opera con palabras de 32 bits).

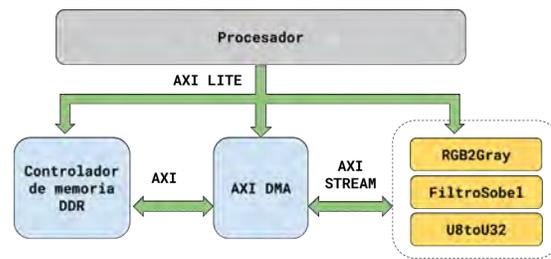


Figura 4: Sistema implementado

Resulta importante mencionar que se utilizaron las directivas de optimización para sintetizar cada bloque del núcleo de procesamiento. En particular, se usó HLS INTERFACE para emplear los estándares AXI Stream y AXI Lite en las comunicaciones de las imágenes y configuración de parámetros. También se utilizó HLS ARRAY PARTITION para mapear arreglos en múltiples recursos de memoria y permitir accesos concurrentes (buffers de línea, ventana deslizante y máscaras). Por último, se incluyó la directiva HLS PIPELINE que permite aumentar la productividad al ejecutar en paralelo funciones y operaciones en bucles [1].

## 4. Resultados Experimentales

### 4.1. Diseño experimental

Las pruebas se realizaron sobre una plataforma ZYBO compuesta por un SoC ZYNQ-7000 de Xilinx. El SoC está integrado por un procesador de doble núcleo ARM Cortex-A9 y una FPGA XC7Z010-1-CLG400C. El sistema completo se diseñó en un entorno gráfico dentro del software VDS. Para la implementación de cada variante, se reemplazó únicamente el bloque IP del kernel Sobel por alguna de sus alternativas (RMS, ABS o GIMP). En cada una de ellas se registró el uso de recursos informado por la herramienta VDS.

Para realizar las pruebas se desarrolló una aplicación *test* con la herramienta XSDK 2019.1 de Xilinx. Además, se seleccionaron tres imágenes de repositorios públicos para realizar las pruebas en cada variante Sobel: *Peppers* de  $512 \times 512^5$ , *Kodim21* de  $768 \times 512^6$  y *Fox* de  $1280 \times 853^7$ . En las pruebas realizadas, se midieron los tiempos de ejecución mediante la librería `xtime.l.h`, informando el promedio de realizar 10 repeticiones de cada una de ellas.

<sup>5</sup> <https://www.hlevkin.com/06testimages.htm>

<sup>6</sup> <https://github.com/lemire/kodakimagecollection>

<sup>7</sup> <https://pixabay.com/es/photos/fuchs-mundo-animal-animales-salvajes-5303221/>

#### 4.2. Resultados del filtrado

En la Fig. 5 se muestran las imágenes originales que fueron seleccionadas para las pruebas junto a las imágenes producidas por las distintas implementaciones desarrolladas. Como se puede apreciar, las tres variantes consiguen detectar los bordes relevantes de cada imagen. Sin embargo, al menos para este conjunto de imágenes, RMS y ABS producen resultados similares mientras que en GIMP los bordes detectados son más débiles.

#### 4.3. Resultados de uso de recursos y rendimiento

La Tabla 1 presenta el uso de recursos y los tiempos de procesamiento obtenidos para las tres variantes Sobel desarrolladas. Los valores de las columnas

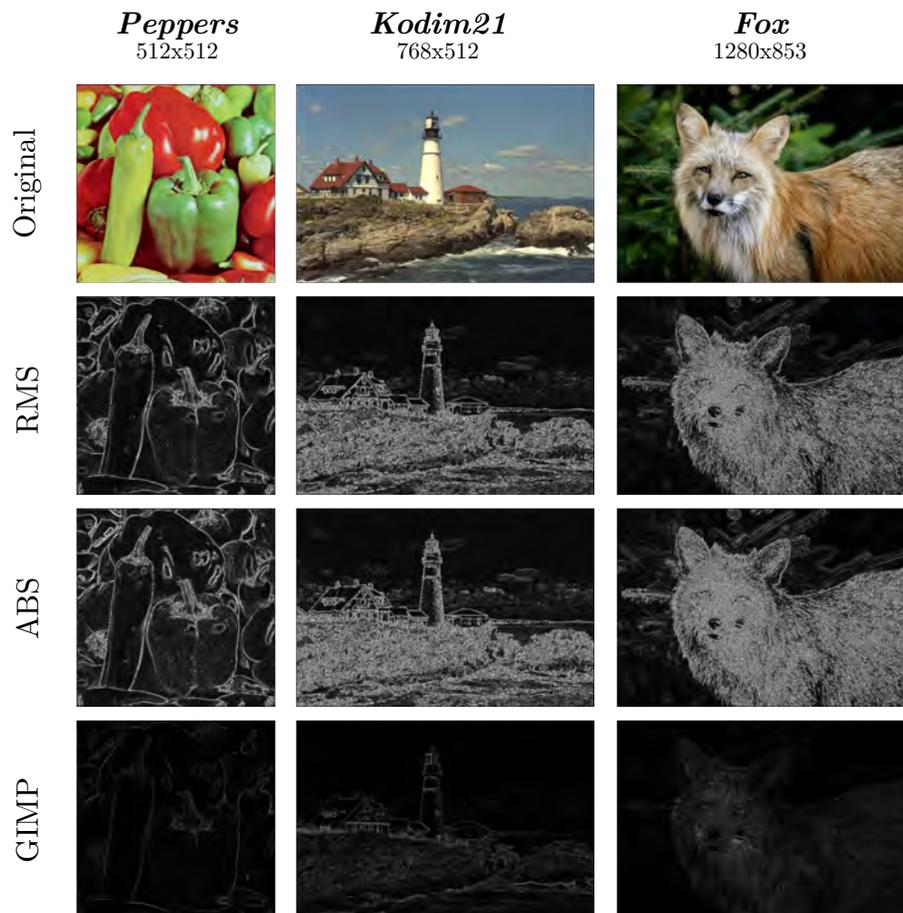


Figura 5: Resultados de aplicar variaciones del filtro Sobel a las imágenes de prueba.

Tabla 1: Uso de recursos y tiempo de procesamiento de las variantes de Sobel implementadas con optimizaciones

Variante	Uso de Recursos (% del total)					Tiempos de procesamiento (ms)					
	Peppers		Kodim21		Fox						
	S.LUTs	S.Registers	F7 Muxes	BRAM	DSPs	Sobel	Resto	Sobel	Resto	Sobel	Resto
RMS	16,3%	9,1%	< 0,1%	1,6%	2,5%	40,96	293,98	40,97	388,59	68,26	984,65
ABS	4,4%	2,4%	< 0,1%	1,6%	0%	40,96	293,98	40,96	388,59	68,26	986,88
GIMP	34,4%	16,3%	< 0,1%	1,6%	2,5%	40,96	294,1	40,97	389,3	68,26	988,34

S.LUTs, S.Registers, F7 Muxes, BRAM y DSPs hacen referencia a los porcentajes de tablas de búsqueda, registros, multiplexores, bloques de memoria RAM y bloques DSPs utilizados, respectivamente.

En cuanto al uso de recursos, se puede notar que ninguna de las tres presenta un consumo excesivo de ellos. Aun así, ABS y GIMP se destacan como las opciones que menos y más recursos requieren, respectivamente. En particular, la versión GIMP demanda  $7.8\times$  más de LUTs y  $6.8\times$  más de Register en comparación a ABS. Por su parte, RMS se presenta como una versión intermedia, consumiendo  $3.7\times/3.8\times$  más que ABS pero  $0.47\times/0.56\times$  menos que GIMP en LUTs/Registers. No se observan diferencias significativas para F7 Muxes, BRAM y DSPs.

Se observa que el tiempo de procesamiento total aumenta a medida que se incrementa el tamaño de la imagen de entrada. Sin embargo, y contrariamente a lo esperado, se puede notar que todas las variantes de Sobel implementadas tuvieron tiempos de ejecución similares independientemente del método empleado para computar la magnitud del gradiente.

Luego de profundizar en el análisis del código y sus resultados, se encontró a dos de las directivas de optimización como las responsables del tiempo de ejecución constante de las tres variantes del filtro para una determinada imagen. La combinación del solapamiento de las diferentes etapas de las operaciones (HLS PIPELINE) con el acceso concurrente a las estructuras de datos (HLS ARRAY PARTITION) conducen a que la diferencia de tiempo en el cómputo de la magnitud del gradiente no tenga peso suficiente como para impactar en el tiempo total de Sobel. Este aspecto se puede apreciar en la Tabla 2, la cual presenta los resultados de uso de recursos y tiempo de procesamiento para las tres variantes de Sobel desarrolladas, pero omitiendo en este caso el uso de las directivas de optimización mencionadas.

Se puede notar que las versiones no optimizadas sí reportan diferencias significativas en los tiempos de procesamiento (Sobel), mientras que el tiempo en el resto del sistema se mantiene estable. La versión ABS resulta ser la más eficiente en este aspecto, ejecutándose hasta  $1.53\times$  y  $1.95\times$  más rápido que RMS y GIMP, respectivamente. Por su parte, RMS se ubica en una posición intermedia, siendo hasta  $1.27\times$  más veloz que GIMP.

Tabla 2: Uso de recursos y tiempo de procesamiento de las variantes de Sobel implementadas sin optimizaciones

Variante	Uso de Recursos (% del total)					Tiempos de procesamiento (ms)					
	S.LUTs	S.Registers	F7 Muxes	BRAM	DSPs	Peppers		Kodim21		Fox	
						Sobel	Resto	Sobel	Resto	Sobel	Resto
RMS	15 %	8,5 %	0,1 %	3,3 %	6,3 %	320,19	293,96	461,34	387,42	1240,27	927,9
ABS	3,5 %	2 %	0,1 %	3,3 %	3,7 %	216,97	294,12	305,89	388,2	806,6	928,2
GIMP	33 %	15 %	0,1 %	3,3 %	6,3 %	400,19	295,27	581,81	389,01	1576,36	930,04

Omitir las directivas de optimización también impacta en el uso de recursos aunque no de forma significativa, ya que se mantienen las tendencias del caso optimizado (GIMP consume más recursos que el resto y ABS es la que menos demanda). Todas las versiones tienen una leve reducción en el consumo de LUTS y Register, pero un pequeño incremento en BRAM y DSPSs. No se observan cambios en Muxes.

## 5. Conclusiones y Trabajo Futuro

El operador Sobel resulta ser un método popular para la detección de bordes en imágenes, existiendo numerosas variantes que ofrecen distintos compromisos entre costo computacional y calidad de filtrado. En este trabajo se presentaron tres variantes de un filtro Sobel sintetizados en lenguaje de alto nivel sobre una plataforma SoC y se analizaron sus prestaciones considerando tiempo de procesamiento, uso de recursos y calidad del filtrado. Entre los resultados experimentales más destacados se pueden mencionar:

- Modificar la fórmula para computar la magnitud del vector gradiente no aumentó la velocidad de procesamiento en los diseños optimizados. Las 3 variantes presentaron tiempos constantes para cada imagen de prueba.
- Sí se encontraron diferencias significativas al utilizar una variante u otra en cuanto al uso de recursos. En ese aspecto, ABS y GIMP se destacaron como las opciones que menos y más recursos consumen. Más allá de eso, ninguna tiene un consumo que implique una restricción de diseño para este sistema.
- Las tres versiones produjeron imágenes de salida similares pero no idénticas para cada entrada, por lo que su elección dependerá del propósito de su uso.

Como trabajo futuro, se espera desarrollar las tres variantes del algoritmo en bajo nivel (HDL) y replicar el análisis realizado a fin de verificar la tendencia de los resultados obtenidos.

## Referencias

1. Introduction to FPGA design with vivado high-level synthesis (UG998) <https://bit.ly/31ICdK>

10 Roberto Millon , Enzo Rucci , and Emmanuel Frati 

2. Vivado Design Suite User Guide: High-Level Synthesis <https://bit.ly/2PHyHuB>
3. Acharya, T., Ray, A.K.: Image Processing - Principles and Applications. Wiley-Interscience, USA (2005)
4. Ben Amara, A., Pissaloux, E., Atri, M.: Sobel edge detection system design and integration on an FPGA based HD video streaming architecture. IEEE, Hammamet, Tunisia (Dec 2016). <https://doi.org/10.1109/IDT.2016.7843033>
5. Chaple, G., Daruwala, R.D.: Design of Sobel operator based image edge detection algorithm on FPGA. IEEE, Melmaruvathur, India (Apr 2014). <https://doi.org/10.1109/ICCSP.2014.6949951>
6. Daoud, L., Zydek, D., Selvaraj, H.: A Survey of High Level Synthesis Languages, Tools, and Compilers for Reconfigurable High Performance Computing. vol. 240. Cham (2014). [https://doi.org/10.1007/978-3-319-01857-7\\_7](https://doi.org/10.1007/978-3-319-01857-7_7)
7. Hong, N., Belleudy, C., Pham, T.: Performance and evaluation sobel edge detection on various methodologies . <https://doi.org/10.12720/ijeee.2.1.15-20>
8. Kowalczyk, M., Przewlocka, D., Kryjak, T.: Real-time implementation of contextual image processing operations for 4k video stream in zynq UltraScale+ MPSoC. <https://doi.org/10.1109/DASIP.2018.8597105>
9. Millon, R., Frati, E., Rucci, E.: Implementación de Filtro de Detección de Bordes Sobel en SoC usando Síntesis de Alto Nivel. In: Actas del Congreso Argentino de Sistemas Embebidos (CASE 2020). pp. 73–75 (2020)
10. Monson, J., Wirthlin, M., Hutchings, B.L.: Optimization techniques for a high level synthesis implementation of the sobel filter. <https://doi.org/10.1109/ReConFig.2013.6732315>, ISSN: 2325-6532
11. Nane, R., Sima, V.M., Pilato, C., Choi, J., Fort, B., Canis, A., Chen, Y.T., Hsiao, H., Brown, S., Ferrandi, F., Anderson, J., Bertels, K.: A survey and evaluation of FPGA high-level synthesis tools . <https://doi.org/10.1109/TCAD.2015.2513673>
12. Nausheen, N., Seal, A., Khanna, P., Halder, S.: A FPGA based implementation of Sobel edge detection (Feb 2018). <https://doi.org/10.1016/j.micpro.2017.10.011>
13. Nguyen, H.T., Belleudy, C., Pham, T.V.: Power evaluation of sobel filter on xilinx platform. <https://doi.org/10.1109/FTFC.2014.6828607>
14. Nosrat, A., S. Kavian, Y.: Hardware description of multi-directional fast sobel edge detection processor by VHDL for implementing on FPGA . <https://doi.org/10.5120/7533-9872>
15. Pavan Kumar, M.: Hardware Acceleration of Edge Detection Using HLS (2019), <https://bit.ly/3i88uBW>, Undergraduate thesis, California State University
16. Rashmi, Kumar, M., Saxena, R.: Algorithm and Technique on Various Edge Detection : A Survey (Jun 2013). <https://doi.org/10.5121/sipij.2013.4306>
17. Rupnow, K., Liang, Y., Li, Y., Chen, D.: A study of high-level synthesis: Promises and challenges. <https://doi.org/10.1109/ASICON.2011.6157401>
18. Sanduja, V., Patial, R.: Sobel edge detection using parallel architecture based on FPGA <https://www.ijais.org/archives/volume3/number4/220-0515>
19. Sikka, P., Asati, A.R., Shekhar, C.: High-speed and area-efficient sobel edge detector on fpga for artificial intelligence and machine learning applications . <https://doi.org/10.1111/coin.12334>
20. Vallina, F.M., Kohn, C., Joshi, P.: Zynq all programmable SoC sobel filter implementation using the vivado HLS tool <https://bit.ly/3h6egD1>
21. Zheng, Y.: The design of sobel edge extraction system on FPGA 11. <https://doi.org/10.1051/itmconf/20171108001>

# Workshop Innovación en Educación en Informática

Coordinadores

Cecilia Sanz (UNLP)  
Beatriz Depetris (UNDTF)  
Marcelo De Vincenzi (UAI)

# Compilador para traducir ubicaciones geográficas en instrucciones atómicas en una aplicación de aprendizaje ubicuo de programación

Denis Acosta<sup>1</sup>, Margarita Alvarez<sup>1</sup>, Elena Durán<sup>1</sup>

<sup>1</sup> Instituto de Investigaciones en Informática y Sistemas de Información  
Facultad de Ciencias Exactas y Tecnologías – Universidad Nacional de Santiago  
del Estero

denislionelacosta@gmail.com, {alvarez, eduran}@unse.edu.ar

**Resumen.** El aprendizaje de los conceptos básicos de programación, como las estructuras de control, son consideradas difíciles debido a su complejidad y al nivel de abstracción requerido por el estudiante. Ante esta problemática, se ha considerado desafiante el desarrollo de una aplicación de aprendizaje ubicuo que asista a los estudiantes en el aprendizaje de estos conceptos. La aplicación, por medio de técnicas de realidad aumentada le muestra un objeto al estudiante, quien se debe desplazar en busca del mismo y como resultado la aplicación le genera un programa con las acciones realizadas. En este artículo se presenta la arquitectura de la aplicación de aprendizaje ubicuo, y en particular se desarrolla el módulo de compilación que permite traducir las coordenadas geográficas obtenidas por el celular del estudiante en instrucciones de un lenguaje de programación. Se muestran además las pruebas realizadas al módulo de compilación, que evidencian la viabilidad de realizar esta traducción.

**Palabras Claves:** Aprendizaje ubicuo, Arquitectura, Compiladores, Estructuras de control, Geolocalización.

## 1 Introducción

Por aprendizaje ubicuo (AU) puede entenderse, aquel que no se circunscribe a un contexto determinado ni viene condicionado por un espacio concreto, sino que la adquisición de nuevos conocimientos puede producirse en cualquier situación en la que se encuentre el sujeto, independientemente del momento y lugar [1]. El AU permite a los estudiantes acceder a todo tipo de información a partir de la interacción con los objetos que lo circundan, ya sean físicos o virtuales [2]. Favorece la colaboración, conecta los espacios formales, no formales e informales. También, resitúa la ubicación del aprendizaje, tanto dentro como fuera del aula y adapta los contenidos, presentaciones y actividades a las características de los estudiantes, y su contexto.

Por otra parte, las actividades de aprendizaje asociadas a la programación han sido reconocidas con alto grado de dificultad. Varios estudios han determinado que las causas que generan dicha problemática se relacionan con determinadas características que suceden dentro del aula y con ciertas habilidades cognitivas que son relevantes al

momento del aprendizaje de los fundamentos de programación. Entre ellas, están la capacidad de abstracción, una buena aptitud lógico-matemática y la facilidad para la resolución de problemas de orden algorítmico [3].

Ante la problemática planteada y las ventajas que proporciona el AU, se ha considerado desafiante el desarrollo de una aplicación que apoye el AU de los conceptos básicos de programación. La misma serviría para iniciar a los estudiantes en el aprendizaje de los conceptos básicos, mostrándole los algoritmos que realiza cuando se desplaza en un determinado ambiente, utilizando para ello el enfoque del microaprendizaje que propone aprendizajes con esfuerzos relativamente cortos y que consumen poco tiempo [4]. Así, la aplicación le muestra un objeto, empleando técnicas de realidad aumentada, que el estudiante debe obtener; el estudiante camina hasta el objeto y como resultado la aplicación le genera un programa con las acciones realizadas. Para ello, la arquitectura de la aplicación prevé un módulo de compilación que toma como entrada los datos de los sensores del celular del estudiante (GPS, Acelerómetro y Giroscopio) y los traduce a un programa escrito en el lenguaje diseñado a tal fin. Esto representó un importante desafío, ya que no se cuenta actualmente con una herramienta que haga posible esta traducción. Además, la interacción con la aplicación le provee una respuesta inmediata, favoreciendo al alumno la experimentación en el mundo real y permitiéndole formar rápidamente un modelo mental gracias a las respuestas que arroja la herramienta.

En este artículo se presenta la arquitectura de la aplicación de AU y el desarrollo del módulo de compilación, con las pruebas realizadas.

En las siguientes secciones se citan antecedentes de trabajos relacionados, se presenta la arquitectura de la aplicación y, el diseño y construcción del compilador. Finalmente, se muestran las pruebas realizadas y se expresan algunas conclusiones respecto al trabajo realizado, y las acciones a futuro.

## 2 Antecedentes

En esta sección se presentan antecedentes de aplicaciones de software para la enseñanza de los conceptos básicos de programación basada en la teoría de microaprendizaje y de realidad aumentada, y también algunos antecedentes de trabajos referidos al aprendizaje de cuestiones informáticas y computacionales mediante AU.

En [5] proporcionaron un medio flexible de AU con una aplicación móvil para que los estudiantes accedan al material. Los autores adoptaron el pensamiento computacional (PC) para ayudar a los estudiantes a desarrollar habilidades informáticas prácticas. Se eligieron tres clases de estudiantes de primer año para el estudio empírico. Se dividieron en tres grupos: dos grupos experimentales (grupo AU&PC y grupo PC) y un grupo de control. Según los resultados de este estudio, los estudiantes que recibieron el tratamiento de AU podrían tener habilidades informáticas significativamente mejores en el uso de PowerPoint y Word que los que no. Sin embargo, el tratamiento de la PC no resultó en un mejor desarrollo de las habilidades informáticas de los estudiantes en esta investigación.

Para alentar a los estudiantes universitarios a tomar cursos de programación de computadoras, en [6] se propone un sistema de apoyo al AU personalizado basado en múltiples fuentes de información. Este incluye una nueva técnica que integra problemas de aprendizaje de los estudiantes, nivel de conocimiento y estilo de aprendizaje para personalizar caminos de aprendizaje que le ofrecen al estudiante para su experiencia de aprendizaje de conceptos de programación.

En [7] se identifican los requisitos necesarios para apoyar el AU en las herramientas de software existentes para enseñar y aprender programación de computadoras a niños de entre 4 y 10 años de edad. Veintidós herramientas fueron analizadas y contrastadas con cinco características conocidas del AU: permanencia, accesibilidad, inmediatez, interactividad y conciencia del contexto. En una nota final, recomiendan agregar características de AU en las herramientas de programación para niños.

Boonbrahmet al. [8] desarrollaron una herramienta para aprender el flujo primario de comandos y las estructuras de control, incluyen la secuencia, las estructuras de selección e iteración. Con el empleo de la herramienta, los estudiantes construyen un diagrama de flujo del programa mediante el uso de marcadores de realidad aumentada. El software desarrollado captura la imagen del diagrama de flujo que el alumno ha construido, procesa el programa y muestra el resultado de la ejecución del comando. La herramienta identifica el comando, el valor de la variable y el operador booleano. Luego, el software simula el resultado de cada comando y así el alumno puede comprobar si la lógica del programa es correcta o no.

Skalka y Drlík [9] presentan un modelo conceptual de un sistema para mejorar las habilidades de programación basadas en la teoría del microaprendizaje en un entorno de aprendizaje móvil. El framework se basa en la combinación adecuada de objetos de contenido e interactivos. El objeto de contenido representa una unidad elemental de información. El objeto interactivo sigue cada unidad de contenido y verifican la comprensión de los estudiantes sobre el contenido educativo presentado.

En la revisión de antecedentes realizada no se encontraron trabajos que presenten experiencias de aplicación del AU en conceptos de programación, ni aplicaciones de software con esta finalidad.

Por otra parte, si bien existen aplicaciones que traducen coordenadas geográficas, no se han encontrado antecedentes de conversión de tales unidades a un lenguaje de programación como el que se propone en este trabajo.

### **3 Arquitectura de la Aplicación de Aprendizaje Ubicuo**

#### **3.1 Descripción de la Aplicación**

La finalidad de la aplicación es que apoye al docente en la introducción de los conceptos de programación referidos a las estructuras de control básicas: secuencial, de selección y de iteración y, que le permita al estudiante su transición a la elaboración de algoritmos más complejos.

Para el diseño de la aplicación se ha considerado el enfoque del microaprendizaje. A tal fin, se ha dividido la enseñanza de las estructuras básicas de control en tres

niveles, de forma tal de introducir en cada nivel una estructura básica de control, incrementando el nivel de complejidad de los ejercicios que puede realizar el estudiante. En el nivel 1 se realizan ejercicios sencillos con el objetivo de aprender el concepto de instrucción y de secuencia de instrucciones. En el nivel 2 se incluyen ejercicios donde aparece la estructura condicional, dando lugar a instrucciones para girar a la derecha y a la izquierda de forma tal de evitar obstáculos. En el nivel 3 los ejercicios introducen el concepto de repetición.

### 3.2 Arquitectura de la Aplicación

La arquitectura de la aplicación (Fig. 1) describe los módulos y las interrelaciones entre ellos.



Fig. 1. Arquitectura de la aplicación

El módulo de interfaz se encarga de la comunicación del estudiante con la aplicación. Determina el nivel de complejidad para ofrecerle las actividades de aprendizaje de acuerdo con el nivel, y por último, muestra al estudiante el programa resultante.

El módulo de sensores, es el encargado de sensar el entorno del dispositivo y coordinar con el Módulo Principal. Cuando el estudiante encuentra un objeto virtual, interviene el Módulo de Gestión 3D enviando el objeto listo para ser procesado por la interfaz.

Una vez finalizada la interacción, el Módulo Principal envía todos los datos recolectados al compilador para que pueda traducirlos y otorgar como resultado el algoritmo (programa) obtenido en el camino hasta alcanzar el objeto. Por último, el módulo de interfaz muestra el programa resultante usando símbolos representativos.

Así, la aplicación determina el nivel de acuerdo con la interacción del estudiante con la aplicación y el aprendizaje alcanzado, y mediante el uso de la cámara del teléfono, captura el lugar y le inserta objetos en realidad aumentada.

El objetivo es que el estudiante llegue caminando hasta el lugar donde se encuentra el objeto virtual. La aplicación le devuelve al estudiante un programa formado por un conjunto de instrucciones atómicas con la secuencia de pasos que realizó hasta encontrar el objeto.

## 4 Diseño y construcción del módulo compilador

Para el diseño y desarrollo del compilador se siguió la metodología que se indica a continuación. Diseño y construcción de las fases de análisis: análisis léxico, análisis sintáctico y análisis semántico. Luego, se desarrolló la fase de síntesis: el generador de código.

La construcción se ha realizado con la herramienta ANTLR (ANother Tool for Language Recognition), un generador de analizadores para leer, procesar, ejecutar o traducir texto estructurado.

### 4.1 Lenguaje fuente y lenguaje objeto

Debido a que un compilador es un programa que lee un programa escrito en un lenguaje fuente y lo traduce a un programa equivalente en un lenguaje objeto, se define en este apartado ambos lenguajes.

Para garantizar un diseño completamente modular e independiente de la aplicación, se ha diseñado el lenguaje fuente con una sintaxis lo más simplificada posible, de forma tal que, el resto de los módulos, no requieran de una interfaz compleja para comunicarse con el compilador.

Los datos de entrada al compilador son cada una de las ubicaciones geográficas (latitudes y longitudes) obtenidas por los diferentes dispositivos capaces de sensar el contexto geográfico (GPS, Acelerómetro y Giroscopio). Por lo tanto, el lenguaje fuente está compuesto del nivel del usuario y el conjunto de puntos geográficos.

El programa resultante de la compilación está escrito en el lenguaje objeto. Se han tenido en cuenta las siguientes características para la selección de las instrucciones o sentencias del lenguaje objeto, considerando principalmente el aprendizaje de los estudiantes:

1- El lenguaje debe ser visible de forma natural, la mayoría de las operaciones realizadas por el actor son producto de cambios visibles en el micro mundo real [10].

2- El lenguaje debe ser atractivo y significativo para la categoría de estudiantes prevista [10].

3- El lenguaje debe tener la capacidad de añadir nuevas características (instrucciones) sin tener que ser reestructurado nuevamente (extensibilidad).

En este caso, el lenguaje objeto está formado por las siguientes instrucciones atómicas: avanzar (representa un paso hacia adelante realizado por el estudiante en su caminata hasta llegar al objeto), retroceder (representa un paso hacia atrás que el alumno realiza para cumplir el objetivo del ejercicio), girar izquierda y girar derecha (permiten desviarse del camino lineal), si hay objeto entonces (representa la bifurcación cuando se presenta o no un obstáculo) y repetir (con esta instrucción el estudiante visualiza una sentencia de control repetitiva).

## 4.2 Construcción de las fases del compilador

Para la construcción del analizador léxico se han definido los componentes léxicos y los patrones o expresiones regulares que generan dichos componentes. Los componentes léxicos definidos son: símbolos de puntuación y constante numérica.

Para la fase del Analizador Sintáctico se define la gramática libre de contexto mediante la notación BNF que es la entrada al generador ANTLR.

```
start:
  LLAVE_A
  CORCHETE_A nivel CORCHETE_C
  CORCHETE_A puntos CORCHETE_C
  LLAVE_C;
nivel: NUMERO;
puntos: tupla (COMA tupla)*;
tupla: PARENTESIS_A latlong COMA latlong PARENTESIS_C;

latlong: NUMERO PUNTO NUMERO;
```

En la fase de análisis semántico la verificación de tipos se realiza a las constantes numéricas para determinar si son de tipo latitud o de tipo longitud.

## 4.3 Generación del programa objeto

Para la generación del programa objeto se realiza el procedimiento descrito en Fig. 2 que aplica tres algoritmos: el de relación, de clasificación y el algoritmo rankeador.

Los algoritmos de relación y de clasificación, se han desarrollado para determinar la relación entre dos pares de puntos geográficos de forma tal de determinar si el estudiante ha realizado un paso hacia adelante, hacia atrás o a girado en alguna dirección.

Cada vez que se almacenan dos elementos consecutivos (proceso que ejecuta el analizador sintáctico cuando recorre el árbol de sintaxis abstracta), se realiza una comparación para determinar cuál es la relación existente entre ellos (vector) y se clasifica dicha relación en su equivalente a las instrucciones atómicas.

Dado que un vector en el plano es un par ordenado (a, b) de números reales perteneciente al espacio  $R^2$  [11], se considera para este caso a “a” como el conjunto de latitudes y el componente “b” como el conjunto de longitudes. Las latitudes y longitudes son la entrada al Algoritmo de Relación y los resultados indican los distintos sentidos que puede tener un vector. A partir de los vectores obtenidos, se aplica el algoritmo de clasificación que establece cuál es la relación existente entre dichos vectores. Los resultados que pueden presentarse luego de clasificar un par de vectores se dividen en dos categorías: paralelos y ortogonales.

De acuerdo con [11] para determinar cuándo un par de vectores pertenece a un conjunto u otro se considera:

Paralelismo de vectores: sean  $u$  y  $v$  vectores no nulos de  $R^2$ , el vector  $u$  es paralelo al vector  $v$  si y solo si existe un escalar no nulo  $c$  tal que:

$$u = c.v. \quad (1)$$



Fig. 2. Proceso para la obtención del programa

*Ortogonalidad:* sean  $u$  y  $v$  vectores no nulo de  $\mathbb{R}^2$ , se dice que  $u$  es ortogonal a  $v$  si y solo si el producto escalar de  $u$  y  $v$  es igual a cero,

$$c. v = 0 \quad (2)$$

*Ángulos entre vectores:* sean  $u$  y  $v$  vectores no nulos de  $\mathbb{R}^2$  y sea  $\cdot$  el producto escalar, entonces existe y es único un  $\alpha$   $[0, \pi]$  tal que:

$$\cos(\alpha) = \frac{u \cdot v}{\|u\| \|v\|} \quad (3)$$

Para la clasificación de los vectores se aplican las siguientes reglas:

- Si ambos vectores calculados son paralelos se asume que el usuario avanzó en línea recta y para determinar si avanzó o retrocedió se considera el sentido de los vectores.
- Si los vectores son perpendiculares, se calcula la intersección entre ellos de forma tal de obtener el ángulo que los une. Si el ángulo es cercano a  $90^\circ$ , se asume que se trata de un cambio de dirección.
  - ✓ Si el vector superior tiene un sentido a la izquierda, se dirá que el usuario giro a la izquierda.
  - ✓ Si el vector superior tiene un sentido a la derecha, se dirá que el usuario giro a la derecha.

Finalmente, se mapea cada una de las reglas con la instrucción atómica equivalente (Tabla 1).

Tabla 1. Mapeo de las relaciones a instrucciones atómicas.

Relación	Sentido	Instrucción Atómica
Paralelos	Hacia arriba	AVANZAR
	Hacia abajo	RETROCEDER
Ortogonales	Izquierda	GIRAR IZQUIERDA
	Derecha	GIRAR DERECHA

Como última etapa, el compilador tendrá la tarea de adaptar el conjunto de instrucciones atómicas en base al nivel que posea el estudiante. Para ello, se aplica el

Algoritmo Rankeador, que toma como datos de entrada el conjunto de instrucciones atómicas y retorna como respuesta un conjunto adaptado en base a su nivel. Ésta adaptación produce una reducción de instrucciones o una agregación de nuevas, tales como: repetir y si hay obstáculo entonces. Para la sentencia condicional, el compilador al detectar la tupla (0,0) infiere que hay un obstáculo. Por ejemplo, si se tiene como entrada tres puntos: (-64.25139158964157,-27.80154927681387), (0,0), (-64.25133660435677,-27.801606219132715) la aplicación deduce que entre el punto 1 y el punto 3 existe un obstáculo.

## 5 Pruebas

Se realizaron pruebas para todos los niveles de complejidad. Se describe a continuación una prueba para cada tipo.

En la tabla 2 se puede observar una prueba realizada para el nivel de complejidad 1. Se observan el conjunto de latitudes y longitudes capturadas por la aplicación cuando el estudiante realiza el recorrido hacia el objeto (Fig. 4.a). También, se muestra la salida del módulo de compilación, programa resultante, y por último la interfaz que se le presenta al estudiante.

En la tabla 3 se puede observar una prueba realizada para el nivel de complejidad 2. En el recorrido que realiza el estudiante se puede encontrar con un obstáculo (Fig.4.b), por lo que el programa resultante introduce el concepto de la estructura condicional.

En la tabla 4 se puede observar una prueba realizada para el nivel de complejidad 3. El estudiante al realizar varias veces la misma operación, por ejemplo, avanzar (Fig. 4.c), el programa resultante contiene la sentencia repetir, introduciendo así, el concepto de estructura de control repetitiva.

**Tabla 2.** Prueba realizada para el nivel 1 de complejidad.

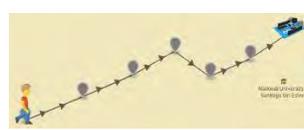
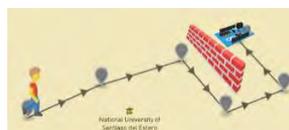
Entrada	Salida	Interfaz
{		
[1]	{	
[(-64.25157129764557,-27.801613336920475),	[AVANZAR,	
(-64.25148211419582,-27.80157715482788),	AVANZAR,	
(-64.25138555467129,-27.8015510562618),	GIRAR DERECHA,	
(-64.25134532153606,-27.801586645213977),	AVANZAR,	
(-64.25128899514675,-27.80155639460536),	AVANZAR,	
(-64.25122663378716,-27.80151843304538),	GIRAR DERECHA]	
(-64.25117500126362,-27.80154868366457)]	}	
}		

**Tabla 3.** Prueba realizada para el nivel 2 de complejidad.

Entrada	Salida	Interfaz
<pre>{   [2]   [(-64.25157129764557,-27.80161570951628),   (-64.25148278474808,-27.801581306871892),   (-64.25139158964157,-27.80154927681387),   (0,0),   (-64.25133660435677,-27.801606219132715),   (-64.25126284360884,-27.801596728748322),   (-64.2511984705925,-27.801581306871892)] }</pre>	<pre>{   [AVANZAR,   AVANZAR,   SI HAY OBJETO   ENTONCES,   GIRAR DERECHA,   AVANZAR,   GIRAR IZQUIERDA] }</pre>	

**Tabla 4.** Prueba realizada para el nivel 3 de complejidad.

Entrada	Salida	Interfaz
<pre>{   [3]   [(-64.25157062709332,-27.801612150622535),   (-64.2514868080616,-27.801581306871892),   (-64.25141505897044,-27.801553428858966),   (-64.25135537981987,-27.8015231782411),   (-64.25130508840084,-27.801555801456097),   (-64.25124809145927,-27.801533854930724),   (-64.25118573009968,-27.801497079662074)] }</pre>	<pre>{   [REPETIR 3,   AVANZAR,   GIRAR DERECHA,   REPETIR 1,   AVANZAR] }</pre>	



(a) Recorrido para el nivel 1. (b) Recorrido para el nivel 2. (c) Recorrido para el nivel 3.

**Fig. 4.** Recorrido del estudiante de las pruebas realizadas.

## 6 Conclusiones

En este trabajo se ha presentado la arquitectura de una aplicación de AU para la enseñanza de los conceptos básicos de programación en un intento por superar los problemas de aprendizaje de los conceptos que incluyen las estructuras de control.

También, hemos descrito el módulo de compilación que permite traducir las ubicaciones geográficas que se registran de las actividades que el estudiante realiza en

el entorno ubicuo y las pruebas realizadas al compilador. Estas pruebas ponen de manifiesto que es factible contar en una aplicación de AU que convierta acertadamente las coordenadas en instrucciones de un lenguaje, que luego los estudiantes podrán usar para programar. Este módulo de compilación resulta útil, novedoso y de amplia aplicación, ya que no existen antecedentes de este tipo de traducción de ubicación geográfica a instrucción de programa, que además se puede visualizar gráficamente, y que podrá ampliarse para generar instrucciones más complejas a posteriori.

Los trabajos futuros se orientan al desarrollo de los otros módulos que conforman la aplicación y a su evaluación en cursos de enseñanza de la programación.

De esta forma, tanto la geolocalización como el AU, pueden resultar en poderosas herramientas que posibiliten la generación de espacios de aprendizaje más ricos y motivadores.

## Referencias

1. Burbules, N.C.: Ubiquitous Learning and the Future of Teaching. *Encounters*, 13, 3--14 (2012)
2. Villalustre Martínez, L., del Moral Pérez, M. E.: Geolocalización y realidad aumentada para un aprendizaje ubicuo en la formación inicial del profesorado. *@tic revista d'innovació educativa*, 21, 40--48 (2018)
3. Insuasti, J.: Problemas de enseñanza y aprendizaje de los fundamentos de programación. *Revista educación y desarrollo social*, 10, 234--246 (2016)
4. Salinas, J., y Marín, V. I.: Pasado, presente y futuro del microlearning como estrategia para el desarrollo profesional. In: *Campus Virtuales*, Vol. III, Num. 2, pp. 46--61. Huelva, España (2014)
5. Chia-Wen Tsai, Pei-Di Shen, Meng-Chuan Tsai & Wen-Yu Chen: Exploring the effects of web-mediated computational thinking on developing students' computing skills in a ubiquitous learning environment. *Interactive Learning Environments*, 25:6, 762--777 (2017)
6. Chookaew S., Wanichsan D., Hwang G-J., Panjaburee P.: Effects of a personalised ubiquitous learning support system on university students' learning performance and attitudes in computer-programming courses. In: *International Journal of Mobile Learning and Organisation* (2015)
7. Hosanee Y., Panchoo S.: The Analysis and the Need of Ubiquitous Learning to Engage Children in Coding. In: Fleming P., Lacquet B., Sanei S., Deb K., Jakobsson A. (eds) *Smart and Sustainable Engineering for Next Generation Applications*. ELECOM 2018. *Lecture Notes in Electrical Engineering*, vol 561. Springer, Cham (2019)
8. Boonbrahm, S., Boonbrahm, P., Kaewrat, C., Pengkaew, P., Khachorncharoenkul, P.: Teaching Fundamental Programming Using Augmented Reality. *International Journal of Interactive Mobile Technologies (IJIM)* (2019)
9. Skalka J., Drlík, M.: Conceptual Framework of Microlearning-based Training Mobile Application for Improving Programming Skills *Interactive Mobile Communication Technologies and Learning*. IMCL 2017. *Advances in Intelligent Systems and Computing*, vol 725. Springer, Cham (2017)
10. Brusilovsky, P., Calabrese, E., Hvorecky, J., Kouchnirenko, A., and Miller, P.: *Mini-languages: A Way to Learn Programming Principles*. *Education and Information Technologies* 2, pp. 65--83 (1997)
11. Kolman, B., Hill, D.: *Álgebra Lineal 8va Edición*. Pearson Educación 216--224. (2006)

# Estrategias para trabajar competencias en una asignatura masiva de Informática. El caso de Seminario de Python.

Claudia Banchoff, Viviana Harari, Sofía Martín

LINTI. Facultad de Informática. Universidad Nacional de La Plata. La Plata,  
Argentina

{cbanchoff, vharari, smartin}@linti.unlp.edu.ar

<https://www.linti.unlp.edu.ar/>

**Abstract.** La mayoría de las asignaturas de las carreras relacionadas a la Ciencia de la Computación, suelen centrarse en la enseñanza de los aspectos específicos asociados a las mismas. Esta forma de enseñanza es lo que tradicionalmente se viene planteando desde hace años y es naturalmente aceptado tanto por los estudiantes como por los planteles docentes de las diferentes asignaturas. En los últimos años, en el terreno educativo y, en función de mejorar los procesos de enseñanza y aprendizaje, se comenzó a hablar de la educación basada en competencias. Este artículo describe cómo fue la adecuación de la propuesta de enseñanza de una materia masiva de la Facultad de Informática de la Universidad Nacional de La Plata para trabajar competencias. No sólo se detallarán las acciones tomadas sino que también se hablará sobre los debates internos que estos cambios generaron al interior del equipo docente de la materia.

**Keywords:** aprendizaje por competencias, evaluación en proceso, cátedra masiva, Python

## 1 Introducción

En el año 2018 la Facultad de Informática de la Universidad Nacional de La Plata (UNLP), propone incorporar en sus planes de estudios las competencias a alcanzar por sus egresados, de acuerdo a los estándares internacionales adoptados por las Universidades de Argentina [1]. A las llamadas Competencias Específicas (CE) en las titulaciones se agregaron dos más llamadas: Competencias Genéricas Sociales, políticas y actitudinales (CGS) y, Competencias Genéricas Tecnológicas (CGT). Las CGS están relacionadas con la capacidad de trabajo en equipo, de comunicación oral y escrita, conocimiento de las normas jurídicas relacionadas con la disciplina, formación en la ética profesional y en la responsabilidad social. Las CGT están relacionadas con los conocimientos, habilidades, pensamientos y valores que debe poseer un profesional de Informática, tales como interpretar problemas del mundo real para resolverlos desde la Informática, gestionar proyectos informáticos, analizar y cuantificar los riesgos asociados con un producto o

2 Claudia Banchoff, Viviana Harari, Sofía Martín

servicio informático, tener capacidad de innovación tecnológica y formación para el trabajo multidisciplinario.

Esta situación plantea la introducción del trabajo por competencias en las asignaturas de sus dos Licenciaturas: Informática y Sistemas, incorporando un nuevo desafío en la elaboración de las propuestas de enseñanza y en la coordinación interna de las cátedras.

A partir de esa decisión, cada asignatura tuvo que hacer una revisión de sus propuestas de enseñanza incorporando los cambios necesarios de acuerdo a las competencias que atravesaban la materia. En particular, la cátedra de Seminario de Lenguajes, opción Python, también llevó a cabo esa tarea.

Seminario de Lenguajes es una materia de segundo año perteneciente a todas las carreras de la Facultad de Informática de la UNLP. En esta materia se proponen varias opciones de dictado, una de ellas es la opción de lenguaje Python.

El Seminario de Python se dicta en el primer semestre, en dos turnos: mañana y tarde. Tiene un formato de tipo taller con un planteo de trabajo final integrador, correspondiente al desarrollo de un software orientado al ámbito educativo. El perfil de quienes lo cursan corresponde a estudiantes que han tenido en primer año una introducción básica sobre programación. Esto hace que, para muchos de ellos, sea la primera vez en la carrera que se enfrentan al desarrollo de un software, cumpliendo todas las etapas: análisis, desarrollo, puesta a punto y prueba del mismo. El plantel docente está integrado por dos profesoras, dos jefes de trabajos prácticos, cinco ayudantes rentados y cuatro colaboradores ad honorem.

En los años 2019 y 2020 la matrícula ha sido de 120 y 132 inscriptos, con un marcado crecimiento respecto a los años anteriores. Esto incorpora un nuevo desafío: la masividad.

La mayoría de las asignaturas de la facultad suelen centrarse en la enseñanza de los aspectos específicos relacionados a las mismas. Las materias sobre desarrollo web, por ejemplo, focalizan sus contenidos en los distintos lenguajes, metodologías y frameworks que permiten abordar la temática, y las que tienen como objetivo enseñar un lenguaje específico de programación, se suelen concentrar en los distintos recursos del lenguaje aplicándolos a desarrollos particulares. Esto es lo que tradicionalmente se ha planteado y es naturalmente aceptado tanto por los estudiantes como por el plantel docente de las asignaturas. Por lo tanto, abordar las competencias, especialmente las CGT no fue un gran problema, dado que muchas de las actividades relacionadas a las mismas ya estaban incluidas en la propuesta de enseñanza desde que se inició con el seminario en el año 2009. Se trabaja en la identificación, formulación y resolución de problemas del mundo real, la utilización de manera efectiva de técnicas y herramientas que permiten realizar una planificación, implementación y testeo de sus programas y, también se trabajan los primeros pasos sobre el análisis, diseño y evaluación de interfaces de usuario.

El abordaje de las CGS en la materia no fue tan sencillo, surgieron otros aspectos que han requerido más atención y coordinación, más aún tratándose de una materia masiva.

A través de este artículo, en las siguientes secciones, se describirán algunas de las modificaciones que se realizaron en la metodología de trabajo de la materia, para que la misma se ajuste a los nuevos requerimientos.

## 2 El aprendizaje por competencias

En el ámbito educativo la definición e implementación de las competencias se encuentran distintas miradas y enfoques e incluso sobre si es factible su aplicación y evaluación. Uno de los interrogantes más importantes es si los requerimientos del mundo actual del trabajo deben o no influir en las propuestas de enseñanzas de las carreras de universitarias que forman profesionales [3]. Existe un debate sobre si le compete o no a la universidad formar en aspectos meramente sociales orientados hacia alguna concepción del ser humano, o solamente en aspectos técnicos de cada carrera, definidas actualmente como competencias sociales y técnicas respectivamente. Estas discusiones se encuentran desde hace tiempo dentro del ámbito académico y han generado distintas posiciones polarizadas [2][4].

A lo largo de los años la innovación en la metodología de enseñanza de los planes de estudio ha sido un desafío constante. Desde los años noventa en general se ha innovado en situaciones de enseñanza con enfoques constructivistas, aprendizaje colaborativo, empleo de metodologías a través de simuladores, entre otros[4]. Introducir estas innovaciones por lo general es complejo, dado que no siempre se cuenta con gran cantidad de respaldo de experiencias de su implementación. Es por esto que es importante su análisis detallado para que las innovaciones se conviertan en acciones pedagógicas reales para su implementación con aciertos y sus limitaciones.

La incorporación de las competencias en una metodología de enseñanza universitaria, pensada como innovación, pone en juego diversas cuestiones. Para entender un poco más el significado de competencias se puede partir desde su origen y los aspectos relevantes que se deben tener en cuenta en la educación superior.

El concepto de competencia surge en el sector empresarial y se amplió al ámbito educativo a solicitud de dicho sector. Se propuso a las instituciones de formación de profesionales que aborden no sólo competencias técnicas sino también sociales, debido a que las mismas serían tenidas en cuenta al momento de la presentación a puestos de trabajos.

En esta interacción entre ámbito laboral y educativo es que surgen las diferentes posiciones antes mencionadas.

La competencia implica saberes, es decir conocimientos, no sólo los relacionados a un campo disciplinario sino también a aspectos sociales. Es un saber que se manifiesta en la acción, en una situación planeada o no. Algunos conocimientos pueden quedar implícitos: son los que hacen referencia a los saberes previos y

4 Claudia Banchoff, Viviana Harari, Sofía Martín

adquiridos. Al ser de tipo implícitos se crean en base a la experiencia, son poco comunicables y difíciles de identificar. Por este motivo, su evaluación plantea un desafío para el docente, dado que no son sencillos de evaluar.

Otro aspecto relevante a tener en cuenta es la importancia de la movilización de la acción adquirida por una determinada competencia, es decir no solamente adquirir el conocimiento sino también poder llevarlo a cabo[5]. La acción de poder aplicar el conocimiento adquirido no puede estar descontextualizada de la situación en que se aplica, dado que en diferentes contextos, la acción a realizar no es la misma. Esto refleja la necesidad de reconocer cómo fue el proceso de aprendizaje dado que no es lo mismo aprender de memoria que comprender lo que se está haciendo.

Los aspectos planteados que involucran una competencia se reflejan en la definición dada por Jacques Tardif: “*Un saber actuar complejo que se apoya sobre la movilización y la utilización eficaz de una variedad de recursos*”[6]. El término de la movilización, se refiere como se dijo anteriormente a que los conocimientos adquiridos no solamente se posean sino que también sean capaces de relacionarlos, combinarlos. En cuanto a los recursos se refiere a todo el contexto que forma parte del estudiante.

### 3 El impacto de las competencias sobre la propuesta de enseñanza del Seminario de Python

Los aspectos antes vistos son considerados importantes para tener cuenta al momento de definir una propuesta de enseñanza que involucre competencias. Es por esto que, en base a las competencias definidas para las carreras de Informática por la RedUnci (CGT y CGS), se ha reformulado la propuesta de enseñanza del seminario.

Algunas de las competencias que se trabajan en la asignatura (en mayor o menor grado) incluyen:

- Identificar, formular y resolver problemas de Informática.
- Utilizar de manera efectiva técnicas y herramientas informáticas.
- Poder comunicarse con efectividad en forma oral y escrita.
- Aprender en forma continua y autónoma.
- Poder interpretar la evolución de la Informática con una visión de las tendencias tecnológicas futuras.
- Trabajar en equipos.

Si bien las dos primeras (CGT) son más sencillas de abordar, ya que se han estado trabajando aún sin enmarcarlas en este contexto desde hace años, la incorporación de las CGS presentó un desafío en la propuesta de enseñanza y en la coordinación interna de la cátedra.

La materia aplica un sistema de evaluación continua en el proceso de enseñanza desde hace dos años y, utiliza un sistema de gamificación para algunas de las actividades. Esto permitió la inclusión y/o reformulación de actividades para enfocarlas en el abordaje de las competencias.

Los estudiantes cursan dos veces por semana en espacios de teoría y de práctica, estrechamente vinculados. En los horarios de práctica se realizan actividades con las computadoras, mientras que en los horarios de las teorías no siempre es así. En todos los casos, las clases son pensadas de forma tal que los estudiantes tengan una marcada participación. Si bien se explican los conceptos relacionados con el lenguaje, se intenta siempre abordar los temas a través de preguntas e indagaciones en donde se estimula permanentemente a los estudiantes a la intervención.

Lo mismo sucede en las clases prácticas, donde muchas veces se plantean problemas y/o desafíos y se incentiva a debatir propuestas de solución. La participación en este tipo de actividades permite guiar a los estudiantes sobre cómo se debe comunicar y/o defender un planteo. En muchos casos, esto requiere que se indague e investigue sobre alguna herramienta o concepto que se deba aplicar. Estas situaciones son el marco propicio para incentivar a los alumnos a realizar búsquedas relacionadas a lo que se está aprendiendo.

Desde el año 2019, se propone un sistema de puntos para la aprobación de la asignatura, dejando de lado la tradicional evaluación de parciales y recuperatorios. La evaluación se basa en la realización de distintas actividades que tienen asociadas ciertos puntajes. Para aprobar la materia, los estudiantes deben alcanzar cierta cantidad de puntos en las dichas actividades y aprobar un trabajo integrador. Estos puntos son de carácter individual mientras que el trabajo integrador se aprueba en forma grupal. En cada espacio, ya sea en la teoría o en la práctica, se buscan distintos mecanismos y estrategias para la realización de las actividades.

A modo de ejemplo, las evaluaciones prácticas, se llevaron a cabo sobre las computadoras, donde los estudiantes debían realizar una modificación menor a sus propios códigos o responder alguna pregunta sobre los mismos. Estas evaluaciones también fueron de carácter individual y se realizaron con la herramienta Virtual Programming Lab (VPL) [7] provista a través de un plugin del EVEA Moodle, utilizado por la cátedra. Esta herramienta, permite realizar una solución programática dentro del mismo entorno o bien, subir el archivo realizado en la computadora como solución del enunciado. Según la configuración adoptada se puede implementar test de la solución, que le permiten al estudiante evaluar su programa según las consignas solicitadas. No sólo permite realizar la entrega, sino que también durante el período de evaluación configurado, admite progresivas entregas, lo que le permite al estudiante corregir errores o realizar mejoras. En este caso la evaluación se realiza sobre la última recibida y además permite visualizar las diferentes versiones realizadas. Otro aspecto interesante de la herramienta es la posibilidad de analizar similaridad de código entre las entregas de los estudiantes. Si bien el análisis no siempre es definitivo al definir una clasificación orienta sobre posibles copias entre las soluciones entregadas del estudiantado. Estas actividades no sólo permiten evaluar lo aprendido por los estudiantes sobre un tema dado, sino que dado que durante la prueba pueden acceder a Internet, pueden acceder y utilizar el código libre que encuentren o el realizado por ellos mismos.

6 Claudia Banchoff, Viviana Harari, Sofía Martín

En los dos últimos años se realizaron talleres enfocados a conocer sobre las nuevas tendencias tecnológicas asociadas al lenguaje. Durante la cursada 2019<sup>1</sup> se trabajó sobre programación de sensores como una introducción a la temática de IoT (por sus siglas en inglés de Internet of Things) y en la cursada 2020 se abordó la ciencia de datos en un taller sobre Pandas.

El trabajo integrador consta de un desarrollo en grupos de no más de tres estudiantes. Se evalúa en dos instancias donde se comprueba que se hayan alcanzado las consignas dadas para cada etapa y la participación de cada integrante del grupo. Este trabajo, en el cual se ponen en juego todos los conceptos vistos en el seminario, también permite abordar el trabajo en equipo, en donde se debe coordinar un desarrollo y su posterior defensa. Cada grupo está coordinado por un ayudante de la cátedra quien lleva un registro de la participación de cada integrante a lo largo del desarrollo. El trabajo se presenta en una exposición oral final y se acompaña de un informe sobre el desarrollo del mismo. Se dicta un taller en donde se dan las guías de cómo se debe realizar la presentación y las pautas generales para redactar un informe académico. Esto último se realiza con la cooperación de la Dirección de Asesoramiento Pedagógico de la Facultad.

## 4 Los Python-plus

Desde el ciclo lectivo 2018, se incorporaron aspectos de gamificación a lo largo de la cursada, otorgando puntos extras, denominados “Python-plus”, a diversas actividades opcionales propuestas. Estos puntos pueden canjearse por beneficios tales como fechas extras de evaluación o una re-entrega para el trabajo integrador.

Las actividades opcionales que otorgan Python-plus surgieron como una forma de mantener a los estudiantes activos a lo largo de toda la cursada estimulando la participación de los estudiantes en las clases. Sin embargo, también son un espacio en donde se pueden trabajar las CGS. A modo de ejemplo, dado que muchos de los estudiantes son reacios a trabajar en grupo, se “premia” con Python plus a aquellos que efectivamente lo hagan.

Mientras que en los dos años anteriores, los estudiantes obtenían estos puntos en cualquier momento ante desafíos o planteos propuestos en las distintas instancias, a partir de la actual cursada 2020, estos puntos fueron enmarcados en seis (6) actividades específicas distribuidas a lo largo de la misma. Esto nos dio la posibilidad de contar con doce (12) instancias específicas que permiten abordar y/o reforzar las competencias trabajadas.

## 5 La coordinación del plantel docente

Como se mencionó anteriormente, el plantel docente está formado por dos profesoras, dos jefes de trabajos prácticos (JTP) y cinco ayudantes rentados y cinco colaboradores.

<sup>1</sup> Si bien la experiencia fue muy buena y estaba planificado repetirlo en la cursada 2020, esto no pudo ser posible debido a la situación planteada por el COVID-19.

Si bien las autoras de este trabajo (profesoras y jefa de trabajos prácticos de la materia) están convencidas de la necesidad de dejar atrás la enseñanza tradicional y, poner la mirada en una nueva metodología de enseñanza, con eso no basta. Para lograr que la metodología propuesta pueda ser exitosa se necesita que todo el equipo docente esté convencido de sus beneficios. La metodología no puede ser aplicada por algunos, debe ser aplicada por todos.

Cuando se tomó la decisión de realizar el cambio se programaron una serie de reuniones de cátedra donde se fue informando sobre los nuevos lineamientos metodológicos. Se discutió y reflexionó sobre el cambio actitudinal que el docente debería tener frente al aula justificando en cada caso el por qué de los cambios. Se definieron una serie de instrumentos que permiten realizar un seguimiento de la evolución de los estudiantes y, los canales de comunicación a utilizar como ser foros provistos por el EVEA, grupos de Telegram, entre otros.

Si bien, cuando se hizo el planteo en forma generalizada, la mayoría de los docentes no manifestaron grandes desacuerdos, cuando la metodología se puso en acción y se comenzó a aplicar, aparecieron ciertas situaciones conflictivas que llevaron su tiempo y esfuerzo revertirlas, requiriendo reuniones específicas, en forma grupal o individual. A continuación se citan las más relevantes:

- Cambio actitudinal frente a la práctica. Dejar atrás el rol pasivo del docente que espera que los estudiantes se acerquen a consultar y, tomar un rol más activo proponiendo pequeños problemas o desafíos que disparen debates sobre posibles soluciones.
- Mantener un informe detallado de los estudiantes a cargo, registrando cómo fue la participación a lo largo de la cursada.
- Detectar estudiantes poco participativos e incentivarlos a participar en diferentes propuestas.

Por otro lado, cabe destacar que este año, en el marco de la pandemia mundial del COVID'19 y, la necesidad de dictar la materia en formato puramente virtual, el trabajo se hizo más complejo. Requirió de mucho más trabajo por parte de los docentes para poder mantener la metodología propuesta. La virtualidad introduce otros aspectos a trabajar que en la forma presencial, quizás no son tan evidentes.

## 6 Los resultados

El hecho de contar con numerosas instancias de participación permite abordar las competencias de diferentes formas, tanto en las actividades establecidas en el cronograma por los puntos de la cursada como las que otorgan Python-plus.

La situación planteada durante la cursada 2020 obligó a que toda la interacción fuese a través de un entorno virtual, generó una mayor participación de los estudiantes en situaciones en las cuales debían expresarse verbalmente frente a todo el grupo conectado en la misma sala virtual. Estas situaciones surgieron al momento de querer explicar un error, plantear una posible solución de un ejercicio o, simplemente hacer un aporte en una clase.

8 Claudia Banchoff, Viviana Harari, Sofía Martín

La Tabla 1 muestra algunas de las actividades propuestas y las competencias.

Table 1: Actividades propuestas y las competencias que permiten abordar.

Competencias abordadas		Actividad realizada					
		A1	A2	A3	A4	A5	A6
CGS	C1	X		X	X		
	C2		X		X(*)	X	X
CGT	C3	X	X	X		X	
	C4	X	X			X	X

(\*) Se les pide que incluyan una sección en donde deben investigar alguna temática específica.

Los items de la Tabla 1 se refieren a:

- A1: ejercicios a entregar.
- A2: desarrollo del trabajo integrador.
- A3: defensa oral del trabajo integrador.
- A4: redacción del informe final.
- A5: participación en talleres extras.
- A6: uso de repositorios y herramientas de software.
- C1: comunicarse con efectividad en forma oral y escrita.
- C2: aprender en forma continua y autónoma.
- C3: identificar, formular y resolver problemas de Informática.
- C4: utilizar de manera efectiva las técnicas y herramientas de aplicación de la Informática.

La cantidad de actividades obligatorias y las de Python-plus de la cursada 2020 son doce (12) a las que se le agregarán dos (2) más producto del canje de puntos. Esta gran cantidad de actividades, que permiten abordar la evaluación como un proceso continuo durante la cursada, requiere un gran esfuerzo de coordinación de todo el plantel docente y una comunicación constante con los estudiantes. Esto también permite abordar el trabajo de habilidades sociales que muchas veces no es tan sencillo en materias con orientación tecnológica. Se puso especial énfasis en mantener actualizado el cronograma de actividades y los puntos acumulados, de manera tal que los estudiantes puedan visualizar su situación frente a la materia.

Si bien esto conlleva a una gran tarea tanto para los estudiantes como para el equipo docente, los resultados al finalizar la cursada son muy buenos, dado que el porcentaje de aprobación de la materia ha aumentado considerablemente. La Tabla 2 muestra un resumen de los números obtenidos en las cursadas en los últimos 4 años en donde puede verse esta mejora y el marcado incremento en la matrícula en los últimos dos años.

De estos datos se desprende que aunque la matrícula aumentó considerablemente en los últimos dos años, el porcentaje de aprobación creció notablemente. Esto se debe principalmente al cambio de estrategia de evaluación planteada. La Fig.1 muestra gráficamente este hecho.

Table 2: Porcentaje de aprobados y desaprobados de las cursadas antes y después de modificar la propuesta de enseñanza.

Año	Inscriptos	Activos	Aprob. sobre los activos	Aprob.	Desaprob.	Ausentes(1)
2017	37	22	59%	35%	25%	40%
2018	80	65	53%	44%	38%	18%
2019	120	101	71%	60%	25%	15%
2020 (2)	132 (3)	106	73%	58 %	22 %	20%

Donde:

- (1) Aquellos estudiantes que abandonan antes de la primera actividad propuesta.
- (2) Datos estimados a la fecha de cierre de este artículo, habiendo evaluado 10 actividades y teniendo en cuenta que 96 estudiantes tienen Python-plus suficientes para canjear por fechas adicionales.
- (3) Corresponde a 126 inscriptos en el seminario y 6 estudiantes que cursan como oyentes provenientes de otras carreras.

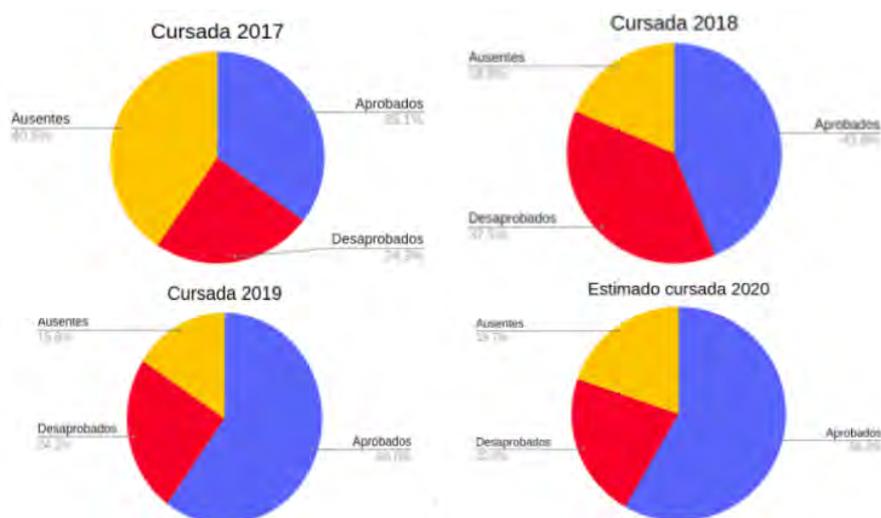


Fig. 1: Resumen de las cursadas 2017, 2018, 2019 e incluye el estimado de la cursada 2020.

## 7 Conclusiones

La evaluación en proceso en cátedras masivas es una tarea muy compleja. Requiere no solamente planificar una serie de actividades bien definidas y un cronograma de cumplimiento estricto, sino también establecer acuerdos entre todo el

10 Claudia Banchoff, Viviana Harari, Sofía Martín

plantel docente. Esta gran cantidad de actividades fueron la base para comenzar a abordar el trabajo por competencias en la materia.

El sistema de Python-plus que fue pensado originalmente como un mecanismo para motivar a los estudiantes a la participación activa en la cursada, permitió encuadrar ciertas actividades no técnicas que permiten abordar las CGS, no siempre tan fácil de trabajar.

El incremento en la matrícula, junto con la resistencia de parte del equipo docente fue una de las dificultades más relevantes a afrontar, pero no la única. En una cursada que tiene una duración de 16 semanas, proponer 12 actividades implica un cronograma muy ajustado. En años en los que abundan los paros docentes, como lo fue el 2019, o en donde algunas de las materias que se cursan en paralelo modifican sus cronogramas, la gestión de la cátedra se complica.

En este trabajo se describió cómo a través de la formulación de distintas actividades y seguimiento de los estudiantes, no sólo se puede abordar las competencias, sino que también se puede realizar una evaluación en proceso con muy buenos resultados. La masividad ha sido un factor constantemente tenido en cuenta, dado que no sólo requiere de la corrección de cientos de tareas, sino que también requiere de una gran coordinación entre el equipo docente.

La actual cursada agregó un elemento más a nuestro desafío: la virtualidad. A la fecha, aún restando dos evaluaciones y la entrega final del trabajo integrador, un 58% de estudiantes ya han obtenido los puntos mínimos requeridos y un 72% tienen Python-plus suficientes para obtener entregas adicionales.

## Bibliografía

1. Propuesta de Estándares para la Disciplina Informática. Propuesta de estándar de la RedUNCI para la disciplina Informática (2018) <http://redunci.info.unlp.edu.ar/docs/Estandares%20RedUNCI%20Septiembre%202018-Libro-2.pdf>
2. González J., Wagenaar, R.: Tuning educational structures in Europe: Informe final fase uno. Bilbao: Universidad de Deusto – Universidad de Groningen (2003)
3. Vallejo B., López Jara N.R: Análisis crítico del modelo basado en competencias en la Universidad (2019)
4. Díaz Barriga, Á.: El enfoque de competencias en la educación. ¿Una alternativa o un disfraz de cambio?. Perfiles Educativos. Instituto de Investigaciones sobre la Universidad y la Educación Distrito Federal, México (2006)
5. Perrenoud, P.: Diez nuevas competencias para enseñar. Barcelona: Graó (2004)
6. Tardif, J.: Desarrollo de un programa por competencias: de la intención a su implementación. Profesorado. Revista de currículum y Formación del Profesorado (2008)
7. Lovos, E., González, A.H.: Moodle y VPL como soporte a las actividades de laboratorio de un curso introductorio de programación. X Congreso de Tecnología en Educación & Educación en Tecnología (2014)

# Actividades de Investigación Científica con docentes y alumnos de grado en Informática y Ciencias de la Computación\*

Germán Bianchini<sup>1</sup>, Paola Caymes-Scutari<sup>1,2</sup>, Patricia Ontiveros<sup>3</sup>, Carina Rotella<sup>3</sup>, Sandra Tagarelli<sup>4</sup>, Sergio Salinas<sup>4</sup>, Mariela Galdamez<sup>1</sup>, Pamela Chirino<sup>1</sup>, Karvin Díaz<sup>1</sup>, Alejo Ponce de León<sup>1</sup>, and Renzo Suarez<sup>1</sup>

<sup>1</sup>Laboratorio de Investigación en Cómputo Paralelo/Distribuido (LICPaD)  
Facultad Regional Mendoza/Universidad Tecnológica Nacional  
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

<sup>2</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

<sup>3</sup>Laboratorio de Gobierno Electrónico, Departamento de Ingeniería en Sistemas de Información

<sup>4</sup>Laboratorio de Analítica de Datos, Departamento de Ingeniería en Sistemas de Información {gbianchini,pcaymesscutari,pontiveros}@utn.frm.edu.ar  
{carinarotella,stagarelli,s4salinas}@gmail,mariela.galdamez@hotmail.com  
{pamelaachirino,karvindiaz,alejo.poncedleon,renzogsuarez}@gmail.com  
<http://www.frm.utn.edu.ar/licpad/>

**Resumen** La tarea de investigar es un proceso sistemático, organizado y objetivo, cuyo propósito es responder a un interrogante para así incrementar el conocimiento y la información sobre algo desconocido. La capacidad investigativa resulta un elemento clave e imprescindible para hacer frente a los retos que los nuevos profesionales de las diversas ingenierías encontrarán en el mundo laboral actual y futuro. Sin embargo, es evidente que esta capacidad no se alcanza por el simple hecho de obtener un título universitario. En el presente trabajo se describe cómo 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 tres grandes áreas: Cómputo Paralelo, Analítica de Datos y Gobierno Electrónico. Dentro del mismo, los estudiantes y docentes participantes lograron planificar, organizar y avanzar en sus desarrollos, a pesar de ser un año particular en virtud de las complicaciones que impuso la pandemia por COVID-19, alcanzando buenos resultados y una activa participación en eventos científicos.

**Keywords:** Investigación, Formación, Alumnos, Docentes, Metodología

## 1. Introducción

De acuerdo a Tamayo y Tamayo [1], la investigación “es un proceso que, mediante la aplicación del método científico, procura obtener información relevante

\* El proyecto cuenta con la acreditación, homologación y financiamiento de la Universidad Tecnológica Nacional a través del proyecto PID TEUTIME0007658TC.

y fidedigna, para entender, verificar, corregir o aplicar el conocimiento”. Es una actividad enfocada en conseguir, mediante observación y experimentación, nueva información y conocimiento necesario para ampliar los diversos campos de la ciencia y la tecnología [2]. Sin embargo, el principal obstáculo que se encuentra en las instituciones educativas universitarias es que las actividades de investigación se encuentran desvinculadas, en general, de las actividades académicas de grado, y por tanto esta disociación mantiene alejados a los alumnos y a muchos docentes de estas tareas de enfoque científico. Por tal motivo, en el presente proyecto se busca 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 tres áreas generales: Cómputo Paralelo, Analítica de Datos y Gobierno Electrónico. Se busca combinar en un mismo proceso la formación teórica y práctica de los investigadores iniciales para que éstos experimenten los pasos involucrados en un ámbito de investigación, recorriendo todas las etapas del proceso científico. Fomentando las competencias investigativas (capacidad para abordar proyectos de investigación y desarrollo, integrando a tal efecto equipos interdisciplinarios en cooperación, o asumiendo el liderazgo efectivo en la coordinación técnica y metodológica de los mismos) se contribuye también a incrementar el número y calidad de las investigaciones realizadas en la institución, así como también a mejorar el pensamiento crítico de docentes y alumnos frente a nuevos retos no sólo durante su formación profesional sino en la vida laboral. Para todo docente y estudiante, ya sea de grado o posgrado, sería recomendable experimentar el proceso de dirigir o participar en un proyecto de investigación, puesto que afianzar los conocimientos por medio de la investigación resulta fundamental para completar la formación profesional. Por tal razón, 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 y formación previa de cada docente o alumno, teniendo en cuenta además para estos últimos el estadio en la carrera.

La investigación debería ser considerada pilar en todas las actividades académicas de nivel superior, y evidentemente esto debería incluir la participación activa tanto de docentes como de estudiantes. En diversas instituciones alrededor del mundo ha comenzado a entenderse esta noción, y paulatinamente se han impulsado planes que mejoran el nivel de la educación superior a través de modelos académicos que contemplan a la investigación científica como un eje fundamental para el desarrollo de una sociedad del conocimiento equitativa y democrática. Por ejemplo, el Centro Universitario de Los Lagos, de la Universidad de Guadalajara [3] (México), se ha propuesto como un centro regional que entiende a la investigación como una actividad irremplazable, necesaria para elevar el nivel y calidad de la educación. Una de sus políticas, en relación con la actividad docente de los investigadores, ha establecido que no debe centrarse en el posgrado, sino que antes debe asegurarse que cada investigador imparta por lo menos una asignatura en los programas de pregrado. De esta manera, se intenta propiciar un

elemento más de conexión entre la docencia y la investigación, sobre todo en las licenciaturas, brindando así la posibilidad de acceder a conceptos o contenidos del ámbito de la investigación [4].

Otro ejemplo es la Facultad de Medicina de la Universidad de Navarra (España) [5], donde se impulsa la actividad investigativa desde distintos departamentos y centros, con el objetivo de repercutir en la formación directa del alumno. La Universidad Maimónides [6] (Buenos Aires, Argentina) también busca acercar a los docentes y estudiantes a la investigación, al desarrollo y la transferencia tecnológica y la creación de empresas de base tecnológica o empresas intensivas en la aplicación de conocimiento. En el caso de la Universidad Nacional Autónoma de Honduras, se ha planteado un Plan Táctico de Reforma Universitaria, que establece una política institucional a nivel de investigación, desarrollo tecnológico e innovación, incorporando a estudiantes y catedráticos en los procesos de investigación a niveles de grado y posgrado [7], sólo por mencionar algunos ejemplos en distintos puntos geográficos. Es de gran importancia que los estudiantes se incorporen a grupos y proyectos de investigación, pero es tan o más importante aún que aprendan de éstos los procesos de la ciencia, de modo que tengan la posibilidad de agregarlo a sus saberes y asimilarlo, y no que simplemente se los convoque para tareas mecánicas o generación de insumos.

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 (salvo aquellos pocos que participan de algún grupo consolidado) están normalmente ajenos a estas actividades, las cuales requerirían una mejor y mayor difusión, con la orientación adecuada para el alumnado, y 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 es costoso, sobre todo considerando que los resultados normalmente se consiguen 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 sustancial para el desarrollo de las instituciones y la sociedad.

En vista de esta realidad, es elemental que los docentes proporcionen a los estudiantes las herramientas necesarias para conocer, interpretar y comprender la situación de la región, del país y el mundo, que se encuentran en constante transformación. Sólo de esta forma éstos podrán intervenir de forma compro-

metida y crítica en su entorno para mejorar la calidad de vida de las personas a través de nuevos aportes a la sociedad.

## 2. Líneas de investigación y metodología

Con el fin de alcanzar la meta global de realizar todo el proceso de investigación científica, se propusieron los siguientes objetivos particulares:

- Promover una cultura de investigación que estimule la vocación científica de alumnos y docentes e iniciar a los estudiantes en los rudimentos del procesamiento paralelo.
- Proponer temáticas adecuadas a los estudiantes y evaluar las propuestas de los mismos en sus áreas de interés.
- Generar e involucrar la participación de los estudiantes en los distintos mecanismos establecidos para la vinculación científica.
- Coordinar y guiar las actividades (individuales y/o grupales) en todo el proceso involucrado en las tareas de investigación.
- Mejorar las habilidades en el uso de las Tecnologías de la Información y Comunicación (TIC's), como una herramienta para realizar investigación por parte de los alumnos.
- Generar intercambio de ideas, discusión y aporte mutuo entre los alumnos con la guía de los docentes de la temática.
- Entender y experimentar el proceso de redacción, publicación y presentación de los aportes académico-científicos.
- Difundir el conocimiento que se genere como resultado de las tareas de investigación de los participantes.
- Promover la participación e incorporación de docentes y estudiantes de grado en el Programa de Categorización de Investigadores.

Inicialmente se plantearon distintos temas de investigación, de los cuales se seleccionaron algunos 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 sugeridas por los alumnos-becarios fueron varias, a saber: Paralelismo en algoritmos de aprendizaje para redes neuronales [8], Transformación automática de imágenes satelitales para su aplicación en sistemas de predicción de incendios forestales [9], Paralelización del método de tránsitos para la detección de exoplanetas [10] y, finalmente, Predicción del avance del dengue en zonas endémicas [11]. Dos de las mismas fueron las elegidas: una, derivada de la temática de Redes Neuronales, abarcando diversos métodos de Inteligencia Artificial, y la mencionada para Predicción del Avance del Dengue, sobre la cual trabajan en la actualidad un equipo de tres estudiantes becados.

Por su parte, los docentes intervinientes de la rama Analítica de Datos orientaron sus temas de investigación hacia Indicadores de procesos y medición, temática vinculada con sus respectivas tesis en curso, dado que ambos (Ing. Sandra Tagarelli y Dr. Sergio Salinas) se encuentran realizando la Maestría en Inteligencia de Negocios de la Universidad HNU de Alemania y la UTN-FRM.

Finalmente, los docentes de la rama Gobierno Electrónico se centraron en la temática novedosa de aplicación de Blockchain Federal (BFA) sobre la gestión de documentos judiciales [12].

Para cada uno de los investigadores iniciales participantes, tanto alumnos como docentes, se ha planteado seguir una misma metodología (inducción-deducción), siguiendo los pasos del método científico. Dicho de forma resumida, comenzar con preguntas generales hasta llegar a enfocarse en un aspecto específico y el diseño de la investigación, para finalmente llegar a la conclusión. El esquema de trabajo que se planteó al inicio del proyecto, vinculado a objetivos de carácter general que pudiesen aplicarse a cada tema específico de investigación dentro del área del Cómputo Paralelo, también puede adaptarse a las otras temáticas, siendo el indicado en el siguiente listado:

1. Estado del Arte: Revisar el estado del arte en lo relativo al tema de trabajo seleccionado; Revisar fundamentos teóricos del tema de trabajo; Revisar fundamentos teóricos sobre procesamiento paralelo; Buscar diversas alternativas de solución al problema planteado.
2. Desarrollo: Proponer y diseñar prototipos; Implementar el diseño; Aplicar métricas para el análisis de la performance (speedup, tiempo de ejecución, eficiencia, escalabilidad, overhead, etc.); Testing, corrección y depuración.
3. Validación: Aplicar el desarrollo sobre casos de estudio, en caso de haberlo, o sobre casos sintéticos; Depurar; Medir y evaluar los resultados obtenidos.
4. Documentación/Publicación: Documentar los resultados y procedimientos; Publicar los resultados obtenidos; Redactar el informe final.

Es importante mencionar que el proceso metodológico es iterativo en dos niveles: se organiza con iteraciones en cada una de las cuatro etapas y también a nivel general para todo el proyecto, ajustando en cada vuelta las aproximaciones para, en función de la etapa que se trate, corregir e ir mejorando los resultados hallados para así lograr un resultado de mayor calidad como salidas entregables de cada etapa.

### 3. Desarrollo de las actividades

La modalidad usual de trabajo en el LICPaD 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 propia del Laboratorio, textos electrónicos, artículos científicos, etc.) y con la documentación continua que se solicita mantengan al día los investigadores, la cual resulta 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 (Fig. 1 y Fig. 2), 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

6 Bianchini G. et al.

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.



**Figura 1.** Seminarios de alumnos investigadores acerca de sus temas de trabajo.

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, de forma casi inmediata, que adecuar la forma de trabajo hacia un esquema que nos permitiese seguir avanzando, posiblemente a un ritmo más lento que el deseado.



**Figura 2.** Charla con alumnos durante una visita al nodo informático de la UTN-FRM y durante las reuniones del grupo.

Las reuniones presenciales se vieron reemplazadas por un esquema de videoconferencias mediante plataformas de videollamadas (Fig. 3). 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[13] 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 (Moodle[14]), a través de mails o por mensajes de whatsapp[15].



**Figura 3.** Encuentros en la modalidad de reunión virtual, mediante la plataforma Zoom, con alumnos investigadores en el marco de la cuarentena por COVID-19.

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 contenidos.

El acceso a los recursos físicos (equipamiento, bibliografía, consumibles, etc.), también planteó una nueva dificultad, ya que poder realizar ensayos y experimentos para dar soporte a las validaciones de los trabajos, sin tener acceso a los espacios de trabajos tradicionales, complejizó las tareas.

En lo atinente al acceso al clúster de investigación de la Facultad, el mismo se ha efectuado mediante acceso remoto, de modo que los alumnos pudiesen tener la posibilidad de 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[16], Google Meet[17], Jitsi Meet[18] o Webex Meeting[19].

#### 4. Resultados hallados

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 (y a las cátedras de la carrera Ingeniería en Sistemas de Información vinculadas como Sistemas de Gestión, Análisis de Sistemas, Modelado de procesos de negocio, Gobierno Electrónico y Computación Paralela), 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 (Workshop de Investigadores en Ciencia de la Computación, llevado a cabo este año en la UNPA, Santa Cruz) mediante cinco trabajos:

- Sintonización de Aplicaciones científico/ingenieriles: un proceso de medición-mejora para incrementar la eficiencia.
- 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.
- Diseño de un sistema de medición de desempeño para Moodle en Educación Superior.
- Visualización de datos en un Tablero de Comando aplicado a plataformas de educación a distancia en el nivel Superior.
- Formación de docentes y alumnos de grado como Investigadores Científicos Iniciales en las áreas de Informática y Ciencias de la Computación

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. Algunos han sido:

- Participación y asistencia en el WICC 2020 (UNPa, Santa Cruz, Argentina).
- Asistencia al IPDPS 2020 (New Orleans, Louisiana, USA)
- Participación y asistencia en el webinar Transformación digital y pandemia: Los riesgos ocultos para la Educación en Ingeniería (CONFEDI, Argentina)
- Participación en el curso Moodle para docentes FRM (UTN-FRM, Mendoza, Argentina)

En la actualidad, los docentes y alumnos investigadores continúan con el avance de sus tareas y se ha planificado la participación en el IDETEC 2020 (UTN-FRVM, pospuesto a último momento para 2021); CONAIISI 2020 (Congreso Nacional de Ingeniería Informática - Sistemas de Información, a realizarse en la UTN-FRSF, Córdoba) y en el IV CICCISI (Congreso Internacional de Ciencias de la Computación y Sistemas de Información, a efectuarse en la UCh, Mendoza).

Como ya se indicó, el objetivo y ganancia a largo plazo al involucrar a los estudiantes en grupos y otras comunidades y redes de investigación, es el desarrollo de proyectos que puedan alcanzar un alto impacto y el tratamiento de temas de interés dentro de las distintas áreas del conocimiento. Este es un proyecto que busca contribuir al crecimiento y desarrollo de investigadores noveles, para lo cual se cuenta con los directores e investigadores formados como guía. Es importante para los estudiantes contar con la posibilidad de acceder a becas de investigación (si bien la remuneración de las mismas resulta muy baja), y por

esta razón, en principio, el proyecto está pensado con alumnos-becarios, lo cual les permite dedicar cierta cantidad de horas semanales a la tarea de investigar de modo tal de llegar a cumplir los objetivos formales establecidos. Además, dicha actividad les permite la realización de proyectos iniciados por ellos mismos, así como contar con fondos mínimos para la compra de equipamiento específico y bibliografía, inscripción en congresos, gastos de traslados y estadías, etc., lo que es extremadamente valioso para aquellos que quieran iniciarse en el camino de la investigación, estén en la etapa de desarrollo de su proyecto de fin de carrera o manifiesten una vocación científica.

## 5. Conclusiones

Consideramos importante que los alumnos de grado tengan una visión concreta de la actividad científica, de modo que, además de incorporar herramientas a sus capacidades profesionales, al concluir la carrera puedan considerar, con conocimiento y fundamento, la posibilidad de continuar una carrera como investigador (ya sea mediante su incorporación a cursos de magister, programas de doctorado, etc.). Por otra parte, también es importante que la planta de docentes/investigadores se vea incrementada y sea objeto de una renovación constante, por lo que deben establecerse medios que permitan la incorporación temprana de los estudiantes a este tipo de actividades, lo que ayudará a la evolución y calidad de los centros y grupos de investigación del Departamento, Facultad y Universidad.

En la era del conocimiento y de los avances tecnológicos, las habilidades del pensamiento crítico y lectura analítica proporcionarán a los jóvenes las herramientas necesarias para enfrentar los desafíos de la actualidad, así como las del campo laboral, dado que el mundo necesita de jóvenes profesionales que puedan resolver tales problemas y que, de la misma forma, presenten nuevos paradigmas. Resulta primordial fortalecer la capacidad de análisis de los alumnos, interpretación de los fenómenos del entorno, sistematización de la información, uso adecuado de la tecnología, con el objetivo general de que sean capaces de construir su conocimiento a partir de la propia experiencia, que es lo que finalmente se ha buscado en este trabajo.

La experiencia llevada a cabo en este proyecto, que se ha detallado en el presente trabajo, es de gran importancia tanto para los investigadores (alumnos y docentes) en formación, como para los investigadores formados, puesto que construye un puente de intercambio y comunicación en donde cada participante aprende y se retroalimenta con la experiencia grupal. En esta oportunidad, consecuencia del brote de coronavirus (Covid-19) declarado Pandemia Global por la Organización Mundial de la Salud (OMS), todos nos vimos inmersos en un contexto social de crisis que exigió la búsqueda de soluciones prontas y pertinentes para continuar los procesos de enseñanza y aprendizaje. Como se comentó en el desarrollo del trabajo, nos encontramos con esta dificultad adicional de vernos obligados a transitar otra modalidad de trabajo que hubo de instalarse y organizarse rápidamente para que las actividades pudiesen continuar.

10 Bianchini G. et al.

La situación actual plantea un reto que, aunque en principio surgió como un hecho evidentemente negativo, a largo plazo debe ser visto como una oportunidad, dado que fomenta y potencia la colaboración interdisciplinar entre científicos, facilitando la circulación de datos como nunca antes y difundiendo la divulgación entre la sociedad.

## Referencias

1. Tamayo y Tamayo, M.: El proceso de la investigación científica. Limusa Noriega Editores (2002)
2. Murguía, E.A.; López Reyes, L.J.: Incorporación temprana a la investigación. Cuadernos de Educación y Desarrollo. Vol.3, N°28 (2011)
3. Centro Universitario de Los Lagos. Universidad de Guadalajara. <http://www.lagos.udg.mx/> (fecha de consulta: abril de 2019)
4. Castelán Rueda, R.: Rector de Centro. Informe Anual de Actividades 2009. Centro Universitario de los Lagos (2009)
5. Facultad de Medicina, Universidad de Navarra. <https://www.unav.edu/web/facultad-de-medicina/la-investigacion-en-la-facultad> (fecha de consulta: abril de 2019)
6. Universidad Maimónides. <https://www.maimonides.edu/investigacion/> (fecha de consulta: abril de 2019)
7. Escalante H., Corrales R., Barahona M.: La investigación científica en los estudiantes de grado y posgrado. Revista Portal de la Ciencia, UNAH. Nro.11. pp.19–25 (2011)
8. 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)
9. Zúñiga A., Arancibia J., Méndez-Garabetti M., Bianchini G., Caymes-Scutari P., Tardivo M.L.: Propuesta de Metodología de Generación Automática de Perímetros de Área Quemada usando Imágenes Satelitales MODIS para su Aplicación en un Sistema de Predicción de Incendios. CONAIISI 2018. Mar del Plata. Publicación on line. ISSN 2347-0372 (2018)
10. Hidalgo J., Galdamez M., Chirino P., Caymes-Scutari P., Bianchini G.: Detección de exoplanetas y cómputo paralelo. In: Libro de Actas 7mo CONNAIISI 2019. pp.1404–1413. Edit. UNLaM, San Justo (2020)
11. 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)
12. Blockchain Federal Argentina. <https://bfa.ar> (fecha de consulta: julio de 2020)
13. Youtube. <https://youtu.be.com> (fecha de consulta: julio de 2020)
14. Moodle - Open-source learning platform. <https://moodle.org> (fecha de consulta: julio de 2020)
15. Whatsapp - Mensajería confiable. Simple. Segura. <https://www.whatsapp.com> (fecha de consulta: julio de 2020)
16. Videoconferencias, Conferencias Web, Seminarios. <https://zoom.us/> (fecha de consulta: julio de 2020)
17. G Suite essentials: la manera más simple de trabajar en equipo. <https://gsuite.google.com/> (fecha de consulta: julio de 2020)
18. Jitsi Meet. <https://meet.jit.si/> (fecha de consulta: julio de 2020)
19. Reuniones en línea, videoconferencias gratuitas. <https://www.webex.com/es/video-conferencing.html> (fecha de consulta: julio de 2020)

## Experiencias de articulación para formar emprendedores en los ciclos lectivos 2019 y 2020.

Sonia I. Mariño y Viviana R. Bercheñi

Departamento de Informática

Facultad de Ciencias Exactas y Naturales y Agrimensura,

Universidad Nacional del Nordeste, 9 de Julio 1449

[simarinio@yahoo.com](mailto:simarinio@yahoo.com), [viviber@hotmail.com](mailto:viviber@hotmail.com)

**Abstract.** El artículo describe intervenciones con miras a iniciar y fortalecer el espíritu emprendedor en la disciplina Informática. Se enfocan en actividades de articulación entre las asignaturas Economía Aplicada y Proyecto Final de Carrera en los ciclos lectivos 2019 y 2020. Se sintetiza el método aplicado. Los resultados exponen dos intervenciones realizadas en los ciclos lectivos 2019 y 2020 en que se aprecian distintos contextos de realización, uno de ellos presencial y otro mediado por tecnologías sincrónicas y asincrónicas. Los resultados indican la necesidad de continuar fortaleciendo la perspectiva planteada orientada a fortalecer el espíritu emprendedor en la formación de grado.

**Keywords:** formación universitaria, formación para el trabajo, TIC, emprendimientos, planes de negocios

### 1 Introducción

Las universidades como espacios de construcción, transferencia y enseñanza de conocimientos innovan constantemente ante una compleja sociedad del conocimiento en que emergen nuevas cuestiones sociales, económicas y culturales que deben contemplar.

En este contexto complejo e incierto el desarrollo de emprendimientos y la formación en esta temática, se comprende como la capacidad que tienen los individuos para crear nuevas empresas que tengan la característica de ser sustentable a través del tiempo. [1]

En [2] sintetizan información relevada a partir de más de un centenar de fuentes provenientes de libros y artículos y destacan que “a nivel general la educación emprendedora ha contribuido significativamente a su entorno y la producción de empresas” en particular en los países en desarrollo.

Diversos autores también tratan la formación emprendedora desde ámbitos académicos como una estrategia ante los problemas de desempleo y exclusión. [2] [3]

En [4] se resume una investigación bibliométrica referida a la intención emprendedora de estudiantes universitarios en el periodo 1996 a 2015 a partir de datos relevados desde la base de datos Scopus. La información producida muestra un

crecimiento exponencial y como nuevos autores abordan la temática desde redes existentes. Un estudio orientado a determinar las diferencias identificadas en la formación en emprendimiento entre dos universidades, una de carácter público y otra de carácter privado se presenta en [5].

Otra cuestión sobresaliente en la formación universitaria en torno al emprendimiento es su carácter de transversalidad en el currículo. A tal efecto se menciona a [7] quienes consideran que se deben diseñar programas de formación que adopten un “aprendizaje holístico e integrado que permite que los participantes se beneficien del trabajo colaborativo interdisciplinar”. Por su parte, en [2] consideran que esta educación debe “trascender su condición optativa, y ser compulsiva dentro de la educación, de modo que se cree un formato de educación transversal”. Las afirmaciones anteriores se reflejan en esta propuesta dado que las experiencias que se sintetizan en el presente artículo resumen articulaciones transversales que se concretan en la formación de licenciados desde dos asignaturas que corresponden a los últimos años de la carrera.

En Ovalle-Toledo et al. [8] se expone un estudio bibliométrico en torno a tópicos de emprendimiento relacionado con “habilidades” o “conocimiento” o “capacidades” o “desarrollo social” localizados en la ISI Web of Science en el periodo 2000 al 2016. Su importancia radica que el enfoque centrado en “habilidades” o “conocimiento” o “capacidades” es viable de identificar en las propuestas de planes de negocio según el abordaje expuesto en este artículo.

Como se expresó en [1], la Facultad de Ciencias Exactas, Naturales y Agrimensura a través de la carrera de Licenciatura en Sistemas refiere antecedentes que se orientan en este sentido en el plan de estudios de la carrera Licenciatura en Sistemas de Información o Plan LSI 2009 [9]. Éste refleja los estándares establecidos por la Resolución Min. 786/09 [10] que establece los contenidos curriculares básicos, la carga horaria mínima, los criterios de intensidad de formación práctica y los estándares para la acreditación de carreras de la disciplina Informática. Así, se incorporaron los Contenidos Curriculares Básicos que corresponden al terminal que la comprende [11].

Como continuidad y producto de la actualización constante de la Disciplina Informática, los estándares establecidos para la acreditación de las Ingenierías, se trabajaron en los años 2012 y 2013 en el documento “Descriptor Curriculares de la disciplina Informática”, desarrollado en el marco de la Red UNCI [11]. Los Descriptores tratan de ser una cobertura de los temas que formarán parte de las recomendaciones curriculares por terminal. En esta etapa no se trató de desarrollar “nuevos estándares”, sino de trabajar en que los descriptores cubran con una granularidad razonable los temas que se dictan en las diferentes carreras de Informática del país. Particularmente, el Área Aspectos Legales y Profesionales destina 50 horas dentro de la currícula de Licenciatura en Sistemas de Información, y se establece como obligatorio el contenido denominado “Innovación y Emprendedurismo”. Esta temática es muy importante en el actual ámbito académico y profesional, reconocimiento que se plasma desde hace unos años atrás.

Se coincide con [2] quienes expresan que no solo se deben crear y desarrollar programas de educación emprendedora sino que estos deben ser mantenidos. En este sentido y en el contexto de esta universidad, cabe mencionar las acciones promovidas por la Agencia de Innovación y Desarrollo, dependiente de la Universidad Nacional

del Nordeste, AGENTIA [12], y las articulaciones diseñadas y ejecutadas desde el año 2017 entre las asignaturas Proyecto Final de Carrera (PFC) y Economía Aplicada (EA).

En el artículo se sintetizan dos intervenciones de articulación concretadas en los ciclos lectivos 2019 y 2020 enmarcadas en la carrera Licenciatura en Sistemas de Información, involucrando a estas asignaturas de lógica compartida.

## 2. Metodología

La metodología propuesta para el desarrollo de las estrategias de intervención articuladora entre ambas asignaturas consiste en:

Fase 1. Identificación de asignaturas del plan de estudio con perspectivas de desarrollar la formación emprendedora con énfasis en TIC. En este caso en particular, se seleccionaron las asignaturas Economía Aplicada y Proyecto Final de Carrera, correspondientes al cuarto y quinto año de la carrera respectivamente.

Fase 2. Diseño de actividades de intervención para fomentar el espíritu emprendedor, orientadas al logro de aprendizajes significativos para aportar a la formación profesional. En esta fase se contempló la planificación de las actividades que involucra la definición de:

- Materiales y equipos necesarios: se ajustan a las circunstancias. En el ciclo lectivo 2019 se realizó presencialmente, considerando el aula y horario en que habitualmente se desarrollan las clases de las asignaturas. En el ciclo lectivo 2020 se redefinieron intervenciones virtuales sincrónicas y asincrónicas atendiendo al aislamiento decretado a nivel nacional y al que la Universidad se adhirió.
- Duración total de la actividad: Se adecuaron a condicionantes en cada ciclo lectivo.
- Cronograma de las clases: en el ciclo lectivo 2019 se planificaron dos intervenciones, en el primero y segundo cuatrimestre. Los resultados sugirieron información de retroalimentación para el ajuste y desarrollo en el ciclo lectivo 2020, precisamente se concretó a fines del mes de mayo de 2020.
- Docentes intervinientes: Expositores docentes responsables de las asignaturas involucradas y en el año 2019 se contó con un referente del tema perteneciente al equipo AGENTIA.
- Necesidades logísticas especiales:
  - o Desarrollo de las actividades previstas en el horario y espacio habituales en que se desarrollan las clases de las asignaturas involucradas, cabe aclarar que un día de la semana coinciden.
- Otras aclaraciones
  - o Se tramita una resolución de aprobación de la instancia de articulación entre asignaturas de la carrera.
  - o Se prevé replicar esta actividad en los próximos ciclos lectivos incluyendo a distintas asignaturas y ajustando según información de retroalimentación resultante del análisis de implementaciones previas.

Fase 3. Realización de la intervención Se abordó la definición y elaboración de un Plan de Negocios viable de realizar. Éste se constituye en un instrumento preliminar que potencialmente orientaría en la definición de un proyecto de fin de carrera. Se

abordó en Economía Aplicada con intervención de Proyecto Final de Carrera, siendo el público objetivo los estudiantes que cursan la primera de estas asignaturas.

Fase 4. Evaluación de la intervención. Finalizadas las actividades, se capitaliza la experiencia para generar la información necesaria que permite ajustar futuras intervenciones.

### **3. Resultados**

Las intervenciones de articulación orientadas al emprendimiento que aportan a la formación profesional y con énfasis en lograr aprendizajes significativos que se describen se concretaron en el período 2019 y 2020 en las asignaturas “Proyecto Final de la Carrera” y “Economía Aplicada”. Como antecedentes inmediatos la organización conjunta se plasmó en el desarrollo de los seminarios que se llevaron a cabo simultáneamente y se denominaron “La Acción Emprendedora” [13] [14] [15].

En este contexto, como recursos didácticos de base se organizan seminarios orientados a fomentar el espíritu emprendedor en asignaturas de cuarto y quinto año y dictados por profesionales del equipo de la Agencia correspondiente de la universidad. Estos seminarios se consideran el elemento disparador para la construcción de un plan de negocios o para completar el formulario idea-proyectos, base para la presentación en distintos programas de fomento a iniciativas emprendedoras [1].

#### **3.1 Contexto de la asignatura Proyecto Final de Carrera**

Las licenciaturas, como carreras universitarias, se caracterizan por requerir el desarrollo de un trabajo final integrador para lograr la titulación. En el caso de la Licenciatura en Sistemas de Información de la FaCENA, este espacio académico se denomina Proyecto Final de Carrera. Su objetivo primordial es completar la formación académica y profesional de los alumnos, posibilitando la integración y utilización de los conocimientos adquiridos durante sus años de estudio para la resolución de problemas de índole profesional, académico y científico.

Con miras de aportar a la definición del proyecto integrador, los docentes de la asignatura PFC realizan numerosas actividades en otras asignaturas previas del plan de estudio –por ejemplo ubicadas en el tercer y cuarto año. Lo expuesto tiene por finalidad de lograr una aproximación a la definición del objeto de estudio que los estudiantes tratarán en el PFC.

En este caso, se prioriza que la idea proyecto emprendedora tratada en la asignatura EA se puede transformar en el PFC. Como se expresó en [1] se trata de un producto de conocimientos y habilidades desarrollados en la carrera y que pueden repensarse, replantearse, profundizarse y actualizarse en el proyecto final.

#### **3.2 Contexto de la asignatura Economía Aplicada**

La asignatura Economía Aplicada es de carácter cuatrimestral y se dicta en el cuarto

año de la carrera LSI. Su objetivo general es realizar el nexo entre los conocimientos técnicos que los alumnos adquieren de las asignaturas específicas de la Carrera, con los objetivos estratégicos de las organizaciones, y de éstas con su entorno. Al aprobar el curso, los alumnos están en condiciones de integrar la tecnología informática con la misión y visión de las organizaciones con fines de lucro.

Desde la asignatura Economía Aplicada, el aporte a la propuesta se realizó a través del desarrollo de trabajos prácticos integradores que se materializan en la presentación de un proyecto de negocio (en equipo con un número máximo de integrantes de hasta cuatro alumnos) con desarrollo teórico y práctico de su rentabilidad proyectada a través de un flujo de fondos.

Se propusieron como actividades relacionadas con el desarrollo del Plan de Negocios lo siguiente:

- Búsqueda de información en línea vinculada a emprendimientos informáticos exitosos en América Latina. La participación en foros de discusión habilitados a tal fin, favoreció el interés de los alumnos por la propuesta.
- Entrevistas a agentes y funcionarios municipales, provinciales y nacionales para identificar costos impositivos inherentes al proyecto.
- Definición de la inversión inicial tentativa, precios, demanda y costos asociados.
- Exposición grupal explicativa de cada plan de negocio de manera presencial en el cursado 2019 y virtual, con recursos sincrónicos y asincrónicos en el año 2020.

Durante el período 2020 debido a la crisis sanitaria generada por el COVID-19, las actividades previstas se realizaron a través de intervenciones sincrónicas y asincrónicas utilizando para estas últimas canales en Youtube habilitados a tal fin. Se utilizó además el aula virtual de la Universidad, (basada en la plataforma Moodle) para disponer a los alumnos la bibliografía necesaria, el material de la asignatura y las presentaciones digitales con audios explicativos de temas vinculados al diseño y desarrollo de emprendimientos. Las intervenciones sincrónicas se desarrollaron a través de Google Meet y Zoom.

### **3.3. Experiencias de intervención en 2019 y 2020**

Dado que se trata de una intervención desde ámbitos académicos, se definen los objetivos didácticos. Así, la inclusión de temas relacionados a Emprendedorismo se orienta a proporcionar competencias y habilidades vinculadas con la práctica profesional y brindar al estudiante una visión más amplia de su carrera y su potencial aporte a la comunidad desde su propia capacidad para autogenerar planes de negocio.

Desde esta perspectiva, siguiendo lo expuesto en [1] [17] se mencionan como objetivos vinculados al desarrollo de conocimientos, actitudes y habilidades vinculados a esta intervención.

En particular los que se relacionan con saberes técnicos y cognitivos que se pretende trabajar durante la propuesta e incluyen aspectos técnicos específicos de las asignaturas en la que se aplica la actividad y plasmados en sendas planificaciones.

- Comprender los fundamentos teóricos que sustentan el desarrollo de emprendimientos y especialmente aquellos de base tecnológica.

- Entender y tomar contacto con la realidad del futuro profesional de sistemas informáticos.
- Conocer y manejar las distintas áreas de la Informática, a partir de los cuales desarrollar un Plan de de negocios de base tecnológica.
- Identificar nichos, necesidades o ideas potencialmente aplicables al diseño de negocios.

En función de estas prioridades, se presentan como objetivos alcanzables por el alumno de 4to año a partir de la presente propuesta didáctica sustentada en las intervenciones de articulación entre las asignaturas mencionadas, y centrada en el abordaje de emprendimientos:

- Un aprendizaje significativo, ya que el alumno usa conocimientos previos para el análisis del problema y la propuesta de soluciones facilitando la generación de los conceptos aplicados en las asignaturas involucradas.
- La coordinación y trabajo en equipo.
- La elaboración de propuestas creativas e innovadoras que generen conocimientos y promuevan la identificación de oportunidades de negocios.
- El desarrollo de habilidades comunicativas: capacidad de explicar, interrogar y de responder; uso de un lenguaje especializado, socializarse, atender y comprender a los otros, interacción con otros estudiantes son una buena preparación para los aspectos humanos de la gestión.
- El desarrollo de habilidades específicas (comprensión lectora, valoración de la información, uso de informaciones y de conocimientos de expertos, roles, toma y argumentación de decisiones, previsión de consecuencias, expresión escrita y oral) entre otras.

Los objetivos procedimentales son los ligados a la aplicación de los conocimientos específicos y a su abordaje en el espacio de la creación de una nueva empresa. Siguiendo lo expuesto en [1] [16] [17] se mencionan: Buscar y localizar información inherente a la temática que permita determinar nichos de desarrollo, Localizar y determinar oportunidades para el desarrollo de emprendimientos, iniciativa y búsqueda de oportunidades, Medir, calcular y decidir la factibilidad de desarrollo asociado al emprendimiento, Persistir con las ideas, proyectos y empresas, Cumplir los compromisos asumidos, Desarrollar las tareas con calidad y eficiencia, Correr riesgos calculados, Fijar metas y objetivos, Planificar sistemáticamente y controlar con miras a implantar y lograr la evolución del emprendimiento, Persuadir y crear redes de apoyo, Desarrollar la creatividad, Dominar los lenguajes habituales de comunicación y aquellas herramientas necesarias para la elaboración y presentación de planes de negocio, Tomar decisiones adecuadas.

Los objetivos ligados al desarrollo de actitudes intentan que los estudiantes reconozcan acciones que permitan aplicar los conocimientos específicos tratados en la carrera en el espacio de una nueva empresa o en el proceso de creación de la misma, y logren así desempeñar el rol emprendedor de manera efectiva. Se retoman los expuestos en [1] [16] [17]: Iniciativa, Sentido de oportunidad, Compromiso (perseverancia y determinación), Facilidad para la toma de decisiones, Autoconfianza e independencia, Razonamiento y actitud crítica, Capacidad comunicativa, Habilidades para las relaciones interpersonales, Optimismo, Creatividad, Innovación, Trabajo en equipo, Tolerancia al fracaso (persistencia), Visión (tener en mente el

modelo de negocio), Manejo de información, Solución de problemas, Predisposición al aprendizaje autónomo y formación continuada, especialmente abordando una visión tecnológica del mundo, Lectura constante como fuente de detección de oportunidades Buena expresión oral y escrita, Capacidad de observación, Capacidad de adaptación a nuevas situaciones o emergentes.

### **Diseño de la intervención en el ciclo lectivo 2019**

En el ciclo lectivo 2019, la intervención articuladora entre las asignaturas mencionadas se realizó en ambos cuatrimestres.

Dado que en el primer cuatrimestre ambas asignaturas se dictan, se realizó un encuentro presencial en donde se contó con la intervención de un especialista en el tema e integrante del equipo AGENTIA que trató en un seminario los planes de negocio en contextos reales. Además, se instaló la cuestión de la importancia referida a como las ideas-proyectos a elaborar en el marco de EA pueden considerarse como un bosquejo preliminar que podría replantearse alternativamente en el marco del Proyecto Final de Carrera.

Durante el segundo cuatrimestre se realizó una convocatoria a los alumnos que habían presentado su Plan de Negocios en la asignatura EA. En una jornada con la presencia de las docentes de ambas asignaturas concretada a fines del ciclo lectivo 2019 se informó a los alumnos a cerca de líneas de financiamiento y/o asistencias financieras no reintegrables disponibles en el mercado para implementar proyectos o planes de negocios de base tecnológica. Los alumnos que concurren se mostraron muy interesados en las alternativas disponibles. Además, se abordaron distintas estrategias para transformar su plan de negocios en el PFC.

La observación realizada en esta segunda intervención permitió inferir que escasamente los estudiantes asisten a actividades complementarias extra-curriculares, aún cuando redundarían en la construcción y realización del PFC. Esta información de retroalimentación implicó la realización de la actividad de articulación durante el cursado de la asignatura EA, es decir, en el primer cuatrimestre del siguiente año.

### **Diseño de la intervención en el ciclo lectivo 2020**

Al inicio del ciclo lectivo se planificó una estrategia didáctica que debió modificarse ante la situación de emergencia sanitaria según el DECNU-2020-297-APN-PRT se dispone el “Aislamiento social, preventivo y obligatorio”,

Es así como se reformuló la estrategia de articulación y se concretó a través de una intervención basada en el uso de herramientas sincrónicas y asincrónicas. Como herramienta asincrónica se recurrió al aula virtual en donde se dispuso material integrado por:

- Formulario ideas proyectos.
- Material orientativo para transformar un plan de negocio en el PFC. La Fig. 1 muestra una interfaz del aula virtual de la asignatura EA donde se despliegan las presentaciones digitales elaboradas por sendas asignaturas, y disparadoras de las actividades propuestas.
- Proyecto de articulación de ambas asignaturas.
- Habilidad de un foro destinado a reflexionar sobre ciertas cuestiones vinculadas al concepto de emprendimiento que sirvió como disparador para reflexionar sobre la propuesta.

A continuación, se concretó la intervención sincrónica mediante una videoconferencia utilizando como herramienta Zoom. En esta actividad, las profesoras responsables de sendas asignaturas realizaron las intervenciones correspondientes con miras a destacar la potencialidad de replantear el plan de negocio presentado como requisito para regularizar o promocionar EA en una idea inicial que podría reformularse o profundizarse para el desarrollo del PFC.

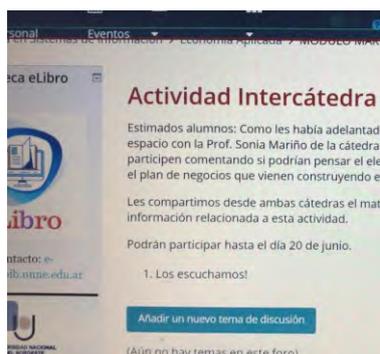


Fig. 1: Espacio habilitado en el aula virtual relacionada con la actividad de articulación.

Los contenidos desarrollados durante los seminarios de articulación se relacionan con ciertas temáticas abordadas por la materia Economía Aplicada y se vinculan con: características de los emprendedores: Cualidades y aptitudes. Concepto de emprendedor y emprendimientos. Identificación de ideas innovadoras para satisfacer necesidades no resueltas. Identificación de ideas de negocios. Elementos constitutivos de un Plan de Negocios. Flujos de fondos. Ingresos y Costos. Estudio de Mercado y Marketing Mix.

Del total de alumnos cursantes, el 11% estarían en condiciones de desarrollar proyectos de negocios de base tecnológica. Por ello, estos resultados indicarían una divergencia entre las características del espíritu emprendedor en la Facultad y los principales sectores económicos de la Ciudad, por lo que el modelo curricular propone una serie de espacios y estrategias para facilitar un diálogo entre la universidad y el medio regional, con miras a lograr un enfoque de los estudiantes atendiendo a las realidades y necesidades de desarrollo del contexto.

## 4 Conclusiones

El artículo sintetizó dos intervenciones de articulación de procesos de enseñanza-aprendizaje significativa orientadas a promocionar el espíritu emprendedor y con énfasis en los planes de negocio de base tecnológica.

Las intervenciones de articulación entre las asignaturas en torno a emprendimientos, se constituyeron en un disparador para los alumnos, quienes iniciaron la identificación de ideas de negocios desde su futura perspectiva profesional. Lo expuesto refleja experiencias de aprendizajes significativos en donde

se vinculan conocimientos disciplinares, metodologías orientadas a la construcción de planes de negocio, y se fortalece el espíritu emprendedor.

Dado que las intervenciones concretadas en la asignatura Economía Aplicada en relación al Proyecto Final de Carrera pueden seguirse y medirse en el tiempo, se propone aplicar los siguientes indicadores de seguimiento a partir de las estrategias didácticas desarrolladas en los ciclos lectivos 2019 y 2020. Por ello, siguiendo lo expuesto en [1] se determinarán:

- Número de ideas de negocios concretadas o generadoras de proyectos de fin de carrera.
- Número de proyectos vinculados a ideas de negocios presentados a la Agencia correspondiente de la universidad.
- Número de proyectos vinculados a ideas de negocios presentados en la incubadora de empresas de la Universidad Nacional del Nordeste.
- Número de estudiantes de la carrera Licenciatura en Sistemas de Información que optaron por cursar la Cátedra Libre para Emprendedores Universitarios.
- Número de estudiantes que, asistiendo a estos seminarios, optaron por cursar la Cátedra Libre para Emprendedores Universitarios.
- Número de planes de negocios que consiguieron financiamiento de distintas fuentes y están en ejecución.
- Número de alumnos que asisten al taller: alternativas de financiamiento para proyectos con base tecnológica, organizado por ambas asignaturas.

Cabe aclarar que respecto al último indicador, en el período 2019 del total de los 53 alumnos inscriptos en Economía Aplicada, seis concurrieron a interiorizarse de las alternativas de financiación disponibles en el mercado para proyectos vinculados a sus ideas de negocios, vale decir, el 11% aproximadamente.

En estos casos, desde una perspectiva académica, integrar y complementar contenidos y esfuerzos hacia la construcción de capacidades emprendedoras constituye un aporte significativo de la Institución Universitaria tendiente a potenciar el espíritu emprendedor, expandiendo el horizonte de planeamiento en virtud del desarrollo futuro del sector privado.

Se espera que intervenciones como las descritas en el artículo se adapten en otras asignaturas del plan de estudios, con la finalidad de asegurar transversalmente una sólida formación en la temática. Simultáneamente, se fortalecen los vínculos entre sectores público – privado y académicos para consolidar ideas de negocios sustentables en el tiempo, como complemento para mejorar el bienestar de la comunidad.

## Referencias

- [1] S. I. Mariño y V. R. Bercheñi, Propuesta de promoción del espíritu emprendedor en la disciplina Informática. Técnica Administrativa, 2020, vol.: 19, núm.: 2; [ISSUE:82], <http://www.cyta.com.ar/ta/article.php?id=190201>
- [2] J. C. Sánchez García, A. Ward, B. Hernández y J. L. Florez, Educación emprendedora: Estado del arte. Propósitos y Representaciones, 5(2), 2017.

- [3] M. Formichella, El concepto de emprendimiento y su relación con la educación, el empleo y el desarrollo local. Buenos Aires, Argentina, 2004.
- [4] A. Valencia Arias, I. Montoya Restrepo y A. Montoya Restrepo, Intención emprendedora en estudiantes universitarios: Un estudio bibliométrico, *Intangible Capital*, vol. 12, núm. 4, pp. 881-922, 2016,
- [5] J. M. Ortiz García; M E Pensado Fernández y A. M. Ortiz Barradas, La formación para el emprendimiento en las IES. Un estudio comparativo entre lo público y lo privado, en: Ezequiel Herruzo-Gómez Brizeida R. Hernández-Sánchez, Giuseppina María, 2019.
- [6] E. Herruzo-Gómez Brizeida R. Hernández-Sánchez, G. M. Cardella, J. C. Sánchez-García (eds.) *Emprendimiento e Innovación: Oportunidades para todos*, Ed. Dykinson, S.L. 2019.
- [7] J. C. Contreras-Velásquez, S. Y. Wilches-Duran, M. E. Graterol-Rivas y M. J. Bautista-Sandoval, Educación Superior y la Formación en Emprendimiento Interdisciplinario: Un Caso de Estudio. *Formación universitaria*, 10(3), 11-20. 2017.
- [8] L. V. Ovalles-Toledo, Z. Moreno Freites, M. A. Olivares Urbina, H. Silva Guerra, Habilidades y capacidades del emprendimiento: un estudio bibliométrico, *Revista Venezolana de Gerencia*, vol. 23, núm. 81, 2018.
- [9] FaCENA, UNNE Plan de Estudios de la carrera Licenciatura en Sistemas de Información. Resol. 1137/09 CS. UNNE. 2009.
- [10] Resolución Ministerial N° 786/09. Ministerio de Educación de la República Argentina, 2009.
- [11] Red UNCI. Documento de Recomendaciones Curriculares de la RedUNCI, 2015, [redunci.info.unlp.edu.ar](http://redunci.info.unlp.edu.ar)
- [12] AGENTIA, Agencia de Innovación y Desarrollo, dependiente de la Universidad Nacional del Nordeste, 2020.
- [13] FaCENA, UNNE. Seminario Corrientes Emprende 2015. Res. 0207/15. 2015.
- [14] FaCENA, UNNE. Seminarios “La Acción Emprendedora – Ciclo Lectivo 2017”, Res. 1688/17 CD. 2017.
- [15] FaCENA, UNNE. Seminarios “La Acción Emprendedora – Ciclo Lectivo 2018”, Res. 191/18 D.
- [16] N. Braidot y R. Cesar. Curso de Posgrado de Formación de Formadores en Competencias Emprendedoras, Curso, Universidad Nacional del Nordeste, 2018.
- [17] S. I. Mariño y M. V. Godoy. Propuesta de un modelo de rol emprendedor en la asignatura Proyecto Final de Carrera, IX Congreso de Tecnología en Educación & Educación en Tecnología, 2014, 75 – 82. 2014.

## Estrategias de evaluación formativa en la enseñanza de programación en modalidad remota de emergencia

Gladys Dapozo, Cristina Greiner, Raquel Petris, María Fernanda Piragine, Ana María Company, María Cecilia Espíndola  
Facultad de Ciencias Exactas y Naturales y Agrimensura, 9 de julio 1449 (3400)  
Corrientes  
Universidad Nacional del Nordeste (UNNE)  
gndapozo, cgreiner, rpetris, mafepiragine, mcespindola, anamacom {@exa.unne.edu.ar}

**Abstract.** 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 este trabajo 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.

**Keywords:** Enseñanza de programación inicial. Enseñanza remota de emergencia. Evaluación formativa. Rúbricas analíticas.

### 1. Introducción

Las condiciones impuestas por el ASPO (Aislamiento Social, Preventivo y Obligatorio) debido a la pandemia por la COVID19, obligó a las universidades a adecuar su estructura de enseñanza presencial a una modalidad de “enseñanza remota de emergencia. Adaptarse muy rápidamente a este nuevo formato implicó tomar decisiones pedagógicas sobre cómo abordar el proceso de enseñanza y aprendizaje en un escenario de incertidumbre, novedoso para todos, estudiantes y docentes.

En este trabajo se expone la experiencia del dictado remoto de emergencia de una asignatura de enseñanza de programación, Algoritmos y Estructuras de Datos I, del primer año de la carrera Licenciatura en Sistemas de Información de la Universidad Nacional del Nordeste. Se destaca la metodología de evaluación formativa implementada con el propósito de obtener “evidencias de aprendizaje”, considerando que se trata de enseñar programación, contenido que, incluso en condiciones normales, presenta dificultades para los estudiantes, máxime cuando recién se inician en la carrera.

Cabe destacar que la metodología de evaluación en el dictado presencial consistía en 2 exámenes parciales con sus correspondientes recuperatorios, que abordaban la resolución de problemas de complejidad similar a los ejercicios dados en las clases prácticas que tenían una

carga horaria de 6 hs semanales, 2 clases prácticas en aula y 1 en laboratorio. Estas actividades aseguraban que el docente pudiera obtener evidencias del aprendizaje mediante la interacción con el estudiante y sus dificultades.

### La evaluación y sus características

La evaluación, entendida como la estimación de los conocimientos, aptitudes y rendimiento de los alumnos, pone de manifiesto el vínculo asumido entre evaluación y aprendizaje.

Elena Barberá [1] considera este vínculo desde cuatro dimensiones: a) evaluación *del* aprendizaje, b) evaluación *para* el aprendizaje, c) evaluación *como* aprendizaje, d) evaluación *desde* el aprendizaje. Señala que la evaluación *del* aprendizaje permite acreditar si el alumno es capaz de saber y de ser competente en un determinado ámbito, en tanto que la evaluación *para* el aprendizaje se sustenta en la retroalimentación y el aprovechamiento que realizan los alumnos y los profesores en este proceso. La evaluación *como* aprendizaje contempla el aprendizaje llevado a cabo por los alumnos en cuanto implica un análisis y reflexión de sus propias prácticas en relación con la dinámica evaluativa. Finalmente, la evaluación *desde* el aprendizaje recupera la idea de que aprender es conectar el conocimiento nuevo con el que se posee, y que esa conexión se debe desarrollar de manera significativa.

Respecto a las funciones de la evaluación, Quesada Castillo [2] rescata la *diagnóstica*, *formativa* y *sumaria*. La *diagnóstica* permite identificar el nivel actual de aprendizaje del alumno, la *formativa* tiene como propósito apoyar al alumno en su proceso de aprendizaje y la *sumativa* tiene una función más de acreditación.

Numerosos estudios describen a la evaluación formativa como parte integral del proceso de aprendizaje, un modo continuo y participativo que permite identificar y analizar dificultades, propiciando la reflexión permanente sobre logros, avances y obstáculos para la toma de decisiones, de estudiantes y docentes.

Bennet por su parte señala que el propósito principal de la evaluación formativa es el de promover el aprendizaje [3]. Para cumplir este propósito es importante que el docente y los estudiantes tengan claro el objetivo de aprendizaje, puedan monitorear el proceso, comparando el desempeño actual con el esperado y tomando acciones para poder alcanzar el objetivo planteado, o para disminuir la brecha. Ambos actores participan en este proceso: el docente realiza las devoluciones y el estudiante las aprovecha para avanzar hacia el objetivo.

Anijovich y Cappelletti [4] visualizan la evaluación como una oportunidad para que el alumno tome conciencia de lo que sabe, recreando sus conocimientos en situaciones nuevas, utilizando lo aprendido para aplicarlo en contextos diferentes, así como reconocer cuestiones que necesita revisar para modificar. En este sentido, cuando las autoras se refieren al concepto de “evaluar para aprender”, lo asocian con el paradigma de la evaluación formativa. Consideran que lo valioso de una evaluación es que el docente pueda utilizar las producciones de los alumnos como evidencia de lo aprendido, y proponer nuevas oportunidades para aprender lo que falta. Una evidencia es una certeza clara y manifiesta de un proceso. Poner en evidencia implica revelar o demostrar algo. Las evidencias constituyen elementos que permiten formular juicios de valor.

La evaluación formativa también se puede entender como la recolección, análisis y uso de evidencias que ayudan a docentes y estudiantes a tomar decisiones encaminadas al mejoramiento, según lo expresa McMillan en [5].

En el proceso de recogida y sistematización de la información es conveniente utilizar una combinación de técnicas e instrumentos, a efectos de poder hacer inferencias a partir de las evidencias obtenidas [6]. La retroalimentación, parte integral de la evaluación formativa, resulta útil tanto a docentes como a estudiantes. Desde la perspectiva del profesor, brinda la oportunidad de rediseñar el entorno de aprendizaje, si fuera necesario, y realizar las devoluciones pertinentes a los estudiantes. Desde el punto de vista del alumno, la retroalimentación le permite reflexionar sobre lo aprendido, lo que debe mejorar, qué aspectos o dimensiones requieren ser reorientadas, para lograr el objetivo de aprendizaje.

En particular, en la educación en modalidad a distancia, la evaluación cobra relevancia y requiere propuestas pedagógicas adecuadas al contexto, dado que en esta situación el alumno adquiere un mayor protagonismo e independencia. En este marco, los conceptos de evaluación formativa contribuyen a favorecer el seguimiento del aprendizaje de los alumnos.

### **La enseñanza de la programación**

Aprender a programar computadoras permite el desarrollo de diversas competencias como el pensamiento crítico y la resolución de problemas. Los estudiantes se inician en el trabajo en equipo y aprenden a colaborar entre ellos, durante el proceso de desarrollar programas ejecutables, mientras se ejercitan en el intercambio de conocimientos y la comunicación de ideas [7].

El desarrollo de habilidades cognitivas como la capacidad de abstracción, una buena aptitud lógico-matemática y la facilidad para la resolución de problemas de orden algorítmico, son muy importantes al momento del aprendizaje de los fundamentos de programación [8].

Aprender programación incluye el aprendizaje de la teoría, es decir conceptos básicos y algoritmos, así como también una parte práctica que incluye dominar la habilidad de usar un entorno moderno para el desarrollo del programa y desarrollar el pensamiento computacional a través de múltiples problemas diferentes siguiendo una metodología y un lenguaje de programación previamente seleccionado [9].

Las actividades de aprendizaje asociadas a la programación de computadoras han sido reconocidas con alto grado de dificultad, según los antecedentes descriptos en [8].

Aprender a programar computadoras es un proceso difícil para los estudiantes que se inician y constituye un desafío a las metodologías de enseñanza-aprendizaje utilizadas por los docentes [10].

Es necesario que los docentes generen diversos contenidos didácticos que permitan orientar al alumno a apropiarse de los conceptos de programación y que lo guíen a resolver problemas de manera autónoma.

### **Evaluación con rúbricas**

Las rúbricas son “guías de puntuación” usadas en la evaluación del desempeño de los estudiantes que describen las características específicas de un producto, proyecto o tarea en varios niveles de rendimiento, con el fin de clarificar lo que se espera del trabajo del alumno, de valorar su ejecución y de facilitar la obtención de *feedback*, útil tanto para el estudiante como para el docente, permitiendo mejorar las propuestas educativas. La literatura sobre el

tema muestra que existe una relación entre el uso de rúbricas de evaluación y un mejor rendimiento académico, así como un mayor aprendizaje profundo en los estudiantes [11].

En este trabajo se aplicaron rúbricas analíticas, en las cuales se evalúa un conjunto de actividades de acuerdo con el tema a analizar, segmentada en subtarear con escala descriptiva.

Se considera un instrumento idóneo para evaluar competencias que permite diseccionar tareas complejas que conforman una competencia en tareas más simples distribuidas de forma gradual y operativa.

A través de los descriptores de las rúbricas se comparten los criterios que se aplicaran para evaluar el progreso en un marco de evaluación formativa y continua, al mismo tiempo que reduce la subjetividad y facilita que diferentes docentes del mismo espacio curricular coordinen y compartan similares criterios de evaluación

El proceso de elaboración de la rúbrica lleva al grupo de docentes a reflexionar sobre cómo enseñar y cómo evaluar, tomando conciencia de los aspectos metodológicos que mejor se adapten a su proyecto y diseñar gradualmente los descriptores mensurables de los objetivos propuestos.

Es oportuno mencionar que el uso óptimo de la rúbrica es inversamente proporcional a la cantidad de alumnos, si el grupo es muy numeroso, la calidad de la evaluación se puede ver afectada, pero al mismo tiempo se puede afirmar que la heterogeneidad en grandes grupos se ve disminuida ante un criterio acordado y plasmado por los docentes con anterioridad [12].

El proceso de la elaboración de rúbricas permite al docente diseñar, orientar y desarrollar contenidos, actividades de formación y de evaluación, vinculados a la enseñanza-aprendizaje, de modo que se valoren los resultados y se elaboren propuestas de mejora.

Asimismo, favorece aplicar estrategias metodológicas de aprendizaje y evaluación adecuadas a los estudiantes, coherentes con los objetivos y los procesos de evaluación, y contemple el uso de las tecnologías de la información y la comunicación para contribuir a la mejora de los procesos educativos.

Permite desagregar y visualizar los aspectos y compromisos ético de cada perfil profesional, donde el alumno gradualmente puede ir tomando conciencia de los diferentes factores que inciden en el hecho educativo y de la importancia de la ética en el rol profesional y el juicio crítico en la ciencia y la deontología [12].

## **2. Metodología de dictado de la asignatura en modalidad on line**

La asignatura inició sus actividades el 17 de febrero, durante un mes se desarrollaron las actividades en forma presencial. A partir del ASPO, el equipo docente acordó la adecuación a la modalidad “A distancia”, mediada por herramientas informáticas. Para este nuevo marco, se propuso seguir la siguiente metodología de trabajo:

### **2.1. Organización de los contenidos**

Para cumplir con los objetivos esenciales de la asignatura se seleccionaron los temas relevantes para cumplir dos objetivos: 1- Programación básica con datos simples, que abarca los temas: Conceptos de Programación por bloques, Resolución de problemas y algoritmos, Tipos de datos y Expresiones, y Estructuras de control. 2-Programación básica con datos

compuestos, que abarca los temas: Funciones y Parámetros, Cadenas, Arreglos, Registros y Archivos.

## **2.2. Atención a los estudiantes (estrategia para interactuar con grupos reducidos)**

Un total de 325 alumnos iniciaron el dictado remoto, divididos en 4 comisiones. Por cada comisión se crearon 2 subcomisiones. En cada una de ellas se crearon grupos de hasta 3 estudiantes. Como medio de comunicación se crearon grupos de Whatsapp, Grupo de Facebook, listas de correos y foros en el aula virtual de la asignatura.

## **2.3. Tecnologías utilizadas**

La asignatura cuenta desde hace más de 5 años con un aula en la plataforma UNNE Virtual, como apoyo a las actividades presenciales. En esta contingencia siguió siendo el principal recurso de comunicación, pero debido a la alta demanda, se optó también por otros recursos complementarios, como Google Drive y canal de Youtube para los videos. Las clases sincrónicas se realizaron con Google Meet.

## **2.4. Cronograma**

Se definió un cronograma muy detallado para que el estudiante se organice, teniendo en cuenta que la autoadministración del tiempo es la parte más difícil para el alumno. Incluía los temas a desarrollar, el material didáctico necesario, la hoja de ruta a seguir, la fecha de entrega de los trabajos requeridos y la fecha de la devolución del trabajo presentado.

## **2.5. Desarrollo de la teoría**

Por cada uno de los temas, se puso a disposición de los estudiantes material escrito en PDF y una presentación en Power Point. Por cada tema, se elaboraron videos cortos, de 15 min a 30 min de duración, indicando en el primero la secuencia de los títulos de los videos correspondientes al tema, y se subieron a Youtube. Se elaboraron cuestionarios de autoevaluación, disponibles en el aula virtual.

## **2.6. Desarrollo de los trabajos prácticos**

Por cada tema, se elaboró una guía de trabajos prácticos, con un enfoque de resolución de problemas, y una guía de laboratorios para realizar actividades en la computadora. Este material se puso a disposición de los estudiantes en el aula virtual y en el Drive. Se ofrecieron videos explicativos sobre el desarrollo de los ejercicios prácticos al iniciar cada tema y cada técnica diferente, que muestra con un ejemplo cómo resolver los problemas propuestos y enfatiza los conceptos involucrados. A partir del ejemplo, los alumnos desarrollaban los ejercicios y consultaban las dudas a sus tutores (el docente responsable de cada subcomisión).

Para supervisar el desarrollo de las prácticas, los docentes realizaron una clase semanal por videoconferencia para atender las consultas y explicar temas puntuales.

Los alumnos a través de Whatsapp, correo electrónico, foros y mensajería del aula virtual realizaron permanentemente consultas a sus docentes tutores.

Se definieron actividades prácticas, a modo de evaluaciones parciales, de frecuencia semanal o quincenal, que abarcaron los conceptos y técnicas abordados, y se fijaron las fechas de entrega. Estas actividades tenían varios objetivos: principalmente mantener el interés de los alumnos y conocer el nivel de participación a través de las respuestas, y realizar una evaluación formativa, en proceso, lo cual cobra mayor relevancia aún que en la presencialidad.

La evaluación de los trabajos se realizó mediante rúbricas analíticas que permitieron determinar el logro de los objetivos de aprendizaje y visibilizar los déficits en la comprensión de los distintos aspectos abordados. El resultado de las rúbricas se publicaba en el aula virtual de manera tal que los alumnos pudieran apreciar cuáles fueron sus puntos fuertes y débiles en el trabajo presentado.

Los docentes contaban con una planilla de seguimiento de los alumnos de su subcomisión con el objetivo de disponer información acerca de la participación y cumplimiento de las actividades de los estudiantes. También detectar a quienes no participaban y contactarlos para tener en cuenta sus dificultades y alentarlos a seguir participando.

## 2.7. Evaluación de los aprendizajes y acreditación

La evaluación formativa se realizó mediante el desarrollo de actividades prácticas de entrega obligatoria, que consistían en resolución de problemas que los alumnos debían realizar siguiendo las consignas dadas y respetando los plazos y condiciones anticipados. El uso de rúbricas contribuyó a la devolución a los estudiantes sobre los aspectos a mejorar y permitió a los docentes evaluar el avance del aprendizaje de los estudiantes, respecto de los objetivos definidos. En la tabla 1 se muestra un ejemplo de rúbrica que permite obtener una calificación numérica en función de las calificaciones conceptuales (**MB**-Muy bueno, **B**-Bueno, **R**-Regular y **DM**-Debe Mejorar), y el valor de ponderación (**Coef**) de cada aspecto evaluado.

Tabla 1. Rúbrica analítica para la evaluación de las actividades prácticas

Aspectos	MB (4)	B (3)	R (2)	DM (1)	Coef.
<b>1. Aplicación del método de resolución de problema</b>	Aplica el método correctamente (Definió una estrategia correcta, la descomposición y los nombres son adecuados).	Aplica el método, pero no completamente (Definió una estrategia adecuada pero la descomposición no es óptima y los nombres adecuados)	Aplica el método en forma deficiente (la descomposición no es óptima y los nombres no son adecuados)	No aplica el método correctamente (No definió una estrategia adecuada)	0,5
<b>2. Estilo de programación</b>	Estilo de programación correcto (Código	El estilo de programación es medianamente	El estilo de programación es deficiente.	No respeta las pautas dadas (El	0,25

	comprensible, indentación correcta, nombres adecuados).	correcto (Código comprensible, no respeta indentación, nombres adecuados).	No respeta indentación, nombres no adecuados).	código es desprolijo y poco claro).	
<b>3. Corrección Funcional (si el código cumple con lo requerido)</b>	El código funciona perfectamente, para todos los casos posible	El código funciona para la mayoría de los casos (más del 70%)	El código brinda menos del 70% de las funcionalidad es requeridas	El código no compila o da error al ejecutarlo	0,5
<b>4. Estructuras de control</b>	Usa correctamente las estructuras de control	Usa las estructuras de control (pero no en forma óptima)	Usa las estructuras de control en forma deficiente	No usa las estructuras de control	0,5
<b>5. Uso de funciones</b>	Usa correctamente las funciones (Prototipo, invocación y definición adecuadas)	Usa medianamente bien las funciones (la definición no es óptima)	Usa funciones, pero en forma deficiente	No utiliza funciones	0,5
<b>6. Casos de prueba</b>	Usa correctamente los casos de prueba (Casos de prueba representativos del problema)	Usa los casos de prueba medianamente bien (No cubren la mayoría de las posibilidades)	Los casos de prueba no son representativos	No escribió casos de prueba	0,25

Para obtener la calificación final de cada trabajo, se multiplica el puntaje de la calificación conceptual (MB-4, B-3, R-2, DM-1) por el coeficiente de ponderación de cada aspecto (Coef) y se suman los valores obtenidos.

En la tabla 2 se muestran como ejemplos, dos trabajos T1 y T2. Cuando el docente ingresa la calificación conceptual para cada trabajo, la planilla calcula automáticamente una calificación numérica entre 1 y 10. Es ilustrativa del grado de cumplimiento de los objetivos de la actividad, para orientar al propio alumno y a los docentes.

*Tabla 2: Cálculo de la calificación numérica de la rúbrica*

Aspectos considerados en la evaluación	T1	T2	T1	T2
1. Aplicación del método de resolución de problema	B	MB	1,5	2
2. Estilo de programación	MB	MB	1	1
3. Corrección Funcional (si el código cumple con lo requerido)	B	B	1,5	1,5
4. Estructuras de control	B	MB	1,5	2
5. Uso de funciones	R	MB	1	2
6. Casos de prueba	DM	B	0,25	0,75
<b>Calificación del trabajo</b>			<b>6,75</b>	<b>9,25</b>

Como ejemplo, para el caso del Trabajo Práctico N°5, se evaluaron 91 trabajos y se obtuvieron los resultados que se muestran en la Tabla 3.

Se puede observar que el aspecto más deficitario que surge de los trabajos de los alumnos está vinculado con la aplicación del método de resolución de problemas. El otro aspecto es el cumplimiento del estilo de programación, El aspecto con mejores valores de ponderación es el de Corrección funcional. Se puede inferir que los estudiantes se concentran en que el código funcione, y no tanto en las cuestiones metodológicas (pasos a seguir) para la resolución de problemas y las recomendaciones para lograr un código, que además de funcionar, sea legible y comprensible.

Tabla 3: Resultados globales de un trabajo práctico

Aspectos para evaluar	MB (4)	B (3)	R (2)	DM (1) debe mejorar	% MB	%B	%R	% D M
1. Aplicación del método de resolución de problema	Aplica el método correctamente	Aplica el método, pero no completamente	Aplica el método en forma deficiente	No aplica el método correctamente.	51%	27%	10%	12%
2. Estilo de programación	El estilo de programación es correcto	El estilo de programación es medianamente correcto	El estilo de programación es deficiente.	No respeta el estilo de programación recomendado.	66%	27%	2%	4%
3. Corrección Funcional	El código funciona perfectamente, para todos los casos posibles	El código funciona para la mayoría de los casos (más del 70%)	Resuelve menos del 70% de las funcionalidades	El código no compila o da error al ejecutarlo	74%	10%	5%	11%
4. Estructuras de control	Usa correctamente	Usa medianamente	Usa en forma deficiente	No usa las estructuras de control	67%	14%	10%	9%
5. Uso de funciones	Usa correctamente las funciones	Usa medianamente	Usa en forma deficiente	No utiliza funciones	67%	15%	5%	12%
6. Casos de prueba	Usa correctamente los casos de prueba.	Usa los casos de prueba medianamente bien	Los casos de prueba no son representativos	No escribió casos de prueba	66%	21%	9%	4%
				Promedio	65%	19%	7%	9%

Para orientar el trabajo de los docentes, divididos en 4 comisiones, se realizó también un análisis considerando los distintos grupos por separado. En la Fig. 1 se puede observar que los distintos aspectos a evaluar tuvieron distintos resultados en las comisiones. Algunas diferencias significativas llevan a pensar en diferentes interpretaciones en la evaluación de

los aspectos o mayor rigurosidad. Cuestiones que deberán ser trabajadas al interior del equipo de cátedra.

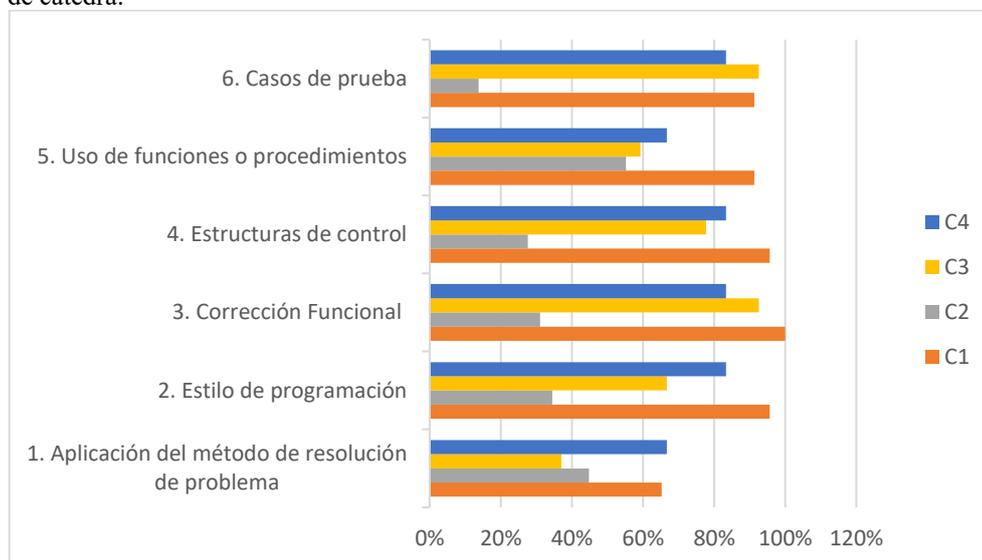


Figura 1. Análisis de los resultados obtenidos en cada comisión

### 3. Conclusiones y futuros trabajos

En el marco de la pandemia 2020, que obligó a un aterrizaje brusco en la enseñanza remota de emergencia, se diseñaron estrategias para lograr los objetivos de aprendizaje de la asignatura. Principalmente, en este escenario de distanciamiento físico del estudiante, el foco se puso en la evaluación formativa a fin de recabar información para obtener evidencias de su proceso de aprendizaje. En este sentido, las rúbricas aportaron información a modo de feedback de los estudiantes, y permitiendo medidas correctivas para lograr los objetivos antes señalados.

Las rúbricas también contribuyeron a que los estudiantes conozcan cuáles son los aspectos relevantes que se consideran necesarios para que se cumplan los objetivos de formación previstos por la asignatura.

De esto modo, se logró que se cumpliera la planificación definida al inicio y se desarrollaran las actividades para el aprendizaje de programación en una asignatura de primer año de carrera universitaria.

Este proceso de aprendizaje, en circunstancias normales, presenta dificultades para un gran porcentaje de alumnos, por el grado de abstracción que se requiere, dedicación, concentración y mucha práctica. En este escenario de virtualidad, se promovió aún más el aprendizaje activo y la motivación.

Aproximadamente un 41% de los estudiantes que iniciaron el dictado remoto, lograron cumplir las condiciones para aprobar la materia. Este valor se considera significativo en el marco de un escenario imprevisto y de gran incertidumbre, que obligó a una rápida respuesta

del equipo docente, mediante la adecuación de las estrategias de dictado de la asignatura, que además implicó la elaboración de nuevos materiales y métodos de evaluación.

A futuro se contempla perfeccionar el uso del recurso Cuestionario de Moodle para los temas teóricos, dado que es una herramienta poderosa para obtener evidencias de aprendizaje que se extraen de los reportes que la plataforma brinda para cada tarea solicitada a los estudiantes. Como así también, revisar y mejorar las rúbricas analíticas, considerando su potencial de ofrecer un *feedback* valioso, tanto para los alumnos como para los docentes.

#### 4. Referencias

- [1] Barberá, E. Aportaciones de la tecnología a la e-evaluación. Revista De Educación a Distancia (RED). 2006. Disponible en: <https://revistas.um.es/red/article/view/24301>
- [2] Quesada Castillo, R. Evaluación del aprendizaje en la educación a distancia. Revista de Educación a Distancia (RED). 2006.
- [3] Bennett, R. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25. DOI: 10.1080/0969594X.2010.513678
- [4] Anijovich, R y Cappelletti, G. La evaluación como oportunidad. Editorial Paidós. 2017.
- [5] Osorio Sánchez, K.; López Mendoza, A. Retroalimentación Formativa en el Proceso de Enseñanza-Aprendizaje de Estudiantes en Edad Preescolar. *Revista Iberoamericana de Evaluación Educativa*, 7(1), 13-30. 2014.
- [6] Canabal, C y Margalef, L. La retroalimentación: la clave para una evaluación orientada al aprendizaje. *Profesorado. Revista de curriculum y formación del profesorado*. ISSN 1138-414X. Julio 2017.
- [7] Insuasti, J. “Problemas de enseñanza y aprendizaje de los fundamentos de programación”. *Educación y Desarrollo Social*, ISSN 2011-5318, ISSN-e 2462-8564, Vol. 10, N°. 2, 2016, págs. 234-246.
- [7] Javier, A.; Jiménez-Toledo, Cesar Collazos, Oscar Revelo-Sánchez, “Consideraciones en los procesos de enseñanza-aprendizaje para un primer curso de programación de computadores: una revisión sistemática de la literatura”. Instituto Tecnológico Metropolitano. *Tecnológicas*, vol. 22, 2019.
- [9] Djenic, S.; Mitic, J. “Teaching Strategies and Methods in Modern Environments for Learning of Programming”. Conferencia Internacional IADIS sobre Cognición y Aprendizaje Exploratorio en la Era Digital. Portugal, octubre 2017. <https://eric.ed.gov/?id=ED579455>
- [10] López Reguera, J.; Hernández Rivas, C.; Farran Leiva, Y. “Una plataforma de evaluación automática con una metodología efectiva para la enseñanza/aprendizaje en programación de computadores”. *Ingeniare. Revista chilena de ingeniería*, vol. 19 N° 2, 2011, pp. 265-277.
- [11] Valverde Berrocoso, J.; Ciudad Gómez, A.: El uso de e-rúbricas para la evaluación de competencias en estudiantes universitarios. Estudio sobre fiabilidad del instrumento. *REDU. Revista de Docencia Universitaria*, 12 (1), (2014) 49-79.
- [12] Masmijá, J., Cuaderno de Docencia Universitaria. Rúbricas para la evaluación de Competencias. Octaedro Barcelona (2013).

# Relevamiento de demandas formativas e innovación educativa en el nivel medio y superior en el DIIT-UNLaM

Santiago Igarza<sup>1</sup>, Bettina Donadello<sup>1</sup>, Mariana Cipicic<sup>1</sup>, Nicolás Kotliar<sup>1</sup>,

<sup>1</sup> Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT)  
 Universidad Nacional de La Matanza, Argentina  
[. asigarza@unlam.edu.ar](mailto:asigarza@unlam.edu.ar), [bdonadello@unlam.edu.ar](mailto:bdonadello@unlam.edu.ar),  
[mcipicic@unlam.edu.ar](mailto:mcipicic@unlam.edu.ar), [nkotliar@unlam.edu.ar](mailto:nkotliar@unlam.edu.ar)

**Resumen.** Introducción: Desde 2006 se realizan en el Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT) de la Universidad Nacional de La Matanza (UNLaM) diferentes acciones de articulación con escuelas secundarias de la zona, dentro de una estrategia más amplia de mejora de la formación de estudiantes de carreras tecnológicas. A tal fin, se llevan a cabo visitas a escuelas secundarias técnicas, talleres vocacionales, charlas para ingresantes, participación en ferias y exposiciones tecnológicas y recientemente, el Proyecto Vincular en sus ediciones 2019 y 2020. Estas acciones son realizadas en forma conjunta por un equipo de trabajo interdisciplinario conformado por docentes, no docentes y asesores pedagógicos del DIIT. Resultados: En el Proyecto de investigación Vincular 2019 se puede vislumbrar una diversificación en la demanda de opciones formativas para el nivel superior por parte de los estudiantes de escuelas secundarias de la zona de influencias, con varias opciones que no habían sido previstas. Un hilo une a los proyectos de investigación Vincular llevados a cabo durante estos dos últimos años: la centralidad otorgada en los mismos a la voz de los estudiantes de los últimos dos años de las escuelas secundarias de La Matanza, con el objetivo de mejorar la oferta académica superior del departamento y la oferta de prácticas profesionalizantes para el nivel medio.

**Palabras clave:** articulación escuela-universidad, despertar vocacional, educación mediada, competencias.

## 1 Introducción

El presente trabajo se elabora con el fin de dar difusión a los proyectos Vincular 2019 y 2020, realizados en el marco del programa de “Articulación Escuela Universidad”, el cual a su vez forma parte del proyecto de mejora de los indicadores académicos del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza (DIIT-UNLaM), localizada en San Justo, Provincia de Buenos Aires.

Desde 2006 se realizan desde este Departamento diferentes acciones de articulación que forman parte de una estrategia más amplia de mejora de la formación de ingenieros y en forma reciente, de arquitectos. A tal fin, se llevan a cabo visitas a escuelas secundarias técnicas, talleres vocacionales, charlas para ingresantes y participación en ferias y exposiciones tecnológicas. Estas acciones son realizadas en forma conjunta por un equipo de trabajo interdisciplinario conformado por docentes, no docentes y asesores pedagógicos del DIIT. Entre estas actividades se encuentran: a) Prácticas profesionalizantes. b) Expo Escuela. c) Visitas a escuelas. d) Talleres de despertar vocacional para alumnos de escuelas secundarias técnicas. e) Participación en ferias y exposiciones. f) Proyecto de investigación Vincular 2019 y 2020. Los objetivos centrales de estas diversas acciones coordinadas por Extensión Universitaria son: el mejoramiento continuo de las prácticas docentes, la flexibilización y adaptación de la oferta curricular de la universidad y la colaboración en el desarrollo de las competencias profesionales de los futuros ingenieros a través del diseño de situaciones de enseñanza-aprendizaje bajo una metodología colaborativa, activa y participativa.

### 1.1 El programa de articulación Escuela-Universidad

El mencionado proyecto de mejora de los indicadores académicos de nuestro Departamento sigue aún el eje estratégico del Plan de Formación de Ingenieros 2012-2016, cuyos objetivos son:

- I. Generar vocaciones tempranas y facilitar el tránsito entre niveles educativos.
- II. Incrementar la retención en el ciclo básico.
- III. Incrementar la retención en el ciclo de especialización.
- IV. Incrementar la graduación de alumnos avanzados.

Dentro de dicho proyecto, las acciones realizadas dentro del *programa de Articulación con Escuelas secundarias* se centran en cumplir con el objetivo I indicado en el párrafo precedente.

Desde la Implementación de este programa, con la participación del DIIT se han realizado actividades con aproximadamente 30 colegios de Nivel Secundario del Partido de La Matanza y otras zonas de influencia, en las que intervinieron estudiantes de los dos últimos años de los colegios y docentes de Matemática, Asesores Pedagógicos y Directivos de ambos niveles, todos capacitados específicamente para la consecución de las actividades. En éstas, se abordaron, fundamentalmente, las problemáticas detectadas en los dos primeros años de las carreras de la Unidad Académica que dificultan la adaptación de los estudiantes a la vida universitaria, las que se vinculan, principalmente, con los conocimientos previos y competencias requeridos para este nivel y con la falta de hábitos de estudio.

La articulación se ha ampliado a más de 70 establecimientos de nivel medio, de orientación técnica en su mayoría, de gestión pública y privada. La zona de influencia se extiende al partido de La Matanza, Morón, Ituzaingo, Merlo y zona sur de CABA.

Las acciones llevadas a cabo entre dichos establecimientos incluyen las siguientes:

- Visitas a escuelas secundarias de zonas de influencia.
- Talleres de despertar vocacional en escuelas secundarias.
- Charlas informativas en el DIIT UNLaM.
- Visitas guiadas por la UNLaM.
- Participación en diferentes Exposiciones (Expo Escuela, Expo Proyecto y otras de la zona de influencia).
- Prácticas Profesionalizantes (PP).
- Proyecto de investigación Vincular 2019/2020

## 1.2 Proyecto Vincular 2019 y la Articulación Escuelas-Universidad

En el año 2006, se establece un convenio marco de cooperación entre la Jefatura Regional del Partido de La Matanza, y la Universidad Nacional de La Matanza, el cual da lugar a un convenio específico, cuyo Programa de Articulación Escuela–Universidad tiene los siguientes objetivos:

- Favorecer el tránsito de los estudiantes de la Escuela Secundaria a la Universidad.
- Sensibilizar a los estudiantes de los dos últimos años del Nivel Secundario acerca de la problemática vocacional.
- Establecer y/o fortalecer vínculos institucionales entre la Unidad Académica y los colegios de nivel medio e Institutos Terciarios.
- Difundir las carreras de la Universidad, su inserción en el medio e importancia para el desarrollo regional y nacional.

El proyecto Vincular 2019 se suma a las líneas de acción recién referidas, y en este sentido, comprende tres objetivos en su plan de trabajo; de los cuales en esta primera etapa del proyecto nos concentramos en el primero: **“conocer las principales áreas del conocimiento que resultan interés de los estudiantes de nivel secundario y demandas de profesionales en el área de influencia de UNLaM”**. En etapas futuras abordaremos los otros dos a partir de los datos generados, haciendo hincapié en las necesidades del sector productivo y también en el desempeño académico de los encuestados en el nivel secundario (cuestión que ya está siendo abordada en el Vincular 2020, al evaluar el impacto de la pandemia en la calidad educativa recibida por los estudiantes).

## 2. Desarrollo de Vincular 2019

### 2.1 Objetivos

Conocer las principales áreas de formación que resultan de interés para los estudiantes de nuestra zona es de vital importancia para poder confeccionar un mapa de oferta formativa que dé cuenta de todas aquellas áreas con una demanda potencial (por parte de los estudiantes y también del aparato productivo del área de influencia) que no está siendo cubierta aún. De esta forma, la universidad apunta a verse en mejores condiciones de evaluar la apertura de nuevas instancias de formación en educación superior, acordes a las demandas de los estudiantes y a los requerimientos del medio económico y social en el cual se encuentra inserta.

Para lograr lo antedicho, partimos de las escuelas del área de influencia de la UNLaM, tomando como base aquellas con las cuales la universidad ya articula en forma sistemática. En resumidas cuentas, el objetivo principal del proyecto será relevar datos que permitan conocer:

- a) Principales áreas educativas de interés de los estudiantes (incluyendo carreras a nivel universitario) y demandas de profesionales en el área de influencia de UNLaM. En particular, registrar la demanda de instancias de formación en oficios.
- b) Características socioeducativas de los estudiantes
- c) Principales dificultades y obstáculos de los estudiantes a la hora de enfrentar estudios de nivel superior

## 2.2 Metodología

Aprovechando que el DIIT se encuentra en permanente contacto con las Escuelas a través de las acciones ya descritas, se diseñó un instrumento de evaluación (encuesta) para tener información de los intereses y necesidades educativas/pedagógicas de los futuros estudiantes del DIIT y de la UNLaM en un contexto más amplio. Se imprimieron cientos de copias en papel de dicha encuesta y se realizó un trabajo extenso de visitas a escuelas en los últimos dos años de educación secundaria (quinto y sexto año) en los cuales la misma se distribuyó y se solicitó a los estudiantes que la completen. Se generó asimismo una web sencilla donde volcar la encuesta y se imprimieron folletos y flyers para difundir el proyecto.

Finalmente, se procesaron los datos en función de los aspectos que se desean evaluar en el presente proyecto, con el objetivo de analizar esa información y posteriormente producir el presente informe, así como un video para difundir desde el departamento de Extensión y las distintas unidades académicas de la UNLaM.

## 2.3 Resultados

Se visitaron 32 escuelas de la zona de influencia de la UNLaM (La Matanza y alrededores), aplicando la encuesta particularmente en las siguientes:

E. Media N° 7, EEST N° 6, EEST N° 1, EES N°37, EEST N° 5 Roberto Noble, EES N° 35, E. Media N°4 Tapiales, Instituto Madero, EEST N°10, EES N° 157, EES N° 41, EES N° 30, EEST N°10 Néstor C. Kirchner, EES N° 103, EES N° 1, EEST N° 4, EEST N° 3, EES N° 20 Escuela Normal de San justo, EEST N° 8 Jorge Newbery, Escuela Evangélica Cristiana, EEST Juan XXIII, Colegio Ward, EEST Instituto Padre Elizalde, EES N°2, EES N° 48, EEST N° 11, EEST N° 2, EEST N° 14, EEST N° 13, Fundación Fangio/Mercedes Benz, Escuela de ed. Secundaria Padre Mario

En total se relevaron 32 escuelas de La Matanza durante 2 meses de visitas, las cuales sumaron un total de 911 encuestas. Vale aclarar que en algunos casos las respuestas incluyen más de una opción por ítem, por lo cual el número de elementos cuantificados excede el número de la muestra de encuestas (por ejemplo, en la pregunta “instituciones elegidas”, podían elegir más de una, por lo cual el número final de elecciones es superior al número de muestra, que es 911).

Los resultados preliminares se resumen en los gráficos y tablas dados a continuación:

**Fig. 1.** Tipo de establecimiento



Fuente: Elaboración Propia.

**Fig.2.** Situación laboral de estudiantes



Fuente: Elaboración Propia.

## 2.4 Demografía y características de la población de estudiantes

Más de un 80% de los estudiantes provienen de escuelas técnicas, y más de dos tercios de los mismos son del sexo masculino, lo cual marca el fuerte sesgo de género de las actividades tecnológicas<sup>1</sup>. La mayoría de los estudiantes no trabaja; una quinta parte sí lo hace, en su mayoría en actividades que no están relacionadas con tareas técnicas. Cerca de la mitad cobra algún tipo de asignación estatal en su grupo familiar.

Por otro lado, es notable el hecho de que la enorme mayoría de los estudiantes (cerca de un 70%) no tiene padres que hayan asistido alguna vez a alguna instancia de formación superior (universidad o instituciones de tipo terciario). Tan solo un 30% lo hizo y dentro de ese porcentaje se encuentran agregados tanto quienes finalizaron estudios superiores como quienes no lo hicieron. Este dato es de relevancia, ya que marca el horizonte de expectativas de muchos estudiantes y luego en la universidad condiciona el proceso de adquirir el “oficio del estudiante” (Viel, Juárez, Kotliar, 2019).

## 2.4 Perspectivas de inserción laboral

Lo primero que se trató de relevar fueron dos cosas: los campos profesionales de interés del estudiante y el ámbito laboral en el que se visualizaba llevando a cabo sus actividades.

Respecto a las dos primeras tablas que recogen estos datos, lo primero que salta a la vista (en cierta forma, lógicamente, dada la fuerte procedencia de escuelas técnicas es la fuerte preferencia de los estudiantes por las carreras del campo tecnológico, las cuales superan en un amplio margen a todo el resto de las opciones. Es llamativa la conjunción del campo “carreras tecnológicas” con el ámbito “oficina” (seguida por el ámbito “fábrica”) y probablemente sea comprensible en vistas a que las carreras tecnológicas dejaron hace tiempo de ser algo eminentemente vinculado al trabajo manual, para formar parte de la sociedad del conocimiento, en la cual las actividades cerebro-intensivas cobran un rol fundamental frente a las manuales (la ingeniería informática es paradigmática, en este sentido), así como por la evolución del rol del ingeniero en la actualidad, claramente dirigida hacia las funciones directivas y gerenciales dentro de las organizaciones (Sobrevila, 2000).

**Tabla 1.** Opciones de formación escogidas

Carrera/Oficio	Nº	Porcentaje
Fotografía	268	4,18%
Electricista	230	3,59%
Ingeniería Electrónica	226	3,53%
Arquitectura	223	3,48%
Ingeniería Informática	215	3,36%
Sonido y grabación	203	3,17%
Programación	192	3,00%
Reparación de celulares	190	2,97%
Desarrollo web	188	2,93%
Diseño aplicaciones móviles	185	2,89%
Mecánica y electrónica automotriz	185	2,89%
Ed. Física	178	2,78%
Seguridad e higiene	170	2,65%
Ingeniería mecánica	169	2,64%
medicina	164	2,56%
Guía de turismo	162	2,53%
Reparación de aires acondicionados	149	2,33%
Tornería/matricera	134	2,09%
Ventas	131	2,04%

<sup>1</sup> Este es uno de los puntos que señala y procura revertir el Programa Mujeres en Ingeniería del CONFEDI

Construcción	131	2,04%
Abogacía	130	2,03%
Carpintería	129	2,01%
Administración	125	1,95%
Instalación de paneles solares	119	1,86%
Ingeniería Industrial	117	1,83%
Analista de sistemas	117	1,83%
Nutrición	115	1,79%
Economía	110	1,72%
Ingeniería Civil	106	1,65%
Mantenimiento	102	1,59%
Comunicación Social	100	1,56%
Contador	97	1,51%
Trabajo Social	95	1,48%
Chapa y pintura	95	1,48%
Herrería	93	1,45%
Gasista	87	1,36%
Reparación de heladeras	84	1,31%
Comercio Internacional	79	1,23%
Trabajo textil (costura, tizado, corte, etc.)	71	1,11%
Relaciones públicas	69	1,08%
Kinesiología	67	1,05%
Instalación de tanques solares	65	1,01%
Relaciones laborales	51	0,80%
Ciencia política	51	0,80%
Cerrajería	51	0,80%
Analista contable	42	0,66%
Calzado	25	0,39%
Ceremonial y protocolo	20	0,31%
Psicología	6	0,09%
Veterinaria	5	0,08%
Marroquinería	5	0,08%
Artes visuales	5	0,08%
Ing. Ambiental	2	0,03%
Traductora do	2	0,03%
Radiología	2	0,03%
Desarrollo de videojuegos	1	0,02%
Total	6408	100,00%

Fuente: Elaboración Propia

Tabla 2. Otras carreras elegidas

Otras carreras	Nº	Porcentaje
Periodista deportivo	8	12,70%
Diseño gráfico	7	11,11%
Deportes	6	9,52%
Profesor	4	6,35%
Militar/Policía	4	6,35%
Cs. Naturales	4	6,35%

<b>Arte</b>	3	4,76%
<b>Psicología</b>	3	4,76%
<b>Derecho</b>	3	4,76%
<b>Animación</b>	2	3,17%
<b>Fotografía</b>	2	3,17%
<b>Prod. Audiovisual</b>	2	3,17%
<b>Futbol</b>	2	3,17%
<b>Turismo/Hotelería</b>	2	3,17%
<b>Cine</b>	2	3,17%
<b>Historia</b>	2	3,17%
<b>Criminalística</b>	2	3,17%

## 2.5 Elección de opciones de formación en educación superior

Como aclaramos previamente, los 911 estudiantes encuestados pudieron en esta instancia seleccionar todas las opciones que les resultasen más interesantes, de manera que se terminaron acumulando un total de 6408 elecciones de las que fueron listadas, más 63 que no estaban dentro de las ofrecidas en la encuesta.

Al momento de agrupar las opciones, podríamos haber separado las opciones de carreras de grado de aquellas instancias de formación más cercanas a la práctica ya los oficios, tales como cursos de seguridad e higiene, reparación de celulares, etc. Sin embargo, una ventaja metodológica de agrupar todas las opciones en la misma tabla es poder dar cuenta del hecho de que los estudiantes no marcaron más las opciones que corresponden a carreras de grado que aquellas que corresponden a instancias de formación más cortas y “no profesionales”, sino que hubo cierta paridad en la elección entre unas y otras. Se puede pensar que esto tiene que ver con aquello que se valora lo cual, como venimos trabajando, tiene cada vez más que ver con la adquisición de ciertas competencias que con la posesión de una certificación académica, en vistas a una pronta inserción laboral. De hecho, las dos primeras opciones más elegidas fueron “Fotografía”, y “Electricista”, ninguna de las cuales corresponde a instancias formativas de grado ni de pre-grado que se encuentren vigentes.

Estas primeras opciones fueron seguidas bastante de cerca por tres carreras del departamento de ingeniería: Ingeniería electrónica, Arquitectura e Ingeniería informática. En lo que hace a las elecciones que les siguen, corresponden a la tecnicatura en sonido y grabación (en oferta), luego dos posibles tecnicaturas o cursos: programación y reparación de celulares y después otras dos tecnicaturas ya existentes en el departamento: Desarrollo Web y Diseño de Aplicaciones Móviles. Dentro del rubro que agrupa a otras carreras no listadas en la encuesta, las opciones fueron por lejos menos numerosas, y agrupan en primer lugar de las opciones a periodismo deportivo, diseño gráfico y deportes.

## 2.6 Instituciones elegidas

Finalmente, al momento de la elección de las instituciones en las cuales los estudiantes llevarían a cabo esas instancias de formación, la institución más seleccionada por lejos, con casi el 40% de elecciones sobre el total, fue la UNLaM, seguida de la UBA y luego la UTN, y recién después la instancia genérica de institutos terciarios de formación docente. Esta última elección es llamativa, dado que el número de veces en que esta opción fue elegida, no concuerda con la cantidad de veces que la opción “profesor” fue seleccionada (de hecho, fue una opción minoritaria y marginal en términos estadísticos). Cabe preguntarse si esta elección de institución corresponde a algo no manifestado o si parte de los cursos e instancias de formación en competencias de oficios fueron pensados a ser cursados en institutos de formación superior

### 3. Conclusiones

Los datos muestran que, dentro de una población estudiantil que posee en su mayoría una formación secundaria técnica, la demanda y valoración de instancias de formación superior de tipo breve y orientada hacia la adquisición de competencias laborales específicas (como tecnicaturas o cursos cortos) está a la par de la demanda de ofertas de formación académica de grado. Esto refuerza, desde el lado de la demanda, la idea de que el giro en educación Superior es, cada vez más, hacia la adquisición de competencias y cada vez menos hacia la adquisición de saberes generalistas que alargan los plazos de formación y graduación en la educación superior.

Conocer cuáles son, en concreto, las especialidades demandadas, permitirá tomar decisiones a nivel departamental y de toda la universidad, que brinden respuestas a estas nuevas demandas y que permitan diversificar la oferta académica de forma de acercar a la comunidad de jóvenes egresados de la educación secundaria, a una universidad más acorde a los intereses y expectativas de los estudiantes y del medio económico y social en el que vivimos. En relación a esto último, queda para una próxima etapa completar este estudio con una investigación que dé cuenta de las necesidades del sector productivo, ya que se requiere tener en cuenta a “ambas partes” para tomar decisiones a nivel de planes de estudio y políticas educativas.

Por otra parte, el DIIT ha comenzado recientemente a implementar la edición 2020 del Proyecto de investigación Vincular. En el mismo, nos proponemos conocer el impacto cognitivo-competencial de los dispositivos de educación mediada desplegados en el contexto de la pandemia de COVID-19 en nuestro país. Esto significa poder evaluar fortalezas, oportunidades, debilidades y amenazas generadas durante este proceso, principalmente en relación a las competencias (saberes-hacer) adquiridas; Especialmente de aquellas que son relevantes para la ulterior adquisición del oficio del estudiante (Igarza, Viel et al, 2019) en una modalidad particular, que se proyecta como semipresencial. También se comienza a evaluar en el mismo aspecto de desempeño académico de los estudiantes secundarios, de relevancia para este proyecto.

Toda esta información permitirá tomar decisiones respecto a la planificación del curso de ingreso en el actual contexto de pandemia y post-pandemia, así como en vistas a ampliar y mejorar la oferta académica del Departamento y la Universidad.

### Referencias Bibliográficas

1. Consejo Federal de Decanos de Ingeniería (CONFEDI) (2018). Propuesta de estándares de segunda generación para la acreditación de carreras de ingeniería en la República Argentina, Mar del Plata: Universidad FASTA Ediciones. Disponible en: [https://confedi.org.ar/download/documentos\\_confedi/LIBRO-ROJO-DE-CONFEDI-Estandares-de-Segunda-Generacion-para-Ingenieria-2018-VFPublicada.pdf](https://confedi.org.ar/download/documentos_confedi/LIBRO-ROJO-DE-CONFEDI-Estandares-de-Segunda-Generacion-para-Ingenieria-2018-VFPublicada.pdf)
2. De Sena, A. (2020). “¿Cómo se está viviendo esta cuarentena en La Matanza?” En Síntesis Clave N° 150, Boletín informativo de la Secretaría de Ciencia y tecnología de la Universidad Nacional de La Matanza.
3. Díaz, Frida, and Hernández, Gerardo (2002). Estrategias docentes para un aprendizaje significativo. Vol. 2. México: McGraw-Hill.
4. Igarza, S.; Viel, P.; Juárez, M.; y Kotliar, N. (2020) “Un aporte a la Construcción del Oficio de Ingeniería”. 7° Congreso Nacional de Ingeniería Informática – Sistemas de Información. Universidad Nacional de La Matanza.
5. March, A. F. (2006). Metodologías activas para la formación de competencias. *Educatio siglo XXI*, 24, 35-56.
6. Sobrevila, Marcelo A. (2000). La formación del ingeniero profesional para el tiempo actual: tesis de las ingenierías de base. Academia Nacional de Educación.
7. SPENCER, L. Y. S., & Spencer, S. (2011). *SM (1993). Competence at work. Models for superior performance.*
8. Teixidó Saballs, J. (2007) "Competencias para el ejercicio de la dirección escolar." Bases para un modelo de desarrollo profesional de directivos escolares basados en competencias. XVIII Jornadas Estatales del Forum Europeo de Administradores de la Educación. Toledo. España

# Didáctica de las Ciencias de la Computación: Experiencias y Percepciones de Docentes de Educación Primaria

Francisco Bavera<sup>1</sup>, Marcela Daniele<sup>1</sup>, Flavia Buffarini<sup>1-2</sup>,  
Teresa Quintero<sup>1-2</sup>, Cecilia De Dominicci<sup>2</sup>

<sup>1</sup>Universidad Nacional de Río Cuarto, Facultad de Ciencias Exactas, Físico-Químicas y Naturales

<sup>2</sup>Instituto Superior Ramón Menéndez Pidal, Río Cuarto

[pancho@dc.exa.unrc.edu.ar](mailto:pancho@dc.exa.unrc.edu.ar), [marcela@dc.exa.unrc.edu.ar](mailto:marcela@dc.exa.unrc.edu.ar), [fbuffarini@exa.unrc.edu.ar](mailto:fbuffarini@exa.unrc.edu.ar),  
[tquintero@exa.unrc.edu.ar](mailto:tquintero@exa.unrc.edu.ar), [cdominici2565@gmail.com](mailto:cdominici2565@gmail.com)

**Abstract.** Este trabajo presenta un análisis de las percepciones y experiencias de docentes de Educación Primaria en su recorrido por el cursado del primer año de dictado de la Especialización Docente de Nivel Superior en Didáctica de las Ciencias de la Computación. Se obtuvieron datos que aportan valiosos resultados preliminares al desarrollo de conocimientos sobre la formación continua de docentes de Educación Primaria en Didáctica de las Ciencias de la Computación. El estudio atrapa parte de la complejidad inherente al proceso dialéctico de construcción del Pensamiento Computacional a partir de la propia experiencia de formación docente y las maneras/modos de construcción, de dicho pensamiento, que se constituyan en nuevas formas de concebir las prácticas educativas en la escuela.

**Palabras claves:** Pensamiento Computacional, Habilidades Cognitivas, Didáctica de las Ciencias de la Computación, Formación docente continua

## 1 Introducción

Sin dudas, el desarrollo del Pensamiento Computacional (PC) [1-4] en la educación integral de ciudadanos críticos, presenta un desafío de relevancia para el sistema educativo. Estudios y tendencias a nivel mundial dan cuenta de la importancia de introducir el PC en el sistema educativo obligatorio. Estonia, Reino Unido, Finlandia, Francia y Australia han sido pioneros en esta decisión. En países de América Latina, como Uruguay, Costa Rica, Perú, Colombia y República Dominicana, han generado diversas propuestas e incluido las Ciencias de la Computación en sus sistemas educativos. En la Argentina, en el año 2015, el Consejo Federal de Educación, destacó la relevancia que reviste en la actualidad la enseñanza y el aprendizaje significativo de la programación “la programación es de importancia estratégica en el Sistema Educativo Nacional durante la escolaridad obligatoria, para fortalecer el desarrollo económico-social de la Nación, conforme lo establecido por el artículo 3° de la Ley de Educación Nacional” (Resolución N° 263/15). Y en 2018, este consejo aprobó los núcleos de aprendizaje prioritarios para la educación digital, programación y robótica (NAP, Res. CFE N° 343/18), imponiendo a las jurisdicciones analizar y llevar adelante su implementación e inclusión en sus documentos curriculares.

La construcción del pensamiento computacional requiere mayor análisis e investigación, explorando sus dimensiones, importancia y beneficios [6]. En la actualidad, es escaso el número de tesis o artículos científicos acerca del tema. En nuestro país, y en particular en nuestra provincia, se hace necesario profundizar en trabajos de investigación que reflejen los resultados obtenidos con educadores, como así también, con alumnos de la escuela primaria, que hayan recibido formación apuntada a la construcción del pensamiento computacional. En la provincia de Córdoba, existen algunas escuelas que tienen una orientación hacia la informática, algunas en TICs, como así también otras en programación. Además, en el año 2014, comenzaron a funcionar las escuelas experimentales PROA (Programa Avanzado en Educación) haciendo foco en las TICs y en el desarrollo de software. Se trata de un proceso innovador para escuelas secundarias, donde la provincia es la primera en Argentina en implementar este tipo de formato educativo que tiene como objetivo la formación en tecnologías de la información y comunicación, otorgando el título de bachiller en desarrollo de software. Por otra parte, la Fundación Dr. Manuel Sadosky, dependiente de la Presidencia de la Nación Argentina, comenzó a desarrollar el programa Program.AR [5] en el año 2013, introduciendo en algunas comunidades educativas de las escuelas de todo el país, nociones de programación y robótica, así como la didáctica de las Ciencias de la Computación.

En la actualidad, el desafío es evaluar en profundidad el impacto de estas iniciativas en la escuela [6], en particular, estudiar si a partir de experiencias de formación los docentes construyen dialécticamente su propio pensamiento computacional y se apropian de *maneras/modos* de construcción que se constituyen en nuevas prácticas educativas que determinan la posibilidad del desarrollo de dicho pensamiento en sus estudiantes de la escuela. Este equipo de investigación, ha comenzado un proceso de estudio [7-10] intentando atrapar parte de la complejidad inherente al proceso de construcción del pensamiento computacional en la escuela, que se identifica a partir de los ejemplos particulares que representa cada grupo -docentes y sus estudiantes de la escuela- en los que se focaliza la investigación. La producción de conocimiento de esta investigación será uno de los insumos para el desarrollo de propuestas de enseñanza situadas para la formación y el desarrollo del pensamiento computacional en docentes y estudiantes.

En el año 2018 comenzó el dictado de la Especialización Docente de Nivel Superior en Didáctica de las Ciencias de la Computación, iniciativa conjunta de la Fundación Sadosky, el Instituto Superior de Formación Docente Ramón Menéndez Pidal y la Facultad de Ciencias Exactas Físico-Químicas y Naturales de la Universidad Nacional de Río Cuarto. Esta instancia de formación destinada a docentes de Nivel Primario tiene como objetivo introducir a los docentes en el Pensamiento Computacional, las Ciencias de la Computación y la programación con el fin de brindarles las herramientas y experiencias necesarias para que puedan replicarlo en sus aulas.

Este trabajo, refleja la investigación realizada a partir del análisis de encuestas y entrevistas, observaciones en clase y de las producciones personales y grupales de los docentes cursantes. Entre el material empírico analizado podemos mencionar: problemas resueltos, programas entregados, planificaciones realizadas y actividades de prácticas docentes. El equipo de investigación, realizó un análisis y se obtuvieron datos que representan resultados preliminares que aportan al desarrollo de conocimientos sobre la formación continua de docentes de Educación Primaria en Didáctica de las Ciencias de la Computación.

## 2 La Especialidad en Didáctica de las Ciencias de la Computación

La primer cohorte de la Especialización Docente de Nivel Superior en Didáctica de las Ciencias de la Computación comenzó el cursado en marzo de 2018. El dictado se realizó en las instalaciones del Instituto Superior de Formación Docente Ramón Menéndez Pidal de la ciudad de Río Cuarto. Y como se mencionó previamente, esta formación es una iniciativa conjunta de dicho Instituto junto con la Facultad de Ciencias Exactas Físico-Químicas y Naturales de la Universidad Nacional de Río Cuarto y la Fundación Sadosky.

La Especialización tiene una duración de 400 horas, divididas en 8 módulos, a lo largo de dos años. Los distintos módulos introducen a los docentes en conceptos de ciencias de la computación, ciudadanía digital, pensamiento computacional, programación y robótica. Durante el primer año se dictaron 4 módulos: (1) Herramientas de comunicación y colaboración, (2) Introducción a la Resolución de problemas, (3) Introducción a los Lenguajes de Programación y (4) Administración y Configuración de Software y Hardware, y en forma transversal se acreditaron horas en el módulo Práctica Docente. Culminando, este primer año, con la *I Jornada de Experiencias Docentes en Didáctica de las Ciencias de la Computación* en noviembre de 2018. Esta jornada constituyó un espacio de integración de conocimientos y producciones donde los docentes cursantes socializaron sus experiencias. La evaluación de cada módulo estuvo centrada en actividades y experiencias áulicas, generadas e implementadas en forma grupal por los docentes, a modo de actividades integradoras en cada uno de ellos.

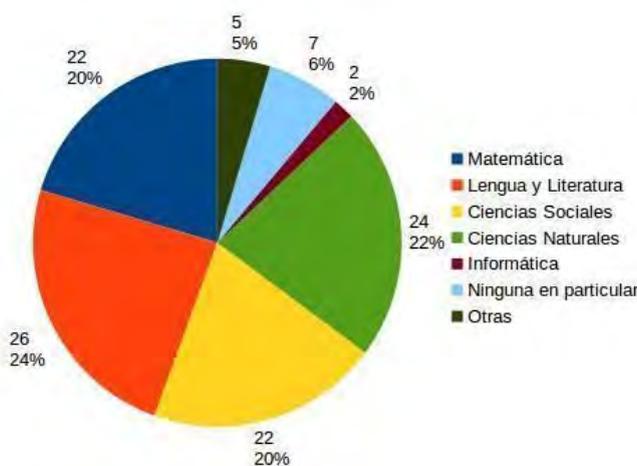
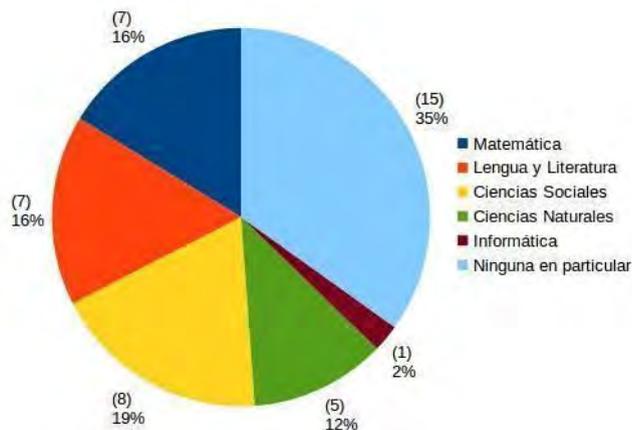
Los temas vistos en estos módulos incluyen: huella digital, ciudadanía digital responsable, derechos de autor y propiedad intelectual. Estrategias para la resolución de problemas, abstracción, reconocimiento de patrones y generalización. Instrucciones, algoritmos, lenguajes de programación. Prueba y depuración de algoritmos. Verificación y validación de algoritmos. Relación entre software y hardware. Conceptos básicos de sistemas operativos y redes.

A lo largo del dictado se utilizaron actividades enchufadas y desenchufadas. Entre las herramientas utilizadas en las actividades enchufadas se pueden mencionar: Lightbot, code.org, Pilas-Bloques, Tortugarte y MBot.

Culminaron el cursado de este primer año cuarenta y tres (43) Profesores de Educación Primaria (40 maestras y 3 maestros), estudiantes de la especialidad, que forman parte de veintiocho (28) escuelas primarias de la ciudad y la región. El 98% de estos docentes de nivel primario son maestras/os de grado y sólo una maestra del área Informática. En el gráfico N° 1 se presenta más detalles sobre las especialidades de la formación de estos docentes. Y en el gráfico N° 2 se puede ver las asignaturas que están a su cargo.

La *media* de edad es 42 años, en un rango que va desde los 23 a los 54 años. Con una *mediana* y una *moda* de 43 y 44 años, respectivamente. La *desviación media* es de 4,5. El 50% son docentes titulares y los restantes son interinos y suplentes, que se desempeñan en los diferentes grados de la escuela primaria. Además, dos docentes no están relacionados con institución alguna y un docente desempeña actividades en forma ad-honorem. Como se mencionó previamente, los docentes pertenecen a 28 escuelas, donde el 80% son instituciones públicas de gestión estatal y el restante 20% instituciones públicas de gestión privada. Un 75% de estas escuelas están localizadas en la ciudad de Río Cuarto, y el restante 25% en localidades de su región de influencia (en un radio de 30 a 150 km de distancia de Río Cuarto).

**Gráfico N° 1: Especialidad en la Formación de los docentes cursantes**



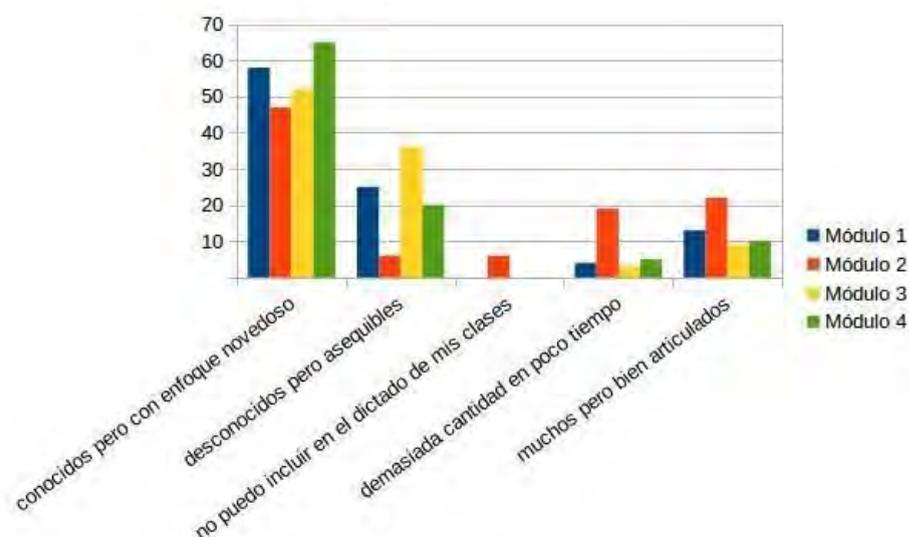
**Gráfico N° 2: Materias dictadas por los docentes cursantes**

El 98% de los docentes informaron, al comienzo de la especialidad, que tenían conocimientos de informática. De los cuales, el 30% proviene de capacitaciones formales, otro 30% recibió capacitaciones informales y el restante 38% se considera autodidacta. Ninguno participó previamente en capacitaciones relacionadas con las Ciencias de la Computación, la programación o la robótica. El 2% restante no tiene conocimiento alguno de informática.

Al indagar sobre las experiencias previas que tuvieron con la informática en el ejercicio docente se pudo recabar que 21 docentes emplearon utilitarios de oficina, 7 aplicaciones de celulares y 9 docentes otras aplicaciones. Mientras que 12 maestras nunca realizaron experiencias relacionadas con la informática en su aula. Sólo 5 de los docentes tuvieron alguna experiencia con lenguajes de programación visuales y 2 con robótica educativa.

### 3 Percepciones de los Docentes sobre la Especialización

Como se puede apreciar en el gráfico N° 3, el balance, con respecto a la percepción de los contenidos por parte de los docentes cursantes, es positivo. Una gran mayoría considera que los contenidos son “conocidos pero con un enfoque novedoso”, “asequibles” y “bien articulados”.



**Gráfico N° 3:** *Cómo evalúan los contenidos de los módulos los docentes cursantes.*

Con respecto a la propuesta didáctica utilizada, los docentes, reconocieron que fue innovadora (para ellos) y que fue un disparador para interpelar sus prácticas y pensar nuevas maneras de dar sus clases. Menos del 5% respondió que no sabría cómo dar sus clases de esa manera (ver Gráfico N° 4). Cabe resaltar que en el dictado se hizo un especial énfasis en *enseñamos cómo aprendemos*. Es decir, se plantearon actividades y formas de trabajo con el fin de que los docentes los apliquen en sus prácticas.

Las distintas herramientas utilizadas a lo largo de los distintos módulos también tuvieron una muy buena respuesta por parte de los cursantes. Como se puede ver en el gráfico N° 5, el 85% expresa que pudo utilizarlas para planificar y dar actividades en sus aulas. Cabe mencionar que estos docentes consideran que son capaces de utilizar herramientas como Lightbot, Pilas-Bloques, code.org y actividades desenchufadas en tareas de introducción a la programación y actividades que fomenten el pensamiento computacional.

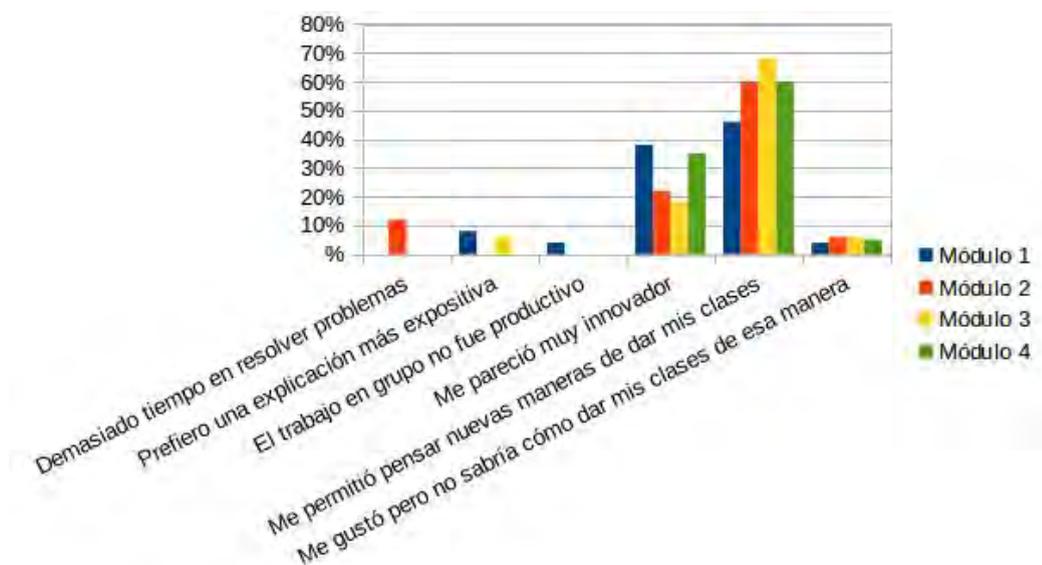
Los docentes resaltan la posibilidad de que los contenidos dados y la modalidad de trabajo aplicada les permitió llevar al aula lo trabajado en los distintos módulos. También recalcan que la mayoría de las actividades pueden ser adaptadas para los distintos grados. La experiencia de trabajo colaborativo fue altamente valorada y puesta en práctica por muchos de ellos con sus alumnos.

Aproximadamente el 65% de los docentes manifiestan poca disponibilidad de equipamiento informático sumado a dificultades en la conectividad en sus instituciones, lo cual, dificulta la implementación de actividades enchufadas en sus escuelas. Así mismo el 85% logró implementar este tipo de actividades en sus aulas en el marco de la Especialidad.

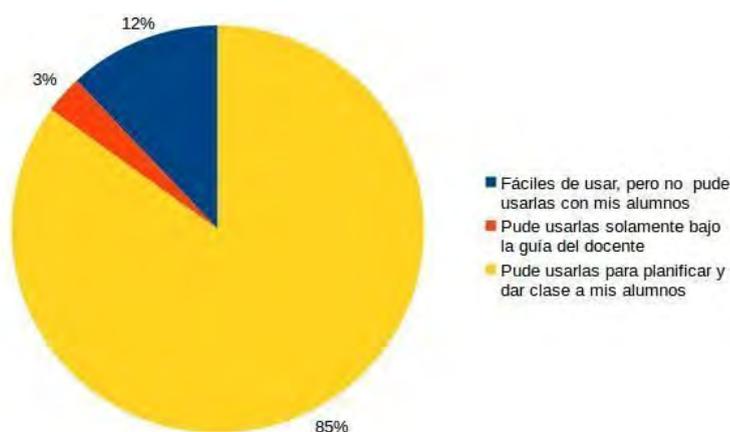
Se observa que este grupo de docentes siente más seguridad (y familiaridad) con las actividades desenchufadas a la hora de implementarlas en el aula. El factor tecnológico y las dificultades por ellos planteadas son el principal motivo de esto.

Es importante resaltar que un alto porcentaje afirma que esta formación les permitió adquirir y/o afianzar ciertas habilidades relacionadas con el pensamiento computacional:

- Abstracción: resolver un problema centrandolo en ciertos atributos y propiedades más importantes (61% de los docentes cursantes).
- Descomposición: Abordar la resolución de problemas con un proceso sucesivo de división del problema en subproblemas más pequeños (61% de los docentes cursantes).
- Reconocer Patrones: Pensar y proponer soluciones que se puedan generalizar para resolver tipos o categorías de problemas similares (45% de los docentes cursantes).
- Solución Algorítmica: instrucciones que ordenadas de una cierta manera permiten obtener la solución a un problema (45% de los docentes cursantes)
- Modelado y simulación: crear modelos o simulaciones para representar procesos (11% de los docentes cursantes).



**Gráfico N° 4:** Cómo evalúan la didáctica de los módulos los docentes cursantes.



**Gráfico N° 5:** Cómo evalúan las herramientas (LightBot, Pilas-Bloques, Tortugarte, code.org, unplugged) trabajadas en el módulo los docentes cursantes.

Es llamativo la falta de conocimientos (previos) con respecto a ciudadanía digital y huella digital. Estos temas despertaron mucho interés y debates muy enriquecedores.

El trabajo realizado sobre resolución de problemas, planteos de estrategias de resolución, búsquedas de patrones, generalización de soluciones y validación y verificación de las mismas usando fórmulas algebraicas, si bien con una resistencia inicial, tuvieron un buen impacto en el resto de los módulos ya que sentaron las bases de la forma de trabajo esperada para la resolución de problemas y el planteo de algoritmos. Estas actividades lograron una gran aceptación y reconocimiento a medida que se fue avanzando en el desarrollo de la especialidad.

### 3.2 Percepción del Impacto sobre sus Prácticas Docentes

De los docentes cursantes el 55% considera que el pensamiento crítico y computacional les permitió definitivamente transformar y mejorar sus prácticas docentes. Mientras que el 45% respondió que cambió y mejoró sólo en parte sus prácticas. Además, es importante resaltar que el 72% asegura que modificó algunas de sus actividades docentes teniendo en cuenta: la programación, el diseño de algoritmos, la abstracción, la descomposición del problema, la definición de modelos y de patrones. Mientras que el 22% afirma que transformó todas sus actividades docentes. Sólo el 6% afirma que no lo hizo, pero que lo realizará en el futuro. Ningún docente respondió que ya lo realizaba antes de cursar el postítulo o que considera que no podría hacerlo.

La recepción, por parte de sus estudiantes, de las actividades de introducción a la programación, pensamiento computacional y Ciencias de la Computación (enchufadas y desenchufadas) los sorprendió gratamente. Ya que los resultados y las respuestas que obtuvieron superaron sus expectativas. Mencionaron los términos motivación, interés, entusiasmo, logros, experimentación y curiosidad de sus alumnos para expresar la respuesta que tuvieron en el aula.

## 4 Evaluación de los Docentes Cursantes

A continuación se presentan los resultados obtenidos por los docentes referidos a las habilidades del PC, la programación, sus prácticas docentes y los trabajos integradores del primer año de cursado de la Especialidad.

### 4.1 Pensamiento Computacional

En trabajos previos [9,10], se presentaron resultados que muestran que los docentes cursantes pueden resolver problemas que implican utilizar habilidades asociadas al PC. Resolvieron correctamente el 75% de los problemas planteados. Dichos problemas son de mediana complejidad y cabe resaltar que estos problemas pueden ser trabajados en el aula, no sólo con el fin de evaluar las habilidades del PC, sino también para desarrollar el PC y las Ciencias de la Computación.

Los mejores resultados se observaron en los problemas asociados a *encontrar un camino* y a *reconocer patrones en el orden* de elementos. Aunque, se observa una leve disminución en la efectividad de resolución, en las actividades donde se debía encontrar *el mejor camino*. Es necesario indagar en las causas que motivan esta diferencia, las cuales pueden ser numerosas y diversas, pero una de las posibilidades es que la mayoría de las respuestas incorrectas pueden deberse a una interpretación errónea de las consignas, a no identificar un dato relevante en el enunciado, o a dificultades relacionadas con las habilidades de abstracción.

Los docentes resolvieron correctamente el 76% y 73% de los problemas que involucran habilidades relacionadas con *descomposición en subproblemas* y *abstracción*, respectivamente. En el caso de *pensamiento algorítmico* respondieron correctamente el 76% de las veces. En cuanto a *reconocimiento de patrones*, se puede apreciar que tienen un 73% de respuestas correctas en total. En el caso de *modelado y simulación*, se puede observar que los docentes respondieron correctamente el 78% de las veces.

Se pudo observar que los docentes manifiestan dificultades significativas para describir y explicar en forma escrita el proceso de resolución de los problemas planteados. Solo un bajo porcentaje pudieron escribir de manera detallada y entendible el proceso. La mayoría lo expresó con frases breves y términos generales. En general, solo nombran una lista de procedimientos, por ejemplo, textualmente expresan: *“observé y descarté”*, *“leí, razoné y reconocí el patrón”*, *“observación, reemplazando patrones, reconocimiento”*, *“por descarté”*, *“comparé”*, entre otros.

Por otra parte, una cantidad elevada de docentes identificaron las habilidades del PC involucradas en la resolución de los problemas, siendo *Reconocimiento de Patrones* la que pudieron reconocer con mayor certeza

### 4.2 Programación

Se percibe que alrededor del 50% de los docentes alcanzó las metas de programación y además autonomía en el desarrollo de programas y la percepción sobre las posibilidades brindadas por la programación. Pero, un 25% de los cursantes tiene mucha dificultad a la hora de programar. El resto puede realizar actividades simples con cierta autonomía.

En lo que respecta al manejo del código, aparecen allí las mayores dificultades, y también las diferencias más pronunciadas entre cursantes. Cabe mencionar que la introducción de un lenguaje de programación textual (javascript), tuvo un grado de dificultad importante, en comparación con código en lenguaje de bloques (por ej. code.org). Los docentes, en su gran mayoría, lo encontraron mucho más complejo y con menos posibilidades de implementarlo en el aula que la programación con bloques. Por lo que se focalizó exclusivamente en lenguajes de programación con bloques.

En general, identificaron errores en programas con mayor precisión y facilidad que la que tuvieron para implementar algoritmos para problemas específicos. Se percibe que los docentes cursantes requieren de mayor cantidad de actividades prácticas de programación y que sienten la necesidad de aprender más sobre este tema.

Con respecto al nivel de reconocimiento y manejo de los conceptos básicos asociados a la enseñanza de la programación, el 60% se encuentra en un nivel básico y el resto se reparte casi en partes iguales dentro de un nivel incipiente y suficiente.

### 4.3 La Práctica Docente

Los docentes lograron trabajar contenidos y habilidades abordados en los módulos de la Especialización en distintas áreas de la enseñanza primaria, aplicación en otras áreas escolares (Ciencias Naturales, Ciencias Sociales, Matemática, Lengua y Tecnología). En la mayoría de los trabajos se observa, la transferencia al aula y el trabajo con sus estudiantes de algunas de las propuestas desarrolladas en la especialización.

Las producciones de los grupos docentes giraron principalmente sobre: situaciones problemáticas, trabajo colaborativo y desarrollo del pensamiento computacional. Las actividades implementadas en el aula se establecieron en torno a la búsqueda y explicitación de estrategias en la resolución de los problemas con posteriores debates y análisis entre sus alumnos sobre las distintas soluciones. Hicieron hincapié en la validación de las soluciones, la detección y solución de los errores. Trabajaron claramente sobre la necesidad de contar con instrucciones precisas y no ambiguas en la definición de secuencias de instrucciones (algoritmos) para expresar las soluciones a los problemas.

En el aula realizaron actividades enchufadas y desenchufadas. El 82% de los docentes trabajaron ambos tipos de actividades con sus estudiantes, el 3% solo realizaron actividades enchufadas y el restante 15% solo realizó actividades desenchufadas.

Sin embargo, aproximadamente el 80% de los docentes cursantes tienen un mayor grado de predisposición a desarrollar actividades desenchufadas. Aunque, es similar el porcentaje que considera que podría introducir a la programación y/o enseñar a programar animaciones o simulaciones simples.

### 4.4 Las Producciones Integradoras

Se desarrollaron un total de quince producciones integradoras del primer año, donde los docentes cursantes se agruparon y elaboraron las diferentes propuestas. Las producciones representaron diferentes dimensiones de la formación docente: cómo habitaron la formación y capacitación sobre un contenido disciplinar, cómo construyeron conocimientos y cómo resolvieron la transposición didáctica en sus aulas de la escuela primaria.

Todos los equipos de trabajo, diseñaron estructuras gráficas para representar la forma en que integraron los conocimientos a la formación docente inicial y a sus prácticas áulicas. Uno de los equipos, describe el recorrido en *forma espiralada*, dando cuenta de una construcción dialéctica y reflexiva del conocimiento. “... representa el crecimiento en nuestros conocimientos, como así también, en las experiencias áulicas realizadas. Reflexionamos sobre cómo este recorrido modificó nuestras prácticas docentes y cómo resignificamos y adaptamos los contenidos y actividades áulicas a partir de nuestro recorrido por el postítulo...” (grupo 10). Otro grupo, lo representa con figuras geométricas, dando cuenta de una construcción integrada y progresiva de los conocimientos. “... la jerarquización del círculo central, se debe a que la enseñanza en los distintos espacios curriculares centrada en el desarrollo de las capacidades fundamentales prioritarias permite a todos los estudiantes participar y potenciar las mismas de acuerdo a sus posibilidades...” (grupo 7).

La mayoría de los trabajos, incluyeron la formación de ciudadanos críticos, “... En el óvalo central se ubicó el Ciudadano Digital, debido a que se pensó que el punto de partida es promover el desarrollo de ciudadanos críticos, responsables, libres e integrados, en el mundo real y digital...” (grupo 14), en el sentido de reconocer la enseñanza de las Ciencias de la Computación en el Nivel Primario con el objetivo de “... formar ciudadanos con conocimientos suficientes sobre tecnología que les permitan participar en debates, con libertad de expresión en la red, capaces de comprender cómo funciona, capaces de intervenir y capaces de crear...” (grupo 4)

Otros equipos integraron los contenidos de los diferentes módulos de aprendizajes, a partir de resaltar la relevancia de los conceptos disciplinares y su relación con el desarrollo del pensamiento infantil. “... *La enseñanza de la programación fomenta el pensamiento computacional. Este proceso de razonamiento estructura la mente y ordena las ideas, permitiendo dividir un problema grande en problemas más pequeños, hasta encontrar la solución...*” (grupo 7).

La mayoría de los trabajos resaltan las relaciones que establecieron entre la formación sobre un nuevo contenido disciplinar y las formas de integrarlos al aula y a sus prácticas docentes. El grupo 15, propone incorporar los conceptos de las Ciencias de la Computación de manera “transversal”, para procurar “... *un rol activo del docente en la formación de los ciudadanos digitales desde la primera infancia...*”. En el mismo sentido el grupo 2, plantea un rol activo del docente en relación a la incorporación de los contenidos en el aula, anticipando la participación del mismo en la evaluación crítica de los recursos y entornos digitales para “... *problematizar e interpretar situaciones reales, a la vez que desarrollar la creatividad y el trabajo en colaboración, como así también contribuye al acercamiento a la tecnología desde un posicionamiento más crítico...*”. El grupo 3, plantea la necesidad de realizar una “transposición didáctica” para el inicio de la construcción de conocimientos y como “... *forma de ampliar el horizonte de nuestras prácticas y los aprendizajes mutuos en esta triada tan conocida de estudiante-conocimiento-docente...*”.

Algunas habilidades que se encontraron en las producciones y que se identifican con las del pensamiento en la era digital, son: dar y seguir instrucciones, secuenciar, ordenar, probar, ejecutar, verificar, analizar, validar, codificar, validar distintas estrategias de solución, identificar situaciones, descomponer en subproblemas, construir algoritmos, reconocer patrones, usar herramientas de colaboración y comunicación, aplicar estrategias, trabajar colaborativamente, buscar y validar información, abordaje y resolución de problemas, programación en acción, robótica, producciones creativas, fundamentación de procedimientos.

## 4 Conclusiones

A partir de la capacitación docente en el marco de la Especialización en Didáctica de las Ciencias de la Computación, los docentes pudieron apropiarse, en diferentes tiempos, de conceptos centrales del PC, de ciencias de la computación y de programación; y ensayar experiencias áulicas, para su análisis y reflexión desde la Didáctica. El diseño y organización de los módulos con horas destinadas a Práctica Docente posibilitó a los maestros integrar progresivamente los contenidos específicos con otras áreas de conocimientos de los Diseños Curriculares para el Nivel Primario.

Respecto a los aprendizajes logrados por los participantes en relación a la programación y a las posibilidades de enseñarla en el aula de primaria, se considera que esta experiencia de formación posibilitó que muchos docentes comprendieran el sentido de las estrategias para la construcción de una solución algorítmica de un problema, y tuvieran una primera aproximación a lo que implica crear un programa, comprender su complejidad, ejecutarlo, encontrar errores utilizando un entorno de programación. A partir de esto, los docentes, lograron realizar sus primeras experiencias con actividades de programación en las aulas (con el acompañamiento de pares y/o de los profesores de la especialidad). La gran mayoría de los docentes pudieron diseñar e implementar (en sus aulas) propuestas didácticas de introducción a la programación utilizando actividades desenchufadas y enchufadas con herramientas, tales como, LightBot, code.org, Pilas-Bloques o similares. En especial, se observó que pueden replicar las actividades realizadas en el cursado de la especialidad en sus aulas sin dificultades. No hay que dejar de mencionar que las actividades realizadas por los docentes en el aula tuvieron una buena recepción por parte de sus alumnos. Estos demostraron mucho interés, entusiasmo y curiosidad. Alcanzaron un alto nivel de motivación y lograron realizar todas las actividades propuestas.

El módulo de Resolución de Problemas configuró un encuadre de trabajo que posibilitó construir sentido didáctico para la incorporación progresiva de los conceptos específicos de los diferentes módulos. La posibilidad de articular los conocimientos de la Didáctica de la Matemática, como sustento para la construcción del Pensamiento Computacional, tuvo resultados relevantes y prometedores en la formación en Ciencias de la Computación de los maestros. Pero, se considera necesario articular más profundamente este módulo con los demás módulos de la carrera, recuperando actividades y rediseñando nuevas secuencias.

Los docentes cursantes se apropiaron de los contenidos sobre propiedad intelectual y diferentes tipos de licencias, comenzaron a identificar las diferencias básicas entre las licencias Copyright, Copyleft y Creative Commons. Además, reflexionaron sobre los derechos de autor y la necesidad de generar ciudadanía digital desde edades tempranas. Se observa en sus producciones, que comenzaron a incorporar en el aula la construcción responsable de la identidad digital y de la huella que cada uno va dejando al ser usuario de la red. Pudieron revisar sus prácticas como docentes, como padres, como usuarios de Internet, y comenzar a partir de ello a actuar de manera diferente y pensar la manera de transferir esos conocimientos a sus comunidades educativas. Los docentes resaltaron la importancia de conocer y poder enseñar en el aula temas relacionados con la seguridad informática, confidencialidad de la información, Gobierno Abierto, como así también, la legislación existente, como por ejemplo, la Ley de Protección de Datos y la Ley de Acceso a la Información Pública.

El uso de herramientas de apoyo a la edición colaborativa, como parte de las estrategias y de las actividades de algunos módulos, resultó novedoso y desafiante para muchos docentes, y permitió reflexionar sobre lo que implica almacenar la información localmente o de forma distribuida, analizar sus ventajas y desventajas para la edición colaborativa, y ayudó a comenzar a entender lo que se pone en juego al utilizar la nube (disponibilidad, costo, relegar privacidad, propiedad de la información y soberanía de los datos, entre otros).

El análisis realizado en torno a las producciones presentadas en el Trabajo Integrador por los cursantes de la especialización puso en evidencia características positivas ligadas a este espacio de formación, sin embargo se considera que la construcción de una *buena* propuesta de formación se retroalimenta a partir de las continuas reflexiones y revisiones sobre el trabajo realizado.

Entre las características positivas se pueden mencionar: trabajo colaborativo entre docentes de diferentes instituciones que formaron equipos de manera espontánea; propuestas de trabajo interdisciplinarias; predisposición de las instituciones a ceder sus aulas para la implementación de las propuestas generadas en torno al trabajo final; estudiantes de escuela primaria muy movilizados e involucrados en el aprendizaje ante las nuevas propuestas, apreciaciones que también manifestaron los docentes responsables; propuestas de trabajo que se construyeron a partir de las actividades abordadas en diferentes módulos de la carrera. Particularmente, se vio reflejada en la implementación la forma de trabajo vivenciada por ellos en las aulas de postítulo; el formato “taller” desarrollado en los distintos módulos favoreció que los docentes implementaran en su aula lo vivido en el postítulo poniendo en evidencia una modificación, casi en tiempo real, de su práctica docente; casi la totalidad implementó actividades desenchufadas; y, la mayoría de las propuestas incluyeron actividades de introducción a la programación.

Luego del análisis de los resultados obtenidos con esta primera experiencia de formación en didáctica de las ciencias de la computación llevada a cabo con docentes de educación primaria, algunas consideraciones importantes a tener en cuenta para una próxima propuesta de formación, giran en torno a aumentar la cantidad de variados problemas que para su resolución aprovechen la fortaleza del uso de herramientas de programación, que hagan explícito la automatización, la fiabilidad de resultados, la exhaustividad, el orden, la posibilidad de abarcar otros casos, la economía de tiempo, etc..

Para finalizar, los autores de este trabajo consideran que profundizar y apropiarse de los conocimientos específicos desde las diferentes dimensiones (epistemológica, metodológica y pedagógica) implica, en cualquier disciplina, mucho tiempo. La Formación Específica (Áreas Disciplinarias y su Didáctica) ocupa el 70% de carga horaria en la Formación Docente Inicial. Por esto, se sostiene que los Docentes que cursaron la Especialización, son capaces de diseñar y desarrollar secuencias didácticas para la implementación de contenidos de la disciplina en las Escuelas de Nivel Primario; pero, para garantizar una transposición didáctica significativa, se requiere del acompañamiento y supervisión de esas experiencias por especialistas (Profesores y Licenciados en Ciencias de la Computación) para dar continuidad y profundizar la construcción de conocimientos disciplinares y didácticos específicos.

## 5. Agradecimientos

Este trabajo no se podría haber llevado a cabo sin la participación de los docentes que cursaron la Especialización de Enseñanza Superior en Didáctica de las Ciencias de la Computación. Un agradecimiento especial a la Fundación Sadosky y al Instituto Ramón Menéndez Pidal por el apoyo y acompañamiento en el desarrollo de esta formación. También un agradecimiento a nuestros colegas que participaron del dictado de la Especialización proporcionando información y experiencias que ayudaron en gran medida a la investigación.

Este trabajo fue realizado con financiación del Ministerio de Ciencia y Tecnología de la Provincia de Córdoba y de la Universidad Nacional de Río Cuarto.

## Referencias

1. Aho, A. V. (2012). Computation and Computational Thinking. *The Computer Journal*, 55(7), 832-835. doi: <https://doi.org/10.1093/comjnl/bxs074>
2. Denning, P. J. (2017). Remaining trouble spots with computational thinking. *Communications of the ACM*, 60(6), 33-39. doi: <https://doi.org/10.1145/2998438>
3. Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-35. doi: <https://doi.org/10.1145/1118178.1118215>
4. Zapata-Ros, M. (2015). Pensamiento computacional: Una nueva alfabetización digital. RED. Revista de Educación a Distancia. Número 46. Consultado el 21/11/18. Disponible en <http://www.um.es/ead/red/46>
5. Program.AR, Disponible en: <http://program.ar/>
6. Bocconi, S., Chiocciariello, A., Dettori, G., Ferrari, A., Engelhardt, K. (2016), Developing computational thinking in compulsory education – Implications for policy and practice; EUR 28295 EN; doi: 10.2791/792158
7. M. Daniele, T. Quintero, F. Bavera, F. Buffarini, D. Solivellas, C. De Dominici (2019). Análisis de producciones de docentes de educación primaria con formación en didáctica de las ciencias de la computación. Aprobado y presentado en las Segundas Jornadas de Didáctica de la Programación, FAMAF, Universidad Nacional de Córdoba, 7 y 8 junio 2019.
8. F. Bavera, T. Quintero, M. Daniele, F. Buffarini (2019). Análisis de prácticas de docentes de educación primaria en el marco de una formación en pensamiento computacional. Aprobado para las 48° Jornadas Argentinas de Informática, SAEI-JAIIO, Universidad Nacional de Salta, 16 al 20 de septiembre 2019.
9. F. Bavera, T. Quintero, M. Daniele, F. Buffarini (2019). Habilidades de Pensamiento Computacional en docentes de primaria: evaluación usando Bebras. Publicado en las actas del XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019). Universidad Nacional de Río Cuarto.
10. F. Bavera, T. Quintero, M. Daniele, F. Buffarini (2020). Computational Thinking Skills in Primary Teachers: Evaluation Using Bebras. *Communications in Computer and Information Science book series (CCIS, volume 1184)*. Springer International Publishing. Computer Science – CACIC 2019. 25th Argentine Congress of Computer Science, CACIC 2019, Revised Selected Papers.

# Workshop Seguridad Informática

Coordinadores

Javier Diaz (UNLP)  
Hugo Ramón (UNNOBA)  
Claudio Aciti (UNCPBA)

## **Riesgos/vulnerabilidades de una red informática en un Laboratorio EDI**

Fabián A. Gibellini, Roberto Muñoz, Analía L. Ruhl, Juliana Notreni, Milagros N. Zea Cárdenas, Ignacio Sánchez Balzaretti, Cecilia Sanchez

Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas de Información/  
Universidad Tecnológica Nacional / Facultad Regional Córdoba  
Maestro M. Lopez esq. Cruz Roja Argentina S/N, Ciudad Universitaria (X5016ZAA)-  
Córdoba, Argentina  
{fabiangibellini, robertmunioz, analialorenaruhl, julinotreni,  
ignaciojsb, milyzc}@gmail.com, csanchezjuriol@hotmail.com

### **Resumen**

La contribución de este trabajo es la presentación de un modelo de riesgos/vulnerabilidades de ciberseguridad en la red informática de un Laboratorio EDI. Este modelo permite ver la relación entre los riesgos/vulnerabilidades identificados y sus respectivos procesos de negocio. Además, el modelado muestra los riesgos más relevantes, categorización de riesgos, medidas de control para cada uno de los riesgos identificados. Estos riesgos/vulnerabilidades son el punto de entrada para identificar indicadores que serán comprendidos dentro las buenas prácticas de ciberseguridad.

Este estudio se realiza en la red informática de un Laboratorio EDI (Educación, Investigación y Desarrollo). Particularmente, es en el Laboratorio de Sistemas de Información (LabSis) de la Universidad Tecnológica Nacional Facultad Regional Córdoba. Los servicios brindados por el LabSis son diversos, y administra una red informática que da soporte a actividades académicas, actividades de investigación (I+D+I) y actividades de desarrollo de software.

**Palabras claves:** ciberseguridad, redes, infraestructura, procesos, riesgos, vulnerabilidades.

## 1 Introducción

Este estudio se encuentra inserto y forma parte del proyecto “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. Cuyo objetivo final es generar buenas prácticas en la ciberseguridad de la infraestructura tecnológica.

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 (EDI).

Un Laboratorio 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. Es indispensable que este tipo de organización cuente con una infraestructura de red informática acorde para afrontar tales desafíos.

En el caso del Laboratorio de Ingeniería en Sistemas de Información (LabSis), de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC) se atienden a todos los servicios de gabinete informático para actividades académicas, desarrollo de software y proyectos de investigación.

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 [1], como así también dar soporte de infraestructura de red a proyectos de investigación como también generar nuevo software/configuración de software para suplir distintas tareas.

El presente trabajo está relacionado al trabajo presentado el año 2019 “Procedimientos de Seguridad en un Laboratorio EDI para la identificación de vulnerabilidades en su red informática.”, en el sentido que en el presente paper se pretende presentar un modelo de riesgos de ciberseguridad del Laboratorio EDI LabSis. [2]

Según la ISO/IEC 270001 es indispensable llevar a cabo acciones que den continuidad a los servicios, protegiendo la confidencialidad, integridad (autenticidad y no repudio de los datos) y disponibilidad de los mismos [3].

Consideremos que este tipo de organización, no solo brinda servicios a un tipo de usuario sino a varios tipos de usuarios (alumnos, docentes, investigadores y desarrolladores) con necesidades muy distintas. En base a esto es crucial considerar la ciberseguridad [4] [5] de una red informática de este tipo de Laboratorio EDI, en parte a la diversidad de datos sensibles que se manipulan y en parte por la masividad de datos que se manejan.

Además, se tiene que tener en cuenta que ataques no solo provienen de atacantes externos sino también puede ser de atacantes internos, los cuales puede interrumpir la información, robar información, manipularla o incriminar al titular de la información en actos que este desconoce [6].

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 estrategia de seguridad alineada, tienen una visión limitada de su madurez de ciberseguridad y poseen practicar insuficientemente para responder a un incidente de ciberseguridad [7].

## 2 Metodología

En el desarrollo de este proyecto se utiliza el método empírico [8][9] para la observación, medición y análisis de los valores que sean obtenidos. La línea de investigación que se sigue es la ciberseguridad en redes de información, el objeto de estudio la red informática del Laboratorio de Ingeniería en Sistemas de Información (LabSis), de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC). Este trabajo está conectado con “Procedimientos de Seguridad en un Laboratorio EDI para la identificación de vulnerabilidades en su red informática.”, presentado en el CACIC 2019.

## 3 Desarrollo

Con el mapa de procesos y el detalle de cada uno de los procesos identificados llevó a cabo un análisis de riesgos del Laboratorio EDI.

Para el análisis de riesgos se establecieron tres categorías al riesgo según su tipo, estas son “Técnico”, de “Negocio” y de “Proceso”.

- En el grupo de tipo Técnico se encuentran los riesgos relacionados a la calidad del servicio.
- Los riesgos de Negocio, los que afectan la viabilidad de la actividad.
- Los riesgos de Proceso a los que afectan a los procesos definidos por la organización.

Otra clasificación que se utilizó para los riesgos fue teniendo en cuenta su origen:

- Internos: aquellos que se originan dentro del contexto de la organización o son generados por personal de la misma
- Externos: que ocurren fuera del contexto de la organización es decir que es generada por un tercero.

Se identificaron los problemas realmente relevantes, o sea el 20% de los riesgos que tiene que tener sí o sí un plan de contingencia [10]. De modo que, para identificar ese 20% de riesgos e idear un plan de contingencia para los mismos se determinó un nivel de exposición para cada uno. Se desarrolló una matriz de probabilidad en la que, primero, se precisó definir frecuencia, en base a valores históricos, como también su impacto para cada riesgo [11]. La frecuencia representa la probabilidad de que ese riesgo ocurra, teniendo como referencia incidentes históricos. 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 base a esto se describieron cuatro niveles de exposición, “Bajo”, “Tolerable”, “Significativo” e “Intolerable”, tal como se muestra en la Figura 1 (Fig. 1).

La determinación del nivel de exposición conduce al nivel de profundidad en la medida de contingencia que se tomará para cada riesgo. Por ejemplo si un riesgo posee un nivel de exposición Intolerable, se debe tener sí o sí un plan de contingencia que contemple al mínimo detalle como mitigarlo, en el caso que el riesgo ocurra. En cambio con un nivel de exposición bajo, a veces no es requerido un plan de contingencia para mitigar el mismo.

El nivel de exposición de todos los riesgos fue calculado a partir del producto de su frecuencia e impacto. La frecuencia de cada riesgo se determinó de forma empírica en base a hechos históricos.

<i>Exposición</i>	<i>Es necesario tomar medidas de control?</i>
<b>Bajo</b>	Riesgo entre 0.00 y 0.19. No es necesario tomar acción para abordar el riesgo
<b>Tolerable</b>	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.
<b>Significativo</b>	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.
<b>Intolerable</b>	Riesgos entre 0.61 y 1. La implementación de medidas de control debe ser inmediata. Las tareas no deberán comenzar o serán suspendidas hasta implementar las medidas de control.

Fig.1 Niveles de Exposición de Riesgos

Una vez definidos los criterios para el análisis de riesgo, se identificaron y analizaron las posibles amenazas que pueden intervenir en el correcto funcionamiento del Laboratorio EDI de LabSis.

A continuación se describen los procesos identificados

1. Administración de aulas
  1. Asignación anual/cuatrimstral de aula para cátedra
  2. Asignación y preparación de aula para parcial o examen
  3. Asignación de aula para práctica libre
  4. Asignación de aula por única vez para cátedra
2. Administración de usuarios
3. Preparación de nuevo equipamiento
4. Administración de software
  1. Mantenimiento de Software
  2. Instalación de software para cátedras
5. Administración de bases de datos
  1. Gestión de bases de datos para cátedras
  2. Gestión de bases de datos del LabSis
6. Administración de backups
  1. Gestión de backups del LabSis
  2. 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 los cuales sumaron un total de 35, podemos observar un resumen de dicho análisis (Fig. 2).

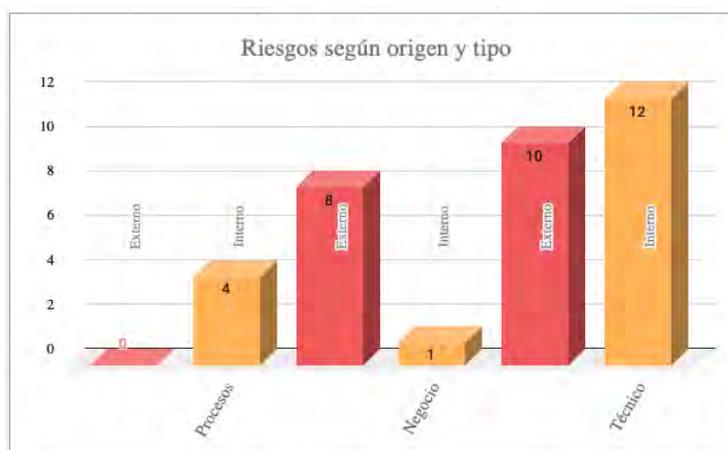


Fig.2 Cantidad de Riesgos según Origen y Tipo

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.

En las Tabla 1, Tabla 2 y Tabla 3 se pueden ver los riesgos técnicos de exposición significativo o intolerable, es decir que si ocurren su impacto puede ser catastrófico. Por lo tanto si o si deben poseer medidas de control planificadas.

**Tabla 1.** Plan de contingencia riesgo R.19

<b>ID</b>	R.19
<b>Amenaza</b>	Inconsistencia en las versiones y configuraciones del software instalado en las aulas
<b>Impacto</b>	Imposibilidad de desarrollar la actividad académica por fallas de software o por no estar disponible la característica que se necesita. Imposibilidad de mover cursos entre aulas.
<b>Frecuencia</b>	0,6
<b>Impacto</b>	0,7
<b>Exposición</b>	0,42
<b>Medidas de control</b>	<ul style="list-style-type: none"> <li>- Realizar una planilla que contenga por aula software instalado y versión</li> <li>- Actualizar aquellas aulas que se encuentran desactualizadas</li> </ul>

**Tabla 2.** Plan de contingencia riesgo R.30

<b>ID</b>	R.30
<b>Amenaza</b>	No realizar pruebas de restauración de los backups almacenados
<b>Impacto</b>	Posibilita la falla de backups a la hora de recuperarlos y pérdida de datos
<b>Frecuencia</b>	0,8
<b>Impacto</b>	0,5
<b>Exposición</b>	0,4
<b>Medidas de control</b>	<ul style="list-style-type: none"> <li>- Realizar capacitación de restauración de cada uno de los backup.</li> <li>- Realizar un instructivo con paso a paso de como realizar la restauración y posibles problemas que se pueden presentar.</li> </ul>

**Tabla 3.** Plan de contingencia riesgo R.34

<b>ID</b>	R.34
<b>Amenaza</b>	No se cambian las contraseñas de los servidores periódicamente
<b>Impacto</b>	Posibilita la filtración de contraseñas que comprometan la seguridad de la infraestructura.
<b>Frecuencia</b>	0,9
<b>Impacto</b>	0,8
<b>Exposición</b>	0,72
<b>Medidas de control</b>	<ul style="list-style-type: none"> <li>- Cambiar la contraseña cada vez que exista rotación de operadores.</li> </ul>

Aunque para el presente se visualizan las medidas de control para los riesgos con nivel de exposición significativo o intolerable, para todo el estudio se identificaron medidas de control para todos los riesgos lo que permite.

#### 4 Resultados, Avances/Discusión

Si bien los procesos descriptos fueron identificados para el LabSis, éstos pueden ser extrapolados a cualquier Laboratorio EDI. Como así también algunos de sus riesgos, con un previo análisis acorde a la adaptación de los procesos de cada laboratorio en particular.

El siguiente paso es identificar los indicadores que nos permitan monitorear y controlar estos riesgos/vulnerabilidades de la red del Laboratorio EDI de LabSis. Una vez identificados estos indicadores se procederá a la etapa de recolección de datos, los cuales generarán la base de conocimiento de indicadores de ciberseguridad de una red

informática. A partir de esta base de conocimiento de indicadores, se podrá determinar cuales serán considerados dentro las buenas prácticas de ciberseguridad

## 5 Referencias

- [1] Dharma, R., Sake, S., Manuel, M. (2013). Backup and Recovery in a SAN. Versión 1.2. EMC2 Techbooks.
- [2] (2017) A Comparison of Cybersecurity Risk Analysis Tools. Elsevier B.V.
- [3] ISO/IEC 27001. “Tecnología de la información”. Técnicas de la seguridad. Sistemas de gestión de la seguridad de la información (SGSI). Requisitos. ISO Ginebra, Suiza 2013
- [4] (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.
- [5] What is cybersecurity?. CISCO. Recuperado de <https://www.cisco.com/c/en/us/products/security/what-is-cybersecurity.html>
- [6] Harmantzis, F., Malek, M. (2004). Security Risk Analysis and Evaluation, IEEE Communications Society, pp. 1897-1901.
- [7] (2020) Strategies for managing cybersecurity risk: Assess and advance your security and compliance posture. IBM Security. IBM Global Services
- [8] Bunge, M. (1998). La ciencia su Método y su Filosofía. Editorial Siglo Veinte. Buenos Aires
- [9] Barchini (2005). G. Métodos “I+D” de la Informática. Universidad Nacional de Santiago del Estero, Argentina
- [10] Figueras A., Morero H., (2006) PARETO COMO CIENTÍFICO SOCIAL: A CIEN AÑOS DEL MANUAL. Asociación Argentina de Economía Política. Anales. Salta.
- [11] (2017) Matriz de Riesgo, Evaluación y Gestión de Riesgos. SIGWEB. Página <http://www.sigweb.cl/>. Recuperado el 03/30/2020.

# Exploración de correlaciones internas de los parámetros temporales generados en dinámicas de tecleo

Nahuel González<sup>1</sup>, Germán M. Concilio<sup>1</sup>,  
Jorge S. Ierache<sup>1</sup> , Enrique P. Calot<sup>1</sup> , y Waldo Hasperué<sup>2</sup>

<sup>1</sup> Laboratorio de Sistemas de Información Avanzados,  
Facultad de Ingeniería, Universidad de Buenos Aires,  
Buenos Aires, Argentina  
{ngonzalez, gconcilio, jierache, ecalot}@lsia.fi.uba.ar

<sup>2</sup> Instituto de Investigación en Informática (III-LIDI)  
Facultad de Informática Universidad Nacional de La Plata  
Investigador Asociado – Comisión de Investigaciones Científicas (CIC)  
whasperue@lidi.info.unlp.edu.ar

**Resumen** El análisis de cadencias de tecleo pertenece al ámbito de la autenticación biométrica por comportamiento. El comportamiento y el estado de ánimo de una persona modifican, sutil pero consistentemente, el patrón de escritura en un teclado de computadora y otros periféricos de entrada equivalentes. En este trabajo se exploraron, para un grupo de palabras elegidas del conjunto de datos de origen, las correlaciones internas entre los parámetros temporales, el rendimiento de clasificación en las tareas de identificación y verificación, y se buscaron grupos naturales de tipistas, entre los que se encuentran diestros y zurdos. Se observó que la influencia de cada parámetro temporal no es necesariamente local e inmediata, limitada a la tecla siguiente o una cercana, sino que pueden encontrarse correlaciones maximales entre letras con cuatro y hasta cinco grados de separación. Los resultados de el rendimiento en la clasificación, tanto en tareas de identificación como de autenticación, resultaron satisfactorios a pesar de la simplicidad del esquema utilizado, por lo que se propone la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas.

**Palabras claves:** keystroke dynamics, biometrics, identificación de usuario, autenticación de usuario, seguridad informática

## 1. Introducción

El análisis de cadencias de tecleo pertenece al ámbito de la biometría por comportamiento. El comportamiento y el estado de ánimo de una persona modifican, sutil pero consistentemente, su patrón de escritura en un teclado de

computadora y otros periféricos de entrada equivalentes. Por medio de estas variaciones individuales es factible verificar la identidad del usuario cuando esta es conocida o incluso identificarlo entre un conjunto de perfiles previamente entrenados con los que se cuente [9]. Lo que es más, ciertas características físicas como la mano dominante [14], o el estado emocional del usuario se pueden inferir de las mismas con precisión aceptable [8]; también se ha investigado la posibilidad de detectar tempranamente en las variaciones del patrón de escritura ciertas condiciones neurológicas de avance progresivo como el Alzheimer [17]. Utilizando únicamente la cadencia de tecleo de los usuarios se logró vincular múltiples cuentas o perfiles que difunden masivamente noticias falsas en Internet y las redes sociales durante la pandemia de COVID-19 [16]. El análisis de cadencias de tecleo en textos libres, con menos de veinte años de antigüedad [10], es un campo más complejo y menos explorado que el análisis de textos fijos como claves. Si bien algunos estudios tempranos han intentado zanjar la brecha entre la verificación de textos fijos y textos libres utilizando  $n$ -gramas [15], sólo se puede hablar de la madurez de la disciplina a partir de la introducción de distancias entre vectores de rango [1,10]. Desde el año 2013, apoyando al surgimiento de las investigaciones que articulan la detección del estado emocional con la dinámica de tecleo en textos libres, desde el Laboratorio de Sistemas de Información Avanzados de la Facultad de Ingeniería de la Universidad de Buenos Aires se comenzó a investigar en el área [11,5,3,4]. La utilización de contextos finitos para el modelado de cadencias de tecleo en textos libres ha sido propuesta en [11] con resultados promisorios en el ámbito de la verificación de identidad, como segundo factor de autenticación. En [7,6] se mostró la utilidad del método para identificar personas reales a partir del patrón de tipeo presentado en campos de texto libre en entornos web, en particular en el contexto de un e-commerce donde hay una gran variedad de tipos de dispositivos de entrada que afectan la efectividad final del método. En la Sección 2 se presenta el problema a investigar. El resto del trabajo se encuentra organizado de la siguiente manera: la Sección 3 describe la fuente de información que será minada en busca de hipótesis y la configuración del método aplicado, la Sección 4 discute los resultados obtenidos, y la Sección 5 presenta las conclusiones y futuras líneas de trabajo.

## 2. Descripción del problema

Si bien algunos periféricos de entrada como las pantallas táctiles de dispositivos móviles permiten obtener información de presión y aceleración, el enfoque clásico para análisis de cadencias de tecleo se ha centrado en los tiempos de escritura pues cualquier teclado convencional carece de medios para capturar parámetros adicionales. Los dos parámetros independientes que caracterizan la escritura en un teclado son el *tiempo de retención* (del inglés *hold time*), que es el lapso transcurrido entre la presión y la liberación de la tecla, y el *tiempo entre teclas* o *latencia* (del inglés *flight time*), que es el intervalo entre sucesivas presiones de tecla. Así, a cada carácter ingresado le corresponderían dos valores y la firma biométrica de un texto de  $n$  caracteres estaría conformado por un

vector de  $2n - 1$  tiempos; no  $2n$  pues el primer caracter carece de tiempo entre teclas. Se han utilizado también otros parámetros temporales, como por ejemplo el *tiempo de espera* que mide el intervalo entre la liberación de una tecla y la presión de la siguiente, o los  $n$ -gramas que agrupan las latencias de varias teclas adyacentes; todos estos, y muchos otros que han sido propuestos y evaluados en la literatura, derivan de los dos únicos parámetros independientes ya descritos.

En [11] se demostró que utilización de contextos finitos para el modelado de cadencias de tecleo aporta una mejora en la efectividad en el ámbito de la autenticación o verificación de identidad. En este trabajo se propone la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas. Se explorarán, para las palabras elegidas del conjunto de datos de origen, las correlaciones internas de parámetros temporales, el rendimiento de clasificación en las tareas de identificación y autenticación, y se buscarán grupos naturales de tipistas.

### 3. Experimentación

#### 3.1. Conjunto de datos

Para este trabajo se ha utilizado un conjunto de datos [13], subconjunto del utilizado en [12], que contiene multitud de sesiones de escritura en un teclado convencional de 158 usuarios, donde se han registrado las teclas presionadas y los tiempos de retención y latencia, junto con la identidad del escritor. Los tiempos se encuentran expresados en milisegundos, y por restricciones de la plataforma de captura en ciertas ocasiones se hallan redondeados al múltiplos de ocho más cercano, o algunos otros valores específicos. Las sesiones fueron capturadas en un entorno realista durante aproximadamente un año con usuarios de ambos sexos en un rango de edad entre 28 y 60 años, cuyas competencias mecanográficas variaban desde teclear con dos dedos hasta el manejo fluido. No hay motivos para suponer que la identidad etiquetada de los usuarios pueda contener errores, pues esta ha sido verificada al inicio de la sesión de escritura utilizando claves y/o otros factores de identificación biométrica. Los textos corresponden a lenguaje natural producido durante el transcurso de la labor cotidiana. Como las sesiones del conjunto de datos anterior varían en longitud y a los fines de construir un conjunto de datos tabular, se ha extraído de las sesiones los parámetros temporales correspondientes a un conjunto de cuatro palabras: QUE, DESDE, CONTROL, y EVOLUCIÓN. Estas palabras han sido elegidas por ser, dentro del conjunto de datos considerado, las que cuentan con mayor cantidad de observaciones para su extensión; las palabras más largas no son frecuentes y por lo tanto no brindarían suficientes observaciones para el análisis. De cada sesión en la cuál aparezca al menos una de las palabras indicadas se han extraído los tiempos de retención y latencia para cada una de las teclas que la conforman, y se ha etiquetado la fila con la identidad del usuario que las ingresó. En el Cuadro 1 se observan como ejemplo un grupo de filas con las columnas a los parámetros temporales de cada tecla de la palabra QUE. Las columnas denominadas HT son tiempo de retención, aquellas denominadas FT son latencias entre teclas

sucesivas, y el subíndice indica a qué tecla corresponde dicho tiempo medido en milisegundos. Todos los valores están expresados en milisegundos.

$HT_q$	$FT_q$	$HT_u$	$FT_u$	$HT_e$	$FT_e$	Usuario
$77 \pm 31$	$198 \pm 22$	$79 \pm 34$	$61 \pm 44$	$94 \pm 15$	$171 \pm 23$	585870
$72 \pm 1$	$156 \pm 78$	$88 \pm 16$	$251 \pm 54$	$68 \pm 8$	$184 \pm 71$	2085887
$83 \pm 3$	$311 \pm 224$	$67 \pm 3$	$121 \pm 20$	$70 \pm 5$	$86 \pm 30$	2038654
$141 \pm 1$	$257 \pm 8$	$86 \pm 16$	$156 \pm 0$	$117 \pm 8$	$149 \pm 7$	470698

Cuadro 1: Experimental Algunas filas para la palabra QUE

En el Cuadro 2 se resume la cantidad de atributos, instancias originales, y cantidad de instancias filtradas de cada palabra del subconjunto de datos seleccionado para el análisis.

Palabra	Atributos	Instancias	Filtradas
QUE	3	5273	5112
DESDE	9	3322	3030
CONTROL	13	4965	4533
EVOLUCION	17	3200	2908

Cuadro 2: Resumen de los subconjuntos de datos

### 3.2. Preprocesamiento

El conjunto de datos de [12] fue preprocesado por medio de una herramienta propia para experimentos de cadencias de tecleo, a los fines de convertir el formato binario propietario en el cuál se encuentra almacenada la información original. En primer lugar, se buscó para cada largo de palabra aquella de mayor frecuencia, y luego se extrajo de cada sesión los tiempos de retención y latencia correspondientes a las elegidas. El resultado de este preprocesamiento fue almacenado en formato CSV para consumo de la herramienta de minería de datos. El archivo CSV resultante [13] incluye 49 columnas, de las cuáles 48 son de parámetros temporales y la restante etiqueta el usuario que los generó; la cantidad de filas asciende a 16760, y 158 usuarios. A los fines de simplificar la evaluación de subconjuntos de los datos por palabras, se produjeron conjuntos de datos reducidos en filas y columnas, para cada una de las palabras consideradas. De esta manera, con las 49 columnas originales se realizó una separación horizontal para generar cuatro conjuntos de datos derivados que representan los tiempos de escritura de las palabras QUE, DESDE, CONTROL, y EVOLUCION con 7, 11, 15, y 19 columnas respectivamente.

La herramienta de captura marca pausas en las sesiones, eliminando el valor de latencia entre teclas, toda vez que este excede un valor en torno a 1500

mseg., que depende parcialmente también de las medias móviles de los tiempos anteriores. Sin embargo, los valores usuales en la escritura se encuentran entre los 100 mseg. y los 500 mseg., en donde valores mayores suelen indicar pausas breves que no deben ser consideradas como parte de la cadencia natural. Para este fin se filtraron los valores extremos. Aproximadamente un 10 % de las instancias son excluidas por presentar valores extremos.

## 4. Resultados

### 4.1. Correlaciones internas en las palabras

Luego de filtrar los *outliers* se filtró el listado de correlaciones para comprender qué teclas se encuentran mayormente correlacionadas con otras dentro de la misma palabra. A continuación se muestran, para la palabra EVOLUCION, aquellos parámetros temporales cuyo coeficiente de correlación es mayor que 0,5. Este valor puede parecer escasamente significativo, pero debe tenerse en consideración el dominio particular del problema, la biometría comportamental, donde los parámetros son siempre muy ruidosos. Para distinguir la primera y segunda letra O de la palabra, en el Cuadro 3 se han utilizado subíndices,  $O_1$  y  $O_2$ .

Parámetro 1	Parámetro 2	Coef. correlación
$HT_I$	$HT_{O_2}$	0,693
$HT_V$	$HT_C$	0,681
$FT_E$	$FT_U$	0,591
$HT_E$	$HT_C$	0,551
$HT_{O_2}$	$HT_N$	0,549
$HT_{O_1}$	$HT_U$	0,526
$HT_I$	$HT_N$	0,524
$HT_E$	$HT_V$	0,510
$FT_C$	$FT_{O_2}$	0,507

Cuadro 3: Parámetros temporales con coeficiente de correlación mayor a 0,5 para la palabra EVOLUCION.

Se recorrió en orden decreciente la lista de correlaciones a los fines de descubrir, para cada parámetro temporal, con cuál otro manifiesta mayor correlación. El resultado para la palabra EVOLUCION puede observarse en la Figura 1, donde los nodos etiquetados con la letra corresponden a los tiempos de retención y los pequeños sin etiquetar a las latencias entre teclas. No es sorprendente que las mayores correlaciones se encuentren en los tiempos de retención pues estos son más estables, y relacionados con la velocidad promedio de escritura, que los tiempos de latencia entre teclas.

Destaca el hecho que la influencia de cada parámetro no es necesariamente local e inmediata, limitada a la tecla siguiente o una cercana, sino que pueden encontrarse correlaciones maximales entre letras con cuatro y hasta cinco grados

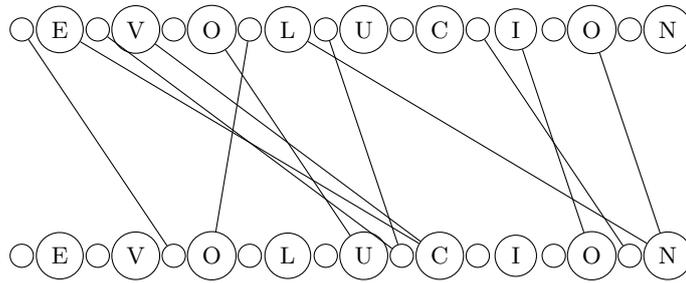


Figura 1: Pares de letras máximamente correlacionados

de separación como la E y la C, o la L y la N. Este fenómeno es relevante para el modelado de cadencias de tecleo utilizando contextos finitos [11] ya que al utilizar esta técnica el tamaño del contexto utilizado debe ser suficientemente amplio para capturar las influencias más significativas de las parámetros temporales de teclas anteriores si se desea obtener un buen rendimiento. La hipótesis de existencia de correlaciones fuertes entre teclas no adyacentes que revela el análisis del conjunto de datos agregado para todos los usuarios es particular también a los usuarios individuales al ser estos considerados por separado.

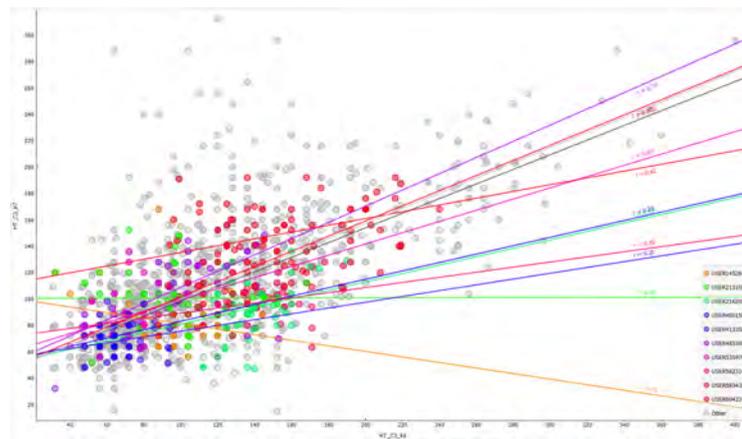


Figura 2: Detalle de correlación entre tiempos de retención de letras I y O en la palabra EVOLUCION

En la Figura 2 se observa el gráfico de dispersión de los tiempos de retención entre las teclas  $I$  y  $O_2$ , donde el eje  $y$  representa el tiempo de retención de la tecla  $I$  y el eje  $x$  el tiempo de retención de la tecla  $O_2$ , y muestra para cada usuario los valores de  $r$  que representan la correlación de las mediciones. Por ejemplo para el usuario 400150 presenta una correlación  $r$  de 0,71 aquellos que presentan

mayor correlación. Casi todos los usuarios presentan similar tendencia, aunque no debe descartarse la existencia de contraejemplo como el usuario 145264 con  $r = 0,25$ .

#### 4.2. Rendimiento de clasificación en las tareas de identificación y autenticación

El análisis clásico de cadencias de tecleo en textos libres ha utilizado indiscriminadamente la información disponible en la sesión, por ejemplo a través de métricas de desorden [1,10]. La ponderación de distintos contextos en base a su capacidad de discriminar ha demostrado mejorar el rendimiento en la clasificación [11], sobre todo al utilizar la fragmentación de las sesiones en fronteras no arbitrarias, sino utilizando inicio y fin de palabras. Aquí se explorará brevemente el rendimiento de clasificación en las tareas de identificación y autenticación en palabras individuales. Para tal fin se compararon cuatro clasificadores (kNN, Random Forest, SVM, y redes neuronales) tanto en una tarea de identificación como de autenticación. En la primera de ellas se intenta asignar la cadencia de tecleo al usuario correspondiente entre todos los que aparecen en el dataset, mientras que en la segunda se trata de un problema de clasificación binaria donde se utilizan tantas instancias del usuario legítimo como de otros usuarios, estos últimos marcados como impostores. Para la primera, de mayor dificultad, es esperable un rendimiento inferior. En el Cuadro 4 pueden observarse los resultados, junto con los intervalos de confianza para asegurar su significación estadística, para las palabras QUE, DESDE, CONTROL y EVOLUCION. Nótese que la diferencia de magnitud en los intervalos para ambos tipos de tareas se explican por la diferencia en la cantidad de filas; mientras que la tarea de identificación utiliza todas las filas del dataset para cada palabra, en el orden de varios miles, la tarea de autenticación utiliza únicamente el doble de la cantidad del usuario evaluado (las del usuario más idéntica cantidad de impostores), en el orden de pocos cientos. No es un error que el intervalo de confianza para las tareas de identificación, que ha sido verificado, sea el mismo en todas las filas. A primera vista el rendimiento en la clasificación es sorprendentemente bueno para las palabras CONTROL y EVOLUCION, pero no debe olvidarse que las palabras elegidas son las más comunes del dataset [13] para el largo dado. Las palabras escritas con mayor frecuencia estabilizan la cadencia de tecleo, en contraste con aquellas más raras que suelen ser más ruidosas y por lo tanto no exhibirían métricas de precisión tan optimistas. Sin embargo, estos resultados son lo suficientemente alentadores como para considerar la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas. Por otro lado, también se puede observar que con palabras muy cortas como QUE o DESDE, resulta imposible identificar unívocamente al usuario. Sin embargo se mantienen resultados satisfactorios en la determinación de autenticación de usuarios, en este orden fortalece el objetivo primario de la investigación presentada.

Palabra	Modelo	Precisión identif.	Precisión autenticación
QUE	kNN	19,3 ± 1,06	72,9 ± 1,28
	SVM	26,3 ± 1,18	74,2 ± 1,26
	Red neuronal	27,3 ± 1,20	68,9 ± 1,33
	Random Forest	25,0 ± 1,17	71,6 ± 1,30
DESDE	kNN	35,3 ± 1,71	83,0 ± 1,38
	SVM	42,1 ± 1,75	93,2 ± 0,95
	Red neuronal	45,7 ± 1,79	88,6 ± 1,19
	Random Forest	43,3 ± 1,76	92,0 ± 1,02
CONTROL	kNN	50,5 ± 1,47	74,1 ± 1,30
	SVM	57,6 ± 1,45	92,1 ± 0,82
	Red neuronal	59,1 ± 1,45	91,3 ± 0,86
	Random Forest	54,7 ± 1,46	93,3 ± 0,76
EVOLUCION	kNN	45,7 ± 1,82	89,1 ± 1,19
	SVM	49,7 ± 1,83	95,4 ± 0,82
	Red neuronal	57,4 ± 1,82	95,4 ± 0,82
	Random Forest	64,0 ± 1,82	94,3 ± 0,90

Cuadro 4: Rendimiento de clasificación en identificación y autenticación para las palabras QUE, DESDE, CONTROL y EVOLUCION.

### 4.3. Grupos naturales de tipistas

Las cadencias de tecleo pueden agruparse naturalmente de diversas maneras. Una división evidente deriva de la mano dominante del usuario, zurdo o diestro. También se puede postular la existencia de estratos, o quizás un espectro, de competencia que abarca desde el que escribe con un único dedo hasta los expertos mecanógrafos. Es más sutil la distinción que puede hacer en base al estado emocional del individuo en el momento de la escritura [5,8,2]. A los fines de evaluar qué tipos de grupos naturales pueden extraerse del conjunto de datos dado, se evaluó la aplicación de mapas auto-organizados SOM, una técnica de agrupamiento y aprendizaje no supervisado. La aplicación de mapas auto-organizados sobre los 158 usuarios del dataset revela un fenómeno inesperado: en la Figura 3 se observan los resultados de dos ejecuciones con una grilla de  $5 \times 5$  e inicialización aleatoria no replicable. Nótese que en ambos casos aparecen dos grupos destacados con casi idéntica cantidad de instancias, seguidos de lejos por diversos grupos menores, entre los cuáles uno sólo parece destacar secundariamente. Los resultados no varían significativamente si se utilizan otros tamaños de grilla u otros tipos de inicialización. Si bien no se cuenta con una clasificación confiable de la mano dominante del usuario en el dataset desestructurado de origen desde el cuál se han obtenido los datos tabulares que aquí se estudian, y que permitiría confirmar esta hipótesis, el tamaño del grupo secundario dominante es consistente con la cantidad de zurdos en la población general, general el orden del 10%.

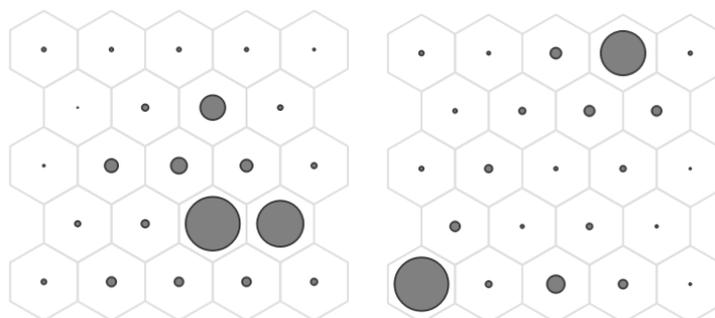


Figura 3: Dos agrupamientos con SOM e inicialización aleatoria

## 5. Conclusiones

En este trabajo se han explorado las correlaciones internas de parámetros temporales, el rendimiento de clasificación en las tareas de identificación y verificación, y se buscaron grupos naturales de tipistas, entre los que se encuentran diestros y zurdos, en un subconjunto del dataset de [12], que contiene multitud de sesiones de escritura en un teclado convencional donde se han registrado las teclas presionadas y los tiempos de retención y latencia. Se observó que la influencia de cada parámetro temporal no es necesariamente local e inmediata, limitada a la tecla siguiente o una cercana, sino que pueden encontrarse correlaciones maximales entre letras con cuatro y hasta cinco grados de separación como la E y la C, o la  $O_1$  y la  $O_2$  en la palabra EVOLUCION. La hipótesis de existencia de correlaciones fuertes entre teclas no adyacentes, planteada en la Sección 2 se mostró particularizable también a los usuarios individuales en el Cuadro 3. El rendimiento en la clasificación, tanto en tareas de identificación como de verificación, resultó sorprendentemente bueno a pesar de la simplicidad del esquema utilizado, particularmente en las palabras más largas —EVOLUCION y CONTROL—, pero no debe olvidarse que las palabras elegidas son las más comunes del dataset para el largo dado y palabras más raras no exhibirían métricas de precisión tan optimistas. Sin embargo, estos resultados son lo suficientemente alentadores como para proponer la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas. Finalmente, la aplicación de mapas auto-organizados reveló un fenómeno inesperado; dos grupos destacados con casi idéntica cantidad de instancias, que no varían significativamente si se utilizan otros tamaños de grilla u otros tipos de inicialización. En el contexto de las futuras líneas de trabajo se contempla profundizar las pruebas de identificación de usuarios, analizar el empleo de mapas auto organizados para la terminación de poblaciones de tipo secundarias de grupos naturales de tipistas como medida para fortalecer la identificación y autenticación de usuarios.

## Referencias

1. Bergadano, F., Gunetti, D., Picardi, C.: User authentication through keystroke dynamics. *ACM Transactions on Information and System Security (TISSEC)* 5(4), 367–397 (2002)
2. Calot, E.P.: Robustez de las métricas de clasificación de cadencia de tecleo frente a variaciones emocionales. Ph.D. thesis, Universidad Nacional de La Plata (2019), <http://sedici.unlp.edu.ar/handle/10915/76652>
3. Calot, E.P., Ierache, J.S.: Multimodal biometric recording architecture for the exploitation of applications in the context of affective computing. In: *Proceedings del XXIII CACIC (La Plata, 2017)* (2017)
4. Calot, E.P., Ierache, J.S., Hasperué, W.: Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations. In: *Advances in Intelligent Systems and Computing*. Springer (2019)
5. Calot, E.P., Rodriguez, J.M., Pereira, G., Iribarren, J., Ierache, J.S.: Métodos adaptativos de educación de dinámica de tecleo centrado en el contexto emocional de un individuo aplicando interfaz cerebro computadora. In: *XVI WICC* (2014)
6. Concilio, G.M.: Aplicación de técnicas de modelado de la cadencia de tecleo para fortalecer la identificación del usuario en el comercio electrónico (Feb 2019)
7. Concilio, G.M., Calot, E.P., Ierache, J.S., Merlino, H.D.: Application of Keystroke Dynamics Modelling Techniques to Strengthen the User Identification in the Context of E-commerce (2018)
8. Epp, C., Lippold, M., Mandryk, R.L.: Identifying emotional states using keystroke dynamics. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 715–724. ACM (2011)
9. Forsen, G.E., Nelson, M.R., Staron, R.J.J.: Personal attributes authentication techniques. Tech. rep., DTIC Document (October 1977), <http://lsia.fi.uba.ar/papers/forsen77.pdf>, final technical rept.
10. Francesco Bergadano, D.G.y.C.P.: Identity verification through dynamic keystroke analysis. *Intelligent Data Analysis* 7(5), 469–496 (2003)
11. González, N.: Finite context modelling of keystroke dynamics in mixed authentication schemes (2014), <http://lsia.fi.uba.ar/papers/gonzalez14b.pdf>, LSIA, FIUBA, Argentina
12. González, N., Calot, E.P., Ierache, J.S.: A replication of two free text keystroke dynamics experiments under harsher conditions. In: *Biometrics Special Interest Group (BIOSIG), 2016 International Conference of the*. pp. 1–6. IEEE (2016)
13. González, N., Concilio, G., Ierache, J., Hasperué, W.: <https://drive.google.com/drive/folders/1T1MUcCivvMrmBTxwTHzpKdW8zZQ3INK?usp=sharing>
14. Idrus, S.Z.S., Cherrier, E., Rosenberger, C., Bours, P.: Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords. *Computers & Security* 45, 147–155 (2014)
15. Monroe, F., Reiter, M.K., Wetzel, S.: Password hardening based on keystroke dynamics. In: *Proceedings of the 6th ACM Conference on Computer and Communications Security*. pp. 73–82. CCS '99, ACM, New York, NY, USA (1999), <http://doi.acm.org/10.1145/319709.319720>
16. Morales, A., Acien, A., Fierrez, J., Monaco, J.V., Tolosana, R., Vera-Rodriguez, R., Ortega-Garcia, J.: Keystroke biometrics in response to fake news propagation in a global pandemic. *arXiv preprint arXiv:2005.07688* (2020)
17. Van Waes, L., Leijten, M., Mariën, P., Engelborghs, S.: Typing competencies in alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior* 73, 311–319 (2017)

## Implementación de Blockchain para aseguramiento de evidencia digital en entornos Forensic Readiness

Javier Díaz<sup>1</sup>, Mónica D. Tugnarelli<sup>2</sup>, Lucas Barboza<sup>2</sup>, Mauro F. Fornaroli<sup>2</sup>, Facundo N. Miño<sup>2</sup>

<sup>1</sup> Facultad de Informática – Universidad Nacional de La Plata

<sup>2</sup> Facultad de Ciencias de la Administración – Universidad Nacional de Entre Ríos  
jdiaz@unlp.edu.ar, {monica.tugnarelli, lucas.barboza, mauro.fornaroli}@uner.edu.ar

**Abstract.** En este artículo se presentan los avances PID-UNER 7059 en la implementación de blockchain aplicada a la preservación de evidencia digital en entornos Forensic Readiness el cual plantea requisitos estrictos de recolección de los datos antes de que se produzca un incidente de seguridad. Se analizan las estructuras y tipos de blockchain disponibles, clasificándolas según diversos parámetros. En base a ese análisis se seleccionaron dos, Ethereum bajo especificaciones de uso de la Blockchain Federal Argentina como blockchain pública y la segunda, del tipo blockchain privada con el framework Hyperledger Fabric. Ambas se utilizarán como entornos de prueba de concepto, dentro de un esquema genérico de funcionamiento y con determinadas restricciones considerando el tipo de datos que se debe resguardar y las características que demanda la cadena de custodia de este tipo de evidencia.

**Keywords:** blockchain, forensic readiness, evidencia digital, cadena de custodia

### 1 Introducción

En este trabajo se presentan los avances del PID-UNER 7059, iniciado en febrero de este año, denominado “*Tecnología Blockchain para aseguramiento de evidencia digital en entornos Forensic Readiness*” que tiene como objetivo principal analizar el impacto de la utilización de la tecnología blockchain aplicada a la preservación, integridad y trazabilidad de evidencia digital, obtenida de activos esenciales de una organización, en un entorno preventivo como lo es Forensic Readiness, también llamado Preparación Forense.

Específicamente, se pretende disponer de un prototipo de blockchain sin criptomoneda asociada, sobre el cual se puedan realizar pruebas que permitan analizar las prestaciones frente a las características demandadas por Forensic Readiness para el proceso de asegurar la evidencia y mantenimiento de la cadena de custodia. Asimismo se busca lograr una adecuada integración de la blockchain con esquemas de recolección de datos con las características que impone la Preparación Forense.

Forensic Readiness propone que la evidencia digital se recolecte y asegure de manera anticipada, es decir, antes de la ocurrencia de un incidente de seguridad. Este término fue enunciado por John Tan [1] quien lo describió principalmente a través de dos

objetivos: maximizar la capacidad del entorno para reunir evidencia digital confiable y minimizar el costo forense durante la respuesta a un incidente.

En este enfoque, que fue analizado en artículos anteriores [2] [3] [4], los datos que se recolectan pueden ser utilizados como insumo para el análisis de incidentes de seguridad y también como prueba legal, lo que involucra el aseguramiento de la prueba a medida que se realiza la recolección activa de los datos, tarea que fue realizada anteriormente utilizando funciones hash para resguardar la integridad de la evidencia digital. Considerando estos requisitos, una instancia fundamental para garantizar su admisibilidad como elemento de prueba es la preservación de la Cadena de Custodia. Esta Cadena de Custodia debe estar claramente documentada y con un registro detallado desde su recolección hasta su almacenamiento, por lo que la aplicación de la tecnología Blockchain resulta de especial interés para cumplimentar esta función.

Por su parte, Blockchain puede describirse como una base de datos distribuida y organizada bajo una estructura de conjunto de bloques que se van encadenando entre sí mediante dos códigos hash que actúan como enlaces: uno para el bloque de datos creado anteriormente y otro para que se comparta y grabe en el bloque que se cree a continuación, de forma tal de obtener una lista enlazada o cadena de bloques. Cada bloque que se encadena se vuelve inmutable y en eso radica la fortaleza de seguridad y verificabilidad de esta tecnología en la que, cuanto mayor es el grado de replicación de los bloques, más sencillo resulta detectar adulteraciones. [5].

Las criptomonedas fueron una de las primeras aplicaciones visibles que tienen como base esta tecnología, pero más allá de ellas, blockchain también permite imaginar múltiples prestaciones relacionadas a transacciones, a la seguridad, a la trazabilidad y a la transparencia.

Como primera etapa del proyecto se relevaron casos de uso a nivel regional y nacional [6], entre las que se destaca la Blockchain Federal Argentina [7] que brinda una plataforma pública para integrar servicios y aplicaciones sobre blockchain siguiendo el modelo de Múltiples Partes Interesadas. BFA mantiene un modelo de gobernanza que asegura la representación de todos los sectores en la toma de decisiones pero, al ser una plataforma pública, su uso no está restringido solo a las organizaciones que integran el consorcio.

Entre los participantes de esta plataforma se encuentra la Facultad de Ciencias de la Administración como miembro parte y encargado del nodo sellador UNER, y dos integrantes de este proyecto, en carácter de responsables designados ante la BFA en representación de la mencionada institución.

## 2 Análisis de estructuras y tipos de blockchain

La segunda etapa del PID se focaliza en el análisis de distintos tipos de blockchain para contar con una base de conocimiento previa a la implementación de los prototipos para la realización de pruebas. A continuación se presenta un breve resumen de esta revisión.

De manera general, las redes blockchain pueden clasificarse como [8], [9], [10], [11]:

- **Públicas:** donde cualquiera puede acceder a los datos de la cadena de bloques. Normalmente son transparentes, los usuarios son anónimos, no existe un administrador de la red y las transacciones se validan mediante un protocolo de consenso.
- **Privadas:** son aquellas donde existe una entidad central que se encarga de controlar la cadena de bloques, definir la lista de participantes autorizados, otorgar permisos, proponer transacciones y validar los bloques. Puede decirse que no existe descentralización ni consenso ya que es una única entidad quien administra la red. Los usuarios finales dependen de una interfaz provista por el administrador para leer o enviar transacciones.
- **De Consorcio o Federadas,** son las que tienen un conjunto de participantes predefinido, tales como empresas u organizaciones, quienes se encargan de la administración conjunta de la red y de asegurar el mantenimiento sincronizado de las copias del registro compartido. El acceso a los datos puede ser público o privado. Son redes parcialmente descentralizadas, adecuadas para aplicaciones en donde se generan grandes volúmenes de transacciones entre entidades con requerimientos de confianza mutua.

Otra clasificación posible surge de considerar quiénes poseen permisos para crear bloques, a saber:

- **Sin permisos (*permissionless*):** cualquier entidad puede procesar transacciones, participar de los protocolos de consenso y crear bloques.
- **Con permisos (*permissioned*):** solo una lista o conjunto de entidades predefinidas, y con identidades conocidas, pueden participar del procesamiento de transacciones, lo que agrega una capa más para el control de acceso e identificación.

En definitiva, las redes públicas pueden o no tener permisos, mientras que las privadas y las de consorcio o federadas suelen ser de tipo permisionadas. Las aplicaciones de blockchain con criptomonedas, como Bitcoin o Ethereum, son ejemplos de redes públicas sin permisos, mientras que el proyecto Hyperledger permite implementar redes de blockchain privadas con permisos.

Por último, desde el punto de vista de la realización de las transacciones, se puede hacer una diferenciación entre off-chain y on-chain, donde:

- **Transacciones on-chain (dentro de la cadena):** son las que ocurren dentro del blockchain, adquieren validez sólo cuando la cadena de bloques se

modifica para registrar la transacción en el ledger distribuido y son visibles para todos los nodos participantes. El tiempo de esta operación depende del volumen de transacciones y la cantidad de nodos selladores que actúan en la red. Por lo general los mineros cobran una comisión por sus servicios de validación y autenticación de transacciones.

- **Transacciones off-chain:** en este tipo de transacciones los datos en cuestión permanecen fuera de la cadena de bloques, quedando dentro de la blockchain el ID y el hash que la identifica. En contraste con las transacciones on-chain, los datos no son de acceso público y las transacciones se ejecutan de manera inmediata. Implican un acuerdo entre las partes para realizar la transacción y, si fuera necesario, la participación de un tercero para validar la operación.

## 2.1. Ethereum versus Hyperledger

De acuerdo a los objetivos de esta investigación se han analizado dos soluciones representativas de blockchain, una pública y otra privada, considerando para esto prestaciones tales como: privacidad, seguridad, velocidad de validación de transacciones, casos de uso, estándar abierto, entre otros ítems. En este sentido:

- **Ethereum** es una plataforma pública, distribuida y descentralizada que funciona en una red virtual llamada Ethereum Virtual Machine (EVM) donde cada nodo participante paga una comisión por cada transacción realizada utilizando para ello la criptomoneda Ether. Cuenta con una billetera, herramientas de línea de comandos, una aplicación GUI y contratos inteligentes (Smart Contract) de uso público y descentralizado para acuerdos entre partes. Como es una red pública no requiere de permisos y es completamente transparente, lo que significa que todas las transacciones registradas en la red blockchain son visibles y accesibles para todos los nodos participantes los cuales implementan algoritmos de consensos llamados Prueba de Trabajo (PoW) para aceptar cada transacción. Con Ethereum, cualquiera puede crear un nodo y cada nodo en la red poseerá una copia de la Blockchain. La red posee una infraestructura de nodos a nivel global. Como el desarrollo está basado en código abierto, toda la comunidad puede participar en las pruebas de concepto existentes para mejorar la plataforma, o tomar todo ese trabajo y adaptarlo a otros contextos y necesidades [12].

En este proyecto se utilizará la infraestructura disponible de BFA que toma el software de Ethereum, utilizando Prueba de Autoridad (PoA), sin criptomoneda asociada.

- Por su parte, **Hyperledger** es un proyecto de código abierto de la Linux Foundation, con la colaboración de empresas y organizaciones, con la finalidad de desarrollar soluciones con frameworks DLTs (Distributed Ledger Technology) conocidos también como blockchains privadas. Cuenta además con motores de smart contracts, librerías clientes e

interfaces gráficas. Tiene un diseño modular, extensible e interoperable que permite identificación de participantes, privacidad y confidencialidad de las operaciones, alto rendimiento para carga y validación de las transacciones y administración centralizada. El framework Hyperledger Fabric demuestra un alto rendimiento en términos de procesamiento de transacciones y baja latencia de confirmación de las mismas, además de permitir la privacidad y confidencialidad de las transacciones y el uso de contratos inteligentes. [13] [14] [15]. En relación al objetivo de investigación planteado se destacan las siguientes características:

- Es una red permissionada privada, con acceso restringido y la identidad de los participantes es conocida.
- El consenso se realiza a nivel transaccional, lo que significa que no es necesario validar todo el bloque lo que aumenta la velocidad de las transacciones.
- No requiere de una criptomoneda para su funcionamiento y prescinde del protocolo de consenso PoW (*Proof of Work*) que es costoso de implementar debido al proceso de minería que requiere y, por ende, consume menos recursos.

## 2.2 Esquema de funcionamiento

En la siguiente figura se presenta un esquema genérico de funcionamiento el cual se utilizará como marco para las pruebas de laboratorio.

En la imagen se describe el proceso de almacenamiento de los hashes de la evidencia recolectada de los activos esenciales determinados, junto a los metadatos de la misma, en el ledger (libro de registro distribuido) para permitir la correlación. De esta manera se resguarda la información de la evidencia en una base de datos con mayor rapidez y seguridad, en comparación a las transacciones on-chain.

El procedimiento es el siguiente:

- Se determina el activo digital sobre el cual se realizará la recolección de datos considerados evidencia.
- Se obtiene el hash o identidad digital de esos datos que junto al sello de tiempo (timestamp) proporcionarán la integridad.
- La evidencia se almacena en una base de datos off-chain o directorio privado
- El hash y los metadatos de la evidencia se registran en la blockchain.

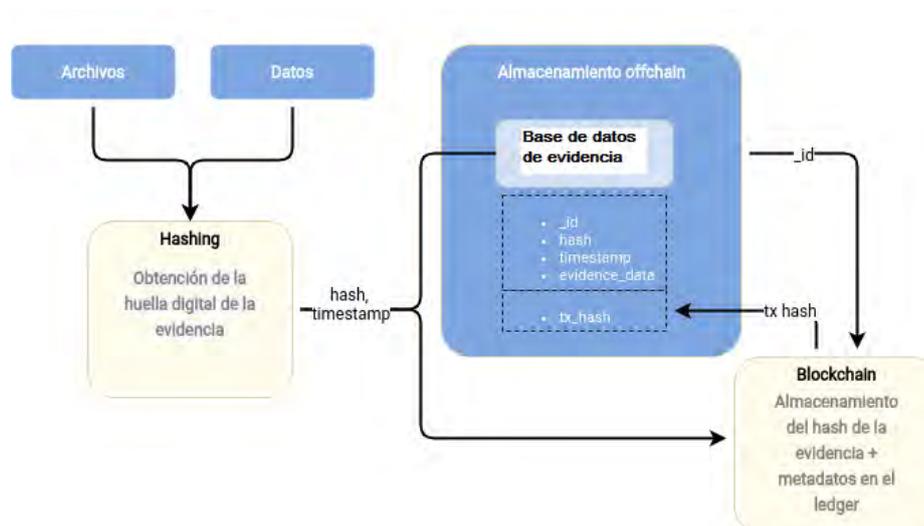


Fig.1 Almacenamiento de evidencia digital en Blockchain con preservación off-chain

Algunos puntos a considerar al momento de realizar las pruebas de concepto:

- a. En vista de los requisitos relacionados con Forensic Readiness, es necesario contar con una blockchain donde los datos, posiblemente confidenciales, puedan almacenarse fuera de la cadena de bloques.
- b. El sello de tiempo es un mecanismo que permite generar una “prueba de existencia” de un archivo digital. En este sentido, BFA ofrece su propio servicio de servicio de TSA (*Time Stamping Authority*).
- c. Es necesario que la base de datos que almacenará la evidencia recolectada tenga como mínimo un backup. Tanto la base como su respaldo deben ser tratados como activos esenciales de la organización. En este punto es indispensable seleccionar una infraestructura adecuada para minimizar los riesgos de borrado, pérdida o destrucción.
- d. Como en Hyperledger Fabric no se necesita el consenso de todos los nodos para validar las transacciones, se utiliza un algoritmo BTF (Tolerancia a las fallas bizantinas) [16] con la posibilidad de usar más de un mecanismo de consenso para resolver los problemas de confianza entre los nodos.
- e. Hyperledger Fabric ofrece mayor customización para casos de uso que requieran estricta privacidad.
- f. Al ofrecer transacciones sin costo, BFA cuenta con una “Destilería de Gas” que regula el envío de ether a nodos transaccionales. Esto puede tener impacto y ser un limitante según el volumen de datos recolectados.

- g. Dadas las implicancias de un proceso legal que requiera el uso de la evidencia almacenada, será necesario contar con una tercera parte que permita validar y asegurar la integridad de las transacciones de forma independiente. La estructura de múltiples partes interesadas de BFA asegura esa independencia. En Hyperledger Fabric puede darse a través de la incorporación de uno o más nodos de confianza a la red blockchain permissionada, los que no sólo tendrán participación en la validación y confirmación de transacciones, sino que además guardarán su propia copia de los registros.
- h. La admisibilidad de una evidencia digital ante un proceso legal requiere la documentación cronológica y detallada de la Cadena de Custodia, por ende será determinante el mantenimiento de los requisitos de integridad, trazabilidad, autenticación y verificación de la misma que se alcance en cada una de las soluciones a analizar.

Para configuración del entorno de trabajo con la blockchain pública se usará la infraestructura disponible con la instalación de un nodo transaccional en la red de prueba (test2), sobre Ubuntu Server, de acuerdo a los requerimientos técnicos especificados por BFA en su Wiki.

Por otra parte también se ha configurado un entorno con Hyperledger Fabric mediante contenedores de Docker [17] para desplegar servicios en una máquina virtual de 4GB RAM, 150GB de disco HDD, 4 núcleos y Ubuntu Server como sistema operativo, sobre una infraestructura propia que utiliza XenServer (XCP-ng) [18] como plataforma de virtualización. En esta estructura mínima se han definido clientes, nodos pares y un nodo que actúa de Autoridad de Certificación para los usuarios.

### 3 Conclusiones y trabajo a futuro

En este artículo se han presentado los resultados de las primeras etapas del PID 7059, avanzando en el análisis de tipos y prestaciones de redes blockchain públicas y privadas para seleccionar la estructura que mejor se adapte a los requerimientos planteados por el entorno Forensic Readiness.

Luego de la revisión se han seleccionado por un lado el framework Hyperledger Fabric como blockchain privada y por otro una implementación de Ethereum de acceso público tal como es la Blockchain Federal Argentina con los cuales se ha instalado en un entorno virtual que será el ambiente de testing a los fines de comparar cómo responde cada una de las blockchains elegidas a la dinámica relacionada con la recolección y resguardo de evidencia digital y la implementación de la Cadena de Custodia exigida.

Asimismo, se ha presentado el esquema básico de trabajo y funcionamiento de la blockchain con la particularidad de almacenamiento con preservación off-chain de los datos.

Las actividades a futuro comprenderán para los dos casos de práctica: la especificación del procedimiento de recolección, almacenamiento y acceso a la base de evidencias, la descripción del procedimiento de autenticación y trazabilidad de la evidencia digital incluyendo la documentación de la cadena de custodia, el correspondiente análisis de performance y escalabilidad, así como también la revisión de aspectos de seguridad relacionados con ambas implementaciones.

### Referencias

- [1] TAN, John. (2001). Forensic Readiness. [http://isis.poly.edu/kulesh/forensics/forensic\\_readiness.pdf](http://isis.poly.edu/kulesh/forensics/forensic_readiness.pdf)
- [2] Tugnarelli, M.; Fornaroli, M.; Santana, S.; Jacobo, E.; Díaz, F.J. Análisis de metodologías de recolección de datos digitales. Workshop de Investigadores en Ciencias de la Computación (WICC 2017). ISBN: 978-987-42-5143-5. <http://sedici.unlp.edu.ar/handle/10915/61343>
- [3] Tugnarelli, M., Fornaroli, M. Santana, S. Jacobo, E. Díaz, J. Analysis of Methodologies of Digital Data Collection in Web Serves. Communications in Computer and Information Science (Springer), Vol. 790, Pag.265. (2018) <https://link.springer.com/content/pdf/bfm%3A978-3-319-75214-3%2F1.pdf>
- [4] Tugnarelli, Mónica; Díaz Francisco Javier. Forensic Readiness: Guía de Buenas Prácticas. Libro de Actas. XXV Congreso Argentino de Ciencias de la Computación CACIC 2019. VI Workshop de Seguridad Informática, pp. 1261-1268. ISBN 978-987-688-377-1.
- [5] Michael Crosby, et. al. BlockChain Technology: Beyond Bitcoin. Applied Innovation Review (AIR). Issue No. 2 June 2016. Berkeley. <http://scet.berkeley.edu/wp-content/uploads/AIR-2016-Final-version-Int.pdf>
- [6] Javier Díaz, Mónica D. Tugnarelli, Mauro F. Fornaroli, Lucas Barboza. Blockchain para Aseguramiento de Evidencia Digital en entornos Forensic Readiness. Workshop de Investigadores en Ciencias de la Computación (WICC 2020). ISBN: en trámite. Exposición virtual de posters: <https://drive.google.com/file/d/1vGiNX9ogumSBjnAnH-4d6N5hn5omJPM4/view>
- [7] Blockchain Federal Argentina <https://bfa.ar>

- [8] Iuon-Chang Lin, Tzu-Chun Liao. A Survey of Blockchain Security Issues and Challenges . International Journal of Network Security, Vol.19, No.5, PP.653-659, Sept. 2017
- [9] Elli Androulaki, Artem Barger, Vita Bortnikov, et al, 2018. Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains . In EuroSys '18: Thirteenth EuroSys Conference 2018, April 23–26, 2018, Porto, Portugal. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3190508.3190538>
- [10] BitFury Group. Public versus Private Blockchains Part 1: Permissioned Blockchains. White Paper. Oct 20, 2015 (Version 1.0). <https://bitfury.com/content/downloads/public-vs-private-pt1-1.pdf>
- [11] BitFury Group. Digital Assets on Public Blockchains. White Paper. Mar 15, 2016 (Version 1.0). [https://bitfury.com/content/downloads/bitfury-digital\\_assets\\_on\\_public\\_blockchains-1.pdf](https://bitfury.com/content/downloads/bitfury-digital_assets_on_public_blockchains-1.pdf)
- [12] Ethereum. <https://ethereum.org>
- [13] Hyperledger. <https://www.hyperledger.org/>
- [14] An Introduction to Hyperledger.V1.1. Published August 2018. [https://www.hyperledger.org/wp-content/uploads/2018/08/HL\\_Whitepaper\\_IntroductiontoHyperledger.pdf](https://www.hyperledger.org/wp-content/uploads/2018/08/HL_Whitepaper_IntroductiontoHyperledger.pdf)
- [15] Hyperledger Fabric: the flexible blockchain framework that’s changing the business world <https://www.ibm.com/blockchain/hyperledger>
- [16] Lei, Kai & Zhang, Qichao & Xu, Limei & Qi, Zhuyun. (2018). Reputation-Based Byzantine Fault-Tolerance for Consortium Blockchain. 10.1109/PADSW.2018.8644933.
- [17] Docker Docs. <https://docs.docker.com/>
- [18] XCP-ng. <https://xcp-ng.org/>

# Track Gobierno Digital y Ciudades Inteligentes

Coordinadores

Elsa Estévez (UNS)  
Ariel Pasini (UNLP)

## Prototipo para la gestión de agua de lluvia para la descarga de sanitarios

Leandro Marín Crespi<sup>1</sup>, Cristian Schiffino<sup>1</sup>, Iris Inés Sattolo<sup>1</sup>, Marisa Daniela Panizzi<sup>1</sup>

<sup>1</sup> Escuela Superior de Ingeniería, Informática y Ciencias Agroalimentarias. Universidad de Morón. Cabildo 134 (B1708JPD). Partido de Morón, Argentina.

[Leandro.Crespi@outlook.com](mailto:Leandro.Crespi@outlook.com), [cristian.schiffino@gmail.com](mailto:cristian.schiffino@gmail.com), [iris.sattolo@gmail.com](mailto:iris.sattolo@gmail.com), [marisapanizzi@outlook.com](mailto:marisapanizzi@outlook.com)

**Resumen.** La población mundial enfrenta un importante problema en relación con el agua, no solamente por tener escasez en la distribución de esta, ya sea potable o no, sino por la falta de concientización que existe en la Sociedad para cooperar en lograr un medio ambiente sostenible. Las TICs constituyen un factor central para el desarrollo económico, social y productivo y su aplicación impacta en todas las actividades que se realizan en una Sociedad. Desde esta perspectiva, se presenta como las TICs contribuyen a la sostenibilidad del uso del agua potable, mediante un sistema de gestión que permita la captación del agua de lluvia destinada para la descarga de sanitarios. Esta solución es aplicable a cualquier tipo de establecimiento donde los sanitarios tienen una alta demanda de uso. En este trabajo se describe la experiencia en la definición, el diseño y el desarrollo del prototipo.

**Palabras clave:** Prototipo de gestión, agua de lluvia, descarga de sanitarios, sostenibilidad del uso del agua potable.

### 1 Introducción

La Organización de las Naciones Unidas, en [1], plantea que la escasez de recursos hídricos, junto con la mala calidad del agua y el saneamiento inadecuado, repercuten en la seguridad alimentaria, los medios de subsistencia y la oportunidad de educación para las familias pobres en todo el mundo.

Los Objetivos de Desarrollo Sostenible, también conocidos como ODS, presentados en [1] están diseñados para mejorar la calidad de vida a nivel mundial; se crearon en la Conferencia de las Naciones Unidas. Los ODS son un conjunto de 17 objetivos que permiten superar desafíos globales como la pobreza, la desigualdad, el cambio climático entre otros, además buscan armonizar tres componentes centrales para el futuro: crecimiento económico, inclusión social y protección Ambiental.

La captación de agua de lluvia es la recolección, transporte y almacenamiento del agua de lluvia que cae sobre una superficie de manera natural o hecha por el hombre. Las superficies que captan el agua en las ciudades pueden ser techos de casas y edificios, techumbres de almacenes y de tiendas, explanadas, etc. El agua almacenada puede ser usada para cualquier fin, siempre y cuando se realicen los tratamientos

adecuados para cada uso. Para usos básicos como limpieza de ropa, de pisos, sanitarios y riego puede usarse un filtro; para aseo personal y para agua que se pretenda beber, se deberá tener un sistema de filtros adecuados para estos fines [2].

En [3] se menciona que la captación de agua de lluvia fue muy utilizada por las sociedades antiguas en todo el planeta y en muchas ocasiones, se utilizó como único procedimiento para el abastecimiento de agua en algunas regiones. Dentro de los potenciales usos del agua de lluvia, se pueden enumerar: Incendio, Ducha, Consumo humano, Descarga de sanitarios, Riego como actividad secundaria, Aseo de superficies, Lavado de Autos y Lavado de Ropa [4].

Con base en esta problemática de la escasez del recurso hídrico para las generaciones actuales y futuras, además de adherir a los lineamientos de los ODS, en esta investigación se propuso la captación de agua de lluvia en aplicaciones en las que no es necesario el uso de agua potable. La aplicación de las TICs como medio para contribuir con el cuidado de medio ambiente, nos permitió definir nuestro objetivo. Este consiste en la construcción de un prototipo de un sistema de gestión de agua de lluvia para la descarga de sanitarios de un establecimiento.

Este artículo se estructura de la siguiente manera; se presentan los trabajos relacionados (sección 2), se describe el proceso de definición, diseño y desarrollo del prototipo de gestión de agua de lluvia para la descarga de sanitarios (sección 3), se describen las pruebas realizadas (sección 4) y, por último, se presentan las conclusiones y trabajos futuros (sección 5).

## 2 Trabajos Relacionados

Se realizó un estudio de mapeo sistemático (SMS, en inglés Systematic Mapping Study) según el proceso propuesto en [5] que permitió dar respuesta a la siguiente pregunta de investigación (PI): *¿Cuál es el estado del arte sobre el mejoramiento de la sostenibilidad del uso agua potable empleando agua de lluvia como sustituto?*

La Tabla 1 sintetiza las tareas realizadas en la actividad planificación del SMS.

**Tabla 1.** Tareas de la actividad Planificación del SMS.

Tareas	Descripción
Definir las Sub-Preguntas de investigación (PI1-PI6).	PI1: ¿Qué métodos existen que permitan recolectar y utilizar agua de lluvia?
	PI2: ¿Qué factores determinan la recolección de agua de lluvia y que finalidad tiene?
	PI3: ¿Qué aspectos se tienen en cuenta en la automatización del proceso de recolección y uso de agua de lluvia?
	PI4: ¿Qué tipo de contribución se ha realizado a nivel académico acerca de la gestión de recolección y utilización de agua de lluvia?
	PI5: ¿Cuáles son los principales usos del agua de lluvia recolectada?
	PI6: ¿Cuáles son los tipos de investigación de los artículos?
Determinar las cadenas de búsqueda	C1: Automatización AND Recolección AND agua lluvia.
	C2: Agua Lluvia AND (Recolección OR Captación) AND Sistema.
	C3: (Utilización OR uso OR consumo) AND Agua lluvia.
	C4: Rainwater AND Electronical AND (Harvest OR Collection).

	C5: Rainwater AND (Harvest OR Collection) AND System. C6: (Utilization OR Use) AND Rainwater.
Determinar los criterios de selección de los estudios.	<p>Criterios de inclusión:</p> <ul style="list-style-type: none"> <li>• Artículos del 2008 hasta diciembre del 2018.</li> <li>• Artículos en el idioma inglés y español.</li> <li>• Artículos que contengan cadenas candidatas en el título y/o en el resumen.</li> <li>• Artículos duplicados: si hay varios artículos de un mismo autor que contemple la misma investigación, se considerara el más completo.</li> </ul> <p>Criterios de exclusión:</p> <ul style="list-style-type: none"> <li>• Artículos que no cumplan los criterios de inclusión.</li> <li>• Artículos sin revisión por pares.</li> <li>• Artículos que estén relacionados a la recolección de agua para ser utilizada en riego agropecuario y forestal.</li> <li>• Artículos que se encuentren en formato de resumen, presentaciones en power point y libros.</li> </ul>
Definir las fuentes de datos.	IEEE Explore, Science, Direct, ACM Digital Library, Scielo, Dialnet, Google Academic, SEDICI <sup>1</sup>
Determinar los tipos de publicación.	Artículos de congresos, artículos de revistas y tesis.
Definir el período.	2008 hasta diciembre del 2018.

Los Estudios Primarios analizados se encuentran en [6] por restricciones de espacio.

Después de analizar 48 estudios primarios, se concluye que:

- Las propuestas de captación de agua de lluvia presentan soluciones para campus universitarios y viviendas familiares en su gran mayoría y en un segundo lugar para desarrollos rurales, plantas industriales y comercios.
- Respecto a los países que presentan soluciones de uso de agua de lluvia como alternativa al agua potable se encuentran Colombia, México, Argentina Perú, Ecuador, Guatemala, España, Italia, Taiwán, China y Papúa Nueva Guinea.
- La mayoría de las soluciones destinadas a campus universitarios se observan en Colombia y en un segundo lugar en México.
- Algunos estudios se focalizan en los aspectos económicos asociados al uso del agua (análisis de costo-beneficio) y en otros en los aspectos ingenieriles para la construcción de los sistemas de captación de agua de lluvia.
- Se destaca el aprovechamiento del agua procedente de la niebla para la construcción de las soluciones.
- De los estudios primarios analizados, el 75 % (36 estudios) corresponden al tipo de investigación "propuesta de solución".

Se analizó la literatura existente sobre la captación de agua de lluvia como alternativa de sostenibilidad del uso del agua potable. Esto permitió evidenciar la necesidad de que las TICs pueden contribuir al aporte de soluciones. Desde esta perspectiva, se plantea la construcción de un prototipo de gestión que permita la captación del agua de lluvia destinada para la descarga de sanitarios de un establecimiento con alta demanda de uso de los sanitarios.

<sup>1</sup> SEDICI: Repositorio digital de la Universidad Nacional de La Plata, website: <http://sedici.unlp.edu.ar/>

### 3 Desarrollo

Dada la tendencia actual de la práctica de la industria del software del desarrollo de software híbrido [7], la cual se adoptó para el logro de la solución. Esto generó la combinación de métodos, prácticas y estándares, los cuales se detallan a continuación.

Para la especificación de los requisitos del prototipo se siguieron los lineamientos de un estándar tradicional, ISO/IEC/IEEE 29148 [8]. La documentación del proyecto se realizó con la herramienta CASE<sup>2</sup> Enterprise Architect (EA) [9]. Para las actividades de la construcción del prototipo se utilizó un modelo de ciclo de vida iterativo-incremental, respetando las disciplinas propuestas en el Proceso Unificado de Rational (RUP) [10]. El modelado del prototipo se logró mediante los diagramas de UML<sup>3</sup> [11] que consideramos necesarios para la definición, diseño y desarrollo del prototipo [6]. Los diagramas logrados se detallan en la Tabla 2.

**Tabla 2.** Diagramas de UML según disciplinas del RUP.

Disciplina de RUP	Denominación del Diagrama.
Ingeniería de Requerimientos	Diagrama de casos de uso.
Análisis y Diseño	Diagrama de casos de uso, Diagrama de comunicación, Diagrama de clases, Diagrama Entidad-Relación <sup>4</sup> , Diagrama de secuencia, Diagrama de estados.
Implementación	Diagrama de despliegue, Diagrama de componentes.

El prototipo para la gestión de agua de lluvia para la descarga de sanitarios cubre las siguientes funcionalidades: Recolección del agua de lluvia en tanques, Utilización del agua de lluvia almacenada para ser utilizada en descarga sanitaria, Automatización del uso del agua potable proveniente de la red de distribución, en caso de que no haya suficiente agua de lluvia recolectada, Agregar desinfectante concentrado al agua a utilizar en la descarga sanitaria, Generación de estadísticas del consumo de agua (lluvia y potable), Generación de alertas sobre el estado general del sistema en tiempo real, Automatización de carga de tanques mediante el uso de sensores, relé y bombas.

#### 3.1 Especificación de requisitos

Los requisitos funcionales especificados en el documento de Especificación de Requisitos de Software (ERS) se presentan en el diagrama del modelo de casos de uso de la Figura 1. Por razones de falta de espacio, los requisitos no funcionales se describen en [6].

Las estadísticas que el prototipo brinda al usuario (administrador) son: 1) Totalizador de litros de agua de red utilizada, 2) Totalizador de litros de agua de lluvia cargada, 3) Totalizador de litros suministrado para descarga sanitaria,

<sup>2</sup> Ingeniería de Software Asistida por Computadora (en inglés, Computer Aided Software Engineering).

<sup>3</sup> UML; Lenguaje de Modelado Unificado.

<sup>4</sup> El Diagrama Entidad-Relación si bien no es un diagrama de UML, ha sido necesaria su construcción para el diseño de la base de datos.

4) Totalizador de caudal de litros por minuto de agua de red, 5) Totalizador de caudal de litros por minuto de agua de lluvia, 6) Porcentaje de ahorro de agua potable, 7) Estado de sistema: Falta de desinfectante/Proceso defectuoso/Procedimiento en curso/En espera. Las consultas de históricos por tipo de agua consumida se pueden obtener mediante rangos de fechas. Las alertas sonoras consideradas en el prototipo notifican: Bajo nivel de desinfectante ultra concentrado y Falla crítica de funcionamiento que impide la carga de agua y/o carga de desinfectante al Tanque descarga sanitaria.

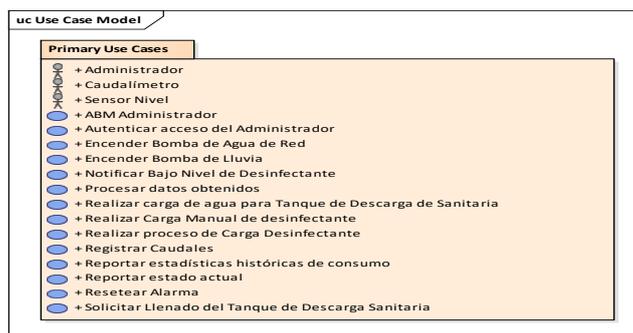


Fig. 1. Requisitos funcionales del prototipo.

Para lograr una precisa conceptualización del prototipo de descarga de agua sanitaria, ha sido necesario construir un modelo conceptual de alto nivel. En el caso de un establecimiento real se requiere que la descarga de sanitarios se realice por medio de tanques de almacenamiento aéreos. El agua de lluvia es recolectada a través del techo, se encauza a tanques subterráneos, y se establecen los procesos básicos para cumplir el propósito del prototipo del sistema. En la Figura 2, se presenta el modelo conceptual de alto nivel.

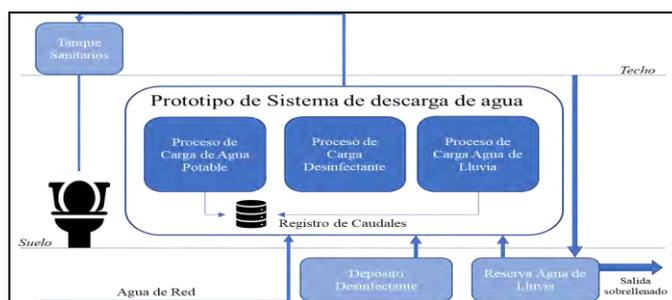


Fig. 2. Modelo conceptual del prototipo de alto nivel.

Para complementar y lograr un mayor nivel de detalle de la solución se construyó un diagrama de interacción de componentes que además de incluir los componentes del diagrama de alto nivel (Figura 2), también incluye el modelado de las interconexiones e interfaces externas y la especificación de componentes electrónicos. En la Figura 3 se presenta el diagrama de interacción de componentes detallando

interfaces, entradas, salidas, procesos y almacenamiento de datos. Las interfaces entre componentes cumplen un rol central en este diagrama. Se visualiza como se concentran los componentes del prototipo, y se incluyen las interfaces de usuario, el software desarrollado y el almacenamiento en base de datos, que se utiliza para generar las estadísticas de uso de agua y estado de este. Es importante aclarar que para una mejor comprensión del prototipo ha sido necesario considerar algunos componentes que se tendrán en cuenta al momento de implementarlo en un establecimiento real (techo y tanques subterráneos).

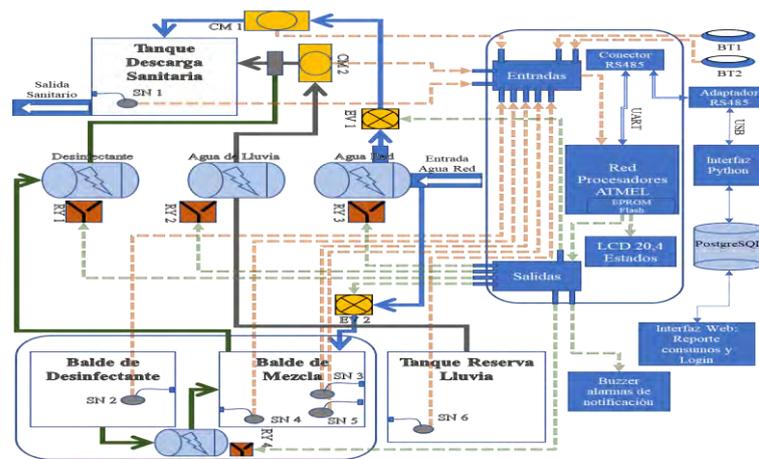


Fig. 3. Modelo detallado de interacción componentes.

### 3.2 Análisis y Diseño

En la Figura 4, se presenta el Diagrama de Casos de Uso del prototipo de descarga de sanitarios. Se presentan como actores del prototipo: el administrador, el caudalímetro, el sensor de nivel. En la Tabla 3, se presentan los roles que cumplen los actores.

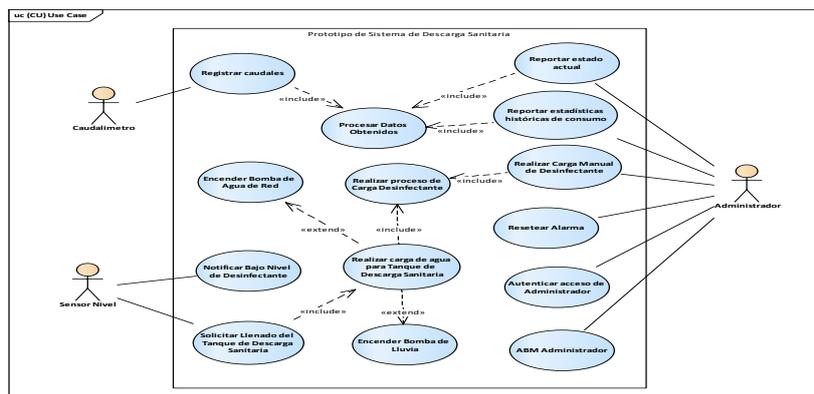


Fig. 4. Diagrama de Casos de Uso.

**Tabla 3.** Actores del prototipo.

Actor	Definición
Sensor Nivel	Informar a los actuadores la interrupción de un evento que podría disparar una determinada acción, como por ejemplo encender o apagar una bomba.
Caudalímetro	informar a las interfaces la cantidad de agua que pasa por una tubería.
Administrador	Solicitar estadísticas de consumos del prototipo y realizar la carga del desinfectante.

Se construyó el modelo de análisis, con el propósito de lograr una especificación detallada de los requisitos hacia una primera aproximación al diseño del prototipo. A modo de ejemplo, se presenta el Diagrama de Comunicación del Caso de Uso - Carga Manual de Desinfectante [6].

Se construyeron los Diagramas de Secuencia para los Casos de Uso presentados en la Figura 4. En [6] se presenta a modo de ejemplo el Diagrama de Secuencia del Caso de Uso: Realizar Carga Manual del Desinfectante y su explicación.

### 3.3 Implementación

Los Diagramas construidos en esta etapa del Proyecto contribuyeron en la decisión respecto a la organización de los componentes del prototipo. Estos son el Diagrama de Componentes y el Diagrama de Despliegue, ambos se encuentran en [6].

El desarrollo del prototipo se realizó mediante el lenguaje de programación C++, que forma parte del lenguaje soportado por las placas Arduino [12]. La interfaz web se desarrolló en PHP 7 [13]. Las interfaces que comunican las placas Arduino con la base de datos están desarrolladas en Python [14].

A continuación, se presentan las interfaces de usuario del sistema (Figura 6), a la izquierda el visor con el estado del sistema, en el centro el caudal de agua de red y de lluvia y a la derecha el consumo de agua de red y de agua de lluvia.

**Fig. 6.** Interfaz del sistema.

En la Figura 7 se presenta una vista cenital y frontal del prototipo para la gestión de agua de lluvia para la descarga de sanitarios.

El prototipo cuenta con 2 tanques de almacenamiento de agua (agua de lluvia y agua mezclada con desinfectante para descarga sanitaria) y 2 baldes de almacenamiento de desinfectante (Balde de desinfectante ultra concentrado y Balde de mezcla con agua).

Los baldes y los tanques poseen sensores de nivel que miden el fluido almacenado y disparan acciones de acuerdo con los niveles de estos, por ejemplo, accionar el llenado de alguno de los tanques o baldes mediante electroválvulas y el aviso de alertas sonoras por vacío del balde de desinfectante ultra concentrado.

El bajo nivel de desinfectante ultra concentrado es una de las operaciones manuales en el funcionamiento del prototipo que requiere intervención del Administrador. En este caso, al detectar el prototipo un bajo nivel de desinfectante ultra concentrado del Balde de desinfectante informa al Administrador mediante un buzzer sonoro. El prototipo, puede continuar el funcionamiento sin cargar de desinfectante ultra concentrado.

Cuando el sensor de nivel detecta un bajo nivel de agua en el Tanque descarga sanitaria, se acciona la ejecución de llenado de este. El Tanque descarga sanitaria tiene como prioridad llenar con agua del Tanque reserva lluvia, y en caso de no haber existencia de agua de lluvia en el tanque correspondiente, se continúa el llenado del Tanque descarga sanitaria con agua de red de forma automatizada.

El prototipo posee dos caudalímetros, uno que realiza la medición de ingreso de agua de Red al Tanque descarga sanitaria y otro que realiza la medición de ingreso de agua de lluvia al Tanque descarga sanitaria. Esta información es almacenada en base de datos para ser consultada por el Administrador y poder visualizar estadísticas.

El prototipo posee una interfaz de usuario con detalle de bajo nivel con los datos obtenidos de los caudalímetros para poder reportar el consumo histórico de los mismos. Esta interfaz de usuario tiene fines estadísticos y detección de errores o inconvenientes por lo que debe ser accesible por el Administrador. De esta forma, el Administrador puede visualizar en tiempo real toda la información necesaria que le permite realizar alguna intervención ante inconvenientes o problemas.

Se permite el ingreso del Administrador mediante nombre de usuario y password. El prototipo desplegará la interfaz web Reportar estadísticas históricas de consumo. También permite realizar el ABM de Administradores si el rol de este lo permite.

## 4 Pruebas

Se planificaron y desarrollaron un conjunto de pruebas unitarias y funcionales. Para cada uno de estos tipos de prueba se diseñaron las plantillas para su registro que se encuentran en [6]. Se realizaron en total 32 pruebas unitarias. Las pruebas funcionales realizadas se corresponden con los 9 requisitos funcionales en la ERS.

En [6] se describe a modo de ejemplo las trazas de las pruebas de integración PI01 del RF8: Autenticar acceso de Administrador y PI02 del RF9: ABM Administrador.

Los tipos de pruebas definidos han sido acertados para verificar el correcto funcionamiento del prototipo.

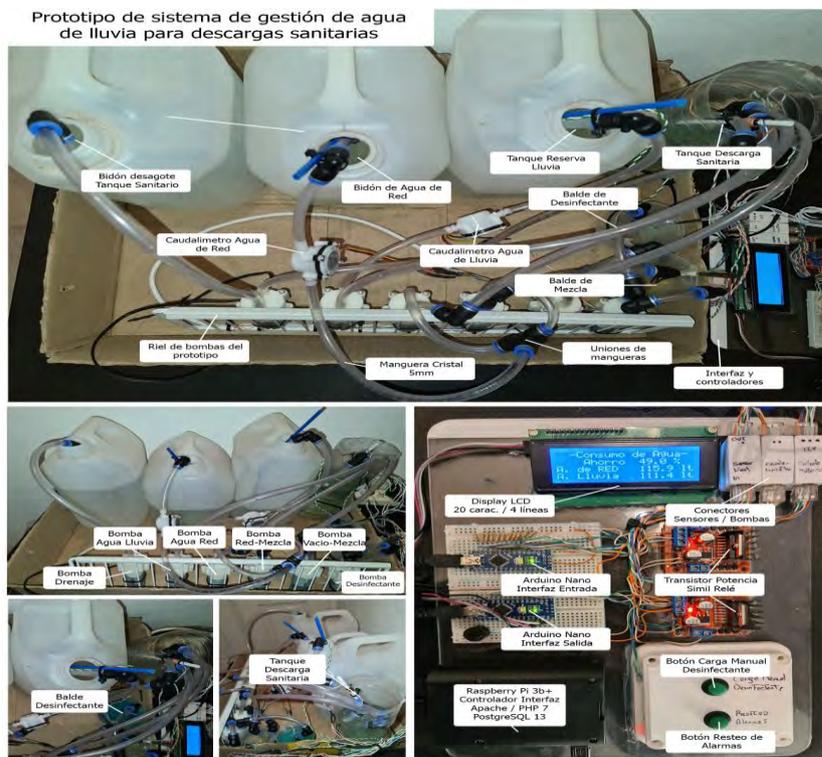


Fig. 7. Prototipo para la gestión de agua de lluvia para la descarga de sanitarios (Vista Cental y Frontal).

## 5 Conclusiones y Trabajos Futuros

En este trabajo se describe la experiencia en la definición, diseño y desarrollo de un prototipo de gestión de agua de lluvia para la descarga de sanitarios. Esta solución es aplicable a cualquier tipo de establecimiento donde los sanitarios tienen una alta demanda de uso.

Las decisiones tomadas al inicio del proyecto respecto a la combinación de métodos, prácticas y metodologías han sido acertadas dado que permitió el desarrollo de la solución. El uso de un estándar de especificación de requisitos tradicional como la ISO/IEC/IEEE 29148 contribuyó de una manera sustancial. Esto permitió reflexionar y revisar todos los elementos necesarios para la construcción de la solución.

Desde el inicio de este Proyecto, nos motivó situar nuestra disciplina a disposición de la Sociedad, en especial con la cooperación en el cuidado del medio ambiente y en particular con la sostenibilidad del uso del agua potable.

Se destacan como trabajos futuros: a) Incorporar en el prototipo la gestión del uso de sanitarios; b) Implementar el prototipo en un entorno real, es decir en un establecimiento considerando las adecuaciones necesarias respecto a la obra civil y c) Extender la funcionalidad mediante la incorporación de un módulo de Inteligencia de negocio que permita aplicar minería de datos sobre consumos de agua.

## Referencias

1. Organización de las Naciones Unidas. Objetivos de Desarrollo Sostenible (2015). Disponible en: <https://www.un.org/sustainabledevelopment/es/water-and-sanitation/> Último acceso: 25 de junio 2020.
2. Adler, I., Carmona, G. & Bojalil, J. A. (2008). Manual de Captación de Aguas de Lluvia – Uso y mantenimiento para su sistema residencial. Centros Urbanos. Instituto Internacional de Recursos Renovables AC. (2008). Disponible en: <http://www.irrimexico.org/biblioteca>. Último acceso: 7 de junio 2020.
3. Carvajal Gomez D. A. Prototipo sistema automatizado de recolección de agua lluvia doméstico. Tesis de Grado de Tecnólogo en Automatización Industrial. Corporación Universitaria Minuto De Dios - Tecnología en Automatización Industrial. Bogotá (2016).
4. Crespi L., Schiffino C., Sattolo I, Panizzi. Impacto de las TICs en la sustentabilidad del uso del agua potable en la República Argentina. En las Actas del XXII Workshop de Investigadores en Ciencias de la Computación. 7 y 8 de Mayo, Calafate, Santa Cruz, Argentina (2020). ISBN en trámite.
5. Kitchenham, B. Budgen D. and Brereton, P. (2015). Evidence-Based Software Engineering and Systematic Reviews. Chapman and Hall/CRC. 1 st. Editon. New York, USA.
6. Crespi L., Schiffino C., Sattolo I, Panizzi. Anexo – Prototipo para la Gestión de agua de lluvia para descarga de sanitarios. Disponible en: <https://doi.org/10.6084/m9.figshare.12935267.v1>
7. Kuhmann M., Tell P., Klünder J., Hebig R., Licorish S., MacDonell S. (Eds.): Complementing Materials for HELENA Study (Stage 2), online DOI: 10.13140/RG.2.2.1.1031.65288, published: 2018-11-28.
8. ISO/IEC/IEEE Std 29148:2011, IEEE Systems and software engineering -- Life cycle processes --Requirements engineering, IEEE Computer Society (2011). ISBN 978-0-7381-6591-2.
9. Sparx Systems. Enterprise Architect 13.5. (2017). Obtenido de <https://sparxsystems.com/>
10. Jacobson, I., Booch, G., & James, R. El Proceso Unificado de desarrollo de software. Addison Wesley (2000).
11. Rumbaugh, J., Jacobson, I., Grady, B. El Lenguaje Unificado de Modelado. Manual de Referencia. Segunda Edición. Pearson (2006).
12. Arduino - Language Reference. Disponible en: <https://www.arduino.cc/reference/en>. Último acceso: 13 de mayo 2020.
13. Minera, F. PHP + MySQL desde cero. Ciudad de Buenos Aires: Dálaga (2014).
14. Python Software Foundation. Disponible en <https://www.python.org/psf>. Último acceso: 9 de mayo 2020.

## Marco de vinculación de datos abiertos aplicado al contexto de datos medioambientales

Juan Santiago Preisegger<sup>1</sup>, Alejandro Greco, Ariel Pasini,

Marcos Boracchia, 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 Comisión de Investigaciones Científicas de la Pcia. de Bs. As. (CIC)

<sup>1</sup> Becario UNLP

{jspreisegger, apasini, marcosb, ppesado}@lidi.info.unlp.edu.ar  
{alegreounlp}@gmail.com

**Abstract.** Los portales de datos abiertos ponen un conjunto muy importante de información a disposición de la comunidad. Los interesados en una temática en particular obtienen los datos sobre dicha temática desde diferentes portales, pero en el momento de querer procesarla en conjunto se dificulta debido a los diferentes criterios de cada portal para publicarlos. Para asistir en este proceso se desarrolló un marco de trabajo que facilita la vinculación de estos archivos a usuarios con poca experiencia técnica en el análisis de datos. Dentro del marco se utiliza una herramienta que permite aplicar diferentes operaciones sobre los archivos, generando gráficos en un panel de control, facilitando a los interesados el análisis de la información. El marco se aplicó a la temática del medioambiente, en particular, sobre la calidad del agua, del aire y la generación de energía.

**Keywords:** Datos Abiertos, Vinculación de Datos, Ingeniería de Software, Gobierno Abierto, Medioambiente.

### 1 Introducción

Una ciudad se nutre del comportamiento de sus ciudadanos. Los ciudadanos, mediante diferentes dispositivos, son capaces de registrar cada vez más información de las actividades que realizan. Hacer un uso inteligente de la información registrada por las autoridades gubernamentales para mejorar la vida de los ciudadanos es de gran aporte a la misma comunidad. Los aportes que se pueden lograr de los datos registrados no solo pueden provenir del área gubernamental, sino también de diferentes organismos o personas físicas, capaces de analizar la información, procesarla y proponer mejoras. Esto es un factor importante en este ciclo de mejora de una ciudad y es lo que la convierte en una “ciudad inteligente sostenible y participativa”. [1]

Para 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 procesarlos y generar aportes en el proceso de mejora de la ciudad [2][3]. 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., las cuales dificultan, y en algunos casos imposibilitan, el análisis de la información.

El marco de trabajo que se propone busca vincular datasets, obtenidos de portales abiertos, de una temática en particular y permitir un análisis en conjunto de estos datos, de una forma sencilla que no requiera de conocimientos técnicos avanzados. Dicho marco se basa en cinco pasos: 1) *Búsqueda*, 2) *Análisis preliminar*, 3) *Carga directa*, 4) *Normalización*, 5) *Vinculación*. Este proceso se apoya en la utilización de la herramienta Indimaker.

Indimaker tiene el potencial de procesar archivos de varios formatos, aplicar diferentes operaciones al contenido y lograr una vinculación entre los archivos, conformando un panel de control de indicadores que facilita la visualización de las operaciones realizadas sobre los archivos.

Para validar el proceso, el marco se aplicó a datos abiertos relacionados con el medioambiente, en particular a conjuntos de datos obtenidos de diferentes portales públicos sobre la calidad del aire, del agua y el uso de la energía.

En la segunda sección se presentan los conceptos de ciudades inteligentes sostenibles y datos abiertos. Luego, en la tercera sección, los conceptos generales sobre los tableros de control de indicadores y la herramienta Indimaker. En la cuarta sección se presenta el marco de trabajo de vinculación de datos abiertos. En la quinta, la aplicación del marco a datos relacionados con el medioambiente y, por último, en la sexta sección las conclusiones y trabajos futuros.

## **2 Las ciudades inteligentes sostenibles y los datos abiertos**

Las ciudades se nutren de la participación de su comunidad. 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 a fin de convertir a una ciudad en una ciudad inteligente.

### **2.1 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 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".

## **2.2 Datos Abiertos**

La sociedad presenta exigencias cada vez más elevadas hacia sus gobernantes y sus entes de gobierno. Entre estas exigencias, se incluye la transparencia y el manejo eficiente de los bienes públicos, la inclusión en la toma de decisiones y la colaboración con distintos sectores de la sociedad [4]. En función de estos requisitos, y con la ayuda de nuevas tecnologías, se generó una nueva forma de gobierno que incluye más al ciudadano, permitiéndole generar aportes a las políticas públicas y participar en la toma de decisiones. [5][6]

A través de la implantación del gobierno abierto se generó la apertura de los datos, la cual consiste en poner a disposición de la sociedad los datos de interés común de la ciudadanía para que, de cualquier forma, éstos puedan desarrollar una nueva idea o aplicación que entregue nuevos datos, conocimientos u otros servicios que el gobierno no es capaz de entregar. [3][7][8]

La variedad de los datos que se ponen a disposición de la sociedad es muy amplia y no solo se realiza desde las agencias gubernamentales, sino que, además, organismos internacionales, ONGs y otras organizaciones impulsan diversas medidas para poner a disposición de la sociedad cada vez más fuentes de datos, no solo referidas a las gestiones de gobierno, sino también relacionados a otros ámbitos, como el uso racional de los recursos y los cuidados del medioambiente.

## **3 Generación de tableros de control con datos abiertos**

Una de las aplicaciones de los datos es la construcción de tableros de control, basados en indicadores generados con estos datos. Para la generación de estos tableros se deben analizar las fuentes de los datos, los datasets y poseer herramientas que nos permitan relacionar los datos, con el fin de construir información relevante que se transforme en un indicador que nos posibilite mejorar la toma de decisiones.

### **3.1 Fuentes de datos**

Los nuevos paradigmas, gestados en los diferentes organismos, están coordinando acciones para mejorarle la calidad de vida a la sociedad, mediante la apertura de datos y las mejoras planteadas en la infraestructura de las ciudades. Se generaron diversas

aplicaciones y herramientas, desde múltiples sectores, para brindarles soporte y automatizar, o mejorar, el proceso de publicación, búsqueda y, en ocasiones, procesamiento de la información a las distintas partes de la sociedad. Entre estas herramientas se destacan los catálogos y portales de datos abiertos, mediante los cuales, los organismos publican los datos sobre diferentes aspectos de su ejercicio y del medioambiente en el cual operan. Por ejemplo, algunos países, provincias, municipios u organismos poseen portales donde unifican datos de las distintas regiones o temáticas en las que se especializan. Los más avanzados en el área, publican sus datos y los describen con esquemas de datos, para hacerlos más descriptivos en cuanto a su contenido.

Una herramienta que utiliza esta información para permitir llevar a cabo la búsqueda de los datos de manera más global es Google Dataset Search: un motor de búsqueda especializado en encontrar conjuntos de datos (datasets) almacenados en la web, a través de palabras clave, siempre y cuando estos utilicen etiquetas de conjuntos de datos schema.org o estructuras equivalentes representadas en el formato Data Catalog Vocabulary (DCAT).[9]

### 3.2 Datasets

El término dataset es un anglicismo. Su traducción al castellano sería “conjunto de datos” y es una colección de datos habitualmente tabulada, es decir, que corresponden a los contenidos de una única tabla de base de datos o una única matriz de datos, en donde cada columna de la tabla representa una variable en particular y cada fila representa un miembro determinado del conjunto de datos.

Los datasets son la pieza principal de los portales y catálogos de datos. Agrupan uno o más recursos de datos y, para su publicación, requieren una preparación previa para poder ser procesados y reutilizados por terceros. Según [10], esto incluye tres actividades: 1) *Documentación*: consiste en definir los metadatos que tendrá cada uno de los datasets a publicar. Los metadatos describen los aspectos básicos del dataset, y se utilizan para organizar, clasificar, relacionar y encontrar los datos necesarios (por ejemplo: título, descripción, institución, licencia, categoría, fecha de publicación, etc.). 2) *Estructuración*: consiste en la preparación del conjunto de datos a publicar en un formato estructurado, sin campos erróneos o vacíos, que permita su reutilización y procesamiento en cualquier software. 3) *Carga de datos*: consiste en su publicación en una plataforma, que permita la organización y fácil acceso por parte de quienes van a reutilizar los datos.

### 3.3 Tableros de control

Un indicador se puede definir como un dato, o conjunto de datos, que ayudan a medir objetivamente la evolución de un proceso o de una actividad, correspondiente a cualquier organización. Los indicadores se pueden organizar y relacionar de manera que conformen un tablero de control. Estos tableros permiten realizar un seguimiento y evaluación del proceso o actividad, de una forma más exhaustiva. Además,

generalmente, permiten ver la evolución de los mismos de forma gráfica facilitando la interpretación de los resultados.

### **3.4 Herramienta para la generación de tableros de control - Indimaker**

IndiMaker es un sistema accesible desde cualquier navegador web, que permite la vinculación entre datasets, para la construcción de indicadores personalizados en tableros de control.

Los tableros de control se generan a partir de los distintos datasets cargados por el usuario en un primer momento, aunque posteriormente se permite agregar, quitar o realizar una combinación entre dos de estos, obteniendo un nuevo dataset listo para poder ser utilizado.

La herramienta permite importar datasets de diversos formatos (.xls, .xlsx, .xml, .ods y .csv), e incluso unir los que tengan columnas con contenido en común, para generar uno nuevo con información más enriquecida. A su vez, posee la capacidad de homogeneizar toda esta información y almacenarla en su base de datos, brindando la posibilidad al usuario de realizar operaciones entre los distintos datos almacenados para la construcción de indicadores.

Los indicadores deben ser generados de manera apropiada, sino, pueden generar información imprecisa, errónea o subjetiva constituyendo un obstáculo para el correcto análisis. Una de las grandes virtudes de esta herramienta es su sencillez a la hora de generarlos, ya que se encuentra diseñada para hacer uso de la información de una forma que resulte fácil para el usuario realizar operaciones complejas entre los diferentes datos, evitando a priori la generación de indicadores erróneos. Los indicadores generados servirán para elaborar una medida cuantitativa que tendrá significado para quien lo analice.

A su vez, la herramienta permite adaptarse a los diferentes dispositivos, ya sea una computadora de escritorio, una tablet o un smartphone, y además puede ser utilizada tanto en idioma español como en inglés.

Al acceder a la herramienta, se puede observar el menú para la gestión de los tableros de control pertenecientes al usuario, con sus fuentes de datos asignadas y una descripción que sirve para comprender el propósito de cada tablero. Para crear indicadores pertenecientes a un tablero, se debe indicar su nombre, tipo de indicador, usuario responsable, periodicidad de medición, descripción y niveles de referencia, además de la operación en la cual se basa. Las operaciones que pueden realizarse entre los datos incluyen sumas, restas, divisiones, porcentajes, agrupaciones de datos y varias comparaciones lógicas, incluida la posibilidad de utilizar expresiones regulares para usuarios más avanzados. La combinación de estas operaciones serán las que le permitan al usuario construir un indicador.

Los indicadores de cada tablero se pueden representar de varias formas, incluyendo diferentes gráficos, lo que permitirá alertar fácilmente al usuario sobre posibles desviaciones de los objetivos previamente establecidos.

## 4 Marco de vinculación de datos abiertos

Al analizar los datos de los diferentes catálogos o portales, aparecen problemas de incompatibilidad entre los formatos de los datos presentados por los diferentes proveedores. Dado este inconveniente, se comprendió que, para lograr una vinculación de datos exitosa, es de vital importancia analizar previamente las fuentes de datos y los formatos de cada una de ellas. En consecuencia, se avanzó en generar un marco de vinculación de datos de cinco pasos: 1) *Búsqueda*, 2) *Análisis preliminar*, 3) *Carga directa*, 4) *Normalización*, 5) *Vinculación*.

1. *Búsqueda*: Realizar la búsqueda de los datasets referidos al área de interés, mediante los portales o catálogos de datos de las organizaciones, o ayudándose con la herramienta Google Dataset Search.
2. *Análisis preliminar*: Analizar los datasets obtenidos, verificando que posean la información en un formato compatible con la herramienta y que todas las columnas posean encabezado. Además, comprobar el formato del contenido (completar filas, columnas, o trasponer filas por columnas en alguno de los datasets, etc.) para lograr una vinculación exitosa.
3. *Carga de archivos*: Cargar el archivo en Indimaker. Al cargar el archivo, la herramienta realizará una serie de verificaciones sobre el contenido. De ser validado, se procederá a la instancia de comparación. En el caso que detecte inconsistencias, se procederá a la instancia de normalización del contenido.
4. *Normalización*: El proceso de normalización puede realizarse de forma manual o automática (utilizando una herramienta externa), dependiendo del tamaño del archivo. En esta instancia se verifica:
  - a. Que los datos en las columnas de los datasets presenten el mismo tipo de dato.
  - b. Que los datos en las columnas posean algún valor.
  - c. Que no se exceda la cantidad de datos que puede poseer cada fila.
5. *Vinculación*: Unificar las tablas en base a un parámetro en común, en caso de ser necesario, para obtener más información y poder relacionar los datos para analizarlos de manera sencilla y directa.

## 5 Aplicación de la vinculación de datos abiertos de medioambiente

Los organismos gubernamentales y no gubernamentales ponen, a disposición de la comunidad, numerosos conjuntos de datos del medioambiente de la región geográfica a la que pertenecen. Al momento de analizar la información en su conjunto para obtener valores regionales, aparecen los problemas de incompatibilidad entre los formatos de los datos presentados por los diferentes organismos. Es de vital importancia analizar las fuentes de datos, y los formatos de cada una de ellas, para permitir realizar un proceso de vinculación de la información.

A continuación, se aplicarán los 5 pasos del marco de trabajo:

## 1. Búsqueda

Se seleccionó como caso de aplicación el medioambiente. Entre las distintas partes que componen la temática, se decidió acotarla a los aspectos que afectan a la sociedad día a día, como son la calidad del agua y del aire, bienes fundamentales para la vida diaria de la sociedad y el futuro de la misma. A su vez, se incluyó a la generación de energía, bien que se relaciona a menudo con la calidad de vida.

Las búsquedas se realizaron bajo las palabras: *Calidad de agua potable*, *Generación de energía*, *Calidad del aire*. En la tabla 1, a continuación, se puede observar la selección de los datasets obtenidos de los diferentes países.

<i>País</i>	<i>Agua</i>	<i>Aire</i>	<i>Energía</i>
<i>EEUU</i>	✓	✓	✓
<i>Chile</i>	✓	✓	✓
<i>Brasil</i>	✓	✓	✓
<i>Uruguay</i>	×	✓	✓
<i>Paraguay</i>	×	×	✓
<i>Bolivia</i>	×	×	×
<i>Perú</i>	×	×	×
<i>Colombia</i>	✓	✓	×
<i>Argentina</i>	✓	×	✓

**Tabla 1.** Datasets obtenidos por país

## 2. Análisis preliminar

A partir de los datos obtenidos se analizó su forma, los formatos disponibles y las estructuras de cada uno de ellos, de forma que, utilizando la herramienta propuesta, se los relacione para poder llevar a cabo un análisis mayor de la situación de las distintas regiones.

### *Datos de aire y agua*

En el caso de la calidad del aire y la calidad del agua, fue posible visibilizar cierto estándar para analizar las magnitudes existentes en distintas características. Se pudo observar que se llevan a cabo casi las mismas pruebas sobre distintas muestras para analizar distintas características y determinar si se encuentran dentro de los márgenes saludables para el consumo humano.

### *Datos de la energía*

En el caso de los datasets del área de energía, la diferencia entre los datos publicados por las distintas organizaciones es aún mayor. Se observó que, algunos países, publican simplemente el porcentaje anual generado y el tipo de obtención de energía en la cual se basan las distintas centrales de generación publicadas, otros países simplemente publican los porcentajes de sus fuentes de generación de energía, sin discernir entre las centrales que poseen, y otros publican las capacidades totales de las centrales, sus tipos y hasta las empresas propietarias de las mismas.

### 3. Carga directa

Con los datos recolectados se procedió a realizar la carga de manera directa, como punto de partida, para analizar si cumplía con el control que la herramienta realiza al momento de la carga de los datos. En este caso se comprobó que no todos los archivos poseían los datos de manera ordenada y sin errores, y solamente un 35% de los datasets pasaron este mínimo control, lo cual indica que más allá de la publicación de los datos realizada por los organismos, hace falta mayor normalización para su procesamiento. Si se analizan los datos por área, se puede observar que la progresión, en base a esa normalización, es: Agua: 80% de forma directa y un 20 % de forma indirecta; Energía: 66% Directa y 33% indirecta; Aire: 50% directa y 50% indirecta.

### 4. Normalización

En las tres áreas se pudo observar que, más allá de la estandarización en cuanto a ciertos datos publicados, los datos están en un formato muy “crudo” que imposibilita un procesamiento óptimo. En muchos casos poseen espacios vacíos, o incluso poseen más datos en las filas de los que se declara. En cada caso podemos ver:

#### *Datos de aire y agua*

Existen diferencias en la estructura de los datasets, las unidades utilizadas entre las distintas magnitudes que se analizan y cómo se almacenan estas en los datasets. Por ejemplo, la variación de columnas a filas o la división de un campo en varios. Esto dificulta la interrelación entre los distintos datasets de estas temáticas.

#### *Datos de la energía*

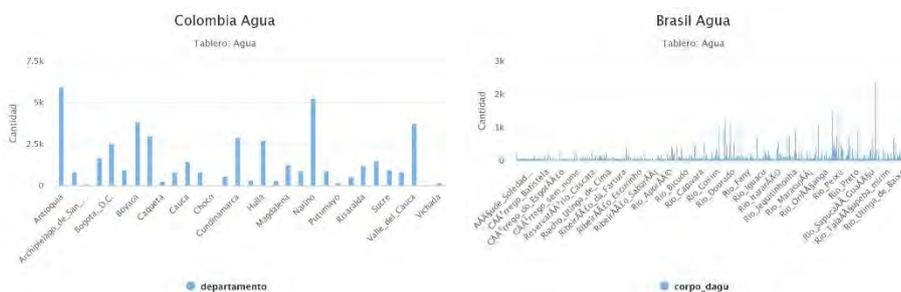
En este caso, se observó que varía mucho la unidad de medición utilizada en los distintos datasets, lo cual imposibilita una comparación lineal de los datos. A su vez, por la gran diferencia en cuanto a los datos publicados por cada organización, es necesario llevar a cabo una selección de ciertos campos en común entre los distintos datasets, sumándolos en base a su tipo, para poder visibilizarlos de manera tipificada. Esto se dificulta debido a la dispersión existente.

### 5. Vinculación

Una vez normalizados los datos, se puede realizar la vinculación de distintos datasets que posean datos en común, permitiendo un análisis regional de la información. Con estos datasets, la herramienta nos permite seleccionar columnas para realizar operaciones y obtener valores con los que se pueden generar diversos gráficos que permiten un análisis lineal de magnitudes de interés. A continuación, se puede visibilizar el análisis que se realizó en cada uno de los conjuntos de datos analizados.

#### *Datos del agua*

Entre la información recolectada sobre la calidad del agua se seleccionaron, a modo de ejemplo, los datos de Colombia y Brasil. Se realizó la suma de la cantidad de mediciones realizadas en ambos países y se generaron los gráficos de barras verticales que se pueden observar a continuación, en la Imagen 1. Se puede observar que la información publicada por Colombia se encuentra sectorizada por estados y, en cambio, la información de Brasil se encuentra sectorizada en función de los distintos efluentes naturales que posee el país. La herramienta permite, de todas maneras, visualizar esta información más allá de sus diferencias.

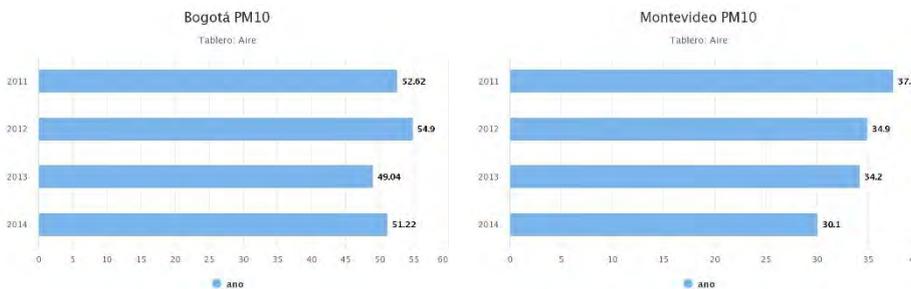


**Imagen 1.** Numero de mediciones de calidad del agua en Brasil y Colombia

En un segundo análisis, se podría relacionar los efluentes con los estados de Brasil y comparar por regiones geográficas.

*Datos del aire*

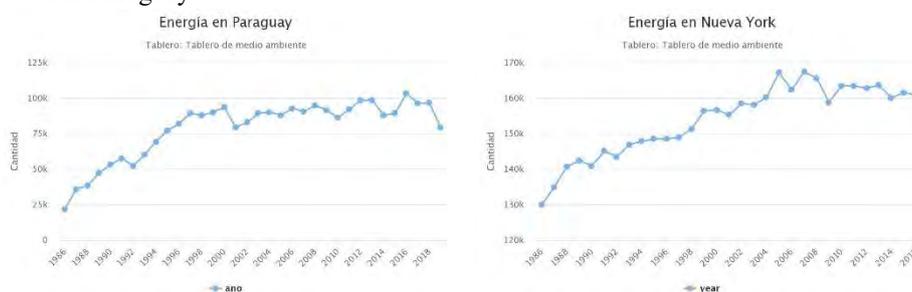
Entre la información recolectada sobre la calidad del aire se seleccionaron, a modo de ejemplo, los datos de partículas en suspensión de la ciudad de Bogotá y Montevideo, se realizó el promedio por año de los valores de Bogotá (de Montevideo no fue necesario porque los brindados ya eran valores promedio) y se generaron los gráficos de barras horizontales que se pueden observar a continuación, en la Imagen 2. En este caso, la información publicada por Colombia estaba más detallada en base a múltiples mediciones por año, por lo cual hubo que procesar más la información para obtener un formato similar al publicado por Uruguay. La herramienta permite, de todas maneras, visualizar esta información más allá de sus diferencias, acotandola en base a los mismos años.



**Imagen 2.** Mediciones de PM10 en Bogotá y Montevideo

### *Datos de la energía*

Entre la información recolectada sobre la producción de energía se seleccionaron, a modo de ejemplo, los datos de Paraguay y la ciudad de Nueva York. Más allá de sus diferencias en cuanto a dimensión y a población, se realizó la comparación con estos datos para poder observar las diferencias en cuanto a los valores de la producción energética, en GWh. La información publicada en este caso era muy similar, aunque Nueva York especificaba, además, las fuentes de producción de energía. Con esta información, se generaron los gráficos de líneas que se pueden observar a continuación, en la Imagen 3, donde se puede apreciar que la ciudad de Nueva York produjo, en el período analizado, más energía que todo Paraguay.



**Imagen 3.** Producción de energía por año en Paraguay y en Nueva York

## 6 Conclusión y trabajos futuros

A lo largo del artículo se introdujeron los conceptos básicos de Ciudades Inteligentes Sostenibles, Datos Abiertos, Fuentes de Datos, Datasets y Tableros de Control. Se hizo hincapié en cómo pueden relacionarse estos conceptos para generar un mayor nivel de información para la sociedad, a partir de la vinculación de los datos que la misma dispone y en base al uso de la herramienta Indimaker.

En función de esta herramienta se generó un marco de trabajo, compuesto por cinco etapas, para la vinculación de los datos puestos a disposición, por diversas organizaciones, de manera de poder compararlos sencillamente.

Como caso de estudio, para la aplicación del marco, se descargaron fuentes de datos abiertos de 10 países sobre la calidad del agua, calidad del aire y la producción energética, obteniendo información en 8 de estos portales de datos.

La información obtenida se procesó, según el marco de trabajo, normalizando la información en los casos que fue necesario y, luego, utilizando la herramienta Indimaker, la cual brindó la posibilidad de realizar operaciones sobre los datos para generar indicadores de forma sencilla. Fue notable la diferencia que se encontró en la información publicada sobre la misma temática, pero, con la aplicación del marco propuesto, la herramienta pudo procesar la información generando indicadores de

manera sencilla para el usuario y diversos gráficos que facilitaron el análisis de cada área. Cabe destacar que el modelo puede extenderse a cualquier área de interés.

A futuro, se espera vincular la herramienta con APIs de diferentes portales de datos brindadas por diversas organizaciones y ampliar el abanico de operaciones que la herramienta brinda, incluyendo la posibilidad de visualizaciones de diversos indicadores en un mismo gráfico.

## 7 Agradecimientos

Project co-funded by the Erasmus+ Programme of the European Union. Grant no: 598273-EPP-1-2018-1-AT-EPPKA2-CBHE-JP.

## 8 Bibliografía

- [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. A. Chun, S. Shulman, R. Sandoval, and E. Hovy, “Government 2.0: Making connections between citizens, data and government,” *Inf. Polity*, vol. 15, no. 1–2, pp. 1–9, 2010.
- [3] G. Concha and A. Naser, “Datos abiertos: Un nuevo desafío para los gobiernos de la región,” *Inst. Latinoam. y del Caribe Planif. Económica y Soc.*, 2012.
- [4] C. Calderón and S. Lorenzo, *Open Government. Gobierno Abierto*. 2010.
- [5] A. Naser, Á. Ramírez-Alujas, and D. R. Editores, *Desde el gobierno abierto al Estado abierto en America Latina y el Caribe: Planificación para el Desarrollo*. 2017.
- [6] J. R. Gil-García and J. I. Criado, *Las Tecnologías de Información y Comunicación en las Administraciones Públicas Contemporáneas*. 2017.
- [7] A. Pasini, J. S. Preisegger, and P. Pesado, “Modelos de evaluación de gobiernos abiertos , aplicado a los municipios de la provincia de Buenos Aires,” *XXIV Congr. Argentino Ciencias la Comput.*, vol. XXIV, pp. 0–10, 2018.
- [8] A. Pasini, J. S. Preisegger, and P. Pesado, “Open Government Assessment Models Applied to Province’s Capital Cities in Argentina and Municipalities in the Province of Buenos Aires,” in *Communications in Computer and Information Science*, 2019, vol. 995, pp. 355–366.
- [9] N. Noy, M. Burgess, and D. Brickley, “Google dataset search: Building a search engine for datasets in an open web ecosystem,” in *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, 2019, pp. 1365–1375.
- [10] A. Naser and A. Ramirez, “Plan de gobierno abierto. Una hoja de ruta para los Gobiernos de la Región,” *CEPAL - Manuales*, vol. 81, p. 80, 2017.

## Validando las dos primeras etapas de metodo\_SCGE en un organismo gubernamental

Marcelo Castro<sup>1</sup>, Claudia Fabiana Pons<sup>2</sup>, Rocío Andrea Rodríguez<sup>3</sup>

<sup>1</sup>Facultad de Ingeniería, Universidad Nacional de Jujuy  
mcastro.fi.unju@gmail.com

<sup>2</sup>Laboratorio de Investigación y formación en Informática Avanzada  
Universidad Nacional de La Plata  
Calle 50 esq.115 1er.Piso  
La Plata, Argentina  
cpons@lifa.info.unlp.edu.ar  
Facultad de Tecnología Informática  
Universidad Abierta Interamericana  
Ciudad Autónoma de Buenos Aires, Argentina  
claudia.pons@uai.edu.ar

<sup>3</sup>Departamento de Ingeniería e Investigaciones Tecnológicas  
Universidad Nacional de la Matanza  
Buenos Aires, Argentina  
rrodriguez@ing.unlam.edu.ar

**Abstract.** En el trabajo se realiza la validación de las dos primeras etapas correspondientes a la metodología que permite analizar, diseñar e implementar soluciones de gobierno electrónico en la gestión pública, a través de servicios computacionales (metodo\_SCGE). En esta metodología se considera el software, hardware y comunicaciones, el análisis de estos componentes permite generar un modelo integral basado en SCGE. El artículo contiene una presentación de algunas metodologías existentes que se pueden aplicar a servicios computacionales, una breve introducción al concepto de servicios computacionales en el ámbito de gobierno electrónico y una descripción de sus componentes. Al final se presenta la validación de las etapas, sobre un servicio específico perteneciente a un organismo gubernamental.

**Keywords:** Servicios computacionales, Metodologías de e\_gov, Aseguramiento de la calidad del Software, Servicios al ciudadano

### 1 Introducción

La tarea que implica formalizar un proceso de gestión de las TIC, en forma global, es demasiado ambicioso; sin embargo resulta muy interesante intentar controlar y estandarizar algunos procesos, como por ejemplo la sistematización del proceso de Gobierno Electrónico (GE). Una gran cantidad de bibliografía relacionada a la gestión de GE, como las guías de los gobierno de Chile [1] intenta proponer alguna secuencia de pasos a seguir para la implementación del mismo.

El objetivo del trabajo es validar la primera y segunda etapa de la metodología metodo\_SCGE, que integra las herramientas de gestión y los servicios al ciudadano, con conceptos de teoría de las organizaciones, como así también la utilización de los aspectos relacionados a la Calidad del Software, agregando una estimación de los recursos computacionales; como procesamiento, almacenamiento y conectividad, a través de los componentes de los Servicios Computacionales de Gobierno Electrónico (SCGE).

Con el desarrollo de metodo\_SCGE, se espera obtener una metodología basada en SCGE que permita analizar, diseñar e implementar soluciones de GE. Para el desarrollo de metodo\_SCGE se tuvieron en cuenta otras metodologías de análisis y diseño de software existentes, que van desde modelos de desarrollo estructurado con metodologías orientadas al flujo de datos, hasta modelos de desarrollo orientado a objetos como el Lenguaje Unificado de Modelado (UML)[2] o más recientemente las Arquitecturas orientadas a Servicios (SOA)[3], la metodología desarrollada propone integrar en un SCGE, tanto los aspectos relacionados al software, como a la estimación de los recursos computacionales.

## 2 Algunas metodologías que resultan aplicables a servicios computacionales

En este apartado se detallan una serie de metodologías que se tuvieron en cuenta al momento de diseñar metodo\_SCGE:

- a. El “Mapa para el desarrollo del GE en los países en desarrollo” [4], realizado por el Pacific Council of International Policy, establece una serie de preguntas fundamentales que los gobiernos deben hacerse antes de encarar algún proyecto de GE.
- d. En relación a la estrategia de un proyecto de gobierno electrónico basado en el servicio, los autores Affisco y Soliman [5] proponen en su trabajo las siguientes etapas: segmentación del mercado, diferenciación de un producto o de un servicio, modelo de servicio para el proyecto gobierno electrónico, generación de valor, rediseño de operaciones, alineación del sistema y promoción de buenas prácticas, entrega de servicios y productos.
- f. En la Universidad de Ciencia y Tecnología de Kaohsiung en Taiwan[6], los investigadores del área desarrollaron un trabajo en el cual los autores invitaron a todas las agencias del sector público involucradas en el estudio, a integrar sus procesos de servicio en un modelo de ventanilla única y establecer un mecanismo colaborativo para proporcionar un servicio personalizado a los ciudadanos.
- g. El proyecto Access-eGov ha desarrollado herramientas de software que permiten la integración de servicios utilizando tecnologías semánticas [7].
- h. En el trabajo “Integración de Modelos de Negocios Electrónicos para Soluciones Gubernamentales: Metodologías y Procesos Orientados a Servicios centrados en el Ciudadano”, los autores Chhabra y Kumar [8] proporcionan una colección única de herramientas esenciales y estrategias alternativas de comercio electrónico que se pueden incorporar a las soluciones de GE.
- k. Dzhusupova, Shareef y Janowski[9] proponen una metodología para la Evaluación de la Preparación para GE que tiene como objetivo abordar los requisitos de información para la planificación estratégica mediante la evaluación de la demanda y las capacidades del gobierno electrónico en contextos nacionales e internacionales, así como el entorno propicio y la infraestructura de las TIC.

Además existen tecnologías como la Arquitectura Orientada a Servicios (SOA)[3] y la Arquitectura Dirigida por Modelos (MDA)[10], entre otras, que realizan aportes significativos en el desarrollo de software. Existen aspectos que podrían complementarse a las citadas metodologías como son la estimación de los recursos computacionales y por otra parte se hace necesaria una metodología específica para el diseño de soluciones de gobierno electrónico.

Teniendo en cuenta lo expuesto anteriormente resulta necesario establecer una metodología al momento de decidir incorporar una TIC y en especial GE, asegurando que el proceso tendrá en cuenta aspectos relacionados no sólo a la tecnología informática, sino también a la teoría de las organizaciones, la gestión pública, como así también a la utilización de los conceptos relacionados a la Calidad y aspectos sociales; como por ejemplo la participación ciudadana.

### 3. metodo\_SCGE: una metodología basada en SCGE

La metodología metodo\_SCGE se encuentra basada en servicios computacionales, por lo tanto se definirá lo que es un servicio computacional de gobierno electrónico (SCGE) en el marco conceptual de este trabajo. Un SCGE es una unidad de servicio que se pone a disposición de una solución de gobierno electrónico. Una solución de gobierno electrónico puede estar compuesta por varios SCGE y estos a su vez conformar varias soluciones de gobierno electrónico. A continuación se describirán los tres componentes que conforman a un SCGE.

#### 3.1 Componentes

Un SCGE está compuesto por:

- Una parte dirigida al análisis, diseño, prueba e implementación del software de gobierno electrónico; denominada Componente Software. Este componente puede poseer una o más funcionalidades específicas del dominio de gobierno electrónico, también contiene la descripción del software de base utilizado, el lenguaje o herramienta de desarrollo utilizada, el sistema Gestor de Base de Datos y la arquitectura de software implementada.
- Un elemento relacionado con la estimación de los recursos computacionales necesarios para soportar el software de gobierno electrónico, denominado Componente Hardware y Conectividad. Este componente se refiere a los recursos de hardware y conectividad asociados a la solución diseñada. Aunque este componente puede considerarse como uno solo, el mismo se dividió en dos, lo que permite separar los recursos necesarios para efectuar el procesamiento y almacenamiento de datos por una parte y la transferencia de datos a través de medios de comunicación por otra.

En consecuencia un SCGE podría graficarse de la siguiente manera (Ver Figura 1):

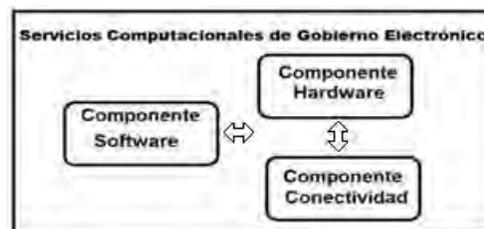


Figura 1. Componentes de un Servicio computacional de Gobierno Electrónico

#### 3.1.1 Componente Software

Este componente es de suma importancia debido a que contiene las funcionalidades concretamente definidas para el SCGE. Estas funcionalidades se encuentran directamente relacionadas con las especificadas en el trabajo “Análisis de las propiedades y atributos propios de sitios de gobierno electrónico”[11] y representan el núcleo del SCGE. También en este elemento se deben especificar el Software de base que utiliza el SCGE, es decir el sistema operativo nativo sobre el cual se encuentra soportado el servicio, el lenguaje de programación utilizado para generar el código fuente o generado por un entorno integrado de desarrollo (IDE), o una herramienta de ayuda para el desarrollo de aplicaciones (CASE) utilizada para el desarrollo del servicio. Conjuntamente se deberá determinar, el sistema gestor de Base de Datos que se utilizará para soportar el almacenamiento, modificación y extracción de la información en la base de datos elegida, además de proporcionar herramientas para añadir, borrar, modificar y analizar los datos contenidos en la misma. Por último se deberá describir la

Arquitectura de Software que soportará el SCGE, es decir si se tratará de una arquitectura host-terminal, cliente servidor o multicapa. En la Figura 2 puede observarse la estructura del Componente Software.

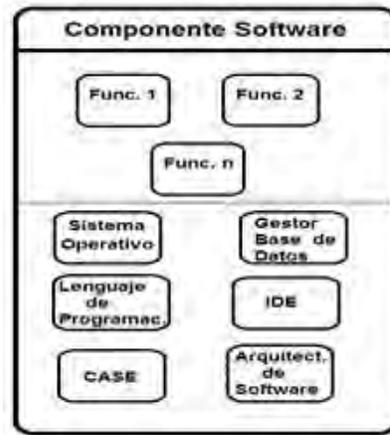


Figura 2. Estructura del Componente Software

### 3.1.2 Componente Hardware

Aquí se deberá definir el equipamiento que soportará el SCGE diseñado. En primer lugar resulta indispensable consignar la unidad central de procesamiento (CPU) que se utilizará, como así también el tipo y la capacidad de la memoria RAM requerida, el tipo y tamaño del soporte de almacenamiento magnético que resulte necesario. Por último es necesario establecer el tipo y características de las interfaces que se usarán. Para poder definir convenientemente este componente resultará adecuado realizar una estimación de los recursos computacionales necesarios, tales como tipo y capacidad de procesamiento y almacenamiento. En la Figura 3 se puede visualizar la estructura del componente descrito.

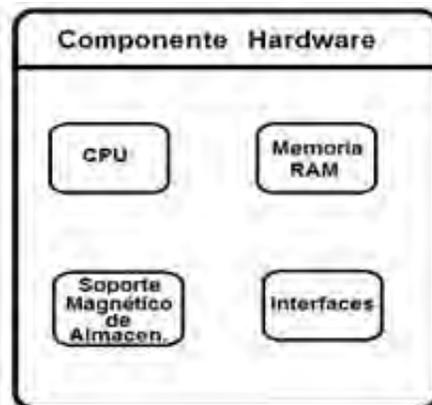


Figura 3. Estructura del Componente Hardware

### 3.1.3 Componente Conectividad

En este componente se define la manera en que el SCGE se encuentra conectado con otros sistemas y servicios computacionales, se deberá definir el tipo de enlace físico a utilizar, la velocidad de conexión, el protocolo de comunicación y el tipo de red. En la figura 4 se puede observar la estructura del componente de conectividad.



Figura 4. Estructura del Componente Conectividad

#### 3.3.1 Características de la metodología basada en SCGE (metodo\_SCGE)

La metodología metodo\_SCGE, se trata de un método iterativo e incremental de desarrollo basado en SCGE. Al igual que las metodologías ágiles tradicionales para el desarrollo de software, como Scrum y XP, metodo\_SCGE realiza incrementos e iteraciones para agregar funcionalidades y características de software, hardware y conectividad para el SCGE que se está desarrollando.

La metodología propone desarrollar un SCGE no como una sola unidad, sino como una serie de incrementos y cada una de ellos incluye una nueva característica de software, hardware y conectividad para el SCGE. En metodo\_SCGE, los procesos de especificación, diseño e implementación están entrelazados, y el SCGE se desarrolla en diferentes versiones. Al igual que las metodologías ágiles metodo\_SCGE, considera al diseño y la implementación como actividades centrales, incorporando a éstas, otras actividades como la adquisición de requerimientos y las pruebas. La iteración ocurre a través de las actividades. En consecuencia la adquisición de requerimientos y el diseño se desarrollan en conjunto, no por separado.

#### 3.3.2 Etapas de metodo\_SCGE

La metodología metodo\_SCGE posee cuatro etapas, las cuales se describen a continuación:

1. Seleccionar características de Componentes SCGE de la versión a liberar: en esta etapa se seleccionarán las características de software, hardware y conectividad que se desean incorporar en la nueva versión que se liberará.
2. Desglosar características de cada componente: en esta etapa se deben determinar claramente las características que se desean desarrollar para cada componente.
3. Desarrollar, utilizar, integrar y probar cada característica por componente: aquí se realizará el desarrollo, la integración y las pruebas de cada característica correspondiente a cada uno de los componentes de un SCGE.

4. Liberar y evaluar versión del SCGE: durante esta etapa se procederá a la liberación y evaluación de la versión desarrollada para el SCGE.

En la figura 5 se puede observar un gráfico de las etapas de metodo\_SCGE.

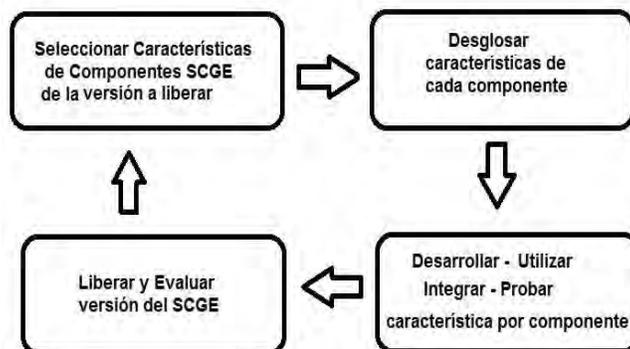


Figura 5. Etapas de metodo\_SCGE

### 3.3.3 Actividades de metodo\_SCGE

En la tabla 1, se detallan las actividades a llevar a cabo en cada una de las etapas de la metodología propuesta:

Tabla1. Actividades por etapas de metodo\_SCGE

Etapa	Actividad/Componente	Etapa	Actividad/Componente
I- Seleccionar características de Componentes SCGE de la versión a liberar	I.1 Seleccionar característica de Software (CS)	III-Desarrollar, utilizar, integrar y probar cada característica por componente	III.1- Desarrollar, utilizar, integrar y probar CS
	I.2-Seleccionar característica de Hardware (CH)		III.2-Desarrollar, utilizar, integrar y probar CH
	I.3-Seleccionar característica de Conectividad (CC)		III.3-Desarrollar, utilizar, integrar y probar CC
II-Desglosar características de cada componente	II.1-Desglosar característica CS	IV-Liberar y evaluar versión del SCGE	IV.1-Implementar y evaluar la versión liberada del SCGE, para los tres componentes
	II.2-Desglosar característica CH		
	II.3-Desglosar característica CC		

## 4. Validación de la metodología propuesta aplicándola a un caso de estudio en un organismo gubernamental.

### 4.1 Introducción

La metodología metodo\_SCGE será validada a través de su aplicación en la Dirección Provincial de Rentas de la Provincia de Jujuy, la cual depende de la Secretaría de Ingresos Públicos perteneciente al Ministerio de Hacienda.

Para la aplicación de metodo\_SCGE se utilizarán algunos de los servicios a brindar al ciudadano y que se encuentra detallados en el Árbol de requerimientos consignados en el trabajo “Análisis de las propiedades y atributos propios de sitios de gobierno electrónico”[11] para la característica “Funcionalidad” de los sitios de GE, y específicamente para la sub-característica “Funcionalidad específica del dominio y contenido”, atributo “Poder Ejecutivo” y sub-atributo “Servicios públicos, finanzas y producción”.

## 4.2 – Desarrollo de la etapa I

Teniendo en cuenta lo consignado en la Tabla 1, a continuación se desarrollarán las actividades previstas para la etapa I.

### **Etapa I: Seleccionar características de Componentes SCGE de la versión a liberar**

#### I.1. Actividad: Seleccionar característica CS

##### I.1.1- Seleccionar y priorizar funcionalidades específicas del dominio:

- Declaración Jurada de Impuestos • Pago de impuestos y tarifas en línea
- Denuncias de evasión Tributaria • Consultar Informe de Deuda Tributaria

##### I.1.2- Seleccionar una herramienta Case y/o entorno IDE

- El entorno de desarrollo integrado (IDE) seleccionado es Eclipse SDK 4.15 (64-bit)

##### I.1.3- Seleccionar un lenguaje de programación en que se generará el código fuente del SCGE

- El lenguaje de programación seleccionado es Java EE versión 8.0.

##### I.1.4- Seleccionar una Arquitectura de Software a utilizar

- La arquitectura seleccionada es una orientada a servicios (SOAP).

##### I.1.5- Seleccionar un Gestor de Base de Datos

- El gestor de Base de Datos seleccionado es Oracle versión 19c

#### I.2- Seleccionar característica Hardware CH

##### I.2.1- Seleccionar una Unidad Central de Procesamiento

- CPU seleccionada Intel® Core™ i7 de 10ma Generación

##### I.2.2- Seleccionar un Tipo y capacidad de memoria RAM

- Memoria RAM 16GB, 2666MHz

##### I.2.3- Seleccionar un tipo y una capacidad de almacenamiento magnético

- Disco Rígido interno, 1TB, Sata III.

##### I.2.4- Seleccionar un tipo de interfaz

- Pantalla táctil LED 18.5 “, TFT, 1366 px x 768 px

#### I.3- Seleccionar característica Conectividad CC

##### I.3.1- Seleccionar un tipo de enlace físico

- Inalámbrico

##### I.3.2- Seleccionar una velocidad de conexión

- 1000Mbs o superior

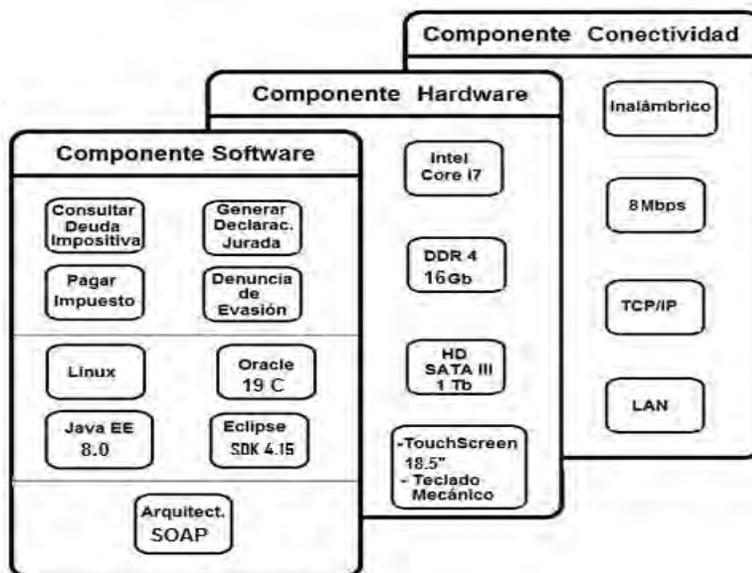
##### I.3.3- Seleccionar un protocolo de comunicación

- Protocolo TCP/IP

##### I.3.4- Seleccionar un tipo de Red

- Red de área local (LAN).

En la figura 6 se puede visualizar el detalle de los ítems seleccionados para los componentes de software, hardware y conectividad.



**Figura 6.** Valores seleccionados para los componentes de software, hardware y conectividad en la etapa I

### 4.3 – Desarrollo de la etapa II

Teniendo en cuenta lo consignado en la Tabla 1, a continuación se desarrollarán las actividades previstas para la etapa II.

#### **Etapa II: Desglosar características de cada componente**

##### II.1. Desglosar característica de Software (CS)

##### II.1.1- Especificar funcionalidades del SCGE:

- **Declaración Jurada de Impuestos:** Se debe consignar información de los plazos para la presentación de declaraciones de impuestos. Además de deberá permitir descargar los formularios a presentar. El objetivo final será efectuar la declaración jurada de impuestos totalmente en línea.
- **Pago de impuestos y tarifas en línea:** Esta funcionalidad es de suma importancia ya que es un servicio indispensable para el ciudadano o empresa y debe permitir realizar la transacción de pago, totalmente en tiempo real; además requiere complementarse con técnicas de pago electrónico a través de diferentes medios.
- **Denuncias de evasión Tributaria:** Esta funcionalidad permite registrar las denuncias realizadas por los ciudadanos a través de correo electrónico y todos los canales de comunicación que posea el organismo gubernamental.
- **Consultar Informe de Deuda Tributaria:** esta funcionalidad deberá permitir realizar consulta en línea e impresión del detalle de la deuda de un contribuyente.

##### II.1.2- Especificar características de la herramienta Case y/o entorno IDE, seleccionadas.

El entorno de desarrollo integrado (IDE) Eclipse SDK 4.15 (64-bit) cuenta con las siguientes características principales:

- Desarrollador: Fundación Eclipse
- Fecha de Lanzamiento: 07/11/2001
- Género: IDE, Java SDK, C/C++
- Sistema operativo: Multiplataforma
- Licencia: Publica de Eclipse
- Idioma: Español
- Autor: Object Technology International
- Última versión estable: 17/06/2020 (Versión 4.16)
- Lenguaje nativo: Java
- Plataforma: Máquina virtual Java

### II.1.3- Especificar características del lenguaje de programación seleccionado

El lenguaje de programación seleccionado Java EE versión 8.0, posee algunas de las características que se mencionan a continuación:

- Versión: Java™ SE Development Kit 8, Update 251 (JDK 8u251)
- Fecha actualización: 14/04/2020 -Full versión string: 1.8.0\_251-b08
- Número de versión: 8u251. -Actualización JDK:8u251 implement JSR 337

### II.1.4- Especificar características de la Arquitectura de Software seleccionada

La arquitectura orientada a servicios SOA es la que mejor se adapta al caso de estudio elegido, debido a que existen algunas partes de software existente que puede ser utilizado. Al tratarse de aplicaciones que se encuentran en un entorno y dominio específico, en el que las operaciones no sufrirán cambios significativos, la implementación de los web services basados en SOAP puede resultar muy conveniente.

### II.1.5- Especificar características del Gestor de Base de Datos seleccionado

El gestor de Base de Datos seleccionado Oracle posee las siguientes características:

- Versión: Oracle Database 19c - Fecha actualización: 10/01/2019
- Sistema operativo: Multiplataforma - Desarrollador: Oracle Corporation
- Cantidad de licencias a adquirir: 300

## II.2- Desglosar característica de Hardware (CH)

### II.2.1- Especificar características de la Unidad Central de Procesamiento

La CPU seleccionada Intel® Core™ i7 de 10ma Generación posee las siguientes características:

- Procesador: Intel® Core™ i7-10700K - Cantidad de núcleos: 8
- Frecuencia básica del procesador: 3.80 GHz - Frecuencia turbo máxima: 5.10 GHz
- Tipos de memoria soportada: DDR4 - Caché: 16 Mb Intel Smart
- Subprocesos: 16 - Gráficos del procesador: Gráficos HD Intel® 630

### II.2.2- Especificar características de la memoria RAM

La memoria seleccionada posee las siguientes características:

- Tipo de memoria interna: DDR4 - Capacidad: 16 GB
- Diseño de memoria: 1 x 16 GB - Velocidad de memoria del reloj: 2666 MHz
- Factor de Forma: 288-pin DIMM

### II.2.3- Especificar características del almacenamiento magnético

Las características del Disco Rígido seleccionado son:

- Disco duro interno - Capacidad de almacenamiento: 1 TB.
- Caché 64 MB. - Interfaz: SATA III
- Velocidad de transferencia interna: 210 Mbps. - Consumo: 5,3 W.
- Dimensiones: 146,99 x 101,6 x 20,17 mm. - Peso aproximado: 400 g.

### II.2.4- Especificar características de las interfaces

Las características de la Pantalla táctil seleccionada son:

- Tipo de panel: Aluminio de magnesio. -Screen: LCD de 18,5", retroilum. por LED.
- Resolución: 1366\*768. - Procesadores: Intel Baytrail J1900 2.0Ghz.
- Tacto proyectivo: capacitivo multipunto. - Grado de protección IP65.
- soporte de montaje: VESA 75/100. - Entrada de 9~24V DC.
- I/O: 4x USB, 2x GLAN, 2x COM, 1x HDMI, Audio.

## II.3- Desglosar característica de Conectividad (CC)

### II.3.1- Especificar características del enlace físico

Las características del enlace físico seleccionado son:

- Tecnología: inalámbrica - Medio: ondas de radiofrecuencia de baja potencia y banda específica
- Uso: privado.

### II.3.2- Especificar características de la velocidad de conexión

Para lograr una buena velocidad se utilizarán las siguientes características:

- Frecuencia: 2,4 GHz
- Velocidad máxima teórica: 11 Mbps
- Estándar inalámbrico: IEEE 802.11b

### II.3.3- Especificar características del protocolo de comunicación

El Protocolo TCP/IP seleccionado es el que mejor se adapta a las velocidades y configuración de la red utilizada.

### II.3.4- Especificar características del tipo de red

La LAN utilizará estructura cableada existente y se adicionarán tres puntos de acceso (AP) Ubiquiti Dual Band 802.11ac 2.4/5.8ghz, para conformar una configuración inalámbrica que soporte el diseño especificado.

## 3 Conclusiones

A pesar que el caso de estudio elegido es de complejidad media, se puede concluir que la metodología para las etapas I y II, se podrá utilizar sin ninguna dificultad en casos de complejidad alta, tanto en aspectos relacionados a la funcionalidad específica del dominio, como así también en los aspectos relacionados a la estimación de los recursos computacionales. En este sentido y como trabajo futuro, se verificarán herramientas que permitan realizar una adecuada estimación de los recursos computacionales requeridos por la aplicación a desarrollar. Por lo expuesto precedentemente, se puede concluir que se cumplió con el objetivo propuesto en el presente trabajo.

El proceso de validación utilizado, permitió demostrar que la metodología metodo\_SCGE resulta de fácil aplicación y sumamente intuitiva para su aplicación en un caso de estudio.

Además se pudo visualizar que las actividades previstas en la etapa I y II, requieren de un conocimiento relativamente simple al momento de especificar los aspectos relacionados tanto al software como al hardware y la conectividad.

El producto final que se obtuvo es un conjunto de especificaciones referidas a los servicios que se deben implementar, los procesos, el equipamiento y la conectividad necesaria para una unidad de organización gubernamental.

## 4 Referencias

1. Gobierno de Chile: Guía metodológica, Sistema de Gobierno Electrónico Programa de Mejoramiento de la Gestión, Metas de Eficiencia Institucional. Disponible en [http://www.dipres.gob.cl/594/articles-51683\\_egov\\_guia.pdf](http://www.dipres.gob.cl/594/articles-51683_egov_guia.pdf) (2009)
2. Booch, G., Rumbaugh, J., Jacobson, I.: El Lenguaje Unificado de Modelado. Addison Wesley (2004).
3. Erl, T.: SOA Principles of Service Design. Prentice Hall (2008)
4. Pacific Council of International Policy: Roadmap for e-government in the developing world. Disponible en <http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan005030.pdf> (2002)
5. Affisco, J., Soliman, K.: E-government: a strategic operations management framework for service delivery. Business Process Management Journal. Vol.12 Issue:1. <https://doi.org/10.1108/14637150610643724> (2006) 13-21
6. Hsin-Pin, F., Tien-Hsiang, C., Pei, C., Chyou-Huey, C.: A collaborative model for service provision by multiple public-sector agencies. Internet Research, Vol. 16 Iss: 4. Disponible en <http://www.emeraldinsight.com/doi/abs/10.1108/10662240610690007> (2006) 365 – 379
7. Hreño, J., Bednár, P., Furdik, K., Sabol, T.: Integration of Government Services using Semantic Technologies. Journal of theoretical and applied electronic commerce research, Res. vol.6 no.1. Disponible en [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-18762011000100010](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-18762011000100010) (2011)
8. Chhabra, S., Kumar, M.: Integrating E-Business Models for Government Solutions: Citizen-Centric Service Oriented Methodologies and Processes., Advisory Board (2009)
9. Dzhupova, Z., Shareef, M., Ojo, A., Janowski, T.: Methodology for e-government readiness assessment-models, instruments, implementation, - Proceedings of the International Conference on Society and Information Technologies (2010) 6-9
10. Mellor, S., Scott, K., Uhl, A., Weise, D.: MDA Distilled, Principles of Model Driven Architecture, Addison-Wesley. Professional (2004)
11. Castro, M.: Análisis de las propiedades y atributos propios de sitios de gobierno electrónico. Tesis de Maestría en Ingeniería de Software, Universidad Nacional de San Luis (2009)

## Improving a Low Cost Surveillance System

Carlos Sebastián Castañeda, María José Abásolo<sup>12</sup>

<sup>1</sup> III-LIDI, Facultad de Informática, Universidad Nacional de La Plata, Argentina

<sup>2</sup> CICPBA, Buenos Aires,, Argentina

mjabasolo@lidi.info.unlp.edu.ar

**Abstract.** The purpose of this work is improving the functionality and usability of a low cost commercial surveillance system. The original system provides simple motion detection and sends alert messages by means of FTP or email. The modified system adds a software layer to the original system for implementing desirable image processing features. Particularly, people detection functionality was implemented by means of OpenCV Oriented Gradient Histograms. Also the modified system adds the use of Telegram messaging service for sending alerts. We compare the performance of both original and modified system regarding the intruder detection. Also the usability of both proposed and original system was tested by means of System Usability Scale (SUS).

**Palabras clave:** computer vision, motion detection, people detection, video surveillance systems, computer vision

### 1 Introduction

Monitoring and control of different areas, stories such as public places, banks, shopping malls, shops, airports has become a growing need in these times. With technological advances in the area of security and the availability of low-cost hardware, the use of security systems has become something of regular use in homes and small businesses. In general, the installation of security cameras does not imply a great economic cost, but the hiring of people who can monitor these cameras 24 hours a day does. One of the main goals is having automated surveillance systems that require little human interaction for their operation and detection of potential threats.

It is possible to acquire fixed security cameras at low cost, but a problem that these devices can present is that the embedded motion detection algorithm [1] is very sensitive to changes in the environment, such as changes in lighting conditions and shadows projections. In general, cameras implement motion detection in base on background subtraction techniques [2-7] and difference between frames [8-10].

In the background subtraction technique, the difference between the current frame and the frame representing the image background is used to detect moving objects. In this method, it is necessary to maintain a model that represents the background of the image. A background model could be obtained, for example, from the average image over a certain training period. Moving objects are easy to detect using this technique,

but the background model must be updated regularly. In its simplest version, this technique is simple to implement but has the disadvantage of being very sensitive to changes in the environment, making it difficult to isolate the interference of the real movement of objects in the image. Bottom subtraction is the most commonly used technique in fixed security cameras. The statistical model based on background subtraction is flexible and fast but the camera must be stationary.

In the algorithms based on difference between frames, the presence of moving objects is determined through the difference between consecutive frames of the image. The absolute difference in each pixel is calculated, between two to three consecutive frames of the video sequence, and a threshold is applied to obtain the moving objects. This method is highly adaptive to dynamic and computationally less complex environments, but is generally not accurate in extracting the full form of certain types of moving objects. In particular, the main causes of incorrect detections are due to ghost images, close-up apertures with objects that have a homogeneous color

On the other hand, there are different techniques or algorithms in the bibliography to detect people in images or video streams. Viola et al. in [11] he proposes to use cascade classifiers, a method widely used in real-time applications. This method has been extended to employ different types of characteristics and techniques, but fundamentally the cascade concept has been used to achieve real-time detection. One of the most popular feature used for human detection are Oriented Gradient Histograms (OGH), developed by Dalal et al. [12].

This work aims to improve a low-cost camera-based surveillance system that implements motion detection based on background subtraction. It is proposed to improve the user interface and add post-processing of images related to alerts, in particular to detect people, being able to filter false positives or increase the level of the alert sent to the user.

The rest of the article is organized as follows: in section 2 the original system tests are described. In section 3 the proposed system for improving the original one is explained. Section 4 shows the tests carried out to evaluate the intruder detection alerts that were incorporated. Finally, section 5 presents the conclusions and future work.

## **2 Low cost commercial surveillance system**

### **2.1 System description**

We analyze a TP-LINK model NC220 camera<sup>1</sup> which has the following features: low cost, wireless connectivity, Wi-Fi signal amplifier, night vision, motion and sound detection that can be enabled and disabled, sending alert notifications by email or FTP, application for live video streaming, for Android, IOS, Windows platforms or through the TP-LINK cloud platform<sup>2</sup>.

---

<sup>1</sup> [http://www.tp-link.es/products/details/cat-19\\_NC220.html](http://www.tp-link.es/products/details/cat-19_NC220.html)

<sup>2</sup> <https://www.tplinkcloud.com/>

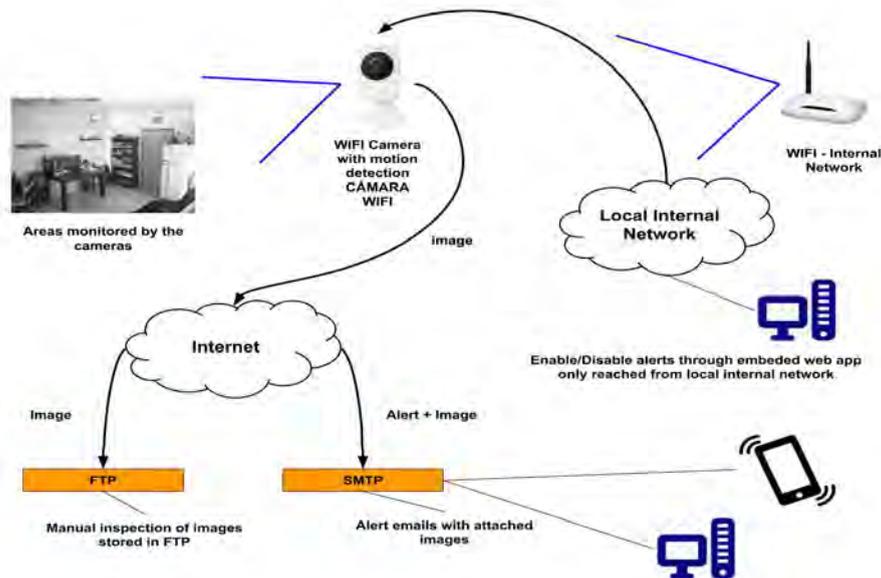


Figure 1. Commercial surveillance system with TP-LINK nc220 cameras (own source)

The camera's embedded motion detection system can be configured with three levels of sensitivity. The execution of motion detection algorithms embedded in cameras and digital video recordings depends on the preselected level of sensitivity which works for all possible environmental conditions. For this reason these systems have an error rate that includes false alarms (false positives) and the omission of true alarms (false negatives). A low sensitivity level can cause more false negatives, while a high sensitivity level can cause more false positives. It is also possible to configure the camera vision in day, night and automatic mode. In daylight mode, the camera shoots in colors, but good lighting conditions are necessary. In night mode the camera shoots in gray scale and is capable of recording video with very little or directly without light in the environment. In automatic mode the camera switches between day and night modes depending on the amount of light in the environment in every instant. In an environment where lighting conditions are not constant, the best option will be to configure the camera in automatic mode.

Figure 1 shows the operating scheme of the original system. The camera is connected to a Wi-Fi internal network through which an integrated application is accessed allows enabling/disabling alerts. Motion detection alerts are sent to an email or to an FTP server which can be accessed directly to inspect the corresponding images.

## 2.2 Test description

For the tests a camera was installed in an apartment, directing the lens to the main entrance. The tests were carried out for 185 days, collecting the images sent by the camera when the department was empty and the system activated, and also forcing the

activation of motion detection by entering the department without disabling.. During this period 1683 images were collected. For each motion detection event the camera sends a sequence of several consecutive screenshots. Each sequence was considered as a single sample, thus reducing the 1683 images to 309 sequences or samples.

### 2.3 Motion detection

The first test carried out consisted of counting the number of true and false positives registered by the camera. We consider positive samples to those sequences where people are observed, or there was a real movement of some object (for example: the opening of the apartment door without the entrance of any people), or also there was a sudden change in luminosity due to turning the light on. In the last type of sequences the images are completely burned by light. The negative samples are those where there was no real movement in the scene, and the alerts were caused by light changes in the environment light, brightness and reflections of the objects in the monitored scene. The result of the classification of the sequences are shown in Table 1. The camera reported 35.60% false positives, with the remaining 64.40% reporting true positives, where 88.44% of them correspond to people detection.

Table 1: Summary of the data obtained in the evaluation test of the embedded movement detection algorithm in the camera

<b>Total number of sequences</b>	<b>309</b>	<b>100%</b>
<b>Positive sequences (true positives)</b>	<b>199</b>	<b>64.40%</b>
People	176	88.44%
real object motion (door open)	11	5.52%
high light change (light were turned on manually)	12	6.03%
<b>Negative sequences (false positives)</b>	<b>110</b>	<b>35.60%</b>
day (natural light change)	86	78.18%
night (object reflection or shine)	14	12.72%
night (alert reason not clear in the image)	10	9.09%

## 3 Proposed system

The objective is to improve the original commercial system described regarding to the following:

- Adding the processing of the images resulting from the motion detection provided by the camera to emphasize alerts;
- Replacing the email based notification system by a messaging notification system.

The new software architecture adds a web application to the original system, and can be seen in figure 2. The original system sends to an FTP server the corresponding image when a motion detection event occurs. The images are processed with an image processing module. Particularly we implement an intruder detection algorithm to confirm the presence of people. The alert is then classified according the results of the

processing (false or true positive), and the image is sent jointly with a text notification to the instant messaging program, particularly Telegram. These alerts can be read from almost any current mobile device. From the instant messaging program it is also possible to send a command to enable/disable the reception of notifications.

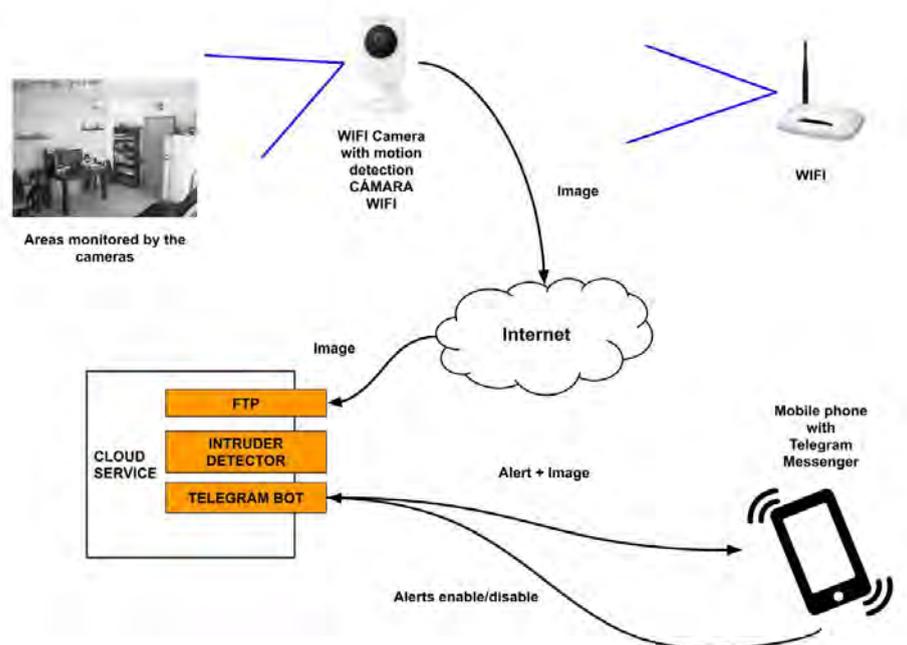


Figure 2. Proposed architecture system that interacts with the TP-LINK nc220 security camera

### 3.1 Web Application

All the modules of the system are contained in a web application implemented with DJANGO<sup>3</sup> which is a framework for the development of web applications in Python, which provides a large number of utilities for developing in an organized and agile way. The implemented Django project consists of four modules:

- Django Admin: this utility is provided by the framework, and makes it possible with a minimum code configuration to register the models defined in the different applications of the project and generate an Administrative Web Application that allows the basic ABM of the entities to be carried out.
- FTP server module
- Image processing module
- Notifications module

The code of our implementation can be downloaded from GitHub<sup>4</sup>.

<sup>3</sup> <https://www.djangoproject.com>

<sup>4</sup> <https://github.com/seba3c/scamera>

### 3.2 FTP server

For our purpose the option of sending the screenshots to an FTP server was set in the original system. This module implements the FTP server that can be run through a command. It defines the model that represents the server users and interacts with the framework's authentication and authorization engine to validate incoming connections.

A mechanism is needed to trigger an event, upon reception of a new image sent by the camera. The `pyftplib` library<sup>5</sup> was used, which allows customizable implementation of an FTP server. This library is developed entirely in Python and allows redefining hook methods that represent the different significant events that can occur on an FTP server.

### 3.3 Notifications module

This module implements the attention to commands from users and sending alerts when an image is received on the FTP server. Telegram is used as the instant messaging application. It has similar features and benefits to WhatsApp but it has the advantage of being open source software that provides a set of APIs to implement, among other things what is known as chatbots<sup>6</sup>. A chatbot is basically a program that responds to users petitions. It can be implemented with different levels of complexity, either from one capable of processing text in natural language and simulating speaking with another person, to the simplest chatbot that can respond to a set of predefined commands.

The notifications module allows receiving commands from Telegram users registered in the system and sending notification after an image is processed. Notifications range from warnings to alerts according the results of the image processing (figure 3).

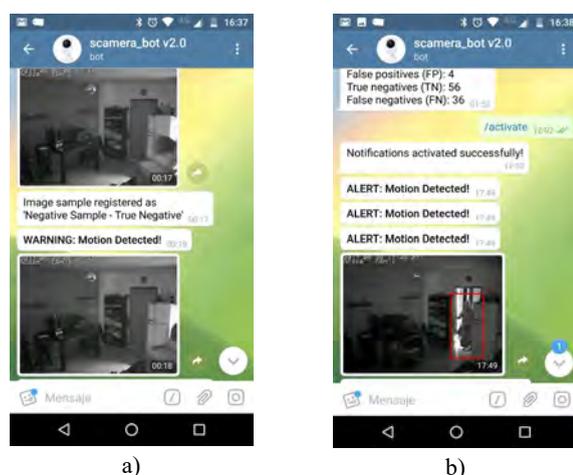


Figure 3. Notifications received in the Telegram chatbot. a) Warning notification is received when people are not detected; b) Alert notification is received when intruder was detected and it is framed in red in the image (own source)

### 3.4 Image processing module

The image processing module receives the images sent by the camera by FTP, processes each image with some desirable image processing algorithm, and enriches the information to be sent to the notification system.

The code was structured in such a way that it is possible to add new detection algorithms. Particularly, our objective was to implement an intruder detector. There are different techniques or algorithms in the bibliography to detect people in images or video streams, among them is the OGH algorithm proposed by Dalal and Triggs [12] in conjunction with Least Square Vector Machina (LSVM). In this algorithm, image descriptors are extracted and a supervised classifier is trained to determine if it contains people. There are available implementations of this algorithm in OpenCV, such as that of Rosebrock [13] using the Python programming language.

When the algorithm result is positive a message of greater relevance is sent to the notification module, together with the enlarged image in which the detected people are highlighted.

At first it was necessary to calibrate parameters of the algorithm, conducting a set of tests adjusting the value of

<sup>5</sup> <https://github.com/giampaolo/pyftplib>

<sup>6</sup> <https://core.telegram.org/bots/api>

each parameter until reaching the best balance between detection accuracy and processing speed. In this way, the parameters of the intruder detector were configured with the initial values suggested by Rosebrock [14], and the algorithm was iteratively executed on the smallest set of samples, gradually varying the values of parameters such as scale, padding, winStrinde and non-maxima suppression threshold. In each case it is observed if the accuracy and the image execution time increase or decrease. For those tests where the best relationship between accuracy and execution time was obtained, the algorithm was run again on a largest set of samples, in order to confirm the results obtained.

## 4 Intruder detection effectiveness

### 4.1 Tests and Metrics

To measure the performance of the algorithm, the metrics of a binary classifier, presented by Fawcett [15], since the algorithm has two possible outputs or categories: the presence or absence of intruders (see Table 2). The parameters of the intruder detector were configured with the initial values suggested by Rosebrock [15], and the algorithm was iteratively executed on a small set of samples, gradually varying the values for the parameters. The tests samples consists of 1289 images, which 786 are positives and 503 are negatives. Table 2 shows the values of the metrics. Accuracy value means that 76% of the total samples was correctly classified. There are 6% of false alarms (FP) and 18% of missed alarms (FN) on the total samples. An image classified as positive generates an alert notification, 87% of which corresponds to true alerts as indicated by PPV, and 13% are false alerts that should be just warnings as indicated by FDR. As indicated by sensitivity (TPR) 71% of the positive samples are correctly classified while its counterpart miss rate (FNR) indicates that 29% of the positive samples are not. An image classified as negative generates a warning notification, 65% of which corresponds to true warnings as indicated by NPV, and 35% are warnings that should be alerts. Specificity (TNR) value indicates that 83% of the negative samples are correctly classified while its counterpart FPR indicates that 17% are not.

Table 2: Metrics of the intruder classification

Metric	Formula	Value
Accuracy (ACC)	$(TP + TN) / S$	0.76
Sensitivity or True Positive Rate (TPR)	$TP / P$	0.71
Specificity or True Negative Rate (TNR)	$TN / N$	0.83
Balanced accuracy (BACC)	$(TPR + TNR) / 2$	0.77
Precision or Positive Predictive Value (PPV)	$TP / (TP + FP)$	0.87
Negative Predictive Value (NPV)	$TN / (TN + FN)$	0.65
Fall-out or False Positive Rate (FPR)	$1 - TNR$	0.17
False Discovery Rate (FDR)	$1 - PPV$	0.13
Miss rate or False Negative Rate (FNR)	$1 - TPR$	0.29
TP: True Positives or hits to detect people		559
TN: True Negatives or correct rejections		420
FP: False Positives or false alarm		83
FN: False Negatives or missed alarm		227
P = TP + FN, all positive samples		
N = FP + TN, all negative samples		
S = P + N, all samples		

#### 4.4 Real Time testing

The next test consists of activating the intrusion detection module and evaluating it in conjunction with the rest of the system. In order to measure the performance of the detection algorithm in real time, the possibility of manually indicating the result of each alert received was added to the Telegram chatbot. Every time the chatbot receives an image, a keypad is displayed with the following options:

- PS-TP (Positive Sample- True Positive)
- PS-FN (Positive Sample - False Negative)
- NS-FP (Negative Sample - False Positive)
- NS-TN (Negative Sample - True Negative)
- DISCARD: this option has been added for certain cases where the sample cannot be evaluated and it is not desirable to affect the metrics. This situation occurs, for example, when the camera sensor cannot adapt to a sudden change in light and the recorded image is completely burned, making it impossible even to visually distinguish people or objects.

Figure 4 shows four types of images that were found in this test, excluding the type of discarded images.

Two different tests were carried out using night and automatic camera mode respectively. A total of 320 images were collected, 161 images in night mode and 159 images in automatic mode. Table 3 summarizes the results obtained from both tests. The accuracy value of 0,76 is very similar to that obtained in the parameter adjustment tests. The results of both samples, with night and automatic camera mode, are very similar.



Figure 4. a) True Positive sample; b) False Positive sample; c) False Negative sample; d) True Negative sample

Table 3: Image classification and metric values obtained in the tests of night and automatic camera mode

	Night	Automatic	Total
<b>Image classification</b>			
TP	61	66	127
TN	60	56	116
FP	4	3	7
FN	36	34	70
Discarded	0	18	18
Total	161	159	320
<b>Metrics</b>			
Accuracy ACC	0,75	0,76	0,76
Sensitivity (TPR)	0,63	0,66	0,65
Specificity (TNR)	0,94	0,97	0,95
Balanced accuracy BACC	0,78	0,81	0,80
Precision (PPV)	0,94	0,96	0,95
Negative predictive value (NPV)	0,62	0,61	0,62
Fall out (FPR)	0,06	0,04	0,05
False discovery rate (FDR)	0,06	0,04	0,05
Miss rate (FNR)	0,37	0,34	0,35

There are 2% of false alarms (FP) on total alarms and 21% of the alarms were missed (FN). An image classified as positive generates an alert notification, 95% of which corresponds to true alerts (PPV) and 5% are false alerts (FDR). As indicated by sensitivity (TPR) 65% of the positive samples are correctly classified while its counterpart miss rate (FNR) indicates that 35% of the positive samples are not. An image classified as negative generates a warning notification, 62% of which corresponds to true warnings (NPV) and 38% are warnings that should be alerts. Specificity (TNR) value indicates that 95% of the negative samples are correctly classified while its counterpart FPR indicates that 5% are not.

## 4.5 Usability tests

Usability is the degree to which software can be used by a set of specific users to achieve specific objectives with effectiveness, efficiency and satisfaction within a context of specific use [16]. There are a wide variety of methods and techniques of varying complexity to evaluate the usability of a software system. Among them is the System Usability Scale (SUS) usability test [17]. SUS is a simple and fast usability test consisting of ten questions, which accept as a response a value within a scale that can range from 1 to 5. The user answers the questions after interacting with the system to later calculate the SUS score that varies between 0 and 100. This score allows obtaining a global vision of the subjective usability ratings of a system.

The usability of both the embedded camera application of the original commercial system and the Telegram based proposed system were evaluated (Table 4). The tests were performed with six users that value using a scale from 1 to 5. A SUS final score of 95 was obtained for the proposed system against 57.08 for the original system, so that users that test the system think that the proposed system improves usability of the original commercial system.

Table 4: SUS usability test carried out to evaluate the usability of both the original system and the proposed system

SUS questionnaire	Original system scores	Proposed system scores
1 I think that I would like to use this system frequently	3.5	4.83
2 I found the system unnecessarily complex	4.0	1.5
3 I thought the system was easy to use	3.0	5.0
4 I think that I would need the support of a technical person to be able to use this system	1.33	1.0
5 I found the various functions in this system were well integrated	3.33	4.83
6 I thought there was too much inconsistency in this system	3.0	1.33
7 I would imagine that most people would learn to use this system very quickly	2.16	4.33
8 I found the system very cumbersome to use	3.66	1.0
9 I felt very confident using the system	4.0	4.83
10 I needed to learn a lot of things before I could get going with this system	1.16	1.0
<b>SUS Final Score</b>	<b>57.08</b>	<b>95</b>

## 5. Conclusions and future work

In this work, we present how to improve a low-cost commercial security camera system which can be useful for monitoring homes and businesses. The characteristics of the equipment were analyzed and some problems were identified in the experience of interaction with the system, in particular: the mechanism to activate and deactivate alerts, and the effectiveness of intruder detection.

The implementation of a software system was proposed, which interacts with the commercial system, adding both a level of image processing and a Telegram based communication between the user and the system. The Python code is open source and it was shared to download.

As a level of image processing, we particularly test the detection of intruders using a people recognition algorithm based on OGH, using an OpenCV implementation. The proposed system improved the alerts reported by the camera by means of the post processing algorithm.

The results of the usability tests, despite having been done with few users, encourage thinking that the added functionality of instant messaging communication improves the usability and interactions of users with the system.

As potential future work, it is proposed to add implementations of other algorithms for image processing. Particularly, methods of detecting people based on neural networks or methods based on machine learning are being incorporated. Also it is intended to incorporate detection of other particular objects, such as vehicles such as fumigation aircrafts.

## References

1. R. S. Shirbhate, N. D. Mishra, and R. Pande, "Video Surveillance System Using Motion Detection: A Survey," *Advanced Networking and Applications*, 2012
2. Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition*, Cambridge, UK, 26 de agosto, 2004.
3. Z. Zivkovic and van der H. F., "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognition Letters*, vol. 27, no. 7, pp. 773–780, May 2006.
4. P. Kaewtrakulpong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real Time Tracking with Shadow Detection," in *Proceedings 2nd European Workshop on Advanced Video Based Surveillance Systems*, 2001.
5. Sobral and A. Vacavant, "A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos," *Computer Vision and Image Understanding*, no. 122, pp. 4–21, 2014.
6. J. Dou, Q. Qin, and Z. Tu, "Background subtraction based on circulant matrix," *J. VLSI Signal Process. Syst. Signal Image Video Technol.*, vol. 11, no. 3, pp. 1–8, 2017.
7. S. Brutzer, B. Hoferlin, and G. Heidemann, "Evaluation of background subtraction techniques for video surveillance" in *IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 1937–1944.
8. N. Singla, "Motion Detection Based on Frame Difference Method," *International Journal of Information & Computation Technology*, 2014
9. S. S. Sengar and S. Mukhopadhyay, "A novel method for moving object detection based on block based frame differencing," in *In 3rd International Conference on Recent Advances in Information Technology*, 2016, pp. 462–472.
10. M. Fei, J. Li, and H. Liu, "Visual tracking based on improved foreground detection and perceptual hashing," *Neurocomputing*, vol. 152, no. C, pp. 413–428, Marzo, 2015.
11. Viola, Jones, and Snow, "Detecting pedestrians using patterns of motion and appearance," in *Proceedings Ninth IEEE International Conference on Computer Vision*, Nice, France, 2005.
12. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, 2005.
13. A. Rosebrock, "Pedestrian Detection OpenCV", 2015. <http://www.pyimagesearch.com/2015/11/09/pedestrian-detection-opencv>
14. A. Rosebrock, "HOG detectMultiScale parameters explained," 16 de Noviembre, 2015. <http://www.pyimagesearch.com/2015/11/16/hog-detectmultiscale-parameters-explained/>.
15. T. Fawcett, "Introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
16. ISO, «Ergonomic Requirements for Office Work with Visual Display Terminals», ISO, 9241-11, Marzo, 1998.
17. Bangor, Aaron, Kortum, Philip T. and Miller, James T. (2008). "An Empirical Evaluation of the System Usability Scale". *International Journal of Human-Computer Interaction*. 24 (6): 574–594

# Desarrollo de un prototipo para estación de monitoreo de un dispositivo híbrido de termotanques solar y a gas natural instalado en la Sede Atlántica de la Universidad Nacional de Río Negro

Javier Valsecchi<sup>1</sup>, Guillermo Malpeli<sup>1</sup>, Juan Cruz Martinez Luquez<sup>1</sup>, Héctor Luis Vivas<sup>1</sup>

<sup>1</sup>Universidad Nacional de Río Negro. Laboratorio de Informática Aplicada. Sede Atlántica  
[javier.valsecchi@gmail.com](mailto:javier.valsecchi@gmail.com) {gmalpeli, jcmluquez, lvivas}@unrn.edu.ar

**Resumen.** El presente trabajo muestra el desarrollo de un prototipo para una estación de monitoreo de un dispositivo híbrido compuesto por dos termotanques, uno solar y otro a gas natural, actuando en conjunto, para el suministro de agua caliente en el comedor del Campus de la Sede Atlántica de la Universidad Nacional de Río Negro. Las variables a controlar son: la temperatura, el caudal de agua y el consumo de gas natural. Estos datos serán capturados por sensores y procesados en una Raspberry Pi como microcontrolador programable y aplicando conceptos de algoritmos concurrentes para la recolección de datos en un mismo instante de tiempo. Las funciones programadas en el microcontrolador permitirán el envío de los datos capturados al servicio Adafruit IO, por internet, como plataforma de visualización. Allí se podrán visualizar los datos capturados en tiempo real y consultar el historial almacenado de las variables.

## 1 Introducción

En la actualidad existe una gran necesidad de reducir la dependencia de los combustibles fósiles[1]. En la Argentina el combustible fósil que más se consume es el gas natural, utilizado por las industrias y por los usuarios domésticos, que asciende al 52% de la matriz energética, valor que resulta el más alto de la región. El promedio latinoamericano es de 26,2% y nuestros vecinos Brasil y Chile tienen un 11,1%[2].

Teniendo en cuenta los futuros desafíos ecológicos y económicos, es necesario plantear la reducción del consumo de combustibles fósiles a favor de las energías renovables que se obtienen a partir del sol, el viento y el agua, entre otras. Los termotanques solares poseen una tecnología que permite la utilización de radiación solar para calentar el agua almacenada, permitiendo así reducir el consumo de gas natural.

Comúnmente es posible instalarlos de dos maneras: (1) el termotanque solar en solitario, generalmente utilizado en viviendas donde no llega la red de gas natural y se torna oneroso o, en algunos casos, difícil su suministro; o (2) instalarlo en conjunto con un termotanque a gas, permitiendo reducir el consumo de gas natural. Se estima que la instalación híbrida entre un termotanque solar y un termotanque a gas, utilizado en caso de que el agua no llegue a la temperatura deseada, permite un ahorro del orden del 70% del consumo de gas utilizado en el calentamiento de agua sanitaria [3].

En otro aspecto, el concepto de Internet de las Cosas (IoT) está cada vez más presente. Desde implementaciones en la agroindustria, medio ambiente, smart homes, entre otras

áreas[4][5]. La tecnología aplicada a este concepto está en continuo desarrollo y estandarización. Cada vez son más las empresas que apuestan a esta tecnología con el objetivo de abarcar parte de este mercado en continuo crecimiento. Se estima que para el 2025, el mercado para América Latina sea de 47 mil millones de dólares[6].

El presente trabajo muestra el desarrollo de un prototipo de implementación de una estación de monitoreo de una instalación híbrida de termotanque solar y termotanque a gas aplicando el concepto de Internet de las Cosas a partir del uso de una Raspberry Pi como soporte tecnológico. El prototipo es capaz de recolectar datos de temperatura del agua en distintos puntos de la red, del caudal de agua del sistema y del tiempo de actividad del termotanque a gas. La estación de monitoreo cuenta con la posibilidad de conectarse a internet para enviar los datos recolectados a través de los sensores y permitir su visualización en tiempo real, como así también, el acceso a datos históricos de su funcionamiento para así tener una estimación del rendimiento o posibles análisis en profundidad de los datos obtenidos.

## 2 Objetivo

El objetivo del proyecto es implementar el monitoreo del rendimiento de un termotanque solar junto a un termotanque a gas en el marco del proyecto “Uso de Energías Renovables en la Universidad Nacional de Río Negro”. El equipamiento fue instalado en el comedor del campus de la Sede Atlántica de la Universidad Nacional de Río Negro[7].

En el análisis de los requisitos necesarios del proyecto se propuso un prototipo que monitoree las siguientes variables:

- Caudal: permite obtener la cantidad de agua utilizada en el sistema. Cuanta más agua se utiliza en un instante, menor será la temperatura en la que se encuentra. Este parámetro permite determinar si está bien dimensionado el equipamiento instalado en cuanto a las necesidades del establecimiento.
- Temperatura del agua entrante: permite obtener la temperatura en la que entra el agua al sistema.
- Temperatura del ambiente: temperatura general de ambiente.
- Temperatura de salida del termotanque solar: variable que determina la temperatura a la cual el termotanque solar calienta el agua.
- Temperatura de salida del termotanque a gas: marca la temperatura de salida del termotanque a gas, permitiendo comparar con la salida del termotanque solar analizando el trabajo que tuvo que realizar para llegar a esa temperatura.
- Tiempo de actividad del termotanque a gas: permite determinar cuánto tiempo estuvo encendido el termotanque a gas.

Las variables seleccionadas son fundamentales a la hora de determinar el rendimiento de la instalación del termotanque solar. Además, el prototipo contempla la posibilidad de poder agregar en un futuro otras variables que permitan recolectar más datos sobre el rendimiento.

### 3 Solución

El microcontrolador principal del presente proyecto es la plataforma electrónica Raspberry Pi 3 Model B+, la cual es una Single Board Computer (Computadora de Placa Única), está basada en hardware libre y software de código abierto, es flexible y fácil de usar. Se desarrolló con el objetivo de construir un hardware económico y libre [8]. Está constituida por un procesador Broadcom BCM2837B0, Cortex-A53 (ARMv8) 64-bit SoC, cuenta con conectividad Wifi, Gigabit Ethernet y Bluetooth. Además cuenta con un grupo de entradas/salidas lógicas llamadas GPIO que permiten conectar una amplia variedad de sensores electrónicos. También existen otras placas electrónicas modulares (conocidas como Shields), que proporcionan funcionalidades extras, permitiéndole además incorporar compatibilidad con Arduino.

La Raspberry Pi permite correr distribuciones linux modificadas especialmente para este dispositivo, ganando así la posibilidad de tener todas las características comunes de cualquier distribución actual de linux, como también la gestión completa de procesos, a partir del uso de los mecanismos de concurrencia implementados. El proyecto está programado en Python utilizando la librería de multithreading, lo que permite tener varios procesos corriendo al mismo tiempo con el fin de obtener datos y procesarlos.

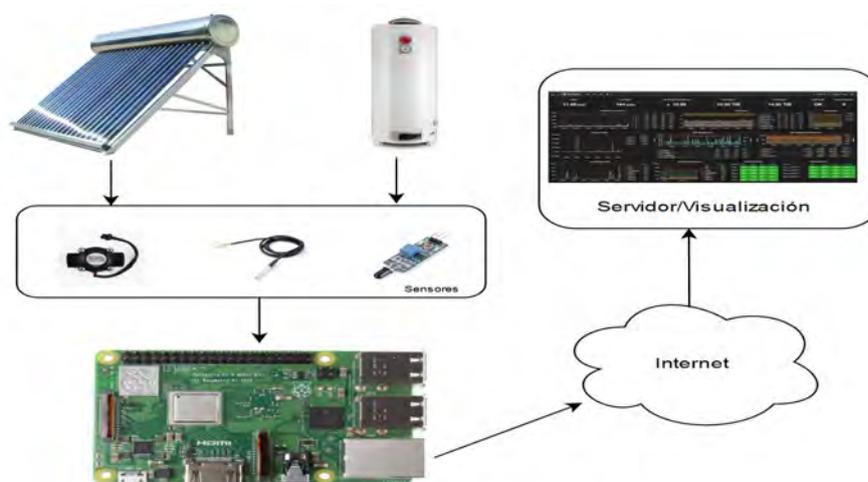
Los datos son obtenidos mediante sensores desplegados en la instalación del termotanque, utilizando sensores DS18B20 para obtener datos de temperatura del agua y ambiente; sensor Ky-026 utilizado para monitorear el encendido del termotanque a gas y el sensor FS300A para contabilizar el caudal de agua utilizado.

Para la visualización de los datos se seleccionó el servicio Adafruit IO ofrecido por la empresa Adafruit, ampliamente conocida por vender sensores, shield y distintos componentes dedicados a la automatización y desarrollos de IoT. El servicio cuenta con distintos planes comerciales, entre ellos uno gratuito y que permite crear tableros con distintos tipos de gráficos de datos, pulsadores y demás componentes necesarios para visualizar los datos enviados desde una Raspberry Pi, un Arduino u otro dispositivo.

#### 3.1 Diseño

La figura 1 muestra el esquema simplificado del diseño de la estación. Se utilizan los sensores de caudal de agua, de temperatura y de flama. Además, cuenta con conectividad a internet para poder enviar los datos recolectados al servidor, donde luego se almacenan y se visualizan.

En caso de que no sea posible enviar la información al servidor, se almacenarán los datos obtenidos de los sensores en la tarjeta SD para su posterior extracción, evitando así la pérdida de información.



**Figura 1.** Esquema de componentes y conectividad

### 3.2 Implementación

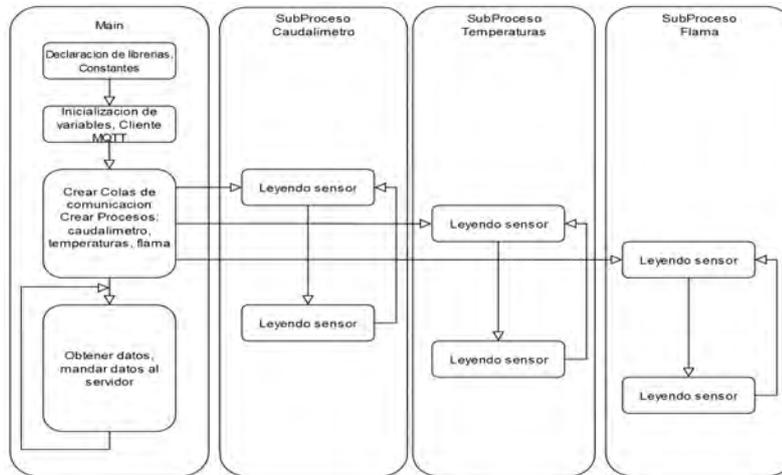
Para el desarrollo del prototipo se integraron los sensores por medio de las interfaces de conectividad que están disponibles en la Raspberry Pi, con lo cual no fue necesario adquirir ningún dispositivo adicional para conectarlos

La Raspberry Pi permite tener conexión a Internet mediante una interfaz Ethernet y una interfaz inalámbrica de doble banda (2.4 GHz y 5 GHz).

Para enviar los datos al servicio de almacenamiento y visualización se utilizó el protocolo MQTT, creado con el fin de ser utilizado para recolectar datos en una infraestructura distribuida. El protocolo se encuentra estandarizado en el documento ISO/IEC 20922 y es ampliamente utilizado para IoT[9].

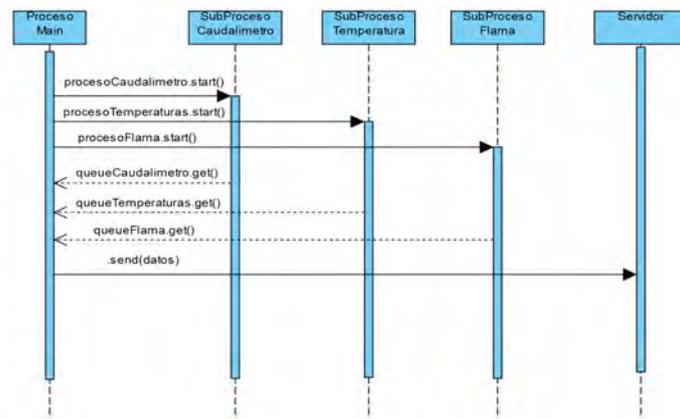
Para este desarrollo se utilizó Python, un lenguaje de programación interpretado, multiparadigma y de tipado dinámico cuya filosofía hace hincapié en una sintaxis muy limpia y un código legible.

En la figura 2 se muestra la lógica del algoritmo desarrollado para la estación de monitoreo.



**Figura 2.** Diagrama de programación

Se decidió que para la implementación del algoritmo se utiliza la librería “multiprocessing” que permite la creación de subprocesos, además del proceso main, para que realicen tareas independientes. En la figura 3 se especifica de manera gráfica la secuencia de mensajes entre procesos. Cada subproceso se encarga de manejar un tipo de sensor específico, el proceso main recolecta los datos obtenidos y los envía al servidor para almacenar los datos y posteriormente permitir visualizarlos.



**Figura 3.** Diagrama de secuencia entre los distintos procesos.

Para el servicio de almacenamiento y visualización de los datos se utilizó Adafruit IO en su versión gratuita, la cual permite crear tableros con distintos tipos de gráficos y pulsadores.

Este servicio tiene un límite de mensajes que puede recibir de los sensores y limitaciones en la cantidad de variables a mostrar en los gráficos. Teniendo en cuenta estas limitaciones, se determinó que es suficiente para probar el prototipo y la conectividad.

En la figura 4 se muestra el diseño del tablero de prueba: el gráfico superior izquierdo muestra si el termotanque a gas se puso en funcionamiento (valor 1 para encendido); el gráfico superior derecho muestra los valores de temperatura leídos en los 4 sensores, el gráfico inferior izquierdo muestra el caudal de agua en L/min; y el gráfico inferior derecho muestra los mensajes recibidos de todas las variables.

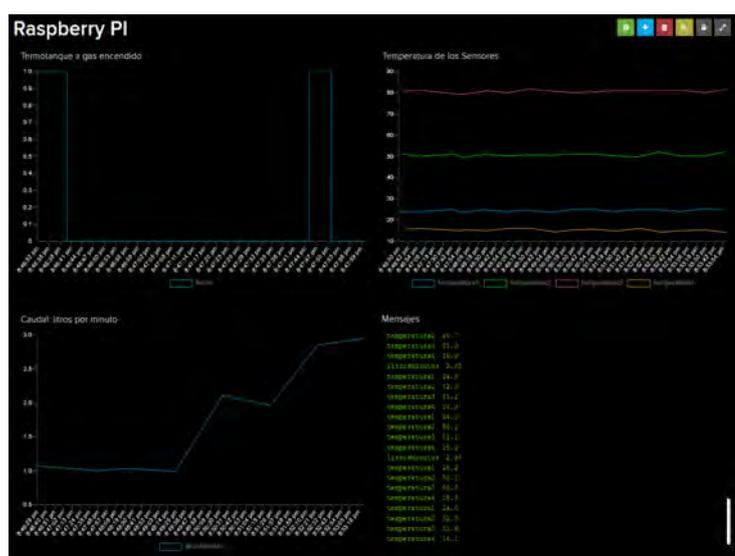


Figura 4. Dashboard con los gráficos de los datos recibidos

Se diseñó la implementación final del prototipo de la estación de monitoreo como indica la figura 5, que incluye el diagrama del sistema de agua caliente con el cual se instaló el termotanque solar, especificando dónde serán colocados los sensores utilizados:

- C1: sensor de caudal destinado a medir la cantidad de agua en L/min que pasan por el sistema.
- T1: temperatura del agua al ingresar al termotanque solar.
- T2: temperatura del agua al salir del termotanque solar e ingresar al termotanque a gas.
- T3: temperatura del agua al salir del termotanque a gas.
- T4: temperatura del ambiente cercano al termotanque solar.
- F1: sensor de flama para determinar en qué momentos se prende el mechero del termotanque a gas.

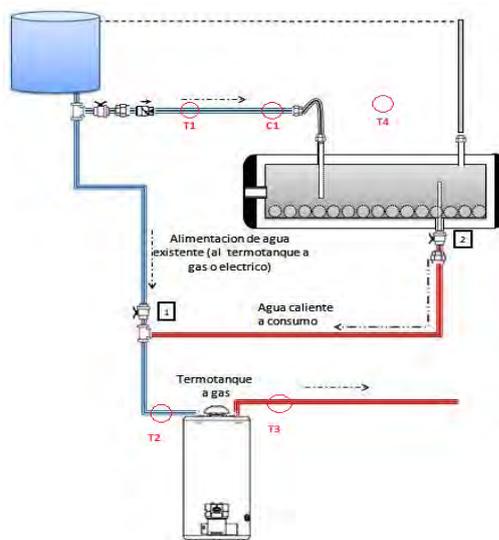


Figura 5. Diagrama de la red de agua caliente con la ubicación de los sensores. Imagen original[10]

## 4 Pruebas

Para determinar el correcto funcionamiento del prototipo desarrollado se llevaron a cabo distintas pruebas de funcionamiento de forma integral, probando los distintos sensores involucrados, así como también el envío de los datos obtenidos y su correcta visualización.

Para probar el sensor de caudal se dispuso a pasar distintas cantidades de agua por el sensor cada 60 segundos. En la figura 6 se muestra el resultado de las pruebas realizadas.

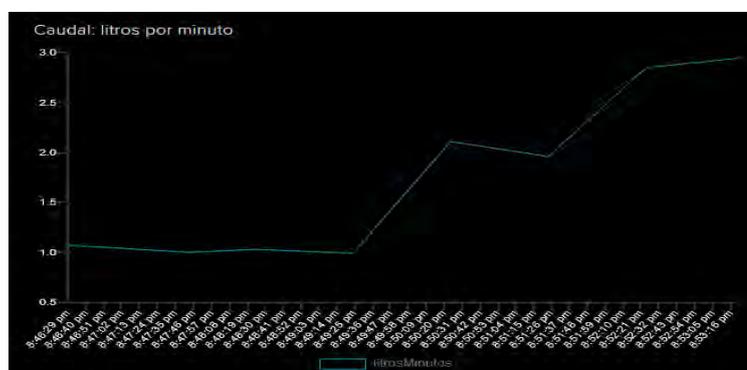
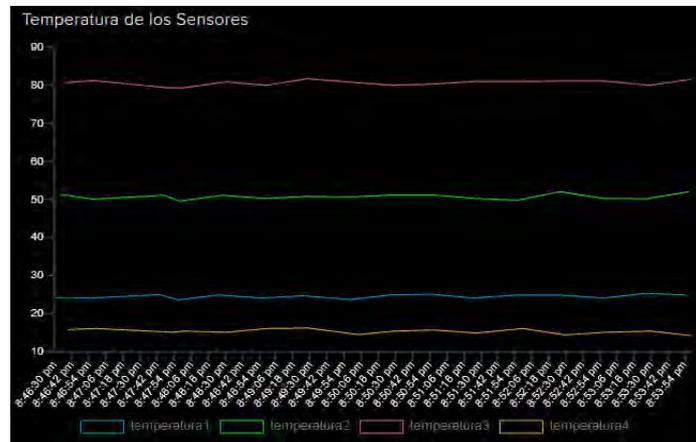


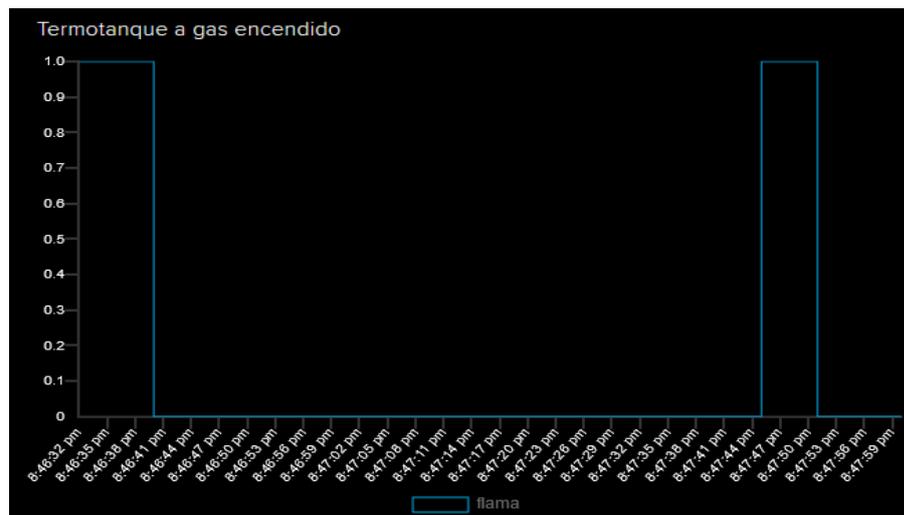
Figura 6. visualización de los resultados de las pruebas de caudal.

Los sensores de temperatura se probaron de forma independiente, teniendo uno a temperatura del ambiente y los otros tres sumergidos en agua caliente a distintas temperaturas, tomando valores cada 30 segundos. En la figura 7 se muestra el resultado.



**Figura 7.** visualización de los resultados de las pruebas de los sensores de temperatura.

Para el sensor de flama se utilizó una fuente de llamas para probar el correcto funcionamiento del sensor y registrar los valores. Se realizaron varios pulsos de prendido/apagado tomando valores cada 5 segundos. Se puede ver en la figura 8 el resultado de dichas pruebas.



**Figura 8.** visualización de los resultados de las pruebas del sensor de flama.

## 5. Conclusiones y futuras líneas de trabajo

Desde el aspecto tecnológico, se desarrolló exitosamente un prototipo que permite el monitoreo de las variables básicas de un sistema de agua caliente que cuenta con un termotanque solar y un termotanque a gas, permitiendo recolectar datos sobre el funcionamiento de dicho sistema. En otro aspecto, la tecnología aplicada para el desarrollo permite fácilmente adaptarse a nuevos requisitos, permitiendo cambios en las formas de conectividad, de recolección de datos, o la inclusión de nuevas variables para monitorear.

La implementación del servicio Adafruit IO como sistema de recolección de datos y visualización permite de manera rápida y sencilla acceder a la visualización en tiempo real e histórico de los datos obtenidos sin complicaciones de instalación de servidores, permitiendo un desarrollo y testeo rápido del prototipo.

Permitir el acceso a los datos, tanto en tiempo real como el histórico, permite un mayor entendimiento de la tecnología y del comportamiento de la zona, permitiendo analizar los datos para generar políticas de incentivo para la adopción en la región de esta tecnología.

El sistema de monitoreo para un termotanque solar junto a un termotanque a gas puede ser aplicado en otras regiones o en otras instalaciones antiguas, así como también en otro tipo de instalaciones, como, por ejemplo, la instalación en solitario de un termotanque solar.

### *Futuras líneas de trabajo*

Utilizar Raspberry Pi como soporte tecnológico permite la incorporación de nuevos requisitos sin tener que cambiar drásticamente el hardware o modificar la tecnología subyacente, lo que haría factible enfocarse en realizar mejoras sobre el aumento de la cantidad de parámetros de la instalación a monitorear, como podrían ser la radiación solar, humedad del ambiente, etc. Esto permitiría reunir aún más datos del comportamiento del sistema donde está instalado.

En relación al apartado de almacén de datos y visualización, se podría utilizar otras tecnologías que permitieran poder utilizar la estación de monitoreo de manera local, en un servidor aparte o en la propia Raspberry Pi, sin la necesidad de tener internet, así como permitir una solución híbrida teniendo acceso tanto local como en la nube, lo que permitiría el acceso a la información de manera más directa.

### Referencias

1. B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds). 2007. CAMBIO CLIMÁTICO 2007 Informe del Grupo de Trabajo III - Mitigación del Cambio Climático. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
2. Joaquín Mario Ortiz. 2018 . Produccion de energia a partir de vinaza en Tucuman: Análisis de las direfentes alternativas. Universidad Nacional deTucumán.
3. A. Lanson1 y R. Righini,1 E. E. Benitez, Et. Al. 2013. HACIA UN USO MÁS EFICIENTE DEL GAS – APROVECHAMIENTO DE LA ENERGÍA SOLAR EN ARGENTINA. Encuentro Latinoamericano de Uso Racional y Eficiente de la Energía - ELUREE2013
4. Fundación de la Innovación Bankinter 2011. “El Internet de las Cosas, En un mundo conectado de objetos inteligentes”.
5. Internet Society. “The Internet of Things (IoT): An Overview”. White paper, 2015.
6. ¿Cuál es el panorama de IoT en Argentina? - <https://www.tynmagazine.com/cual-es-el-panorama-de-iot-en-argentina/> (verificada el 25/02/2020).
7. Dr. Pablo Fabián CARRANZA. (2017). “Uso de Energías Renovables en la Universidad Nacional de Río Negro”. UNIVERSIDAD CULTURA Y SOCIEDAD – SPU.
8. Raspberry Pi Foundation - About us - <https://www.raspberrypi.org/about/> (verificado el 27/02/2020)
9. Politi, M. I. M., Cruz, T. A., Acosta, T. C., González, J. L., & Laiz, H. DESARROLLO DE SISTEMA PARA INTEROPERABILIDAD DE INVERSORES FOTOVOLTAICOS EN PROYECTO DE RED INTELIGENTE CON APORTE DE FUENTES RENOVABLES. PRIER-ARMSTRONG (SANTA FE, ARGENTINA).
10. Portale-cologica.com - <http://www.e-cologica.com.ar/test/wp-content/uploads/2017/09/INTEGRACION-SERIE1.jpg> - (verificada el 02/03/2020)

# Innovación en el Sector Público para Ciudades Inteligentes Sostenibles

Rocío Muñoz <sup>1</sup>, Ariel Pasini, 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 Comisión de Investigaciones Científicas de la Pcia. de Bs. As. (CIC)

<sup>1</sup> Becario Postgrado UNLP

{rmunoz, apasini, ppesado}@lidi.info.unlp.edu.ar

**Abstract:** Se presentan los conceptos de Ciudades Inteligentes Sostenibles (SSC) y los pilares que las componen. Se realiza un análisis de lo que implica innovar en el sector público, presentando los conceptos de laboratorios de innovación y metodologías para la innovación pública. De esta forma, se seleccionan dos laboratorios de innovación de Argentina y se analizan 12 casos de éxito de cada uno, categorizándolos según a cuál o cuáles de los pilares que componen una SSC logran satisfacer y qué metodología de innovación es utilizada. Finalmente, se realiza un análisis con los datos obtenidos.

**Keywords:** Ciudades Inteligentes Sostenibles, Servicios Públicos, Innovación, Metodologías de innovación

## 1 Introducción

Al hablar de Ciudades Inteligentes Sostenibles (SSC), se hace referencia a aquellas ciudades que buscan, de manera innovadora y mediante la utilización de las TIC, mejorar la calidad de vida de sus ciudadanos y ayudar a satisfacer sus necesidades brindando servicios públicos eficientes.

Las SSC se componen de cinco pilares: *Social*, relacionado con factores tales como la salud, educación y seguridad de los ciudadanos; *Económico*, que tiene en cuenta los aspectos relacionados con el crecimiento económico responsable y la generación de empleos; *Ambiental*, relacionado con la protección ecológica del medio ambiente; *Gobernanza*, referido a la administración de recursos y políticas para la regulación de procesos e *Infraestructura Urbana*, que tiene en cuenta tanto la infraestructura física como digital para lograr satisfacer las necesidades de la comunidad.[1]

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.[2]

Los laboratorios de innovación son nuevas formas de trabajo en la gestión pública que nacen para apoyar a las administraciones en la búsqueda de soluciones innovadoras. Su crecimiento ha sido exponencial en la última década y presentan características que los diferencian de los organismos públicos tradicionales. Estos laboratorios utilizan diversas metodologías que les permiten contribuir a la resolución de problemas públicos, siendo algunas de ellas: *Ciencia de datos*, *Ciencia del comportamiento*, *Simulación y modelaje*, *Inteligencia colectiva* y *Design thinking*. [3]

En Argentina existen diversos laboratorios de innovación y organizaciones afines que buscan, a través de soluciones innovadoras, colaborar con el crecimiento de las ciudades o regiones a las que pertenecen. Realizando un análisis de los diferentes laboratorios de innovación identificados por la Red Innolabs, se determinó que aquellos con mayores avances y resultados presentados fueron dos laboratorios. Para los mismos, fueron seleccionados 12 de sus casos de éxito con el fin de presentar una categorización en cuanto a con qué metodología son llevados a cabo y a cuál o cuáles de los pilares de una SSC satisfacen.

En la sección dos se describe el concepto de Ciudades Inteligentes Sostenibles (SSC) y los servicios públicos que se prestan teniendo en consideración los pilares que componen las SSC. En la tercera sección se define el concepto de innovación y se presentan los laboratorios y metodologías para la innovación en el sector público. Luego, se realiza un relevamiento sobre los diferentes laboratorios de innovación presentes en Argentina y se analizan algunos de sus casos de éxito. Finalmente, se presentan las conclusiones de este artículo.

## 2 Ciudades Inteligentes Sostenibles

Una Ciudad Inteligente Sostenible (SSC) 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. [1]

Las SSC se componen de cinco pilares: *Social*, *Económico*, *Ambiental*, *Gobernanza* e *Infraestructura Urbana*, como puede observarse en la Figura 1, para los que requiere la prestación de servicios públicos que satisfagan las necesidades de sus ciudadanos. [1]

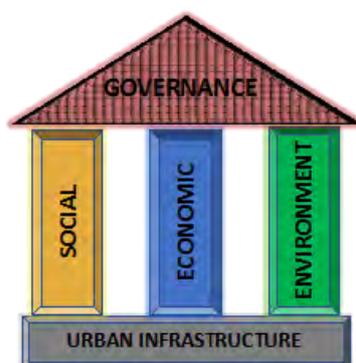


Figura 1. Pilares de una SSC

## 2.1 Servicios públicos en SSC

Un servicio público es un resultado producido y entregado para beneficio del público. Generalmente llevado a cabo por los organismos del Estado o bajo el control y la regulación de estos.

El pilar **Social** cubre los aspectos relacionados con las personas y comunidades para garantizar la calidad de vida, teniendo en cuenta factores tales como la salud, educación, seguridad, entre otros.

En cuanto a los aspectos relacionados con la salud, una gestión inteligente de la atención médica convierte simples datos en información clínica, permitiendo a los especialistas mejorar el servicio prestado a los pacientes. Una SSC debe permitir comunicaciones e intercambio de información seguros. Además, se deben brindar oportunidades de forma equitativa a todos los ciudadanos y proveer un servicio de información eficiente para mantener a la población informada sobre medidas como: programas de vacunación, programas de prevención de enfermedades, etc. [4]

La educación es un componente crucial de los servicios de ciudades inteligentes. A medida que el mundo se globaliza rápidamente, una de las únicas formas de mantenerse competitivo es desarrollar habilidades basadas en el conocimiento a través de la educación. Esto incluye tanto el conocimiento inicial como el aprendizaje permanente. El papel de las escuelas y universidades es, por lo tanto, un elemento clave a considerar en el diseño de soluciones educativas inteligentes. En una SSC, es necesario generar políticas para proveer acceso a los dispositivos y la capacitación de los ciudadanos para reducir la brecha digital. [4]

Teniendo en cuenta los aspectos relacionados con la seguridad, esta es considerada la puerta de entrada al concepto de ciudades inteligentes. Una SSC supone una ciudad donde sus ciudadanos se ocupan de realizar sus actividades cotidianas con tranquilidad, sin preocuparse por su protección personal. Deben proveerse servicios que estén

relacionados con el monitoreo y control de la seguridad de los ciudadanos. Existen, por ejemplo, herramientas analíticas que ayudan a detectar, responder y resolver incidentes.

El pilar **Económico**, por su parte, cubre los aspectos relacionados con el crecimiento económico responsable y sustentable y la generación de oportunidades laborales.

Una SSC debe promover el desarrollo económico y la innovación para mejorar la calidad de vida de sus ciudadanos. Se deben generar condiciones apropiadas para que las empresas promuevan empleos de calidad y se fomente el emprendedorismo regional de forma sustentable.

En cuanto al pilar **Ambiental**, este se relaciona con la utilización de prácticas ecológicas, la protección y la restauración del medio ambiente.

A nivel mundial, las ciudades representan un 75% del consumo de energía y un 80% de las emisiones de CO<sub>2</sub>. Una SSC debe garantizar el uso racional de los recursos naturales, disminuir la contaminación ambiental, promover el uso de energías renovables, el uso eficiente del agua y el manejo de los residuos. Se considera fundamental reducir la contaminación que producen los combustibles fósiles.

La **Gobernanza** se refiere a la capacidad de administrar recursos y políticas e involucrar a diferentes partes interesadas, proponiendo mecanismos y procesos regulatorios y de cumplimiento bien equilibrados de manera estandarizada y continua.

En una SSC, la gobernanza incluye normas sociales, personas, políticas, prácticas, datos, información y tecnología para: equilibrar las dimensiones sociales, económicas y ambientales utilizando la infraestructura urbana y las TIC para conectar los elementos de una región; gestionar las relaciones a largo plazo con las partes interesadas dentro y entre los sectores gubernamentales; definir y realizar caminos apropiados para el desarrollo de territorios inteligentes; asegurar un fuerte apoyo institucional para las iniciativas ambientales.

El gobierno de una SSC debe enmarcarse en un gobierno abierto, transparente y participativo. Los representantes del gobierno y las instituciones deben dar lugar a los ciudadanos a colaborar y participar en la creación de políticas públicas para garantizar la transparencia de estas. Proveer a los ciudadanos servicios para que puedan participar es el primer paso.

La **Infraestructura Urbana**, se refiere tanto a la infraestructura física como digital, las cuales son herramientas esenciales para construir SSC que satisfagan las necesidades de su comunidad.

La infraestructura física hace referencia a las carreteras, transportes, edificios públicos, puentes, etc., esenciales para el funcionamiento de una ciudad. Una SSC debe comprometerse a brindar servicios de transporte eficientes, que reduzcan la contaminación ambiental fomentando el uso de energías renovables, proporcionar alternativas para la circulación, como por ejemplo bicisendas, brindar servicios inteligentes para el estacionamiento de vehículos en la ciudad, evitando así la circulación innecesaria en búsqueda de un lugar para estacionar, etc.

Al hablar de infraestructura digital, se hace referencia a una infraestructura de comunicaciones informáticas para la realización del procesamiento computacional de

los datos. Una SSC debe proveer una sólida infraestructura digital proporcionando los equipos de hardware necesarios para el procesamiento y almacenamiento de la información de los ciudadanos, con las medidas de seguridad que el sistema requiera. Además, debe garantizar enlaces entre sus instituciones y proveer el acceso a fuentes de conectividad. [4]

### **3 Innovación**

La innovación es la creación o modificación de un producto y su introducción en el mercado. Las organizaciones deben innovar para asegurar persistencia a lo largo del tiempo. En función del tipo de organización, podemos diferenciar la innovación en el sector privado y en el sector público.

La innovación privada se concentra en la necesidad de proponer soluciones competitivas a los clientes, los cuales deciden si compran o no un producto o servicio. Al hablar de innovación en el sector público, el receptor generalmente no tiene la opción de dejar de consumir el servicio y debe adaptarse al servicio prestado por la organización.

El proceso de innovación en el sector público busca satisfacer las ideas que tienen los políticos acerca del servicio prestado, omitiendo en muchos casos los detalles sobre las necesidades de los ciudadanos. Con el avance de la tecnología, los ciudadanos cada vez poseen más herramientas para comunicarse con los organismos gubernamentales y hacerle conocer su punto de vista. Estos aportes de los ciudadanos son fundamentales en un proceso de innovación, permitiendo generar un ambiente en que los usuarios y los prestadores de servicios puedan desarrollar ideas conjuntas, que satisfagan a todos los interesados.

Para lograr un proceso de innovación constructivo es fundamental un equipo de trabajo multidisciplinario. Estos equipos aseguran diferentes puntos de vista, por ende, la solución abarca un espectro mayor de usuarios, logrando satisfacer las expectativas de un mayor número de interesados y fortalecer la innovación en el sector público.

Innovar implica dar impulso a mecanismos de participación ciudadana, colaboración y transparencia. Estos elementos, componentes esenciales del paradigma de Gobierno Abierto, conllevan un cambio no sólo en cuanto a las prácticas de la administración pública, sino también de la acción estatal, su vínculo con la sociedad y viceversa. La innovación pública debe aspirar a producir herramientas útiles que vuelvan permeables las fronteras de las organizaciones, en un intento por validar y legitimar (mediante la experimentación y la co-creación con la ciudadanía) respuestas efectivas a los desafíos sociales ante escenarios complejos, diversos e inciertos. [2]

#### **3.1 Laboratorios de innovación**

Los laboratorios de innovación pública son nuevas formas flexibles, abiertas y horizontales de trabajo en la gestión pública que nacen para apoyar a las administraciones en la búsqueda de soluciones innovadoras. Su crecimiento ha sido

exponencial en la última década, propiciado por la necesidad de las administraciones públicas de incorporar la innovación en su gestión [3]

Los laboratorios enmarcan una agenda más amplia que la de “modernización” del Estado. En ella confluyen el gobierno abierto, el gobierno digital, la búsqueda de hacer un mejor uso de los recursos públicos y la necesidad de volver capilares a las instituciones del estado frente una sociedad que exige mayor efectividad, eficiencia en la asignación y transparencia. Los laboratorios de innovación brindan estímulo para hacer las cosas de una nueva manera, rompiendo los esquemas tradicionalmente aceptados, buscando nuevas herramientas y metodologías en la resolución de problemas y desafíos públicos.

Existen algunas características que diferencian a los laboratorios de los organismos públicos tradicionales [2]:

- son entidades pequeñas y dinámicas con relativa independencia y mayor espacio para el fracaso que los organismos tradicionales.
- se recurre a metodologías de experimentación, adopción rápida de proyectos piloto y evaluación de impacto.
- se requiere cierto *know-how* en tecnologías digitales y ciencia de datos.
- suelen estar compuestos por un plantel multidisciplinario, lo que permite ampliar la visión tradicional de las organizaciones gubernamentales, combinar enfoques y visiones de la realidad, y aumentar los niveles de comprensión ante la diversidad social.
- suelen incentivar espacios de colaboración, tanto entre diferentes carteras gubernamentales, como entre estas y actores del sector privado, la sociedad civil y la academia.
- buscan involucrar voces ciudadanas en procesos de diseño de políticas públicas.

### 3.2 Metodologías para la innovación pública

La metodología que utilizan los laboratorios de innovación para contribuir a la resolución de problemas públicos constituye un elemento distintivo de su labor. Los laboratorios se pueden clasificar en: a) aquellos que utilizan metodologías orientadas al diseño; b) los que se especializan en metodologías de gobierno abierto y el uso de datos; c) los que utilizan metodologías basadas en evidencia, y d) los que emplean metodologías mixtas.

Entre las metodologías utilizadas para la innovación en el sector público se pueden distinguir [3]:

1. *Ciencia de datos*: Este conjunto de metodologías se ha explotado significativamente en los últimos años. Involucra análisis estadísticos, análisis de datos y expertos del dominio para extraer información de grandes volúmenes de datos. La información obtenida sobre un dominio determinado permite realizar tomas de decisiones certeras. Estas herramientas permiten combinar diferentes fuentes de datos, tanto estructurados como no estructurados, combinar patrones de solución y modelos estadísticos, análisis y visualización de indicadores, información georreferencial, que facilita el análisis en función de un lugar específico, etc.

2. *Ciencia del comportamiento*: Estas metodologías intentan comprender los problemas públicos, en general, y el comportamiento humano, en particular, con el fin de diseñar intervenciones que sirvan de evidencia para la toma de decisiones en políticas públicas. Este tipo de técnica se ha utilizado para realizar intervenciones en el área de salud, educación, participación ciudadana, medio ambiente, entre otros. Las ciencias del comportamiento tratan de entender el comportamiento real de las personas en función del contexto, sus motivaciones y sus creencias sociales, que puedan influir en sus decisiones. En base a este análisis se pueden proporcionar herramientas que permitan diseñar soluciones a diferentes problemas públicos.
3. *Simulación y modelaje*: El método de simulación y modelaje permite comprender un sistema social. El resultado de la simulación, por ejemplo, permite ver el resultado de aplicar una política pública. Se realizan modelos computacionales complejos para analizar fenómenos sociales y simular el efecto de posibles políticas públicas antes de su implantación.
4. *Inteligencia colectiva*: Grupos de individuos que actúan colectivamente de maneras que parecen inteligentes. La metodología consiste en que un grupo de personas con conocimientos en la problemática a tratar, con características personales diversas y distintas visiones del problema, juntas por consenso generen nuevas ideas sobre la solución al problema. Existen herramientas como: consulta a expertos, métodos participativos, modelo estructurado, análisis de actores e identificación de variables claves, que facilitan el proceso necesario para aplicar esta metodología.
5. *Design thinking*: Es una metodología orientada al desarrollo de políticas y servicios públicos con base en los usuarios. Permite idear soluciones colaborativas que satisfagan las demandas de las ciudades. Se presenta un método de cinco etapas: empatizar con el usuario; definir el problema; generar soluciones alternativas; creación de prototipos; implantaciones pilotos. Esta metodología no sigue un proceso lineal ya que, en muchas ocasiones, cada una de las etapas que la integran pueden realizarse de manera paralela y con múltiples iteraciones.

## 4 Laboratorios de Innovación en Argentina

La Red Innolabs [5] se conforma como un espacio de conexión de los laboratorios de innovación iberoamericanos para compartir ideas, conocimiento y experiencias, planteando líneas de trabajo conjuntas de mejora de métodos de trabajo para enfrentar problemas comunes a los importantes retos de la sociedad.

En un relevamiento llevado a cabo en 2018 por la Red Innolabs de los laboratorios de innovación y organizaciones afines del continente Europeo y el continente Americano [6], se identificaron 6 laboratorios en Argentina: LABgcba, Santalab, Buenos Aires LAB, NQNLAB, PoliLab UNR y CreaPolis-CoLabs.

Para el desarrollo de este artículo se recolectó información sobre los laboratorios identificados en Argentina como laboratorios de innovación, con el fin de seleccionar aquellos que demuestren un mayor avance y presenten mayor cantidad de resultados.

De los laboratorios identificados por la Red Innolabs en Argentina, se encontró que algunos de ellos han dejado de existir y algunos otros tienen poca participación activa

en la actualidad (mayoritariamente en redes sociales). Por estos motivos, se seleccionaron dos laboratorios.

#### 4.1 Casos de éxito Laboratorio 1

Se analizaron las diferentes líneas de acción que presenta el Laboratorio 1 y se seleccionó un conjunto de casos exitosos llevados a cabo en los últimos años.

La Tabla 1 presenta 12 casos de éxito identificados e incluye:

- una descripción del caso
- una clasificación según la metodología con la que se deduce que fue llevado a cabo, teniendo en cuenta lo propuesto en la sección 3.2
- una clasificación según a cuál o cuáles pilares de una SSC, presentados en la sección 2, se satisface

Tabla 1. Casos de éxito Laboratorio 1

Caso de éxito	Descripción	Metodología	Pilar SSC
1	Desarrollo colaborativo de ciclistas para edificios públicos de la ciudad.	Ciencia del comportamiento	Infr. urbana / Ambiental
2	Co-creación de un manual con aspectos básicos y necesarios para espacios de exhibición audiovisual.	Inteligencia colectiva	Social
3	Estudiantes de escuelas técnicas desarmaron residuos de aparatos electrónicos y recuperaron piezas que se reunieron y clasificaron para la creación de prototipos y dispositivos tecnológicos de baja complejidad.	Design thinking	Ambiental
4	Motor de búsqueda colaborativo que utiliza información de sitios de origen público y permite la colaboración para mejorar los resultados.	Inteligencia colectiva	Social
5	Mapa colaborativo para recuperar la identidad de los barrios. Fomenta la participación de sus habitantes más jóvenes, a través del uso de las nuevas tecnologías.	Inteligencia colectiva	Social
6	Mapa global colaborativo que da cuenta de las calles con nombre de mujer en Latinoamérica y España.	Inteligencia colectiva	Social
7	Mapa analógico con las iniciativas de distintas organizaciones y temáticas.	Ciencia del comportamiento	Gobernanza
8	Mapa que separa edificios de la ciudad (bares, bancos, universidades, escuelas, etc.) en tres niveles, teniendo en cuenta la accesibilidad en silla de ruedas.	Design thinking	Infr. urbana
9	Iniciativa que promueve el trabajo colaborativo entre el Estado y organizaciones de la sociedad civil en búsqueda de transparencia, participación y co-creación de políticas públicas con los ciudadanos.	Inteligencia colectiva	Gobernanza
10	Herramienta digital para la participación de la ciudadanía en las votaciones de presupuesto participativo.	Ciencia de datos	Gobernanza / Económico
11	Creación de prototipos utilizando la tecnología para la resolución de problemáticas sociales en el contexto educativo.	Simulación y modelaje	Social
12	Mejora la experiencia de vacunación en infantes mediante un video interactivo reproducido en cascos virtuales.	Design thinking	Social

## 4.2 Casos de éxito Laboratorio 2

Siguiendo la metodología de trabajo aplicada en la sección 4.1, se estudiaron y analizaron los trabajos realizados en los últimos años por el laboratorio 2

De este modo, se seleccionó un conjunto de 12 casos de éxito y se completó una tabla que incluye los mismos datos que los mencionados en el apartado anterior. En la Tabla 2 se puede observar la información analizada para los casos seleccionados.

Tabla 2. Casos de éxito Laboratorio 2

Caso de éxito	Descripción	Metodología	Pilar SSC
1	Organiza y coordina la infraestructura informática de telecomunicaciones y de los sistemas de información.	Ciencia del comportamiento	Infr. urbana
2	Permite cargar solicitudes, reportes o denuncias y hacer el seguimiento de cada requerimiento.	Inteligencia colectiva	Social / Gobernanza
3	Permite realizar, en forma online, la fila para un trámite y enterarte cuándo toca tu turno.	Ciencia de datos	Gobernanza
4	Jornadas cuyo objetivo general es pensar ideas enriquecedoras para encontrar soluciones reales.	Design thinking	Social
5	Mapa que permite calcular la tasa de éxito de un negocio en base a información detallada por rubro y zona de la Ciudad.	Ciencia de datos	Económico
6	Mapa que permite consultar los hechos delictivos de la ciudad por comuna o barrio.	Ciencia de datos	Social
7	Mapa que brinda información geográfica de interés y herramientas para la localización y búsqueda de recorridos en transporte público, auto, bicicleta o a pie.	Ciencia de datos	Infr. urbana
8	Diseño e implementación de un nuevo dispositivo para las bancas de la Legislatura.	Design thinking	Infr. urbana
9	Solución amigable y efectiva para la normalización automática de direcciones en conjuntos de datos.	Ciencia de datos	Social
10	Colaboración para la divulgación, intercambio y co-creación de metodologías y enfoques para la innovación en políticas públicas.	Inteligencia colectiva	Gobernanza
11	Trabajo colaborativo con el fin de encontrar soluciones para un problema histórico: la baja tasa de donantes registrados en el sistema de donación de órganos.	Design thinking	Social
12	Talleres que brindan herramientas para pensar nuevas formas de proveer servicios públicos y desarrollar políticas públicas centradas en el ciudadano.	Design thinking	Social / Gobernanza

## 5 Conclusiones

Se presentó el concepto de Ciudades Inteligentes Sostenibles (SSC) y se clasificaron los servicios públicos destinados a satisfacer las necesidades de los ciudadanos teniendo en cuenta los distintos pilares que componen una SSC.

Se definió el concepto de innovación, haciendo énfasis en la importancia de innovar en el sector público en conjunto con los ciudadanos, dando impulso a mecanismos de participación ciudadana, colaboración y transparencia. Además, se presentaron los

laboratorios de innovación y las metodologías que se utilizan para contribuir a la resolución de problemas públicos.

Para el presente artículo se seleccionaron, de los laboratorios de innovación identificados por la Red Innolabs, los dos que demuestran mayores avances en Argentina y se llevó a cabo un relevamiento de 12 casos de éxito de cada uno de ellos, permitiendo categorizarlos según qué metodologías se considera que fueron utilizadas para su desarrollo y a cuál o cuáles de los pilares que componen una SSC satisfacen.

De la categorización realizada se puede observar que en ambos laboratorios el pilar *Social* es el más beneficiado en los procesos de innovación llevados a cabo, seguido por los pilares *Gobernanza e Infraestructura urbana*. Por su parte, el pilar *Economico* solo registra una iniciativa en cada uno de los laboratorios y el pilar *Ambiental* solo se encuentra presente en iniciativas propuestas por el Laboratorio 1.

En cuanto a las metodologías de innovación que se considera que fueron utilizadas por los laboratorios para el desarrollo de las iniciativas seleccionadas, se observó que Laboratorio 1 aplicó, en la mayoría de sus propuestas, la metodología *Inteligencia colectiva* (5), seguido por *Design thinking* (3). Por otro lado, analizando las propuestas seleccionadas del Laboratorio 2, la metodología mayormente utilizada fue *Ciencia de datos* (5), seguida también por *Design thinking* (4).

## 6 Agradecimientos

Esta publicación fue realizada en el contexto del Proyecto CAP4CITY – “Strengthening Governance Capacity for Smart Sustainable Cities” ([www.cap4city.eu](http://www.cap4city.eu)) co-financiado por el Programa Erasmus+ de la Unión Europea. Acuerdo Número 598273-EPP-1- 2018-1-AT-EPPKA2-CBHE-JP. Número de proyecto: 598273

## 7 Referencias

- [1] E. Estevez *et al.*, “Desarrollo de Recursos Humanos para la Gobernanza de Ciudades Inteligentes Sostenibles – Proyecto CAP4CITY,” 2020.
- [2] F. Long, “Laboratorios de gobierno para la innovación y burocracias públicas,” *RedInnolabs*, p. 80, 2020.
- [3] F. Rojas Martín, “10 ideas para la creación de un laboratorio de gobierno para la innovación pública,” *RedInnolabs*, 2020.
- [4] ITU-T Focus Group on Smart Sustainable Cities, “An overview of smart sustainable cities and the role of information and communication technologies,” 2014. [Online]. Available: [https://www.itu.int/en/ITU-T/focusgroups/ssc/Documents/Approved\\_Deliverables/TR-Overview-SSC.docx](https://www.itu.int/en/ITU-T/focusgroups/ssc/Documents/Approved_Deliverables/TR-Overview-SSC.docx).
- [5] “RED DE LABORATORIOS DE INNOVACIÓN PÚBLICA EN IBEROAMÉRICA.” <http://www.redinnolabs.org/>.
- [6] Exequiel Rodríguez, “Laboratorios de Gobierno para la Innovación Pública,” *RedInnolabs*, p. 80, 2020.

# New Technologies for Innovation in Public Services: A Review of Case Studies

Ana Inés Dennehy<sup>1,2</sup>, Elsa Estévez<sup>1,3</sup>, and Pablo R. Fillottrani<sup>1,2</sup>

<sup>1</sup> LISSI

Departamento de Ciencias e Ingeniería de la Computación  
Universidad Nacional del Sur, Bahía Blanca, Argentina

<sup>2</sup> Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina

<sup>3</sup> Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina  
inesdennehy@gmail.com, {ece,prf}@cs.uns.edu.ar

**Abstract.** This paper analyses case studies of Artificial Intelligence and Big Data Analytics applications for innovation in digital public services. Recently, many governments at all levels are using these technologies to improve service provision to citizens. However, they are not always aware of their potential for disruption and of the emergent ethical questions. In order to have a better insight of their applications and provide a initial point for policy makers, we review several case studies and analyse how these new technologies are being used to meet specific needs.

**Keywords:** artificial intelligence, big data, digital governance, public services

## 1 Introduction

New technologies are being adopted in the private sector, and citizens quickly accept them as they are familiarised with the use of social networks and mobile applications. Their access to innovative services such as the ones used in banks and e-commerce platforms, has improved in recent years. Because citizens are conscious on how to receive new services that better satisfy their needs, they put pressure on governments to innovate. Many governments are using and applying Artificial Intelligence (AI) and Big Data and Data Analytics (BD) techniques in order to innovate in such services[1].

This paper analyses case studies of AI and BD applications for innovation in digital public services, being its main objective to provide a landscape review of good practices. There is little experience in using these new technologies which are generally driven by cost efficiency, but also involve the potential for disruption and usually raise ethical questions. In this sense it's essential to understand how emerging technologies are being used to meet specific government and citizens' needs, how government organisations drive their approaches to innovations, and which innovation capabilities are necessary within the government, and this review aims at contributing to this end. The cases presented are intended to be illustrative, by no means complete, and to provide a sense of current activities

and future directions that can especially help local governments in public policies for the adoption of AI and BD aiming at innovation of public services.

The rest of this paper is structured as followed. Section 2 provides a background to two main concerns for the adoption of AI and BD - ethical issues and privacy, as well as explains government-citizens interactions; and Section 3 states related work. Section 4 and 5 explain the possibilities of innovation with BD and AI respectively, each presenting according study cases. Finally, Section 6 concludes the paper.

## 2 Background

In this work, we assume AI and BD as extensive terms since there is no single precise definition. AI refers to devices, software and systems that are able to learn and to make decisions in almost the same manner as people [24]. AI allows machines, devices, software, systems and services to function in a sensible way according to the task and situation at hand. Big Data refers to large volumes of data that cannot be processed effectively with the existing applications. The processing of Big Data begins with the raw data that isn't aggregated, and there is the need to add semantics to it. Data Analytics involves applying an algorithmic process to derive insights and, for example, running through several data sets to look for meaningful correlations between each other[26]. Both together, BD is used to make better decisions as well as to verify or refute existing theories or models.

Applying AI and BD raises different types of issues, like ethical, privacy, and enhance interactions between government institutions and service recipients.

### 2.1 Ethical issues

The public sector differs from the private sector in its capacity for innovation, and it seems to stay behind in matter of new technological advances. For the society, the former has a certain ethical obligation to citizens, safeguarding personal data that can be used to identify a person directly[2–4]; mitigating the possibility of misuse of collected data for commercial purposes; mitigating the possibility of discrimination caused by the technology, for example elderly or disabled people that cannot use digital services[2, 5]; mitigating the possibility of discrimination in algorithmic decisions, caused by possible stigmatisation, racism, etc. [2, 5, 6]; and deciding who is responsible for accidents caused by autonomous cars, or even ensuring adherence to cultural principles in its algorithms.

### 2.2 Privacy

Privacy, which means the selection of which personal information is revealed to the world, is considered as a human right in many jurisdictions. The large amount of data that are currently handled in electronic systems means that people are exposed to greater levels of scrutiny while more and more aspects of

their everyday lives are captured as data[3], [4], these being names, usernames, passwords, emails, phone details, credit card numbers, Smart ID cards[7], car plaques, or even faces[3]. Another important concern is capturing movements, images of private property, or even maintaining personal data without citizen's consent. It is important to highlight the damage that prediction systems can do, given that they make decisions based on private user information[3, 8]; and the threat of the enormous amount of user data stored by the emergence of Internet of Things[9].

### 2.3 Interactions

Government agencies interact with many stakeholders, including other government entities, citizens, business, tourists, non-government organizations, employees, and others. Thus, there is a vast variety of services in which governments can innovate[10]. First and foremost, interaction with other governments can happen between the same or different government level, and within the same country or with other countries. Such interactions allow to perform collaborative strategies, exchange information and mutually provide services, for example share data to improve national policies, or improve global airport security. In fact, the potential benefits of these type of interactions is immense[4, 10].

Another type of interactions is those with businesses (Government to Business, G2B), which can benefit from many public services, such as registering new companies or paying taxes, while providing immediate information and maintaining digital communication active. Efficiencies of these type of communications can be achieved with reduction of paperwork. Another interesting type of G2B interaction consists in the public sector providing a platform in which businesses and other actors contribute in the co-creation of public services[4, 10, 11]. Additionally, there is the interaction with non-profitable organizations, which includes providing information, defining regulations and providing financial support to said organizations, including associations with social organizations, charities, and political parties[10].

One of the most important interactions is with citizens (Government to Citizens, G2C). The objective of this kind of interactions is to establish or maintain a direct relationship with citizens, while a variety of services is offered[4]. Another objective is to strengthen relationships between public authorities and citizens that live under their jurisdictions. G2C interactions through technology includes direct exchange of instant messages with public administrators, electronic vote, online tax declaration, online payment of city utilities, and the provision of digital services, such as change of residential addresses, and renovations of driving licenses[10]. For example, in Bahía Blanca (Argentina) several services can be accessed through the internet<sup>4</sup>. Another example is the case of Austria, in which they tried to implement an electronic ID, even though it was not very well accepted[7].

---

<sup>4</sup> [www.bahia.gob.ar/servicios](http://www.bahia.gob.ar/servicios)

Furthermore, there are interactions between the public sector with visitants tourists or visitors (Government to Visitors, G2V). G2V interactions include informational services to explain how to move within a country or territory, and provide information for visa applications and issuing<sup>5</sup>, booking tickets to national parks or cultural events, among others. Technological innovation has been used extensively on this domain, specially through providing online information and digital marketing to future tourists[10]. Some destination campaign examples include Your Singapore<sup>6</sup>, or Experience Catalunya in Spain<sup>7</sup>.

Finally, we must also consider interactions between government and its employees (Government to Employees, G2E), which are very similar to those between a business and its employees. The objective is to offer a variety of tools, documents and data that help employees to maintain communication and coordinate work in their offices. For example, government agencies can use an intranet to provide information to its employees. Public administrators can keep online record of personal information of their employees or create shared platforms for internal documentation to promote paperless interactions[4, 10].

### 3 Related work

AI and public governance is an emerging area of research. The systematic literature review in [20] acknowledges the lack of theorisation of most of research articles, and that most research is exploratory, conceptual, qualitative and practice-driven. Critical issues related to AI in the public sector, like ethical considerations, fairness, explainability, bias, privacy, safety, and societal impact among others, remain with relatively little understanding. In this sense, we present three works describing and analysing AI case studies in the government, and another three works introducing policy recommendations.

Several AI applications are categorised and described in [21]. Seven types of applications are presented: monitoring, discovering, predicting, interpreting, interacting with the physical environment, interacting with people, and interacting with machines; and the domains are ample, ranging from agriculture to education, from environment to business operations.

Three cases of AI applications in government are presented in [22], describing drivers, goals, barriers and risks. The authors highlight the governments' dilemma of the obligation from protecting citizens and the temptation to increasingly adopting algorithms in the sake of efficiency, and advocate for a common framework to evaluate the impact of the use of AI in the public sector. The paper shows that even seemingly trivial application of AI by the government can be an instrument of exercising control over the citizens, even in democratic regimes.

---

<sup>5</sup> Examples are India ([www.indianvisaonline.gov.in/visa/tvoa.html](http://www.indianvisaonline.gov.in/visa/tvoa.html)) and USA ([esta.cbp.dhs.gov/esta](http://esta.cbp.dhs.gov/esta))

<sup>6</sup> [www.yoursingapore.com](http://www.yoursingapore.com)

<sup>7</sup> [experience.catalunya.com](http://experience.catalunya.com)

In the following we describe related works in the area of policy recommendations. Six strategies to help governments start off with AI are presented in [23], in the context of describing AI applications. The strategies are: make AI a part of a goals-based, citizen-centric program; get citizen input; build upon existing resources; be data-prepared and tread carefully with privacy; mitigate ethical risks and avoid AI decision making; and, augment employees, do not replace them. The aim is that governments can approach the use of AI in citizen services with a focus on building trust, learning from the past, and improving citizen engagement through citizen-centric goals and solutions.

The government of Finland has recognised the utilisation of artificial intelligence as an important driver for creating conditions for strong economic growth and a higher rate of employment, and the Steering Group of the Artificial Intelligence Programme within the Ministry of Economic Affairs and Employment developed in [24] a strategy consisting in eight actions in order to adopt the necessary transformations. The actions are: enhancing the competitiveness of companies through the use of artificial intelligence; utilising data in all sectors; speeding up and simplify the adoption of artificial intelligence; ensuring top-level expertise and attract top experts; making bold decisions and investments; building the world's best public services; establishing new cooperation models; and making Finland a trendsetter in the age of artificial intelligence. The importance of digitalisation of all sectors, education by improving learning in the society and building skills for the employees; and an ethical perspective of AI service were recognised in the policy recommendation document [25].

## 4 BD for innovation

Historically, there has been a production of big clusters of data in the public sector, such as national census. Nevertheless, said data is not acquired continuously, and are complemented with small polls, study cases, interviews, or focus groups, which are all very focused, time- and place-specific, and significantly expensive to generate and analyze[12]. The term Big Data does not have an academically agreed definition. However, there are some agreed main characteristics: great volume, great speed, exhaustive scope, possibility of indexing, relational and flexible[12]. All these are characteristics that would help solve the problem mentioned above, which is why it is clear that Big Data techniques can be a great assistance to the public sector.

Among communication technologies that allow the collection of data are RFID to label objects; Wireless Sensor Networks (WSNs); wireless networks such as WiFi or Bluetooth; and mobile phone networks such as 4G, LTE, LTE-A, or 5G[13]. Another use for this kind of technologies, is responding to behavior of people with whom governments need to interact. For example, in the same way that sentiment analysis has been applied to extract sentiments and feeling from citizen tweets[14], this kind of analysis can also applied as a means of producing some feedback to citizens.

Transport is a sector in which many data applications have been developed in order to provide a better service[15–18]. For area, in fact, is very complicated to analyze data, given that if, for example, someone would like to make a mobile application for citizens that analyzes information in their context, some challenges should be addressed, such as people always being in movement; relevant information to a citizen will depend on their proximity to a specific place; there has to be a guarantee that the application will be real time; communication unreliability will need to be supported; and data will have to be stored and accessed efficiently[16].

#### 4.1 Case studies

CityOS, which is used in Dubrovnik, among other cities, is an open source library, APIs, and applications that allow to create platforms for smart cities. The frameworks and templates allow easy customization of application and hardware corresponding both to open data visualization as well as citizen applications.

There are several cases in the transport domain. MobyPark, which is used in Amsterdam, Paris, Brussels, Madrid, Barcelona, Milan and Düsseldorf, is a platform where people can look for parking spaces in a certain destination, book them on a certain period of time, and pay for them. Smile, which is used in Wien, is a platform in which the user can access public transport information, such as transport in their area, different options for getting from one point to another, and ticket pricing and booking.

Smart Traffic Lights, from Wien, is a project in which traffic lights are interconnected, and use certain algorithms to detect if a person intends to cross the road. Weather and environmental sensors also allow traffic lights to contribute to improving air quality through intelligent traffic flow control. The system also communicates with navigation systems in order to exchange valuable traffic related information. ITS (Intelligent Transport Systems), which is used in Copenhagen, is a product made by State of Green, a non-profit public-private partnership from Denmark. Its objective is the better coordination of traffic flow, prioritizing buses, and optimizing cycle paths, while also providing citizens with updated information about delays and alternative routes. Transit, used in Barcelona, is a mobile app for citizens, where they can consult traffic congestions, and set alerts for a specific area. App&Town Public Transport, used in Barcelona, Madrid and Laval, is another mobile app that provides the user with efficient and real-time routing information, taking into account any possible incidents that may occur and calculating the corresponding alternative routes. C-MobILE, used in Barcelona, Bilbao, Bordeaux, Copenhagen, Newcastle, North Barbant Region, Thessaloniki, and Vigo, is a program in which cooperative intelligent transport systems (C-ITS) and services are designed to deal with specific mobility challenges and deployed across Europe. All these cases are summarised in Table 1.

**Table 1.** Summary of BD case studies.

Name	Cities	Summary	Reference
CityOS	Dubrovnik	Open source framework for deploying smart city applications and hardware.	<a href="http://cityos.io">cityos.io</a>
MobyPark	Several european cities	System for booking and paying parking places.	<a href="http://www.mobypark.com">www.mobypark.com</a>
Smile	Wien	System providing public transport informational services and enabling ticket purchase.	<a href="http://smile-einfachmobil.at">smile-einfachmobil.at</a>
Smart Traffic Lights	Wien	Traffic lights communication and collection of data to improve traffic flow.	<a href="http://smartcity.wien.gv.at/site/en/smart-traffic-lights">smartcity.wien.gv.at/site/en/smart-traffic-lights</a>
ITS	Copenhagen	System for the coordination of traffic flow, providing useful information to citizens.	<a href="http://stateofgreen.com/en/partners/ramboll/solutions/its-intelligent-transport-systems">stateofgreen.com/en/partners/ramboll/solutions/its-intelligent-transport-systems</a>
Transit	Barcelona	System providing information and alerts about traffic congestion.	<a href="http://play.google.com/store/apps/details?id=com.tsol.transit&amp;hl=es">play.google.com/store/apps/details?id=com.tsol.transit&amp;hl=es</a>
AppTown	Barcelona Madrid, Laval	Public transport real-time routing information	<a href="http://www.appantown.com/index.html">www.appantown.com/index.html</a>
C-Mobile	Several european cities	C-ITS design and deployment platform.	<a href="http://c-mobile-project.eu">c-mobile-project.eu</a>

## 5 AI for innovation

Another technology that has emerged in the last years is the one enabling the processing of large volumes of data to achieve results normally generated by humans. Case studies of G2C services fall into one of the following categories: answering questions, scanning and filling documents, classifying petitions, translation, and document creation. These types of applications can do governmental work more efficiently while freeing employees time to build better relationships with citizens[19]. AI applications designed for answering citizens questions is the most common, whether it is via direct contact with citizens to solve common questions through chatbots, or applications relying on, e.g. IBM Watson, that can be used in diverse areas, such as finance, judicial decisions, or even client management[19].

Another category refers to scanning and filling documents, which is used specially in the legal field, where extensive and numerous documents need to be analyzed<sup>8</sup>. IBM Watson, for example, can also be used to scan and analyze documents, saving millions of dollars to companies of any kind<sup>9</sup>[19].

Lastly, creating and assembling documents is used in diverse areas, both in public and private sectors. Document assembly can be made through natural

<sup>8</sup> [emerj.com/ai-sector-overviews/ai-in-law-legal-practice-current-applications](http://emerj.com/ai-sector-overviews/ai-in-law-legal-practice-current-applications)

<sup>9</sup> [www.ibm.com/blogs/watson/2016/07/10-industries-using-big-data-win-big](http://www.ibm.com/blogs/watson/2016/07/10-industries-using-big-data-win-big)

**Table 2.** Summary of AI case studies.

Name	State/Country	Summary	Reference
Facebook chatbots	North Carolina	Answering simple questions like resetting passwords to free up IT help desk personnel.	<a href="http://www.govtech.com/computing/Chatbots-Debut-in-North-Carolina-Allow-IT-Personnel-to-Focus-on-Strategic-Tasks.html">www.govtech.com/computing/Chatbots-Debut-in-North-Carolina-Allow-IT-Personnel-to-Focus-on-Strategic-Tasks.html</a>
KeyReply	Singapore	Chatbot company used by private and public sectors.	<a href="http://www.keyreply.com">www.keyreply.com</a>
MySurrey	Surrey, British Columbia	Mobile app powered by IBM Watson that answers common questions in natural language.	<a href="http://www.surrey.ca/city-services/15161.aspx">www.surrey.ca/city-services/15161.aspx</a> , <a href="https://medium.com/cognitivebusiness/watson-assists-cities-with-311-3d7d6898d132">medium.com/cognitivebusiness/watson-assists-cities-with-311-3d7d6898d132</a>
DoNotPay	US, Canada, UK	Robot lawyer that helps citizens with legal problems, and sorts and pre-fills forms.	<a href="http://donotpay.com">donotpay.com</a> , <a href="http://www.theguardian.com/technology/2017/mar/06/chatbot-donotpay-refugees-claim-asylum-legal-aid">www.theguardian.com/technology/2017/mar/06/chatbot-donotpay-refugees-claim-asylum-legal-aid</a>
AI Translate	US	AI translation to fulfill the needs of residents who are not native English speakers.	<a href="http://aitranslate.com/government">aitranslate.com/government</a>
UEBA	Mexico	Data analysis and decision making in the justice system.	<a href="https://justice-trends.press/es/el-uso-de-la-inteligencia-artificial-en-el-analisis-del-comportamiento-de-usuarios-y-entidades-en-el-sistema-de-justicia-mexicano">justice-trends.press/es/el-uso-de-la-inteligencia-artificial-en-el-analisis-del-comportamiento-de-usuarios-y-entidades-en-el-sistema-de-justicia-mexicano</a>

language generation (NLG), which is a method widely used by news companies to make previews of automated news[19].

## 5.1 Case studies

Most use cases are given as chat services. Facebook chatbots are used in North Carolina to simulate human conversations for regular questions, like resetting passwords, which make up about 80-90% of the tickets submitted to the IT help desk, and aiding internal IT help desk personnel, freeing them up to focus on more strategic tasks. KeyReply, whose service was used in Singapore, is a chat automation company that has helped many industries, like insurance, finance, healthcare, e-commerce, and logistics, connect with their customers, and also connect governments with citizens. MySurrey, from Surrey, British Columbia, is a mobile app powered by IBM Watson that allows citizens to access important information about the city, answering questions in natural language. It addresses 65 percent of 311 calls that have answers on the city websites, but are too difficult for residents to find, which can cost each between four and five U.S. dollars.

DoNotPay, used in the US, Canada and the UK, is a robot Lawyer that aids citizens with many kinds of legal problems, like cancelling service or subscriptions, appealing parking tickets, and it even helps refugees seeking asylum, and pre-fills the corresponding form. AI Translate, which has coverage in many US government agencies, is a company that provides AI translation to many systems so that residents who are not native English speakers may fulfill any essential requirements such as getting healthcare, safety, and any legal informa-

tion. UEBA, or User and Entity Behavior Analytics, is used in Mexico justice system to improve efficiency and security of prisons, analyzing data and creating patterns that can improve decision making such as individuals classifying, social reinsertion, physical security, logical security and cybersecurity. These use cases are summarised in table 2.

## 6 Conclusions

This work provides a review of AI and BD technological innovation landscape in government public services. Eight BD applications, mostly in the transport sector, and six AI applications, mostly intelligent chatboxes, were summarized. Main obstacles to technology adoption are the complex interactions between government and citizen, and other stakeholders, which could be worsened by emerging ethical and privacy issues. Only clear understanding of the technologies and citizen education, not an only efficiency argument, may help a wider adoption and benefit by the government. Our future work comprises increasing the number of case studies to analyze and proposing a conceptualization of the use of new technologies for public service innovation based on them.

## References

1. Bertot, J., Estevez, E., Janowski, J.: Universal and contextualized public services: Digital public service innovation framework. *Gov. Inf. Q.*, 33 (2), 211–222 (2016)
2. Domanski, R., Estevez, E., Styrin, E., Alfano, M., Harrison, T. M.: Toward an ethics of digital government. In: 19th Annual International Conference on Digital Government Research, pp 123:1–123:4. Delft, The Netherlands (2018)
3. Kitchin, R.: The ethics of smart cities and urban science. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.*, 374(2083), 20160155 (2016)
4. Hiller, J.S., Belanger, F.: Privacy Strategies for Electronic Government. *E-government*, 200(2001), 162–198 (2001)
5. OECD : Recommendations on Digital Gov. Strategies for Transforming Public Services in the Welfare Areas. <https://www.oecd.org/gov/digital-government/recommendation-on-digital-government-strategies.htm> (2016)
6. Eubanks, V.: The Policy Machine. <https://slate.com/technology/2015/04/the-dangers-of-letting-algorithms-enforce-policy.html> (2015)
7. Aichholzer, G., Straub, S.: The citizen's role in national electronic identity management: A case-study on Austria. In: 2nd Int. Conf. Adv. Human-Oriented Pers. Mech. Technol. Serv. - CENTRIC 2009, 45–50. IEEE (2009)
8. Crawford, K., Schultz, J.: Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms. *Boston College Law Review*, 55 (1), 93 (2014)
9. Weber, R.H. : Internet of Things - New security and privacy challenges. *Comput. Law Secur. Rev.*, 26 (1), 23–30 (2010)
10. Kalbaska, N., Janowski, T., Estevez, E., Cantoni, L. : When digital government matters for tourism: a stakeholder analysis. *Inf. Technol. Tour.*, 17 (3), 315–333 (2017)
11. Janssen, M., Estevez, E.: Lean government and platform-based governance-Doing more with less. *Gov. Inf. Q.*, 30 (1), S1–S8 (2013)

12. Kitchin, R.: The real-time city? Big data and smart urbanism. *GeoJournal*, 79 (1), 1–14 (2014)
13. Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., Chiroma, H.: The role of big data in smart city. *Int. J. Inf. Manage.*, 36 (5), 748–758 (2016)
14. Hubert, R. B., Estevez, E., Maguitman, A., Janowski, T.: Examining government-citizen interactions on Twitter using visual and sentiment analysis. *Digit. Gov. Res. Pract.*, 1 (2), 15:1–15:10 (2018)
15. Anastasi, G., Antonelli, M., Bechini, A., Brienza, S., D’Andrea, E., De Guglielmo, D., Ducange, P., Lazzerini, B., Marcelloni, F., Segatori, A.: Urban and social sensing for sustainable mobility in smart cities. 2013 Sustain. Internet ICT Sustain. (2013)
16. Dobre, C., Xhafa, F. : Intelligent services for Big data science. *Futur. Gener. Comput. Syst.*, 37, 267–281 (2014)
17. Schnitzler, F., Liebig, T., Marmor, S., Souto, G., Bothe, S., Stange, H.: Heterogeneous stream processing and crowdsourcing for traffic monitoring: Highlights. *LNAI vol. 8726*, 520–523. Springer, Heidelberg (2014)
18. Horng, G.J., Li, J.P., Cheng, S.T.: Traffic congestion reduce mechanism by adaptive road routing recommendation in smart city. 2013 3rd Int. Conf. Consum. Electron. Commun. Networks, CECNet 2013. 714–717 (2013)
19. Mehr, H. : Artificial Intelligence for Citizen Services and Government. Harvard Ash Cent. Technol. Democr., Technical Report 19 (2017)
20. Zuiderwijk, A., Chen, Y.C., Salem, F.: Public Governance in the Age of Artificial intelligence: State of the Art, a Research Agenda and a Call to Action. Submitted to *Gov. Inf. Q.*
21. Castro, D., New, J.: The Promise of Artificial Intelligence. Technical report, Center for Data Innovation (2016)
22. Kuziemski, M., Misuraca, G.: AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 44, 101976 (2020).
23. Mehr, H. : Artificial Intelligence for Citizen Services and Government. Technical report, Asch Center for Democratic Governance and Innovation, Harvard Kennedy School. (2017)
24. Finland : Finland’s Age of Artificial Intelligence: Turning Finland into a leading country in the application of artificial intelligence. Technical report, Ministry of Economic Affairs and Employment, Finland. (2017)
25. Finland : Work in the age of artificial intelligence: Four perspectives on the economy, employment, skills and ethics. Technical report, Ministry of Economic Affairs and Employment, Finland. (2018)
26. Tsai, C., Lai, C., Chao, H., Vasilakos, A.: Big data analytics: a survey. *J. Big Data*, 2, 21 (2015)